Assessing and Monitoring Technical Performance in Youth Football Players: Design and Validation of the Technical Performance Assessment (TEEM)

Stuart Macfarlane

A thesis submitted in partial fulfilment of the requirements of Liverpool John Moores University for the degree of Professional Doctorate of Applied Sport & Exercise Science (DSPORTEXSci)

March 2021

Table of Contents

ABSTRACT

An important component in many modern professional football clubs is their academy structure. Professional football clubs invest substantial amounts of revenue into their academies to discover and develop potential talents which possess the necessary attributes to succeed in the first team or beyond. One important aspect in talent identification and development (TiD) is the assessment and monitoring of technical performance. Research investigating technical performance has been somewhat under-represented compared to other aspects of performance. In recent years, modern technology has facilitated the data collection process at the professional level and now there exists an abundance of data relating to player and team technical performance. However, despite the availability of this data at the professional level, this technology has not yet filtered down to most academy levels due to financial and operational constraints. Traditionally, the player appraisal and evaluation processes are subjective in nature, (i.e., at the academy level) and are predicated on the subjective opinion of coaches with little or no objective data to support their assessment. The availability of objective technical data would provide a more data-driven and in-depth assessment to support subjective coach opinion and in turn aid player development. To date, no 'gold standard' method of assessing technical performance in youth footballers has been globally accepted. This thesis therefore aims to develop a contextually relevant, valid and reliable tool for assessing and monitoring technical performance in youth football players, which will offer a contribution to the field to enhance our understanding.

Study 1 (chapter 3) attempts to establish content validity of the Technical Performance Assessment (TEEM) and provide a framework for tool design. This study aimed to determine what technical attributes were perceived to be most important for success at the elite level. The study adopted a qualitative Delphi method which required data collection by means of multiple rounds of questionnaires. Following each round, responses from participants was filtered down, summarised and presented back. This cycle was repeated until consensus was reached between participants. Participants in this study were the most highly qualified and experienced coaches from a Scottish Premier League academy. The next stage of this study involved designing a contextually relevant tool based on the results of the Delphi process and by adapting and applying various components of previously developed assessment tools.

Study 2 (chapter 4) involved establishing the measurement properties for tool validation. This study attempted to establish inter- and intra-observer reliability, test/retest reliability, typical

error (TE) and the smallest worthwhile change (SWC). The purpose of this study was to test the TEEM's reliability and establish associated measurement error to facilitate the interpretation of 'true' changes in performance. Results revealed moderate to good reliability in 2 of the 4 selected key performance indicators (KPIs), poor reliability in 2 of the 4 selected KPIs and a wide variation of reliability strengths in sub-contexts of these KPIs.

Study 3 (chapter 5) involved establishing criterion-based validity. This study used a correlation analysis to investigate the relationship between performance in the TEEM protocol and performance in competition. Results revealed trivial-strong correlations in the four main KPIs with a wide range of correlation strengths in the sub-contexts of these KPIs.

Study 4 (chapter 6) aimed to establish the tools sensitivity to longitudinally monitor changes in performance over time and explore the influence of maturation on performance. This study involved measuring performance in the newly developed TEEM over a period of 12 months. Furthermore, the study compared differences in performance between players at different stages of biological maturation. Results revealed that the TEEM lacked the sensitivity to identify changes in performance over a 12-month period. However, results revealed that stage of maturation was a significant predictor of performance and discriminated between players at earlier stages of maturation compared with players at later stages of maturation.

In summary, the results presented throughout this thesis demonstrate the difficulty and complexity of monitoring technical performance in football. Due to the random and unpredictable nature of the game, technical performance is associated with substantial variation between observations. The TEEM developed in this study offers one possibility for assessing skill proficiency which can be easily applied within an academy environment. For its ease of application, feasibility and ability to discriminate between players of stages of maturation, the TEEM offer's a welcome addition to a club's assessment protocol where there is often very little or no objective data to support coach opinion, however, the tools limitations must be taken into consideration prior to administration. In addition to the research carried out throughout this thesis, individual professional aims were outlined at the beginning of the process and the journey through which these aims were achieved is intertwined throughout. The professional doctorate process enabled simultaneous researcher and practitioner development which was invaluable for professional development.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank my supervisor Dr Allistair McRobert who has guided me through this process. My professional doctorate journey has been a challenging one and Allistair has provided me with the support and guidance I required at every stage along the way. Allistair's enthusiasm, passion and expertise has motivated me at times when mines lacked.

I would also like to thank Professor Barry Drust for providing me with the opportunity to enrol on the professional doctorate programme and also for providing support and advice along the way.

Lastly, I would like to thank my employers Montrose Football Club for providing the financial support to fund my professional doctorate and St. Johnstone Football Club for granting me access to their participants and data.

CHAPTER 1: General Introduction

1. General Introduction

The purpose of any Professional Doctorate programme is to facilitate the development of a researching practitioner through the combination of experiential knowledge, whilst subsequently developing professional practices and competencies that are essential within an applied environment (Fulton et al, 2012). The Professional Doctorate programme should induce a deeper analysis of an individual's current professional competencies that in turn provides a structured and detailed professional development plan, whilst simultaneously presenting relevant questions specific to the practitioner's research area. The structure of the Professional Doctorate programme can be split into two distinct learning outcomes: researcher development and professional development. Therefore, the remainder of this introduction will outline the aims and objectives from both a research and a professional perspective.

1.2. Research Background

Modern football has experienced an exponential surge in the growth of sports science research aimed at enhancing sporting performance over the previous two decades (Drust, 2019). Consequently, this has contributed to the evolution of performance over a prolonged period (Barnes et al, 2014). Whilst, the field of sports science encompasses a wide variety of disciplines, a significant proportion of the research has focussed on athletic development and the monitoring of training and competition status (Drust, 2019). The generation and application of this research has undoubtedly made a significant and positive contribution to the development of professional football, however, there remains many important aspects of performance that are unexplored (Kirkendall, 2020). One of these aspects, which is underrepresented within the scientific literature, is technical performance and how it influences match success (Ali, 2011). Of the new and existing research available on the role of technical proficiency in performance, it has been demonstrated that superior technical performance is associated with overall team success (Rampinini et al, 2009, Filetti et al, 2017). Rampinini et al (2009) reported that 'more successful' teams completed a higher number of passes, successful passes and shots on target compared with 'less successful' teams as determined by their league ranking. Furthermore, Filetti et al (2017) reported an increased probability of winning (125%) in teams which performed better in a skill efficiency index composed of various aspects of technical performance. Considering these findings, one would suggest that the development and monitoring of technical performance should be a critical element within both an elite level and youth development training structure.

The advancement of modern technology and its application within professional football has made in-depth analysis of technical performance commercially available and is now commonplace throughout many levels of world football (Filetti et al, 2017). Despite these technological advancements within the professional game, availability is restricted to most youth academy structures due to associated limitations such as financial restrictions, feasibility of technology application and operational contraints. Consequently, there remains a lack of objective technical performance data to aid the development process within youth development. Traditionally, performance appraisal within an academy setting has relied on subjective coach opinion (Sieghartsleitner et al, 2019). Whilst this information and tacit knowledge is valuable, issues around validity and reliability have been identified within the existing research suggesting that objective data is a necessary supplementary component of the subjective appraisal process (Bergkamp et al, 2018; Johannson & Fahlen, 2017; Sieghartsleitner et al, 2019). One of the main limitations to subjective player appraisal by coaching staff is coaches seem to have pre-conceived beliefs and (un)conscious biases towards what they perceive as 'talent' and rarely rely on generally accepted talent models resulting in high variability between observers (Johansson & Fahlen, 2017). One example of selection bias which is particularly prevalent within youth football, and a result of subjective coach assessment, is maturation selection bias (Hill et al, 2019). The concept of maturation selection bias is characterised by the selection of 'early maturing' players who possess superior motor skills and physical competencies, which are developed through the process of advanced biological maturation, ahead of their 'later' maturing peers at earlier stages in their physical development (Towlson et al, 2017). Players who showcase the best physical attributes in relation to their chronological age group during competition, are often perceived by coaches to be better performers than their later maturing counterparts due to the positive impact these players have on the overall performance of the team and subsequently results (Hill & Sortiriadou, 2016). The consequence of this type of selection based on physical attributes is perhaps reflected in the study by Ostojic et al (2014) who longitudinally monitored a cohort of 14-year-old youth footballers over an 8-year period through to adulthood. Ostojic et al (2014) reported that when the adult level attained by these youth football players was assessed and categorised as either 'elite' (top 5 European leagues) or 'non-elite', later maturing players represented 60.1% of the players competing at the elite level compared with just 11.8% being represented by early maturing players. Results of this study support the premise that the presence of maturation selection bias at youth level could potentially elicit far reaching consequences if 'late' maturing players who are also relatively younger, are overlooked and deselected from an academy programme and subsequently 'slip through the net' after being given insufficient time to develop. Therefore, objective feedback relating to the current (on-time), biological age-matched performance level of youth players provides valuable information for identifying later maturing players who are technically and tactically competent which will in turn will provide more comprehensive information to support subjective coach opinion and the decision-making process during selection/de-selection. In addition to this quantitative performance measure, objective technical performance data could provide additional contextual information regarding the longitudinal development of performance, through monitoring and establishment of benchmark data for the purpose of TID. The question remains as to which method of assessment/analysis is most accurate, feasible and cost effective within an academy environment.

Previous research in the field of technical performance assessment and monitoring within youth footballers has focussed around three main methodological constructs: (1) 'closed skill' isolated performance tests, (2) measurement within competition, and (3) measurement within small-sided game protocols (Bergkamp et al, 2019). However, many of the research studies involved in the development of such protocols often fail to report essential measurement properties that relate to reliability and validity (Bergkamp et al, 2019; Robertson, 2017). Therefore, it is essential that any technical performance assessment tool developed for use within a practical setting has withstood the scientific scrutiny of establishing the appropriate measurement properties and adhered to the necessary framework for developing new measurement tools (Brewer et al, 2002). This in turn will ensure the validity and reliability of results during dissemination.

1.3. Research Aims and Objectives

The research aims and objectives of this thesis are as follows:

1. To design and develop a contextually relevant assessment protocol for assessing and monitoring technical performance in youth footballers.
2. To establish the reliability of the newly developed assessment tool. This will consist of three subsequent parts: establishing inter- and intra-observer reliability; establishing test/retest reliability and establishing associated measurement error and smallest worthwhile change for longitudinal monitoring.

3. To establish criterion-based validity of the newly developed measurement tool. This will explore the relationship between test performance in the newly developed assessment protocol and competition performance
4. To establish the assessment protocols responsiveness/sensitivity to change over time. This represents a key stage in test development for monitoring change in technical performance over time. A secondary objective for this section of the thesis will be to investigate the effect of biological maturation on technical performance.

1.4. Professional Background

Enrolment onto the Professional Doctorate in Applied Sport and Exercise Sciences (DSPORTEXSci) at Liverpool John Moores University (LJMU)occurred at the early stages of my professional career. I was fortunate to be involved in professional youth and senior football since the onset of my master's degree study and began my practitioner journey as a volunteer coach with my current employer. Since then, my career has followed a natural upward trajectory as depicted in Figure 1. This natural progression culminated in a full-time role as the Academy Head of Football Science and Medicine and facilitated my development as a practitioner due to the multi-disciplinary nature of the role. At this stage in my organisation's development, I was the first sports science practitioner to be employed by the club at the academy level. This in turn, provided me with the autonomy to design and implement a sports science curriculum which previously did not exist. For this I am grateful, as the knowledge I acquired from this process significantly broadened my understanding of the complexities of the holistic long-term development of young footballers. During the ongoing process of curriculum design and refinement, enrolment on the DSPORTEXSci was motivated by a research question presented to me by my Academy Manager, which I ultimately intend to answer within this thesis. I am optimistic that the knowledge gained, and professional competencies developed from the professional doctorate process will benefit both my organisation and me as an applied practitioner in the field of sport and exercise science.

1.5. Professional Aims and Objectives

One of the main benefits of this professional doctorate process is the simultaneous development of relevant professional competencies and doctoral level research skills. Prior to the commencement of the research process, an important step is to conduct a reflective self-evaluation of researcher,

Figure 1. Academic and career path to date

professional and behavioural competencies (Boud, 1995). Identification of individual strengths and weaknesses will facilitate the formation of professional aims and objectives to be achieved during the professional doctorate process. To engage in the process of self -evaluation, I utilised recognised and relevant industry tools: Vitae Researcher Development Framework (2005); the British Association of Sport and Exercise Scientist (BASES) Competency Profile and a Behavioural Profile Report (PDA International, 2004). Details of the full process can be found in the Training Plan (8004SPOSCI) illustrated in Appendix 1.

Completion of these processes highlighted the following areas for development:

1. To develop my research and technical skills
2. To improve my communication (specifically public speaking) and dissemination of information for various audiences
3. To improve my self-evaluation and critical reflection skills to highlight areas for continued professional development.

In addition to these main areas for development, I aim to achieve the following professional objectives:

1. To develop transferable knowledge, concepts and processes that can be utilised in other contexts

2. To practically apply knowledge and insights formulated from the process into my club's curriculum to aid player development
3. To engage in regular public speaking to disseminate research findings
4. To develop my ability to manage and complete multiple simultaneous projects
5. To develop a long-term professional development process through continued reading and engagement in continued professional development (CPD) events.

CHAPTER 2: Literature Review

2. Literature Review

2.1 Introduction

Scientific research within football has experienced a significant explosion in recent decades (Kirkendall, 2020). An exponential increase in the volume of research available to practitioners, coaches and players has contributed to increased physical and technical demands of modern football (Barnes et al, 2014). Research investigating facets of physical performance and its application to training and monitoring has dominated the literature, with the physical demands of performance being well documented (Reilly & Gilbourne, 2003). In comparison, only a small proportion of scientific research has investigated the complex and multi-factorial aspects of technical performance (Ali, 2011). This is somewhat surprising given the significance of technical performance in elite level football (Rampinini et al, 2009; Filetti et al, 2017). Barnes et al's (2014) longitudinal study spanning seven consecutive seasons in the English Premier League reported that players now perform 40% more passes, with the percentage of players with a pass completion rate of below 70% dropping from 26% in 2006-07 to just 9% in 2012-13. Furthermore, Dellal et al (2011) compared technical performance between two top European leagues and suggested that players should possess a minimum pass completion rate of 70% in order to meet the technical demands of their competitive standard. It should however be noted however that both studies lacked supporting contextual data (such as position-specific demands or behavioural habits) associated with the chosen key performance indicators, which may provide a deeper analysis of the technical performance demands (Yi et al, 2018).

This marked improvement in technical quality observed over an extended period highlights the importance of technical performance, which in turn emphasises the increased requirement for technical proficiency. Clemente et al (2019) observed that running performance at various speeds did not differentiate between teams when classified in terms of their final season ranking in the Spanish top division, La Liga. Furthermore, Di Salvo et al (2009) observed a superior high intensity activity output from bottom ranked teams compared with top ranking teams in the English Premier League (919 vs. 885m at velocities > 19.8 km/h respectively). Together, results presented in these studies suggest that physical performance may not play such a dominant role in discriminating between successful and unsuccessful teams. It is clear that players competing in these top domestic leagues should possess a high level of physical attributes in order to meet the demands of the modern game (Barnes et al, 2014), however, it is clear that other components of performance, such as technical performance could explain the disparity between successful and unsuccessful teams. For example, Rampinini et al's (2009) study of 416 elite level Italian football matches found that more 'successful teams', as judged by their final league rankings, completed:

more short passes; more successful short passes; dribbles; shots; shots on target and had more involvements with the ball than unsuccessful teams. In addition, a comparison between the best foreign and domestic players in the Chinese Super League reported that successful forward passes and shot success were discriminating factors, whereas no differences in physical performance were observed (Gai et al., 2019).

The summary of research presented here provides a foundation for further emphasis on the promotion of technical focussed training during player development. In support, Castillo et al (2018) and Matinez-Santos et al (2016) demonstrated that physical performance capabilities did not determine promotion from the last stage of academy player development (U19) to professional elite football in the Spanish Top Division over an 18-year period, with the exception of players playing the centre back position, where neuromuscular (speed and jump) performance was observed to be a competitive advantage. Therefore, consistent with the previously mentioned research regarding the lack of prognostic power of physical attributes for determining success in elite level football, these studies contribute to the formation of a stronger argument supporting a greater emphasis on the technical development.

One possible reason for the comparatively smaller volume of research studies investigating technical performance, especially within an academy environment, is that technical performance is difficult to assess due to its highly complex nature. Consequently, even with a number of developed measurement tools within previous research such as the Loughborough Passing Test (McDermott et al, 2015) and the Game Performance Assessment Instrument (Oslin et al, 1998), methodological concerns identified within these studies which relate to the establishment of appropriate validity and reliability measurement properties limit their application in a football academy environment (Bergkamp et al, 2019). The generation and availability of technical objective data for talent development purposes remains limited until the successful development of a measurement tool that withstands the scrutiny of methodological validation. Due to the significant importance of technical performance in modern elite level football, the author suggests that further investigation into the development of a new valid and reliable assessment protocol could provide invaluable objective data, which will aid player development. The following section will review the current literature surrounding the measurement of technical performance.

2.2 Current Literature Investigating the Measurement of Technical Performance in Football

Association Football is the most popular sport in the world and attracts global media attention (FIFA, 2018). Many young players aspire to be elite professionals and compete with the world’s top teams. Similarly, the world’s top teams aspire to develop the best young talent through their academy systems to compete on the international stage (European Club Association, 2018). Clubs invest substantially in staff and infrastructure to create the highest quality environments for developing talent. If clubs can develop and maintain strong academy structures and practices, talent development can be a profitable business model capable of generating substantial income, making academy systems fully self-sustainable (KPMG, 2020). The process of talent development is highly complex and multifactorial and can be separated into two distinct processes - talent identification and talent development (Vaeyens et al, 2008). In recent years, there has been growing interest within the scientific community in developing new research to aid the identification and development process, such as the design of foot-mounted inertial measurement systems for tracking technical and physical actions during training and competition (Waldron et al, 2020). Within the existing scientific literature, it is important to distinguish the nuances between the two processes as they both involve different methodological constructs. From this point onwards, talent identification will be defined as "the process of recognising current participants with the potential to excel in a particular sport" and talent development will be defined as "providing the most appropriate learning environment to realise this potential" (Vaeyens et al, 2008, p 703). Recently, the need for longitudinal study designs within the context of TID have been highlighted (Lehyr et al, 2018; Sieghartsleither et al, 2019). Furthermore, recent research teams have highlighted the requirement for developing appropriate assessment protocols that better reflect competition performance and have better prognostic capabilities for long term player development (Bergkamp et al, 2019). Research investigating the assessment of technical skill proficiency in sports performance has been evident in literature for decades (Oslin et al, 1998; Grehaigne, 1997). Research teams have used various methods to pioneer tools used for the assessment of technical performance, which are applicable to one or more sports. When selecting a performance test for assessing technical performance, practitioners should understand the strengths and limitations of the two main types of protocols commonly utilised within the scientific research.

One frequently used method of assessing technical performance is through the use of 'closed skill' assessment protocols (Lehyr et al 2018; Zuber et al, 2016). 'Closed Skill' protocols involve the execution of a specific football action (such as passing, dribbling, shooting etc.) under test conditions that measure skill accuracy and/or speed of execution. These tests are typically isolated from the competition context and performed as an individual. Two of the most common protocols used within original research articles are the General Soccer Ability Skills Test (Mor & Christian, 1979) and the Loughborough Passing and Shooting Tests (Ali et al, 2007). Both tests primarily assess aspects of football specific technical performance by isolating a specific part of it (i.e., passing or dribbling) and design specific test parameters to measure accuracy and speed of execution. Overall test performance is measured by completion of the task in as little time as possible with a low number of errors. Previous research has demonstrated that these assessment tools are capable of successfully discriminating between elite and non-elite players (Borges et al, 2017; Huijgen et al, 2013; Keller et al, 2016; McDermott et al, 2015), and longitudinal tracking of changes in performance capabilities (Lehyr et al, 2018; Zuber et al, 2016), however they lack criterion-based validity (Roberts et al, 2019; Rubajczyk & Rokita, 2015).

In a position paper by Bergkamp et al (2019), a significant methodological issue identified within the existing body of research relates to the 'fidelity' of the predictors of performance. The fidelity of specific assessment protocols designed for measuring performance is defined as "the extent to which the performance task mimics the criterion behaviour in content and context" (Bergkamp et al, 2019 p. 1327). Based on this analysis, Bergkamp et al (2019) proposed a scale on which performance assessments are rated between low and high fidelity, where high(est) fidelity relates closely to the criterion method (in the case of technical performance assessment - competition performance). One major limitation of tests that assess skill in a 'closed' situation is their context in relation to actual game performance (low fidelity). Technical performance in football is described as a process which requires communication with surroundings (information gathering), decision making (what to do) and the execution of a skill (Praca et al, 2015). By isolating a specific skill during such a test, the external focus of attention is removed, thereby eliminating one of the vital components required during competitive match play. For example, Praca et al (2015) reported that Brazilian youth football players performance in 'closed skill' tests correlated poorly with performance in small-sided games (ICC = -0.252 – 0.367). Furthermore, Rubajczyk and Rokita (2015) reported a weak correlation (ICC = -0.325 - 0.452) between performance in soccer specific tests and performance in a small-sided game situation. In the study by Praca et al (2015), the General Soccer Skill Ability Test was used as the method of technical performance assessment. This test involves the football actions of dribbling, passing and shooting. For dribbling performance assessment, the participant is required to dribble the ball through a pre-determined track as fast as possible. In the assessment of passing and shooting, participants are required to accurately pass or shoot the ball at targets within an allocated timeframe. Similar to the dribbling test utilised by Praca et al (2015), Rubajczyk and Rokita (2015) adopted a similar 'soccer specific' skill assessment which included a dribbling test with a turn and had to be completed as fast as possible with minimal errors. Both studies concluded that disparity exists between performance in a 'closed skill' technical test and performance in actual match play and that the perceptual and cognitive demands of competition require both internal and external focus for successful skill execution. Therefore, practitioners should exercise caution when deciding to include such a test within a test battery due to validity concerns and additional time commitments required to carry out a high number of assessments in an academy environment. The time commitments required to administer such a test battery should be evaluated against the value of the information gained. In conclusion, it is pertinent that the development and utilisation of an appropriate assessment tool that closer replicates match performance would provide more valuable information to coaches, practitioners and players.

As research investigating the measurement of technical performance in football has evolved over time, the use of 'game-related' assessment protocols (high fidelity) have gained more interest within the literature. This method of assessment involves the assessment of technical performance within a game-like scenario or within competition itself. The concept of assessing technical performance within game context has been previously investigated in earlier research. Two of the earliest models were developed by Oslin et al. (1998) and Grehaigne (1997) and were named the Game Performance Assessment Instrument (GPAI) and the Team Sport Assessment Procedure (TSAP). The GPAI and TSAP both adopted an observation and hand notation system through performance analysis to assess game performance in real time. Oslin et al (1998) reported that the GPAI can discriminate between high- and low-level performers (ES = 0.23-1.93) and both studies demonstrated acceptable levels of intra- and inter-observer reliability (ICC = 0.73-0.97). One limitation apparent in both studies lies within its applicability to 'elite' or professional level football. The coding system developed in these studies was taken from the perspective of general team sport games. It is possible that this reduces the value of the information presented in the context of specific football performance. For example, both studies report the success rate of technical actions (such as passing), however, there is no context to provide a more detailed analysis (such as direction of pass, where on the field, and length of pass), which is of great significance when analysing technical performance at the elite level, especially in the case of talent development.

Another limitation to both these studies (Grehaigne, 1997; Oslin et al, 1998) is the observation system employed within the research methods. Both studies rely on the successful and accurate recording of key performance indicators by observers using the naked eye in real time. It is possible that due to this observation method, actions may be missed or incorrectly interpreted due to the subjective nature of the coding process (Memmert & Harvey, 2008). James et al (2005) identify various potential sources of measurement error. These sources of error include events that are coded incorrectly due to coding mistakes made by the observer or ambiguous operational definitions; issues with player identification if the observer is unfamiliar with test participants and events that are missed in the coding process. Furthermore, a further example in the case of the GPAI is that each observer is instructed to observe a player for a 10-minute period during a small-sided game situation and manually record the number of decisions made, skill executions and supporting movements. It is possible that actions recorded during this short time period do not provide a complete reflection of player behaviour as 10 minutes in the context of a full match could be considered only a small cross-sectional snapshot. Several steps can be taken to minimise the potential risk of measurement error such as video capture for subsequent analysis; however, it is impossible, if not, extremely difficult to eradicate it entirely. In light of this, it is of critical importance that practitioners establish any measurement error associated with the selected performance test to allow for meaningful changes in performance to be identified (Robertson et al, 2014).

Following the work of Oslin et al (1998), more recent studies adopted the methodological approach of video recording small-sided and competitive games for subsequent coding and analysis (Cobb et al, 2018; Clemente, 2014; Garcia-Lopez et al, 2013; Moreira et al, 2017; Walron & Worsfold, 2010). Video recording games and analysing post-event drastically reduces the potential for human error compared with recording data in real-time. Despite this, there are several common limitations evident that need to be considered. Firstly, these studies were designed for talent identification and report successful differentiation between elite and non-elite players, therefore do not take into consideration the use of the tools for talent development. Although this may not have been the aim of the studies in question, the assessment tools developed may have more practical relevance if they had the power to longitudinally track player performance over time. Cross-sectional study designs are a common theme throughout the literature, and it is apparent that future research is required into the development of an assessment tool for performance monitoring. Secondly, no study performed test/retest analysis to assess reliability and variance in performance between two assessment procedures. Due to this limitation, it is difficult to say with confidence that the assessment tools provide a reliable data set. Thirdly, within the previous research studies there is a wide variety of coding procedures implemented which range from a general, simplistic team sports system to complex, football specific systems.

Moreover, only Gracia-Lopez et al (2013) attempted to establish content validity for the development of their coding framework by seeking the opinion of physical education teachers and coaches. Despite the in-depth coding framework developed by Garcia-Lopez et al (2013), it still lacks contextual information around the technical actions being assessed, which has been identified as an important aspect of technical performance (MacKenzie & Cushion, 2013). For example, Garcia-Lopez et al (2013) code pass completion success rate but do not take into consideration where on the field the pass was played from or what direction the pass was played. Therefore, it is possible that a player has a high pass completion rate but habitually opts for 'safe' passes as opposed to penetrating passes. This is a common limitation to the coding frameworks developed in previous technical assessment tools. Lastly, reporting the technical actions as a percentage of success results in greater variability between trials due to the unpredictable nature of team sports, especially when the number of technical actions are low (Memmert & Harvey, 2008). Memmert and Harvey (2008) suggest that a method to overcome this issue is to implement a 'performance index' equation which demonstrates less variation between performances. The performance indices suggested by Memmert and Harvey (2008) can be used to minimise variation in performance scores due to the influence one successful or unsuccessful skill execution has on the overall performance score in 'percentage success' method vs. the 'performance index' method (Memmert & Harvey, 2008). For example, in using the hypothetical situation whereby player A completes 3 out of 4 passes vs. player B who completes 2 out of 4 passes, the corresponding performance score for passing using the 'percentage success' method would be determined as 75% and 50% respectively. The argument being presented by the author of this thesis is, can we be confident that player A is 25% more efficient in skill execution compared with player B with a difference of just one pass? When adopting the performance index model proposed by Memmert & Harvey (2008) and using the same hypothetical example outlined above, the corresponding performance scores would be 54.2 and 50 respectively and therefore the associated variation is of a lesser magnitude and consequently reflects our confidence in the true difference in performance.

After considering the strengths and limitations within the previous research presented in this section, the author suggests that by manipulating the parameters of previously developed assessment tools and combining the best aspects of each, future research should investigate the development of a new, valid and reliable, football-specific assessment tool with 'high fidelity', so that it provides contextual performance data that can be effectively implemented and practically applied in an academy environment.

2.3 Development of New Measurement Tools

Football is an unpredictable sport with technical and physical actions occurring randomly as game play develops. This unpredictable nature means that assessing the performance characteristics of players becomes difficult due to the multi-factorial determinants of performance (Hughes & Bartlett, 2002). Only recently has there been an increased emphasis on the provision of scientific process to aid the development of technical performance. Methodological limitations to existing research designs mean that no 'gold standard' instrument for assessing technical performance have been accepted to aid the coaching process (Robertson et al, 2014). The scientific literature describes a multitude of instruments and methodologies utilised for the assessment of technical performance in football (e.g. Borges et al, 2017; Moreira et al, 2017; Waldron & Worsfold, 2010). After reviewing the various instruments adopted or designed throughout the scientific literature, two consistent limitations become apparent. Firstly, most research studies adopt a cross-sectional study design which take a 'snapshot' of performance in a given moment. Very few research studies have implemented a longitudinal study design which facilitates the development of performance rather than assesses it at one particular time point. Secondly, a significant number of research teams demonstrate little or no evidence of appropriate methodological components of the scientific process required for designing and validating new tests or assessments (Brewer & Jones, 2002; Robertson et al, 2017). Hughes et al (2002) reported that 47 of 67 research studies investigating notational analysis for performance assessment did not demonstrate evidence of the necessary reliability and validity tests required to validate a measurement tool in scientific research. Until the rigorous process of validation is conducted on any given observation or measurement tool, the application of such a tool within future research should be approached with caution.

When designing or selecting performance tests for application within sports science, Robertson et al (2017) demonstrated the measurement properties that are considered as being of critical importance. These properties, as interpreted by a panel of experts within the field of sports science, can be seen in figure 2. It should be noted that the development of tests for scientific enquiry within the field of sports science should follow the same systematic process for developing valid and reliable tools as in any other scientific discipline. Brewer and Jones (2002) proposed a 5-stage process for establishing contextually valid observation tools for assessing performance. The first stage proposed by Brewer and Jones (2002) involves observer training using a previously designed tool as a framework. The purpose of this stage is to familiarise observers with the analysis process which will subsequently enhance reliability. The second stage of the process involves amending an existing instrument to the context of the intended environment (e.g., within a professional football academy). A key stage during this process is developing operational definitions for key performance indicators used within the analysis. This process establishes content validity. Stage 3 involves establishing face validity within the instrument by determining whether the test measures what it is perceived to measure. For example, does the test correlate with the performance variable we intend to measure? Finally, stages four and five involve establishing inter- and intra-observer reliability. A summary of operational definitions for the measurement properties previously mentioned within this paragraph and to be addressed throughout the remainder of this thesis can be seen in table 1. When reviewing the 5-stage model proposed by Brewer and Jones (2002) and the significant measurement properties outlined by Robertson et al (2017), there appears to be several nuances between the stages. Therefore, it could be suggested that by incorporating additional stages, such as re-test reliability and testing sensitivity, the robustness of any tool design could be further enhanced.

From a talent development perspective, the importance of developing valid and reliable tools for evaluating and monitoring technical performance is well recognised amongst coaches and practitioners (Aquino et al, 2017). The availability of objective data to provide more information on player attributes can only be seen as a positive reinforcement to aid the decision-making process around player retention or the refinement of individual player learning objectives. As previously highlighted, numerous methods of developing tools for assessing technical performance have been utilised. The methodological processes of selected research studies can be seen in table 2. The research studies presented within table 2 were selected as they were empirical based models that could be adopted as a potential framework for developing a new assessment protocol as identified in stage 1 by Brewer and Jones (2012). From table 2, it is clear that all studies report methods of inter and intra-observer reliability measures. However, a significant limitation to these studies is the absence of test/re-test measures. Of the six studies selected, only two employ test/ re-test measures. When utilising observation tools for talent development, it is imperative that the smallest worthwhile change and the test sensitivity to change is identified. In terms of validity measures, four of the six studies selected incorporated content validity measures. Furthermore, the determination of associated test variation, specifically within-subject variation, and the magnitude of this variation will enable practitioners to see through the 'noise'. With this mind, an appropriate measurement property for establishing within-subject variation is typical error (TE). Swinton et al (2018) describe the typical error as the 'standard deviation for repeated tests' and is necessary during test/retest measurements. Establishing content validity ensures the key performance indicators used within the measurement tool are contextually relevant (James, 2006). Contextually relevant content is important as it ensures the measurement tool is generating meaningful and important data which is comparable to the criterion. For example, is the number of backward passes completed an appropriate key performance indicator to assess when 5 out of 5 highly experienced and qualified coaches do not perceive this action as being important or even relevant? Three of the six studies reported measures of discriminant validity which measures "the extent to which results from a test relate to results on another test which measures a different construct (i.e., the ability to discriminate between dissimilar constructs)" (Robertson et al, 2017, p.6). From the studies reviewed, it is clear that all design attempts have their individual strengths and limitations and collectively, the work done here provides a solid foundation upon which future design attempts can be further developed. It should be mentioned however, that the context in which the assessment tool is used will determine the extent to which methodological robustness is required. From the perspective of talent development within a professional football academy, where decisions are made based on the

Table 1. Table 1 shows the operational definitions for the measurement properties discussed in this and subsequent chapters in this thesis. All definitions were citied from Robertson et al (2014)

MEASUREMENT PROPERTY	OPERTATIONAL DEFINITION
Reliability	
Test-Retest Reliability	The consistency of performer(s) scoring over repeated rounds of testing.
Intra/Inter-Observer Reliability	Inter-observer: level of agreement between scoring/assessing when undertaken by two or more observers Intra-observer: defined as the agreement among two or more trials administered or scored by the same observer
Validity	
Content Validity	How well a specific test measures what it intends to measure. Do the items included in the test cover the entirety of those relevant to assessing a particular skill outcome measure?
Criterion-Related Validity	The ability of a test to show good agreement with an external measure of performance
Discriminant Validity	The ability of the test to discriminate between performers of different abilities
Generalisability	The extent to which the results can be transferred and used in other contexts
Responsiveness (sensitivity to change)	The ability of a test to detect worthwhile and 'real' skill improvements in its intended population between initial bout of testing and subsequent rounds
Smallest Worthwhile Change (SWC)	Information relating to the smallest meaningful change in performance

perception of player playing ability, it is of critical importance that information generated from any assessment tool is precise due to the highly competitive nature of 'elite' level academy football.

In summary, work completed by the various research teams previously highlighted should not be discredited but used as a framework to guide the development of any future models. By taking a concept from an existing model and further enhancing the robustness of its methodological processes, we can contribute to the continually growing body of research in the area of technical performance assessment within football. Ultimately, this will help provide supplementary, objective data to assist the coaching and decision-making process.

Figure 2. The measurement properties regarded as important when designing or selecting a performance test within sports science as determined by a panel of experts in the field. Level 1 properties refer to items which should be required for all performance tests. Level 2 properties refer to items which should be considered depending on the context of application (Robertson et al, 2017).

2.4 The Use of Performance Analysis in Football and its Relevance in Performance Assessment

Performance analysis in football has become a fundamental component in top level performance and player development (Carling et al, 2005). Performance analysis has been widely accepted by coaches, players and sports science practitioners as a method of providing factual, objective feedback to support the dissemination of player and team performance. It is commonplace within the sport in so that clubs invest and develop full-time performance analysis departments to improve player performance to gain a competitive advantage over their opponents (Carling et al, 2005). Furthermore, recent times have seen the emergence of academy category status (e.g., Elite Player Performance Plan - English FA; Project Brave - Scottish FA), whereby minimum standards are required to achieve accreditation. Performance analysis represents one aspect of performance

Table 2. A summary of reported measurement properties within a selection of studies which represent potential frameworks for the development of a new observational tool.

Study	Test/ retest Reliability	Intra-Observer Reliability	Inter-Observer Reliability	Content Validity	Discriminant Validity	Responsiveness/ Sensitivity	Smallest Worthwhile Change	Interpretability	Familiarity Required	Duration
Oslin (1998)	YES	NO	YES	YES	YES	NO	NO	NO	YES	YES
Grehaigne (1997)	YES	NO	YES	YES	YES	NO	NO	NO	YES	YES
Garcia-Lopez et al (2013)	NO	YES	YES	YES	YES	NO	NO	NO	YES	YES
Waldron & Worsfold (2010)	NO	YES	YES	NO	NO	NO	NO	NO	YES	YES
Moreira etal (2017)	NO	YES	YES	YES	NO	NO	NO	NO	YES	YES
Cobb et al (2018)	NO	YES	YES	YES	YES	NO	NO	YES	YES	YES

which is an essential requirement for accreditation, and in turn highlights the importance of this growing role within sport. Performance analysis is a dynamic discipline which can encompass a creative element in the way performance data is analysed and presented. For example, in a team context, match analysis can provide a detailed analysis of an opponent's tactics and playing style which can subsequently influence the structure of a training week in the lead up to the game. From an individual player perspective, performance analysis can provide a detailed breakdown of skill success via notational analysis or a review of a player's positional habits to aid development.

Recently there has been an interest in the utilisation of performance analysis for the purposes of TID (Van Maarseveen et al, 2017; Thomas et al, 2009; Waldron & Worsfold, 2010). Traditionally, player feedback within a youth academy environment has centred around the subjective opinion of academy coaches based on observation with no or little objective evidence to support their player appraisal (Nicholls & Worsfold, 2016). This subjective method of player appraisal increases the risk of bias towards level of player performance, which can be susceptible to judgement based on individual coach philosophy and beliefs. Nicholls and Worsfold (2016) demonstrated that during competitive match situations, experienced 'elite' academy coaches were only capable of re-calling 38.8% of technical actions, the majority of which occurred around key incidents such as goal scoring opportunities. Furthermore, Nicholls and Worsfold (2016) reported that a pre-determined bias towards certain players from coaches based on their own opinion subsequently under or over-estimated skill proficiency. Results suggest that for player development to be stimulated and improved, development of a reliable process whereby accurate appraisal of performance can be assimilated is required.

One method of performance analysis used for talent development from a skill proficiency perspective in football is notational analysis (James, 2006; Thomas et al, 2009; Van Maarseveen et al, 2017). "Notational analysis is primarily concerned with the analysis of movement, technical and tactical evaluation and statistical compilation" (Hughes & Franks, 2004, p59). In modern times, it has become possible to retrospectively analyse performance through the application of video capture, which has significantly improved reliability compared with previously adopted hand-notation systems used during live performance (Grehaigne, 1997; Waldron & Worsfold, 2010). Despite the potential usefulness of notational analysis for assessing and monitoring technical performance in football, several methodological limitations have been identified which need to be addressed before adopting or designing an appropriate assessment protocol (James, 2006; MacKenzie & Cushion, 2013). Firstly, the study sample size utilised during the design of the assessment protocol must be adequate to establish any variation across multiple testing events. This will ensure the test is robust enough to distinguish between true change in performance and change in performance due to measurement error when monitoring performance over time (signal vs. noise). Secondly, it is essential that operational definitions for the coding procedure are clear and published within the research to ensure the assessment protocol can be easily repeatable over multiple trials. Thirdly, contextual information about performance or performance variables being measured (such as pass direction, direction of first touch or type of ball manipulation) must be considered to provide a more in-depth insight into performance capabilities. Lastly, previous research has primarily adopted a cross-sectional study design. A longitudinal study design will improve validity as this will ascertain whether the test procedure is sensitive enough to track change in performance levels over time. Despite these methodological limitations, notational analysis is a promising method of monitoring technical performance, however, research employing this process for talent development is limited and requires further investigation.

Thomas et al. (2009) and Van Maarseveen et al (2017) both employed notational analysis to analyse technical performance in football with acceptable inter and intra-observer reliability scores ($r = 0.987 - 0.997$). Furthermore, both studies provided detailed operational definitions of coding events and procedures to allow for a repeatable assessment protocol. Whilst both successfully provided objective feedback for player development, the methodological approaches adopted within studies varied significantly. Van Maarseveen et al (2017) elected to use a simulated 3v2, attack v defence situation whereby various attributes such as dribbling and passing were assessed. One methodological concern with analysing performance in this situation is that this only represents one phase of match play which could occur during competition. Furthermore, it could be questioned how relevant a 3v2 situation is in the context of actual competitive match play. Therefore, the criterion-related validity of this protocol could be questioned since assessing only one match phase may not be representative of real-life competitive match play performance. In addition, the sample size was low, which has been identified as a common limitation within notational analysis research (James, 2006). In contrast, Thomas et al (2009) implemented a technical performance analysis taken from competitive match play. It is important to consider potential limitations when providing regular on-going player assessment using full competitive match scenarios. The first to consider is the variability in performance during competition due to external factors such as team tactics or opposition quality, which may influence on-field behaviour (Hughes, Evans & Wells, 2004). Fernandez-Navarro et al (2018) described substantial variation in team playing style depending on three influential contextual factors relating to the specific match. The authors of this investigation reported that match status, venue and quality of opposition significantly influenced team playing style and subsequently accounted for high match-to-match variation. For example, if a team is instructed to play a specific way, the natural behaviour of the individual player may be influenced. This in turn could alter a player's habitual behavioural habits and/or skill execution proficiency. Secondly, the time and labour demands required to analyse full match performance could be considered a significant limitation. For example, academies need to analyse multiple players over a short period of time which might not be practically feasible. Therefore, the assessment protocols need to be effective and efficient for them to be integrated into practice. Lastly, the number of technical actions performed during a large-sided game compared with a small-sided game has been shown to be significantly lower (Owen et al, 2013). Liu et al (2016) reported that certain technical actions display a high level of match-to-match variability, which makes it difficult to interpret player match performance and specifically performance of these actions. Therefore, after considering the results of the research investigating variation in competition performance, a more stable and repeatable situation should be considered for assessing and monitoring technical performance whereby a high number of skill repetitions can be performed. With this in mind, perhaps a small-sided game format would provide a better opportunity to analyse a higher number of technical actions in a shorter time period. A higher number of technical actions would allow for the key performance indicators to stabilise thus making them more sensitive to change. A common limitation identified in research (Thomas et al 2009; Van Maarseveen et al, 2017), is the absence of a test/retest reliability procedure meaning it is impossible to determine the magnitude of variation in performance between trials (Robertson et al, 2014), and consequently the ability to distinguish between 'signal and noise'. Furthermore, a test/retest protocol would allow the establishment of appropriate measurement properties regarding meaningful changes in performance such as the smallest worthwhile change (SWC). This lack of information renders it difficult to determine if player improvement has occurred across multiple testing events.

In summary, it is evident that using notational analysis to analyse technical performance provides valuable objective information for facilitating player development. However, several limitations must be addressed to improve the robustness of the procedure: 1) studies should employ a large sample size to improve reliability; 2) studies should include a test/retest procedure to determine variation across multiple trials; and 3) the assessment format must establish criterion-related validity and take into consideration the multi-factorial facets of performance such as time and space demands, phase of game being assessed (e.g., full game, 3v2) and the number of technical actions per assessment.

2.5 Longitudinal Monitoring and Assessment of Performance

Procedures used to assess and monitor skill within talent development in youth football have been under-represented in scientific research compared to studies investigating physical performance (Bergkamp et al, 2019). Furthermore, studies which longitudinally track improvements in technical performance over time are scarce. From a physical perspective, player development has been well-researched with several models being proposed. One recognised and well cited model is the Long-Term Athlete Development Model (LTADM), which proposes various stages, or 'windows of opportunity', for developing athletic potential (Balyi & Hamilton, 2004). In comparison, little research investigating how technical skill or motor performance develops longitudinally throughout childhood and adolescence has been conducted. The Development Model of Sport Participation proposed by Còtè et al (2007) outlines three defined pathways of sport participation and the development of sport-specific skill. At the end of each pathway lies an alternative potential outcome and resulting level of adult sport participation. The activities, and time spent in each of these activities, performed during each phase in the chosen pathway ultimately determines the adult level attained which is either 'elite' (or professional) performance or recreational participation. Ford et al (2012) described the developmental activities through early childhood and adolescence of elite youth football players from seven prestigious footballing countries. The authors presented the average number of hours spent in practice, play and competition relative to each individual country using retrospective memory recall. One consistent finding among all countries, is that early childhood is dominated by playful sport-specific activities in comparison to structured practice and competition. However, although this research provides valuable information about the developmental activities from early childhood through to late adolescence in 'elite' youth footballers, it does not provide any objective information about how sport-specific skill develops throughout this time. Research in this area would undoubtedly provide further valuable information regarding the patterns of skill development, which in turn, could assist in long-term player development through the generation of benchmark data.

One reason for the disparity in objective longitudinal skill development research may lie in the fact that technical skill is difficult to objectively measure due its complex and multi-factorial nature, especially in team sports. Ultimately, the main aim of any professional academy is to develop footballers who can perform in its first team. Castillo et al's (2018) longitudinal study, demonstrated that physical performance capabilities did not significantly differ between players at the last stage of their development (reserve team level). They concluded that technical ability was one important determining factor which separated elite and non-elite players, highlighting the importance of technical performance and its influence on high-level competition. Therefore, it is imperative that professional academies: 1) are aware of the technical ability demands required to be successful at the highest level; and 2) understand how technical ability develops from childhood through to late adolescence and beyond. This knowledge could allow academies to develop technical performance and develop benchmarking data for longitudinal monitoring and evaluation in youth footballers. In addition, by identifying the technical demands of elite competition, academies can make informed decisions about specific, age-appropriate coaching content within structured curriculums thus guiding individual player development.

Although research investigating the longitudinal development of technical performance is scarce, several studies have examined this utilising populations specific to professional football academies (Honer & Votteler, 2016; Huijen et al, 2013; Huijgen et al, 2010; Lehyr et al, 2018; Zuber et al, 2016). These studies demonstrated that when comparing elite to non-elite players (or selected vs. non-selected), as judged by performance level reached following the last measurement point (e.g., national, regional, local), the rate at which both elite and non-elite players developed over time followed similar patterns (figure 3). These findings could suggest that skill levels attained prior to the U12 age group could be a determining factor in subsequent skill development. This theory is supported by Deprez et al (2015) who reported that dribbling skills at the U10 and U12 age-group discriminated between future club players and players who dropped out of the game. Furthermore, longitudinal studies demonstrate that 'elite' players performed better across all measurement points in isolated, technical performance tests compared with 'non-elite' players (Honer & Voetteler, 2016; Huijgen et al., 2013; Lehyr et al., 2018). In summary, the research suggests that a higher skill level at the first measurement point (U12) could correspond to a higher skill level, and thus higher playing level, attained at the final measurement point (U15+).

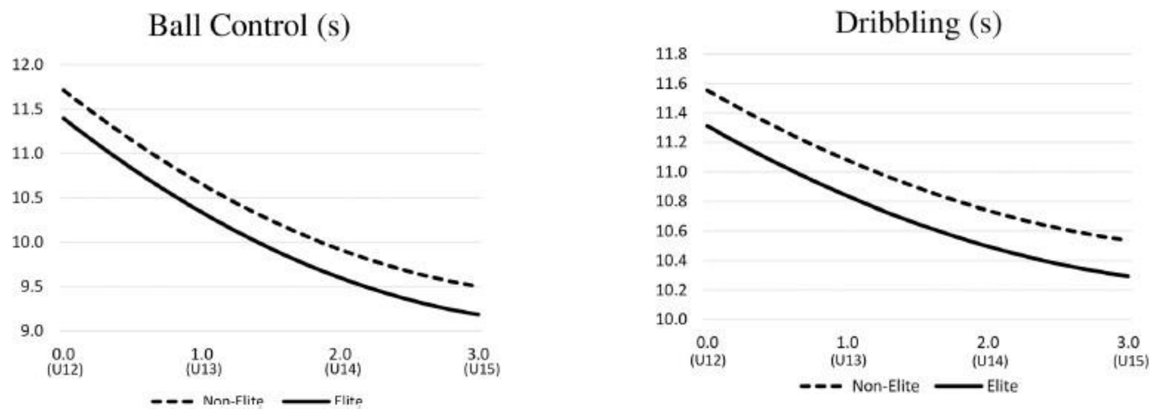

Figure 3. Players' motor performance development from U12 to U15 separated by adult performance level. Note. The x-axis represents the time (in years) from the first measurement point in U12 (time = 0) to the last assessment in U15 (time = 3) (Lehyr et al, 2018).

Although these results provide novel information regarding longitudinal skill development in youth football players from childhood through to adolescence, there are several limitations to consider. Firstly, there is no direct correlation between junior and adult success reported. For example, a longitudinal study of German Youth National Team players reported that the retention of players selected for the U16 National Team level through to U19 was only 33% despite players being identified through the selection procedure and monitored from the U11 age group (Gullich, 2014). Furthermore, Lehyr et al (2018) reported that performance in selected technical skill tests over a 3-year development period did not correspond significantly with the adult performance level attained, regardless of how high a playing level was achieved during youth level (e.g., international, national, regional). One explanation for this may be related to the technical test selection within the research methodology (Honer & Votteler, 2016; Huijen et al, 2013; Huijgen et al, 2010; Lehyr et al, 2018; Zuber et al, 2016). Research teams chose to utilise 'closed skill' tests such as the Loughborough Passing Skills Test and the Dribbling Slalom Test which isolate technical performance from the competitive environment of game situations and lack the perceptual-cognitive demands associated with match performance. In reflection, the authors concluded that these tests differ considerably from competitive match situations due to a lack of external focus during the tests such as opposition pressure and movement of teammates (Honer & Voetteler, 2016; Huijgen et al, 2013). Furthermore, authors suggest that future research should focus on the development of an assessment procedure which requires "more complex diagnostics that enable the assessment of a wider range of multidimensional and more representative characteristics" (Honer & Voetterler, 2016, p 2276). By developing such an assessment tool, practitioners and coaches will be able to effectively identify longitudinal skill development patterns which better reflect the demands of competitive match play for assisting player development.

As previously mentioned, one model proposed in relation to longitudinal motor skill and athletic development, is the Long-Term Athlete Development Model (LTADM) which has been widely investigated within scientific literature (Balyi & Hamilton, 2004). The model theorises certain 'windows of opportunity' for developing specific physical and motor skills in relation to an individual's Peak Height Velocity (PHV) (figure 4). The LTADM proposes that the trainability of motor skills and coordination tasks is enhanced between the ages of 8-12 – prior to the most rapid period of growth during childhood and adolescence (PHV). From the perspective of football specific motor skills, recent motor skill literature has demonstrated that proficiency in gross functional movement skills, such as horizontal jumping and catching, presented a strong correlation with game-specific motor skills in highly trained U12 youth footballers (Kokstejn et al, 2019). Authors concluded that the mastering of basic functional movement skills in early years development provides the building blocks for the subsequent learning of game-specific skills. The 'trainability' theory surmised by Balyi and Hamilton (2004), is supported in a longitudinal study by Fransen et al (2017) who reported an accelerated development of motor abilities prior to predicted age at PHV and subsequent plateau in Belgian youth soccer players. Following this rapid acceleration and subsequent plateau in motor abilities, it could be possible that the process of biological maturation plays an important role in some aspects of performance.

Research investigating the effect of growth and maturation on general and game-specific motor skills yields contradictory findings (Figueiredo et al, 2009; Kokstejn et al, 2019; Malina et al, 2005; Moreira et al, 2017; Vandendriessche et al, 2012). A brief summary of the studies investigating the effect of biological maturation on football-specific skill can be seen in table 3. Of the 5 studies referenced in table 3, only 1 reported a significant effect of biological maturation on football-specific motor skills. Moreira et al (2017) reported that early maturing players exhibited significantly more ball involvements and completed a higher number of passes and successful passes. However, one aspect of their design that should be considered is the assessment protocol selected. Moreira et al (2017) was the only research team who selected an 'open skill' as opposed to an isolated 'closed skill' performance test. Moreira et al (2017) selected a small-sided game-based

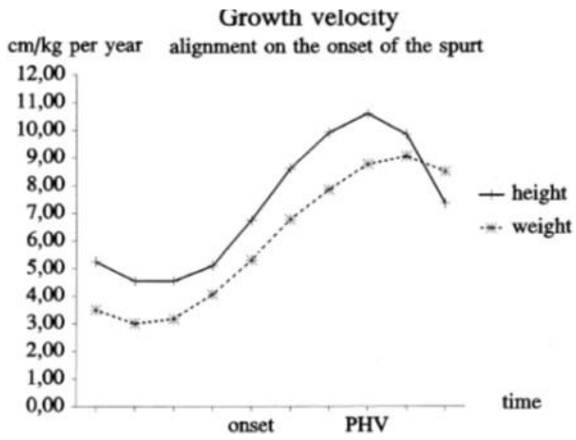

Fig 4. The rate of growth during childhood and adolescence. *PHV = Peak Height Velocity.* (Visser et al, 1998).

protocol and retrospective notational analysis of video recording data. Assessing performance in this manner may be considered more representative of competitive match performance, however, the small-sided games were not adjusted for biological age which would have resulted in a high between-subject variation in physical capabilities due to current stage of maturation. This in turn, could have resulted in more physically dominant players having more involvement with the ball and subsequently completing more passes. If Moreira et al (2017) had chosen to adjust for biological maturation, it could be possible that alternative results may been reported.

Despite this possible lack of association between maturation and football-specific skills, Visser et al (1998) reported a temporary decline in motor skill performance as assessed by the Movement Assessment Battery for Children (ABC) in a group of elementary school boys. Following this stage of 'adolescent awkwardness' experienced during PHV, motor skill performance again improved and continued to develop beyond pre-PHV levels into late adolescence and adulthood (Huijgen et al, 2010). Taking this into consideration, it is of critical importance that practitioners, coaches, and those involved in talent development understand the pattern of individual technical skill development. The cross-sectional study design of football-specific research does not allow for definitive patterns of skill development to be observed throughout PHV (Figueiredo et al, 2009; Kokstejn et al, 2019; Malina et al, 2005; Moreira et al, 2017; Vandendriessche et al, 2012). Furthermore, the between-age group comparison utilised in studies do not provide any information regarding the longitudinal, individual pattern of development. Until further research is conducted, which objectively demonstrates the individual pattern of development whilst adopting growth and maturation as an independent variable, we can only assume that football skill development follows the same path as other generic motor skills (Visser et al., 1998). Consequently, further research

Table 3. Table 3 shows a brief summary of the studies investigating the effect of biological maturation on football-specific motor skills.

STUDY	WAS THERE AN OBSERVED EFFECT OF MATURATION ON FOOTBALL - SPECIFIC SKILL?	ASSESSMENT PROTOCOL USED	PARTICIPANTS
Figueiredo et al (2009)	No	'Closed Skill' Tests – ball control with the body; ball control with the head; dribbling speed; dribbling with a pass; passing accuracy and shooting accuracy	Portuguese youth football players (Aged 11-15)
Kokstejn et al (2019)	No	'Closed Skill Tests' – dribbling test (Bangsbo and Mohr, 2013)	High level Czech football players (U12)
Malina et al (2005)	No	Closed Skill' Tests – ball control with the body; ball control with the head; dribbling speed; dribbling with a pass; passing accuracy and shooting accuracy	High level Portuguese youth football players (aged 13-15)
Moreira et al (2017)	Yes	'Open Skill' test – small-sided game protocol with notational analysis	High level Brazilian youth football players (aged 15)
Vandendriescche et al (2012)	No	'Closed Skill' test – UGent dribbling test	High level Belgian football players selected for National team programme (U16 and U17)

in this domain would help prevent selection\ deselection mistakes and allow us to provide a fair, between-subjects, comparison of technical ability based on biological maturation.

CHAPTER 3: Exploring the Opinions of Experienced and Qualified Academy Coaches on Technical Performance Attributes Perceived to be Important in Elite Level Football

3.1 Introduction

Elite performance in football is a complex phenomenon that is determined by a multitude of interrelated factors including physical, tactical, psychological, sociological and technical capabilities (Vaeyens et al, 2006). The identification and development of young players who have shown a potential aptitude to progress into an elite performer, is an important process within professional clubs (Williams & Reilly, 2000). To date, TID research has had a strong focus on assessing physical, physiological and anthropometric characteristics for predicting successful performance in youth footballers (Coelho e Silva et al, 2010). Only recently has there been an enhanced emphasis on the importance of technical performance within top level football (Barnes et al, 2014; Liu et al, 2016; Rampinini et al, 2009). Despite the recognition of the importance of technical ability to performance, there has been comparatively little research investigating the measurement of it when compared to literature focussing on the assessment and development of physical attributes (Ali, 2011). The assessment of technical performance allows for the generation of objective data to support the traditionally subjective assessment of playing ability. Consequently, the provision of valid and reliable objective data can aid the decision-making and coaching process by including a more systematic, evidence-based approach to compliment subjective coach opinion.

Previous research has investigated a plethora of different methods for assessing technical performance in football (Ali, 2011). These range from isolated performance tests in a 'closed' environment (McDermott et al, 2015) to the analysis of performance within the context of a game scenario (Garcia-Lopez et al, 2013). Inherent differences exist between 'closed skill' assessments and assessments carried out within the context of competitive match play. The main difference between the assessment protocols is the requirement for both internal (successful completion of the skill) and external focus (such as movement of teammates and opposition) during match play situations, compared with the requirement for only internal focus during 'closed skill' procedures (Rubajczyk & Rokita, 2015). 'Closed skill' assessments generally involve the completion of a football-specific task in a pre-determined pattern as quickly as possible to determine accuracy and speed of skill completion. An example of this is the Loughborough Passing Test (McDermott et al, 2015), whereby a specific number of targets must be hit by passing the ball against them as quickly as possible. In comparison, an assessment protocol measured within the context of match performance involves video capture during competitive performance situations and subsequent evaluation of skill proficiency through notational analysis (Garcia-Lopez et al, 2013).

Due to the complex nature of technical performance and its interaction with the numerous internal and external confounding variables, no method has been globally recognised as the 'gold standard'. It has been demonstrated that isolated performance tests which require no external focus, correlate poorly with game performance (Lehyr et al, 2017; Praca et al, 2015; Rubajczyk & Rokita, 2015). This suggests that future research should continue to focus on the development of an appropriate assessment protocol that is more specific to the demands of competition. Following this, it has been suggested that the use of small-sided games could be an attractive and practically efficient method of assessing and monitoring technical performance due to the similar perceptual-cognitive demands compared with match performance (Cobb et al, 2018; Garcia-Lopez et al, 2013; Unnithan et al, 2012). However, consistent with the development of any new experimental procedure in other scientific disciplines, a structured and rigorous process must be followed when considering the application of a new, previously unvalidated assessment tool (Brewer & Jones, 2012; Robertson et al, 2017).

Within current research, several concerns with the methodological processes adopted are consistent throughout the literature. Brewer and Jones (2002) proposed a 5-stage model for developing any contextually valid observation tool. The first stage proposed by Brewer and Jones (2002) involves observer training using a previously validated observation tool as a framework. The purpose of this stage is to familiarise observers with the analysis process that will subsequently enhance reliability. The second stage of the process involves amending an existing instrument to the context of the intended environment (e.g., within a professional football academy). A key stage during this process is the development of operational definitions for key performance indicators so that content validity can be established. It is of critical importance that the definitions provided are created through scientific process to ensure the tool is contextually relevant. The third involves establishing instrument criterion-based validity with the aim being to determine whether the instrument measures what it is perceived to measure. For example, does the instrument correlate with the performance variable measured. Finally, stages 4 and 5 involve establishing objectivity and reliability by checking inter- and intra-observer reliability. In addition, Robertson et al (2017) highlighted several other key stages in the development of a new observation tool not included in the Brewer and Jones (2002) framework. These key stages include measures of test/retest reliability and test sensitivity through the calculation of the smallest worthwhile change. From the perspective of talent development and longitudinal tracking of performance over time, an important procedure in any holistic talent development programme, the steps outlined here are of critical importance (Bennett, Vaeyens & Fransen, 2018). Therefore, to be truly confident about the robustness of any measurement tool within sports science, establishing appropriate measures of validity and reliability need to be achieved. Until this process is adhered to, results generated from any observation tool should be interpreted with caution.

One of the early stages outlined by Brewer and Jones (2002) is creating definitions of key performance indicators to be measured during the analysis process. An important component of this process should include validation of these key performance indicators through scientific process to ensure the assessment tool is contextually relevant within its intended environment. Therefore, the first aim of this study is to establish content validity using the Delphi method to explore current perceived opinions on the technical attributes that are most important for successful performance in high level competition. This initial step in the development of a new observation tool will provide a framework upon which performance will be measured. Due to the subjective nature of performance, it is of critical importance that consensus can be reached regarding the selected key performance indicators to ensure contextual validity. Furthermore, this process will help conceptualise how technical performance is viewed and will provide a building block for subsequent stages in the development of the Technical Performance Assessment (TEEM). The second aim of this study is to combine the information gathered from the exploration of ideas and models adapted from previous research to design an observation tool for assessing and monitoring technical performance which can be easily administered within a football academy environment.

3.2 Methods

3.2.1 Study Design

The Delphi method is a structured and systematic qualitative research method used for exploring professional opinions of a selected panel of field experts on any given research question (Eubank et al, 2016; Hasson et al, 2000; Larkin & O'Connor, 2017). The Delphi method consists of multiple rounds of questionnaires, with each subsequent round providing consolidated and summarised data from previous rounds. In the first instance, the aim of the Delphi method is to establish a broad opinion of the subject matter. Following this, each subsequent round is filtered, consolidated and confirmed until consensus between experts is reached (Diamond et al, 2014). This method provides a valid and reliable gathering of data at the convenience of the participant without the need for a long, structured interview process. Ethics approval was granted by the Liverpool John Moores University Ethics committee and all participants provided written consent prior to participation.

3.2.2 Participants

The Delphi method requires careful consideration of participant selection for inclusion during the data collection process. One of the main strengths of the Delphi method is the ability to select a panel of 'experts' who are considered as having 'domain-specific expertise' in the subject area.

Participant recruitment was based on the bespoke requirements of a Scottish Premier League Football Academy. By restricting the participant recruitment procedure to only individuals employed by the club, the contextual validity of the results generated could be improved. Each participant possesses a detailed knowledge and understanding of the academy playing philosophy and therefore, the results of the Delphi process should reflect the technical attributes that the academy perceives to be most important for success at senior club level. The participants recruited for this study were the most highly qualified and experienced coaches within the club. Five coaches were approached to participate in the study. Each coach possessed the highest coaching qualification in their line of expertise. Two of the five coaches possessed an UEFA Pro License, and three of the five coaches possessed a UEFA Elite Youth A License. All participants had high level playing experience of more than 15 years and were employed at the club's first team or academy.

3.2.3 Procedures

First Round of Questionnaires. To explore the general opinion of the technical attributes required to be a successful professional football player, the first round of questionnaires provided the participants with a series of open-ended questions (e.g., What aspects of technical performance do you consider being most important to elite technical performance without taking into consideration positional differences?). The questions were designed to be thought-provoking to elicit varied responses (e.g., Please provide personal comments which you feel are important when considering important attributes for technical performance). At this stage, all responses were welcomed in the hope that all important attributes were identified. Questionnaires were administered face-to-face to improve compliance. Furthermore, this provided the participants the opportunity to read over the questionnaire and ask any questions about issues they were unsure about. Researcher contact details were provided should any questions have arisen prior to the second round.

Second Round of Questionnaires. Round 2 of questionnaires began with the presentation of a summary of responses generated from Round 1 in a visual, easy-to-understand format. All responses complied by the coaches in the first round of questionnaires were included in this summary. Following this, the panel were asked 'closed' questions to filter down the responses towards consensus. The participants were asked to rate each individual attributes' importance on a Likert scale ranging from 1-5. The 1-5 scale was based on the number of responses each attribute received in round 1 (e.g., If an attribute received 4 out 5 responses from coaches in the first round, then the attribute was assigned a rating of 4, with 1 = not important and 5 = very important). The current rating (based on number of responses) was provided, and the coaches were asked if they would like to change the rating of each attribute in an accompanying box. This method was designed to quantify the subjective data generated in the initial round for analysis. Similar to round 1 questionnaires, round 2 questionnaires were administered face-to-face to improve compliance and provide the participants an opportunity to ask any questions.

Third Round of Questionnaires. The third round of questionnaires attempted to establish consensus. Consensus was considered reached when 80% agreement between experts was established (Diamond et al, 2014). In attributes which invoked 0-60% agreement, consensus was deemed to have not been reached and attributes were omitted from the analysis. In attributes that reached between 60-79% agreement, further exploration was considered warranted and were proceeded to round 3. The panel were presented with the summarised list of attributes generated from subsequent rounds of questionnaires and their current percentage of agreement. The panel were simply asked if they agreed or disagreed with the inclusion of each attribute in the assessment procedure. Following this stage, if the 80% agreement level was reached, the attribute was included in the analysis. If the agreement level was below the 80% threshold, the attribute was removed. Furthermore, during this round of questionnaires, the panel were presented with operational definitions of the key performance indicators to be assessed. Again, the panel were asked if they agreed or disagreed with the definitions and a box was provided to make any necessary amendments.

3.3 Results

3.3.1 Delphi Questionnaires

First Round of Questionnaires

Responses from the first round of questionnaires provided a broad range of attributes that the coaches perceived to be important to technical performance. As the initial stage was designed to explore opinion, attributes which may not be directly related to technical performance were highlighted. A total of 24 attributes were identified following the first round of questionnaires (see figure 5). Of the 24 attributes identified, 14 were directly related to technical performance. The remaining 10 attributes related to cognitive, tactical or physical performance attributes. Attributes which were not directly related to technical performance (e.g., speed) were excluded from the analysis. Furthermore, attributes which were unable or difficult to measure were also excluded from the analysis (e.g., 2 footedness, play at speed, vision)

Second Round of Questionnaires

Results of the second round of questionnaires are shown in table 4. Following the first round of questionnaires: consensus was reached on 2 attributes; consensus was possibly reached on 4 attributes but required further exploration in round 3 and consensus was not reached on 5 attributes. If consensus was not reached, the attribute was omitted from the analysis.

Third Round of Questionnaires

Of the 11 attributes identified by the panel of experts, 5 reached consensus and qualified for inclusion in the assessment protocol. Six attributes did not reach the 80% agreement level set before the questionnaires were administered. Results from round 3 of questionnaires can be seen in table 5. The attributes which reached consensus and were included in the analysis of technical performance were ball manipulation, shooting, passing, first touch, decision-making and two footedness. Furthermore, a finalised list of operational definitions for the key performance indicators can be seen in table 6.

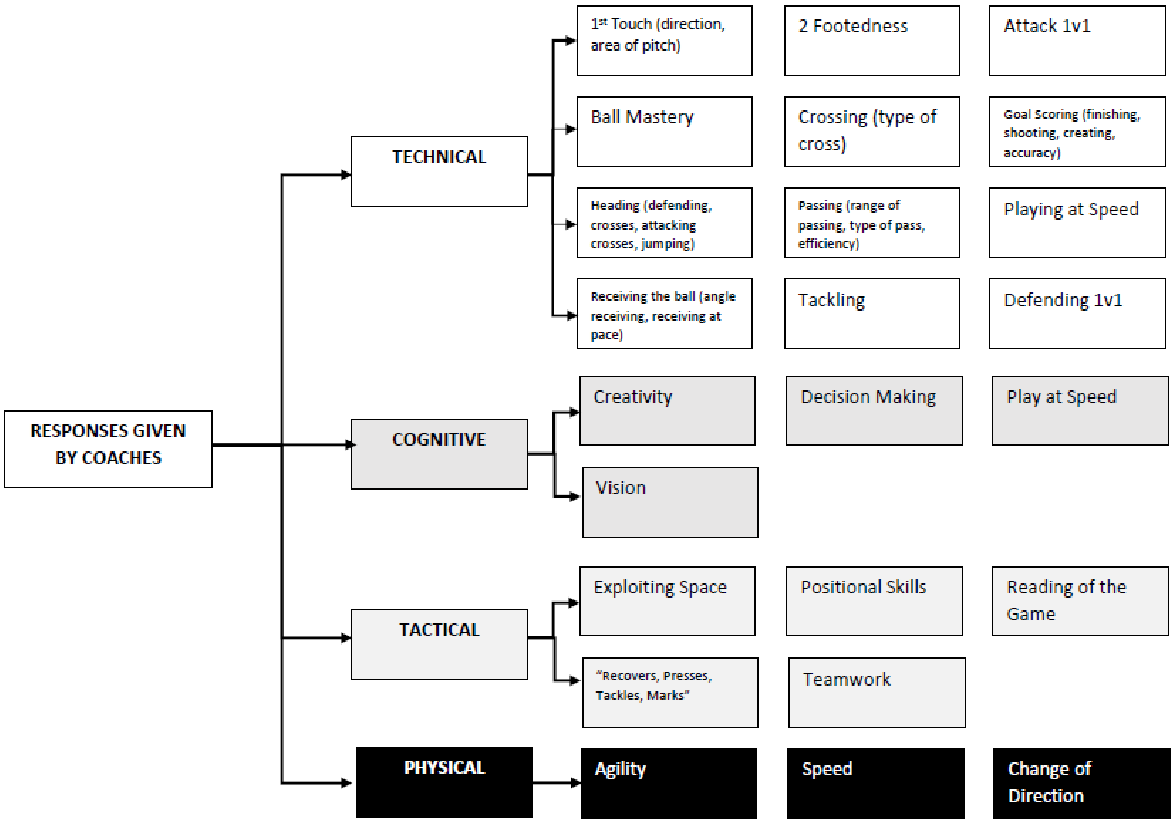

Figure 5. The diagram shows all the responses given by coaches in round 1 of questionnaires. The list of attributes was classified into 'technical', 'cognitive', 'tactical' or 'physical'. *(information in brackets describes the additional contextual information provided with the responses)*

3.3.2 Development of the Assessment Tool

Following the establishment of key performance indicators from the Delphi process, which was determined after reaching coach consensus, the operational definitions for each technical action provided a framework for tool design. After the observation tools content validity was established, a combination of previous assessment protocols (Cobb et al, 2018; Unnithan et al, 2012) were used as a framework to develop a new tool. Previous research was identified, and the parameters of tools developed within these research studies were manipulated to conform to the context of the club's academy environment (Cobb et al, 2018; Unnithan et al, 2012). Due to the practicality and ease of application within a training environment, small-sided games were used as the format for technical performance assessment (Bennett et al, 2017). Furthermore, previous research reported large variability in performance during competitive matches probably due to external confounding variables such as team tactics, quality of opposition and environmental conditions (Liu et al, 2016). This variation in match performance suggests that performance assessment within a small-sided game scenario may provide a more stable and consistent environment for evaluation.

3.3.3 Test Protocol

After these considerations, a 6 v 6 (+2 goalkeepers) small-sided game format was selected for the assessment. To make small-sided game conditions as specific to real life match performance as possible, an individual playing area of $100m^2$ / player (outfield only) was applied (Frauda et al, 2013). In line with Frauda et al (2013), 10m from each goalkeeper to the playing area was also applied. This resulted in pitch dimensions of 50m (length) x 40m (width). The assessment protocol consisted of 6 x 6-minute games. Normal game rules were applied except for the offside rule and the exclusion of throw-ins, corners and free kicks. If the ball went out of play or a foul was committed, the game started with the goalkeeper in possession. Players were given a number, and teams were selected at random using an online random team generator. After each game, teams were changed randomly to avoid any bias in team selection (Unnithan et al., 2012). This resulted in players being required to play with different players and against different opposition each game. To ensure the games were competitive, players were informed that this was an assessment, and each individual player would receive 2 points per win, 1 point per draw and 0 points per loss. Players also received 1 point for every goal their team scored. The aim was to develop a competitive environment to ensure there was adequate pressure on players during technical actions to resemble match performance as closely as possible. Games were video recorded for retrospective analysis of the frequency of technical actions performed during the assessment protocol.

3.3.4 Score Reporting of Technical Actions

Following the collection of frequency data, a 'skill efficiency performance index (SEPI) (Memmet & Harvey, 2008) was calculated for the 4 main key performance indicators (KPI's) identified by the panel of coaches during the process of establishing content validity in chapter 3. The SEPI was selected over percentage success due to its inclusion of a starting score of 10 (a constant) for each of the KPIs. Memmet and Harvey (2008) highlighted a limitation to reporting skill success as a simple percentage. The authors observed that when the frequency count of a specific action is low, a simple percentage success calculation may not provide a fully accurate representation of skill success due to the possibility of completing a successful action by chance. For example, if the total frequency count of an action was 3 and the participant successfully completed 2 actions, then this would give a percentage success of 66%. In this situation, it is possible that 1 (or even 2) actions could have occurred due to luck. When using the SEPI for the same example, a score of 52.2% would be generated. Therefore, with the SEPI, a higher score is generated as the element of chance decreases with a higher frequency of actions. The main key technical performance indicators were identified as first touch, ball manipulation, passing and shooting (figure 5). The SEPI was calculated using the following formula, where SA = successful actions and UA = unsuccessful actions:

$$(10 + SA)/ ((10 + SA) + (10 + UA)) * 100$$

For the shooting SEPI, successful and unsuccessful actions were replaced by 'shots on target' and 'shots off target'. For youth footballers, shooting accuracy was determined by the number of shots successfully hit on target as opposed to conversion rate, therefore, ON = shots on target and OFF = shots off target:

$$(10 + ON)/ ((10 + ON) + (10 + OFF)) * 100$$

Table 4. Table 4 shows the ratings of perceived importance for each attribute generated from round 2 if DELPHI questionnaires.

SUMMARISED LIST OF ATTRIBUTES	PERCEIVED IMPORTANCE (1-5)					AVERAGE	% AGREEMENT	Consensus Reached	Action Taken
	EXPERT 1	EXPERT 2	EXPERT 3	EXPERT 4	EXPERT 5				
1st Touch - (direction, movement, area of pitch touch)	4	4	4	4	5	4.2	84	Yes	Include in Analysis
Passing - (range of passing, type of pass)	5	5	5	5	4	4.8	96	Yes	Include in Analysis
2 footedness	4	3	3	4	4	3.6	72.0	Possibly	Proceed to Round 3
Heading - (defending cross, defending long ball, attack cross, jumping, timing)	3	4	4	4	3	3.6	72.0	Possibly	Proceed to Round 3
Decision Making	3	4	4	4	4	3.8	76.0	Possibly	Proceed to Round 3
Ball Manipulation	3	2	4	4	4	3.4	68	Possibly	Proceed to Round 3
Crossing - (type of cross)	3	4	3	2	2	2.8	56.0	No	Omit from Analysis
Defend 1v1	1	4	3	3	3	2.8	56.0	No	Omit from Analysis
Shooting - (type of shot, accuracy)	3	3	3	2	3	2.8	56	No	Omit from Analysis
Tackling	1	3	3	4	2	2.6	52.0	No	Omit from Analysis
Creativity	3	1	4	3	2	2.6	52.0	No	Omit from Analysis

Table 5. Table 5 shows the responses given by the panel of experts from round 3 of DELPHI questionnaires. Experts were asked if they agreed or disagreed with the inclusion of the attributes identified in the assessment

SUMMARISED LIST OF ATTRIBUTES	PERCEIVED IMPORTANCE (1-5)					AVERAGE	% AGREEMENT	Consensus Reached	Action Taken
	EXPERT 1	EXPERT 2	EXPERT 3	EXPERT 4	EXPERT 5				
Ball Manipulation	Agree	Agree	Agree	Agree	Agree		100	Yes	Include in Analysis
Shooting - (type of shot, accuracy)	Agree	Agree	Agree	Agree	Agree		100	Yes	Include in Analysis
Passing - (range of passing, type of pass)	5	5	5	5	4	4.8	96	Yes	Include in Analysis
1st Touch - (direction, movement, area of pitch touch)	4	4	4	4	5	4.2	84	Yes	Include in Analysis
2 footedness	Agree	Agree	Agree	Disagree	Agree		80	Yes	Include in Analysis
Decision Making	Agree	Agree	Disagree	Agree	Agree		80	Yes	Include in Analysis
Heading - (defending cross, defending long ball, attack cross, jumping, timing)	Disagree	Agree	Disagree	Agree	Agree		60	Yes	Omit from Analysis
Crossing - (type of cross)	Disagree	Disagree	Agree	Agree	Agree		60	Yes	Omit from Analysis
Defend 1v1	Disagree	Agree	Disagree	Agree	Agree		60	Yes	Omit from Analysis
Creativity	Agree	Disagree	Disagree	Disagree	Agree		40	Yes	Omit from Analysis
Tackling	Disagree	Disagree	Disagree	Disagree	Agree		20	Yes	Omit from Analysis

Table 6. Table 6 details the operational definitions for each of the key performance indicators identified during the Delphi process and following the third round of questionnaires

PERFORMANCE VARIABLE	DEFINITION
PASSING	
Successful Pass	The pass was able to be controlled by a teammate
Unsuccessful Pass	The pass was not able to be controlled by a teammate or possession was lost
Forward Pass	A pass played in the direction of the opposition goal
Side Pass	A pass played sideways in the same third of the pitch
Backward Pass	A pass played in the direction of their own goal
Short Pass	A pass of approximately 25 metres or less
Long Pass	A pass approximately than 25 metres or more
Pass in Defensive 3rd	A pass played in the defensive 3rd
Pass in Middle 3rd	A pass played in the middle 3rd
Pass in Attacking 3rd	A pass played in the attacking 3rd
Penetrating Pass	A penetrating pass played through the midfield or defensive lines to a teammate. Can be in the air or on the ground. The pass 'takes out' 2 or more players.
1ST TOUCH	
Successful 1st Touch	When receiving the ball, the next action is set up and possession is retained
Unsuccessful 1st Touch	When receiving the ball, the ball is not under control following the first touch resulting in lost possession
1st Touch Forward	1st touch is in the direction of the opposition goal
1st Touch Side	1st touch is taken sideways in the same 3rd of the pitch
1st Touch Back	1st Touch is taken in the direction of their own goal
1st Touch in Defensive 3rd	1st Touch in the defensive 3rd of the pitch
1st Touch in Middle 3rd	1st Touch in the middle 3rd of the pitch
1st Touch in Attacking 3rd	1st Touch in the attacking 3rd of the pitch
BALL MANIPULATION	
Successful Ball Manipulation	The player in possession of the ball moves it into available space; to beat an opponent, create space for an attempted pass, cross or shot, or to escape opposition pressure by running with the ball
Unsuccessful Ball Manipulation	Loss of possession or being tackled when attempting the above
1v1	Beating and opponent in a 1v1 situation resulting in a shot, cross or pass to a teammate
Skill/ Move	Escaping from opposition pressure using a skill or move
Drive	Running with the ball into space towards the opposition goal
SHOOTING	
Successful Shot	Scoring a goal
Unsuccessful Shot	Failure to score a goal

Smash	A shot hit with power with the instep
Finish	A shot hit with precision and attempted accuracy rather than power alone
On-Target	A shot which forces a save from the goalkeeper or hits the goal area including the crossbar and posts
Off-Target	A shot which does not hit the goal area including the goal crossbar and posts
Shot with Instep	A shot struck with the laces
Shot with Inside of Foot	A shot struck with the inside of the foot
1st Time Shot	A shot taken with the players 1st touch
Short	A shot taken from within approximately 16 metres from the goal
Long	A shot taken out with approximately 16 metres from the goal
PRESSURE	
Pressure	Distance between a player with the ball (first attacker) and an immediate pressing player (s) (first defender), excluding the goalkeeper with the defending player making it difficult to complete the current or next action
Pressure-Loose	When the first defender is estimated to be more than 1.5 metres away and is NOT making it difficult for the player in possession to complete the current or next action
Pressure - Tight	When the first defender is estimated to be within 1.5 metres, and is making it difficult for the player in possession to complete the current or next action
DECISION MAKING	
Best Decision	Execution of a skill which results in the best possible outcome for the team in the given situation e.g. scoring a goal, a key pass, keep possession, creating an opportunity. This includes situations when the best decision is made but the execution of the skill is incorrect
Ok Decision	Execution of a skill which results in keeping possession of the ball but having better options available e.g. passing backwards when a forward pass was on. Includes an unsuccessful skill execution when the player fails to select the best decision
Poor Decision	Choosing an inappropriate skill execution in the given situation or the wrong option which results in a loss of possession
ASSIST	Directly creating a goal using a pass, cross or skill l

3.4 Discussion

3.4.1 General Discussion

This current study provided a systematic method for establishing content validity when designing a new instrument to reliably assess and monitor technical performance in youth footballers. This study contributes to the understanding of the perceived attributes required for high performance in elite football within the context of a Scottish Premier League club. The findings indicate that from a general perspective of successful football performance at an elite level, 1st touch, ball manipulation, passing, shooting, decision-making and 2 footedness were the technical attributes which were perceived to be most important. Furthermore, heading, crossing, 1 v 1 defending, creativity and tackling were not considered as being fundamental technical attributes for successful performance. Results presented in this study are consistent with the findings reported by Larkin and O'Connor (2017). They reported that high-level national coaches and performance directors from Australia confirmed 1st touch, ball striking (defined as passing and shooting) and decision-making as some of the most important attributes in football. In addition, they reported that technique under pressure was considered an important technical attribute, which may possibly have been overlooked in the current study. Also consistent with the results, Larkin and O'Connor (2017) reported that defending ability (defined as defensive 1 v 1's, tackling and heading) was omitted following the first round of results. One possible explanation for this is that defensive ability is an attribute which can be more easily improved with age and practice. Furthermore, it is possible well-developed physical attributes and anthropometric characteristics may significantly contribute to strong defensive ability in youth footballers when differences in maturational stages are considered (Towlson et al, 2017).

An important attribute highlighted in both the current study and Larkin and O'Connor (2017) is decision-making. Although this attribute is not measurable as a skill execution, decision-making plays a pivotal role in the process (Praca et al, 2015). The process of skill execution involves 3 stages (Williams, 2000). The first stage involves information gathering from the surrounding, external environment and context of the game moment. Factors influencing the success of this first process are visual scanning capabilities for identifying team-mate and opponent movements, and tactical knowledge. The second stage during skill execution relates to the decision-making process whereby fast decisions on the appropriate skill selection and execution based on the information gathered from the first stage need to be applied. Lastly, the ability to execute the selected skill is the final stage of skill execution whereby technique plays a critical role. Therefore, it is essential that decision-making is included in any skill proficiency analysis due to the extent to which it influences the success of any technical action. One difficulty in measuring decision-making within performance highlighted in previous research is its subjective nature due to the inability to observe the process taking place i.e., the process is internal within the player (Lorains et al, 2013). However, by providing detailed operational definitions, the subjectivity of assessing decision-making can be reduced (James, 2006; Lorains et al, 2013). Therefore, by including detailed and appropriate operational definitions (table 6), decision-making can, and should, be included in the analysis of technical performance.

One attribute which was not identified in the study but was highlighted by Larkin and O'Connor (2017) was skill success under pressure. Due to the varying levels of opposition pressure in specific pitch locations and tactical instructions that might alter time and space demands, (e.g., high press tactics and pressure on ball in opposition penalty area) the author recognises that this is an attribute which should be considered for inclusion in the analysis of skill proficiency. Like the issue highlighted in the analysis of decision-making, opposition pressure is highly subjective. Also, in this instance, the subjectivity of opposition pressure analysis can be reduced, but not entirely eradicated, with the provision of detailed and clear operational definitions (Tenga et al, 2010). Therefore, the authors conclude that due to its substantial external influence on skill success, opposition pressure should be included in the analysis of technical performance.

One novel finding from the current study which did not emerge in previous work (Larkin & O'Connor, 2017) is the perceived importance of contextual information which accompanied each technical attribute. Table 4 shows a list of responses given by coaches and in brackets the accompanying range of descriptions that were supplied along with those responses. For example, the coaches perceived the direction and type of pass as important information alongside passing efficiency. This finding is consistent with the opinions in performance analysis research (James, 2006; McKenzie & Cushion, 2013). The inclusion of contextual information as a descriptor to technical actions provides the opportunity to gain a better understanding of player behaviour. Following the identification of player behaviour through objective feedback, coaches can attempt to manipulate player behaviour and subsequently improve performance. For example, a shift from a high percentage of sideways and backwards passes to a greater percentage of forward penetrating passes. The use of contextual information in player technical performance evaluation also provides a more in-depth analysis that can add valuable information to individual development plans. By identifying behavioural habits through notational analysis in relation to a player's technical performance, coaches should be able to positively manipulate behaviour through coaching to improve performance. Furthermore, by matching a player's technical performance characteristics to player technical profiles designed by the club, a position-specific assessment of performance can be utilised. However, until appropriate validity and reliability measures are conducted within any assessment tool, the results generated must be interpreted with caution. Therefore, it is important that the criterion-related validity (i.e., the extent to which test performance correlates to competitive match performance) of any assessment tool is established so practitioners and coaches can be confident that manipulating behaviour in a performance test will transfer to match performance.

The use of small-sided games for the assessment of technical performance in football has gained popularity in recent years (Cobb et al, 2018; Unnithan et al, 2012). Within research, it has been generally accepted that performance in small-sided games could potentially better reflect match performance compared to isolated, 'closed-skill' performance tests (Praca et al, 2015; Larkin & O'Connor, 2017; Rubajczyk & Rokita, 2015). One possible explanation relates to perceptual-cognitive demands during game-like situations, where information gathering, and decision making are critical components of skill proficiency. In contrast, assessment protocols which are performed in isolation (such as a target-based passing tests or dribble slalom track) do not correlate well to match performance. Therefore, the utilisation of a small-sided game scenario as an assessment tool was preferred to one, or a combination of, isolated skill assessments. In support, Ford et al (2009) put forward the Expert Performance Approach which provides a framework for learning and skill acquisition. They propose a 3-stage model for skill acquisition in which the first stage is characterised by the capture of expert performance in environments closely resembling the criterion performance measure (competition performance). Performance within a small-sided game requires the integration of perceptual, cognitive and action skills that is comparable to the criterion performance alluded to by Ford et al (2009). The presence of opposition players and teammates within the field of play contribute to the multifactorial facets of technical performance. With this in consideration, research towards the use of small-sided games for assessing and monitoring technical performance is gaining momentum (Bergkamp et al, 2019). Furthermore, small-sided game protocols are easy to administer within an academy environment and are more time efficient since data from many participants can be collected within one testing session. One issue which remains unexplored, to the authors knowledge, is the relationship between technical performance in small-sided games and match performance. To date, research investigating the relationship between small-sided game performance and real-life match performance is limited and warrants further exploration. It is important for further enhancing criterion-based validity that the relationship (if any) between the performance being measured within our assessment protocol and the performance we want to directly influence is established. Perhaps one limitation which could influence the results of establishing this relationship is the variation reported in match-to-match performance due to external factors such as team tactics, quality of opposition and environmental conditions (Liu et al, 2016). However, further research is warranted in this area to expand current knowledge.

3.4.2 Limitations

While this study provided valuable information on the technical attributes perceived to be important to adult success within a Scottish Premier League academy, several limitations are identified which should be considered. Firstly, data was collected from a panel of coaches within one Scottish

Premier League academy and therefore, the results obtained must be interpreted with caution when transferring across contexts. Assumptions should not be made regarding the generalisability of the results as various factors could contribute to the perception of attributes deemed important within clubs or organisations. These factors include the individual clubs playing philosophy, coach bias or the club's position-specific player performance profile. Secondly, the attributes identified within the current study provide a general framework for technical skill competency. On reflection, the differences in position-specific technical demands could be considered to add further context to the player assessment. For example, the relevance of ball manipulation ability for central defenders. However, a counter argument to this limitation could be the need for a foundation level of technical competency across all areas of performance to be successful at the highest level. Future research could investigate the position-specific intricacies of technical performance which could further enhance the player evaluation process. Thirdly, a limitation to the Delphi process is the potential for coaches to alter their responses based on what they think should be the right answer rather than what they believe. In this instance, the test administer is relying on the responses to be truthful and honest to provide reliable data. Fourthly, in the current study, only 5 coaches met the selection criteria and were selected to participate on the panel. A small number of participants in the panel increases the chances of responses being overlooked during the initial round of questionnaires. However, as mentioned previously, findings in the current study are consistent with the findings by the similar study by Larking and O'Connor (2017) and therefore the authors are confident that all aspects of technical performance in football have been identified. Lastly, when applying the Delphi method in the data collection process, it is important that researchers consider the sample size being recruited for participation. Whilst the Delphi method is a scientific process of using subjective data to systematically ensure content validity, reliability and a reduction in cognitive bias, two main contraindications are apparent for inappropriate sample sizes. A small sample size limits the reduction in bias by having a lower number of 'opinions', whereas an excessively large sample size makes it difficult to reach consensus due to larger variation in beliefs. Therefore, the researcher/ practitioner must find a balance to ensure all valuable responses are being collected and consensus is reached.

3.4.3 Conclusion

This study provided a systematic and scientific approach to establishing content validity for the development of a new assessment tool for assessing and monitoring technical performance in youth footballers within a Scottish Premier League academy. A novel finding in this study is the recognition of the importance of contextual information to add detail to the analysis process. This in turn can provide a more in-depth player analysis which can aid coaches in manipulating player behaviour to positively impact performance. By designing a contextually relevant assessment tool for player evaluation and monitoring, players and coaches can be provided with objective data to support subjective intuition which can be used to further enhance player development

3.5 Personal Reflection and Professional Skills Development

This stage of my professional doctorate journey opened my eyes to new research skills, which I had previously no exposure to. Firstly, the necessary reading around the process of developing new tools to be used within science highlighted the importance of establishing appropriate measurement properties such as test/retest reliability; criterion-based validity and sensitivity to change. When considering my professional development, the process of acquiring this knowledge positively influenced the way in which I now read and critique scientific journal articles. Previously, I would quickly skim over the methods section and rush to the results and discussion section which I now realise could have resulted in me overlooking important methodological issues. I now understand the importance of robust scientific methodology and feel more confident when critically analysing literature which facilitates the process drawing more informed personal conclusions. I feel this deeper understanding of research methodology is an important learning step in my development as a practitioner as research-informed decision making is an essential industry skill. In addition to this, I was given exposure to qualitative research methods. During development of the Delphi questionnaires, I decided to pilot the questionnaire by reaching out to a number of coaches/practitioners with whom I was connected with via a social media platform. I sent around 40 people a message explaining the purpose of my research and issued the questionnaire to them (which typically took around 10 minutes to complete). Of the 40 people I had sent the questionnaire to, I received only 8 responses. This opened my eyes to the limitations and potential issues I may face if I decided to collect data in this manner. Following this, I decided to keep my data collection in-house from coaches I knew personally and had previously established personal relationships. I felt that this method of 'convenience sampling' would dramatically improve the response rate and compliance. I understood that this would limit the generalisability of my results but concluded that my research is being conducted for the specific benefit of my organisation and not a greater population.

At this early stage of my project, I had my first opportunity to practice my dissemination skills following the first round of questionnaires. After receiving the responses from the first round of questionnaires, I had to decide upon the best way the present back the results to the participants in the next round of questionnaires. The participants themselves were not from academic backgrounds and therefore I did not want to disengage them from the process by reporting mathematical or scientific terms. I feel that this was a good decision which made the process of data collection easier. This extract from my reflective journal demonstrates my thoughts during the process:

> *"Was happy with the initial responses to the open-ended questions. Most agreed that SSG's were a good way to assess technical performance. Got several varied responses. I think this was always going to happen given the subjective nature of coaching and the fact that everyone has their own opinion. However, this is good as it allows for further elaboration in subsequent Delphi rounds*
>
> *Presentation of results*
>
> *I decided not to go 'too statistical' as the panel selected for questionnaires were not from an academic background. I think this was a wise decision as it get them engaged and didn't think it was a waste of time. I wanted to make the presentation of results simple and clear and didn't want to use too much jargon.*
>
> *Consensus*
>
> *I used the experts rating 1-5 (5 being most important) to gain insight into what they regarded as important attributes. If I was to change anything, I would have considered using a larger scale (e.g., 1-10). This might have allowed consensus to be reached easier. For example, instead of requiring a 4 out of 5 for 80% agreement, 8 out of 10 would equal the same. Not sure if that would have made a massive difference or not. I may have also considered using a larger panel. Would generalisability have been improved if I branched out and sought responses from coaches from other clubs? The downside to this would have been that compliance would have dramatically decreased whereas in this instance, compliance was 100%. Furthermore, the current methods make it bespoke to my club."*

In line with my research and professional skills development aims outlined in chapter 1, the following learning outcomes from this stage were as follows:

- To develop research and technical skills
- To develop communication skills (specifically public speaking) and dissemination of information for various audiences
- To develop transferable knowledge, concepts and processes which can be utilised in other contexts

CHAPTER 4: Establishing the Reliability of a Newly Developed Assessment Tool for Assessing and Monitoring Technical Performance in Youth Footballers

4.1 Introduction

The use of notational analysis in football is becoming an increasingly popular method of assessing performance (Aquino et al, 2017). Recent research has attempted to develop observational tools that incorporate digital video capture and subsequent analysis with computer-based software, which can assess and monitor technical performance within match and training contexts in youth football (Cobb et al, 2018; Garcia-Lopez et al, 2013; Moreira et al, 2017; Waldron & Worsfold, 2010). Objective data on specific components of performance provides key stakeholders with information to aid the traditionally subjective coaching and development process. Although this research represents a positive step towards supporting the development of young players, one limitation to the methods employed is a lack of consistency in the methodological processes utilised for assessing performance. Due to the complex nature of technical performance in a random and unpredictable team invasion sport, no globally accepted method of assessment has been identified (McKenzie & Cushion, 2013). It is of critical importance that any tool utilised for assessing technical performance is subjected to the same appropriate processes of establishing validity and reliability as with developing measurement tools in any other scientific discipline, processes which have been under-reported previously (Robertson et al, 2014).

Brewer and Jones (2002) proposed a 5-stage model for developing new, contextually valid, systematic observation tools for the assessment of performance. This model provides a structured and logical framework for establishing content validity, criterion-based validity and inter/intra-observer reliability. More recently, Robertson et al (2017) explored practitioner opinion on the measurement properties perceived to be important when selecting or developing performance tests within the field of sport and exercise science. Results revealed additional measurement properties overlooked in the model proposed by Brewer and Jones (2002), which should also be considered during test selection/development. These additional measurement properties include tests to establish test/retest reliability, responsiveness, discriminant validity and the smallest worthwhile change. By incorporating these measurement properties and procedures outlined by both Brewer and Jones (2002) and Robertson et al (2017) in the development of a new observation tool, the scientific robustness of such a tool should be improved and subsequently support the justification of its selection for use in the field. Furthermore, a tool that has withstood the rigour of methodological scrutiny has a greater chance of being adopted as a legitimate, valid and reliable method of performance assessment within the field.

The use of any measurement tool during scientific enquiry has a degree of error associated with it. This error can be a result of random and/or systematic error. When selecting a measurement tool, it is essential that its user is aware of the magnitude of this error and whether this magnitude is acceptable enough as to be considered appropriate for its intended use. When investigating within-subject variation over repeated tests in a measurement tool, the most appropriate variable for calculation is the typical error (TE) (Swinton et al, 2018). The typical error represents the variation in observed scores caused by measurement error when an individual performs repeated tests. In addition to TE, another measurement variable used to investigate performance changes over time is the smallest worthwhile change (SWC) (Hopkins, 2004). The SWC represents any 'real' change in performance that lies outside the associated test variation. Calculation of the SWC allows researchers to identify any 'real' change in performance and the magnitude of this change. By identifying both TE and SWC, practitioners can not only evaluate the reliability of the measurement tool but can establish meaningful values which can facilitate the longitudinal monitoring of performance over time.

Therefore, the aims of this study are to 1) establish test/retest reliability (repeatability) of a newly developed observation tool for assessing and monitoring technical performance in youth footballers and 2) to establish inter and intra-observer reliability. The author hypothesizes that the test, when administered in a consistent and controlled environment, would provide a reliable assessment of technical performance that would provide objective data for the benefits of player development.

4.2 Methods

4.2.1 Video Recording Data

Digital video recording data from habitual training sessions conducted within a Scottish Premier League Academy was used in data collection. Video recording data was collected by the academy performance analyst. Written participant and parent consent were obtained by the club as part of their registration process and complied with the regulations set by their National Governing Body regarding child protection procedures. Participants in the video data were 33 youth football players (age: 14 ± 2.5 years) who participated in a series of small-sided games (as described in chapter 3) as part of their habitual training routine. Participants trained at the academy for an average of 8h per week for approximately 10 months per year. Gatekeeper consent was granted following the issuing of a detailed information sheet and consent form.

4.2.2 Tagging Procedure

Digital video recording data was uploaded to the performance analysis software Sportscode (Elite Version 10) on a laptop computer (Apple MacBook Pro, United States). A specifically designed 'code window' was used to collect frequency data of the key performance indicators (KPIs) described in chapter 3 (table 6). Detailed operational definitions of KPIs were provided to the observer to enhance reliability (chapter 3). The frequencies of each technical action and accompanying contextual data were collected using the 'matrix' feature in the software. Every time a player touched the ball, their action was recorded. The software allowed the video recording to be paused, re-wound and slowed down to ensure the appropriate action was correctly recorded and to limit the possibility of an action being missed.

4.2.3 Establishing Reliability

Test/retest reliability was established by assessing technical performance in the small-sided game format in the first instance, then again within 14 days. Retests were not performed on subsequent days to account for the effects of any accumulated fatigue and within 14 days to avoid any real change in performance due to training and practice. The same participants took part in both sessions with teams being randomly generated as per the procedure described in chapter 3. Assessments were conducted at the same time of a day as per the previous assessment and under the same conditions as part of the teams habitual training session. Players were asked to refrain from any intensive training prior to each technical performance assessment. Both the frequency data and the individual SEPI scores were used for test/retest reliability analysis. Inter-observer reliability was established by comparing the frequency of actions observed between observer 1 (OB1) and observer 2 (OB2). The same video footage was used for comparison between OB1 and OB2. Intra-observer reliability was established by comparing the frequency of actions recorded in an initial performance assessment by observer 1 and the frequency of actions recorded in the same assessment by the same observer, 4 weeks following the initial assessment. Both observers were proficient in the use of the performance analysis software and had 3+ years of previous experience as a performance analyst

4.2.4 Statistical Analysis

Frequency data was used for statistical analysis. The individual SEPIs were also assessed for reliability in the test/retest measure. The SEPIs are calculated using frequency data as outlined in chapter 3. An intraclass correlation coefficient (ICC) was applied for test/retest reliability and inter- and intra-observer reliability measures (Swinton et al, 2018) to ascertain if any differences existed between measurement points and between observer observations (v.16; SPSS Inc., Chicago, IL). ICC scores were interpreted as: <0.5 = poor; 0.51-0.75 = moderate; 0.76-0.90 = good and > 0.90 = excellent (Koo & Li, 2016). The mean difference score between test 1 and test 2, and the 95% confidence intervals (CI) were calculated as a preliminary step for calculating typical error (TE) using the recommended protocol by Swinton et al (2018) to establish measurement error. As the assessment tools intended use includes the monitoring of technical performance over time, the smallest worthwhile change (SWC) was calculated to enable the monitoring of technical performance over time. The SWC was calculated by using Cohen's effect sizes (Cohen, 1988), where: SWC = effect size

(0.2, 0.5 or 0.8) * between-subject coefficient of variation (CV). Calculation of the SWC provides meaningful results generated from the technical assessment that will enhance the feedback process during player development.

4.3 Results

4.3.1 Inter-and Intra-Observer Reliability

Figure 6 shows the inter- and intra-observer reliability between A1 and subsequent observation following a period of 4 weeks (A2) in OB1 and the inter-observer reliability between OB1 and OB2 performing the same analysis on one assessment. Results of the analysis report ICC's for the intra and inter-observer reliability of 0.963 and 0.908 respectively. Intraclass correlation coefficients of ICC > 0.90 indicate excellent intra- and inter-observer reliability within the observation tool (Koo & Li, 2016)

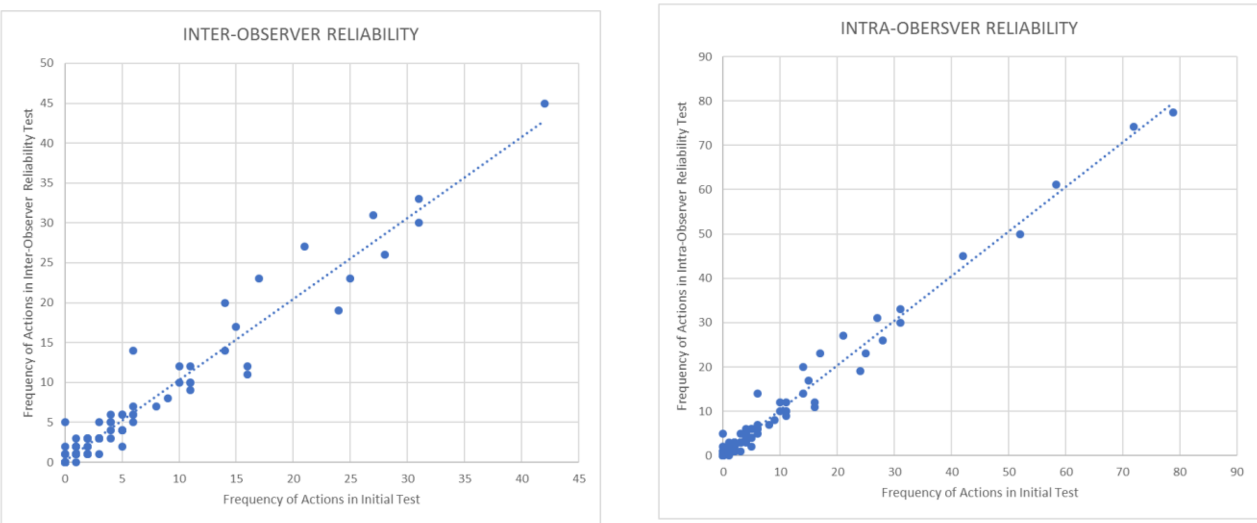

Figure 6. Figure 6 shows the Inter (difference between Observer 1 and Observer 2) and Intra-Observer (difference between Observer 1 and Observer 1 within 14 days of first analysis) Reliability Statistics between Observer 1 and Observer 2 *(ICC = 0.908 and ICC = 0.963 respectively)*

4.3.2 Test/Retest Reliability

Results of the test/retest reliability ICC revealed a significant correlation for frequency of actions between assessment 1 (A1) and assessment 2 (A2) that were separated by period of 14 days (*ICC = 0.966*) (Figure 7). The analysis of individual SEPI reliability revealed moderate and good reliability between A1 and A2 for the 1st touch and passing SEPI's *(ICC = 0.67 and 0.78 respectively)* and poor reliability for the ball manipulation and shooting SEPI's *(ICC = 0.43 and 0.38 respectively)* (figure 8).

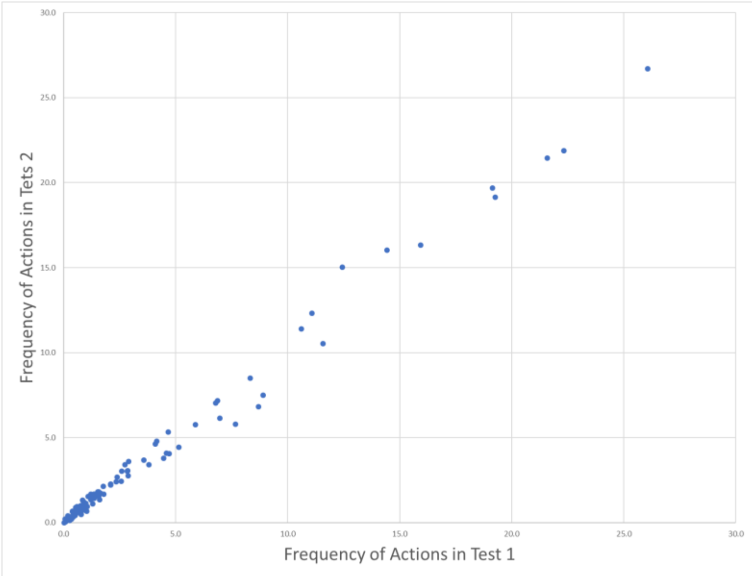

Figure 7. Figure 7 shows the frequency of technical actions recorded by the observer in test 1 and in the re-test within 14 days following the initial test (*r = 0.966*).

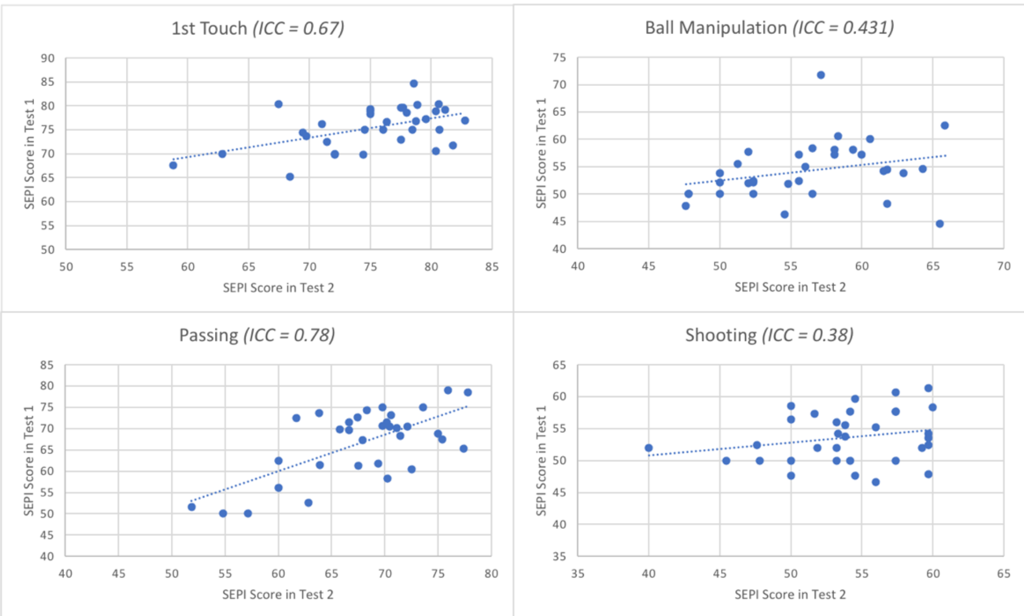

Figure 8. Figure 8 shows score for each individual SEPI between assessment 1 and assessment 2 separated by a period of 14 days

Table 7. Table 7 shows the measurement error calculations for analysis of test / re-test reliability. *SD = standard deviation; TE = typical error; CI = confidence intervals; SWC = smallest worthwhile change; SWC + TE = smallest worthwhile change + typical error*

Key Performance Indicator	Test 1 Mean (± SD)	Test 2 Mean (± (SD)	TE	95% CI Lower (dif)	95% CI Upper (dif)	SWC Small	SWC Medium	SWC Large	SWC+ TE Small	SWC+TE Medium	SWC+ TE Large
1st Touch SEPI	**75.2 ± 2.2**	**75.4 ± 4.4**	**3.55**	**-4.83**	**4.40**	**1.00**	**2.51**	**4.02**	**4.55**	**6.06**	**7.57**
Forward	16.5 ±2.8	17 ± 7.2	4.02	-5.94	4.50	1.14	2.84	4.54	5.15	6.86	8.56
Side	4.5 ± 1.5	5.1 ± 1.9	1.84	-2.89	1.88	0.52	1.30	2.08	2.35	3.13	3.91
Back	6.3 ± 1.9	6.1 ± 4.4	2.43	-2.95	3.36	0.69	1.72	2.75	3.11	4.14	5.17
Ball Manipulation SEPI	**56.1 ± 5.2**	**54.2 ± 5.3**	**4.43**	**-3.92**	**7.60**	**1.25**	**3.13**	**5.01**	**5.69**	**7.57**	**9.45**
1v1	1.3 ± 1.3	1.1 ± 1.8	1.14	-1.40	1.56	0.32	0.80	1.29	1.46	1.94	2.42
Skill/ Move	4.6 ± 2.1	4.4 ± 4.2	3.26	-4.09	4.39	0.92	2.31	3.69	4.18	5.57	6.95
Drive	0.9 ± 0.4	0.7 ± 0.9	0.72	-0.71	1.18	0.21	0.51	0.82	0.93	1.24	1.55
Passing SEPI	**67.9 ± 6.3**	**66.7 ± 8.1**	**4.30**	**-4.43**	**6.76**	**1.22**	**3.04**	**4.87**	**5.52**	**7.35**	**9.17**
Forward	11.9 ± 3.7	12.2 ± 6.7	5.00	-6.79	6.21	1.42	3.54	5.66	6.42	8.54	10.66
Side	7.8 ±2.0	8.3 ± 4.1	2.66	-4.11	2.81	0.75	1.88	3.01	3.41	4.54	5.67
Back	7.2 ± 2.1	6.7 ± 3.9	2.61	-2.78	4.00	0.74	1.84	2.95	3.35	4.45	5.56
Short	25.7 ± 4.6	26.2 ± 9.5	6.16	-8.51	7.51	1.74	4.46	6.97	7.90	10.52	13.13
Long	1.00 ± 0.9	1.00 ± 1.2	1.29	-1.57	1.78	0.36	0.91	1.46	1.65	2.20	2.75
Penetrating Passes	2.34 ± 1.4	2.18 ± 2.0	1.89	-2.34	2.58	0.54	1.34	2.14	2.43	3.23	4.03
Shooting SEPI	**54 ±4.7**	**53.6 ± 4.0**	**3.84**	**-4.64**	**5.33**	**1.08**	**2.71**	**4.34**	**4.92**	**6.55**	**8.18**
On Target	2.7 ±1.4	3.2 ± 2.4	1.79	-2.09	1.74	0.51	1.26	2.02	2.29	3.05	3.81
Off Target	1.4 ± 0.8	1.4 ±1.3	1.17	-1.61	1.43	0.33	0.83	1.37	1.55	2.06	2.57
Smash	2.7 ± 1.1	2.9 ± 2.2	1.45	-2.22	1.55	0.41	1.03	1.64	1.86	2.48	3.09
Finish	1.5 ± 1.2	1.7 ± 1.9	1.67	-2.36	1.99	0.47	1.18	1.89	2.15	2.86	3.57
Side Foot	1.0 ± 0.9	1.3 ± 1.3	1.21	-1.85	1.28	0.34	0.85	1.37	1.55	2.06	2.57
Instep	3.2 ± 1.3	3.3 ± 2.4	1.63	-2.33	1.91	0.46	1.15	1.85	2.09	2.79	3.48
1st Time	1.1 ± 0.7	1.5 ± 1.3	1.02	-1.65	1.02	0.29	0.72	1.16	1.31	1.75	2.18
Not 1st Time	3.1 ± 1.3	3.2 ± 2.3	1.67	-2.35	1.99	0.47	1.18	1.89	2.14	2.85	3.56
Decision Making											
Best	23.6 ±4.3	24.2 ± 9.6	6.01	-8.86	6.77	1.70	4.25	6.80	7.71	10.26	12.81
Ok	9.7 ±3.2	8 ± 4.5	4.00	-3.47	6.92	1.13	2.83	4.52	5.13	6.82	8.52
Poor	4.5 ± 2.2	5.6 ± 3.0	2.64	-4.41	2.44	0.75	1.86	2.98	3.38	4.50	5.62
Skill Success Under Pressure (%)	79.4 ± 6.2	76.1 ± 12.8	8.24	-7.80	13.62	2.33	5.83	9.32	10.57	14.07	17.57
Unforced Errors (%)	11.7 ± 3.7	13.4 ± 8.0	4.83	-7.92	4.64	1.37	3.42	5.47	6.20	8.25	10.30

Results from the measurement error reliability analysis are presented in table 7. In addition to reported ICCs, the generation of the TE further enhance our understanding of the stability of technical performance in this TEEM protocol between trials *(TE = 0.12-8.24)*. The highest measurement error (TE) in the current study was observed in the *'decision making'* and *'skill success under pressure'* measurement variables (6.01 and 8.24 respectively).

4.4 Discussion

4.4.1 General Discussion

This study aimed to establish the reliability measurement properties of the newly developed TEEM protocol. Results demonstrated that assessing technical performance in a small-sided game format using video and subsequent notational analysis is a reliable method of performance measurement for certain aspects of technical performance in youth footballers. Intra- and inter-observer reliability are the most widely utilised reliability measures, and the high ICC's reported for frequency data are consistent with previous studies (Cobb et al, 2018; Garcia-Lopez et al; 2013; Grehaigne, 1997; Moreira et al, 2017; Oslin, 1998). The moderate to strong reliability reported in two technical actions suggest that when test conditions are consistent and controlled, as in this instance, the TEEM protocol described in chapter 3 is appropriate for assessing and monitoring certain aspects of technical performance in youth football players. Results of this study show poor reliability of ball manipulation and shooting performance between trials. Due to the random and dynamic nature of association football and the match-to-match variation observed in competition performance (Liu et al, 2016), a reliable and easily administered assessment protocol is highly desirable within youth football to provide accurate objective feedback to aid the player development process.

The high test/re-test reliability reported in the frequency data may be partly explained by the nature of the data being collected and analysed in the assessment protocol. Due to the constraints applied to the TEEM protocol (i.e., the consistent time, pitch size and player number) and the fact that the type of data being collected is a frequency count of on-the-ball events, it is probable that the number of on-the-ball events for any individual player within the given timeframe of a 6-minute game will remain consistent between trials with only minor fluctuations occurring possibly due to contextual factors such as motivation and fatigue (Badin et al, 2016). In addition, it is unlikely that habitual player behaviour will change in such a short time-period (e.g., 14 days) due to the long-term nature of technical performance development (Lehyr et al, 2016). Therefore, the author surmises that a combination of a limited and consistent time period in which to perform on-the-ball technical actions and habitual player behaviour will result in similar frequencies of technical actions being observed between trials, and thus partly explains the high ICC's reported in this study.

Furthermore, the lower number of frequency counts recorded in the ball manipulation and shooting actions, as well as in some of the accompanying contextual information for each SEPI, result in more comparable counts between trials and consequently may offer an additional explanation to the high ICC's reported in the frequency data. In contrast to the high test/retest reliability of frequency data reported in the current study, Clemente et al (2019) concluded that small-sided games elicited a high degree of variation between measurement points separated by one week and were therefore not appropriate for reproducing consistent technical stimuli. These results are consistent with certain technical actions in this study (mainly the global ball manipulation and shooting SEPI's). Clemente et al (2019) reported high coefficient of variation values which ranged between 52.2 and 133.8%. However, this high variability as determined by the coefficient of variation perhaps presents a misleading picture due to an important aspect of calculation. The coefficient of variation calculation involves inclusion of the sample mean and standard deviation. It is expected that that the variation in technical performance in U11 football players is high and therefore this would be reflected in a sample standard deviation. Consequently, this would result in a high coefficient of variation since it is a product of the standard deviation and mean. Swinton et al (2018) suggest that a more appropriate measure of within-subject variation is calculation of the TE based on the difference scores between the two trials, as adopted in the current study, which in turn would provide a more complete assessment of test reliability.

In the case of inter- and intra-observer reliability measure, one possible explanation for the ICC values reported in the current study is again the low number of technical actions performed in certain contextual elements of the individual SEPI's and their simple operational definitions that are easily identified by the observers. For example, in the case of the ball manipulation SEPI, one contextual descriptor of a ball manipulation action is an attacking 1v1 situation. From the perspective of the observer, this situation is easily identifiable and is performed less frequently (1.3 ± 1.3 occurrences per assessment) than other technical actions such as a sidewards pass (8.3 ± 4.1 occurrences per assessment). This low number of occurrences, combined with the simple identification of the event, should result in little variation between observers and multiple measurement points within observers. A high number of technical actions with similar properties, would result in almost a perfect match when comparing the same dataset, and consequently in a high ICC, as observed in the current study. In contrast to this, there are also present key performance indicators which have a high number of occurrences and more subjective operational definitions which are open to interpretation. Examples of these are 'decision making' (Gonzalez-Villora et al, 2015) and skill 'success under pressure'. In these circumstances, one would expect a higher degree of variation within and between observers which is reflected in the fact that highest

TE values reported in this current study were associated with decision making and skill success under pressure (table 7).

During the design of this assessment tool, operational definitions were created to reduce the subjectivity associated with the coding procedure. However, it could be argued that assessing decision making in a random and chaotic sport, where any number of options regarding skill execution could be considered the 'best', perhaps the definitions provided are too simplistic to provide an accurate assessment of such a complex key performance indicator. Lorains et al (2013) and Garcia-Lopez et al (2013) both proposed observation tools for assessing decision making in team sports using performance analysis and reported acceptable inter- and intra-observer reliability statistics. However, in the case of Lorains et al (2013), the coding framework and number of measurement variables involved in the analysis would significantly increase the time and labour demands of the observation tool developed, which in turn, would reduce the practicality of the tool in the intended environment. Furthermore, both Lorains et al (2013) and Garcia-Lopez et al (2013) failed to report any measure of test/retest reliability and therefore, further investigation is required before practitioners can be certain that the objective data generated is truly accurate. Similarly, the *'skill success under pressure'* measurement variable also poses the same problem. The subjectivity associated with what constitutes 'opposition pressure' may explain the higher TE observed within this key performance indicator. Competition constraints, such as 'tight' or 'loose' defensive pressure, have been considered during previous reports when designing tools to assist with technical performance analysis (Tenga et al., 2009). Despite the acceptable reliability statistics reported by Tenga et al (2009), it remains difficult to ascertain whether a player feels 'under pressure' from an opposing defender as this is likely to vary from individual to individual. Furthermore, the proximity of an opposition player to the player in possession may not be the best indicator of opposition pressure. The opposition players intention to recover the ball, block passing angles and their intensity of pressing are all variables which should be considered during analysis. Further investigation into this measurement variable is required to construct a clear definition with appropriate contextual parameters, which in turn could improve the reliability of this key performance indicator. However, due to the potential and significant impact that opposition pressure could have on technical performance, the authors conclude that including it in the analysis process is warranted.

Results provide reliability measures which have been identified as critical stages in the development of new observational tools for measuring technical performance (Robertson et al, 2017). The generation of essential reliability statistics in this study could be a welcome addition to the existing body of research in this topic. Only a few previous studies involving the development of new observational tools for assessing performance have reported reliability statistics (Robertson et al, 2014). Previously, Oslin et al (1998) and Grehaigne (1997) reported test/retest reliability measures using an intraclass correlation coefficient. They provided early indications of the potential performance analysis has for assessing technical performance within competition. In contrast, more recent studies have focussed the attention of their reliability measures on the inter- and intra-observer reliability of their newly developed observational instruments (Cobb et al, 2018; Garcia-Lopez et al, 2013; Moeira et al, 2017; Waldron & Worsfold, 2010). Although these are essential measurement properties, they are only a single step in the necessary requirements for establishing scientifically robust measurement tools. One feature that is consistent in the research studies cited here is a failure to report measurement error associated with the observation and or the smallest worthwhile change (SWC) for key performance indicators. These statistics provide essential information for monitoring performance over time, which is a key objective of the proposed observational tool in this study. By performing the same test procedure at regular points during player development over a prolonged period, we can identify any improvements (or regressions) in technical performance. The availability of this objective information can provide coaches and practitioners with valuable data regarding the development of each individual player and even the impact of the coaching curriculum on a team basis. Reporting of the TE and SWC allows for the identification of any meaningful changes in technical performance over time – a crucial process in the development of young academy footballers. Establishment of the TE allows the assessment tool user to recognise the measurement error associated with it and subsequently, by combining TE with the SWC, conclusions about meaningful changes in technical performance which lie outside the TE + SWC parameters can be drawn.

In summary, results of the reliability analysis conducted in this study reveal moderate to good reliability in the 1st touch and passing SEPI's respectively and poor reliability in the ball manipulation and shooting SEPIs. Additionally, the intra- and inter-reliability analysis demonstrated excellent reliability within a repeated observation of the same video footage by the same observer and between observers analysing the same video footage. The frequency count data demonstrated excellent reliability between two measurement trials separated by 14 weeks, however, these results should be interpreted with caution given the previously stated potential explanations for this regarding low frequencies in some technical actions. This study, to the authors knowledge, is one of the first to report measurement error (TE) and SWC in a newly developed observational tool for assessing and monitoring technical performance in youth footballers. In this study, only the large SWC is greater than the TE in most variables. In an ideal situation, the SWC must be greater than the

TE to ensure any change in performance is a result of actual improvement in performance rather than due to measurement error or variation. This may suggest that measuring technical performance in the TEEM is insufficient for detecting small changes in performance. By taking this into consideration, the authors propose that either the large SWC is used to interpret results or the TE and SWC are combined to ensure that changes beyond the measurement error are likely to be meaningful changes in performance in the landscape of long-term player development, the author suggests that data is collected over an extended monitoring period where greater improvements (or decrements) in performance will occur. Further research is required to establish if the current observational tool is sensitive enough to detect changes in technical performance over time.

4.4.2 Limitations

Despite the good reliability statistics reported in certain aspects of the TEEM protocol in this study, several limitations have been identified. Firstly, the random team re-configuration after each match may result in players being required to play in un-natural positions on the field. This in turn may affect the players behaviour when in possession of the ball which may give an inaccurate portrayal of player habits. Perhaps, the configuration of the teams should involve a process whereby playing positions are accounted for to ensure a well-balanced team structure. This would give players the opportunity to play in their preferred positions during all games even if personnel are rotated. As a result, variation in the contextual parameters of skill execution, such as direction of passes, could be reduced. Secondly, more work should be conducted to reduce the subjectivity of some key performance indicators such as decision making and skill success under pressure. Further work is required to develop clearer operational definitions, which contain more contextual information that cover all situations observed during match play. For example, when coding for decision making, additional information on whether the decision was ‘best, ok or poor’ due to incorrect option chosen, incorrect selection of skill or any variation of these behaviours could provide more meaningful feedback to aid player development. Lastly, one important factor that should be considered is the individual biological maturity of the players involved in the assessment. Moreira et al (2017) demonstrated that hormonal status plays an influential role in technical performance in youth footballers. One possible explanation for this is players who are at earlier stages of maturation may alter their behaviour when playing amongst players who are at a later stage of maturation. For example, a smaller player whose natural behaviour is to take players on in 1v1 situations may have to alter their behaviour due a deficit in physical power when accelerating away after a skill/ move. Therefore, one adjustment to the current protocol which could alleviate this issue is by grouping players according to their biological age as opposed to their chronological age (Bradley et al, 2019).

As a result, a possible reduction in variation between trials may be observed, however, further investigation into this methodological approach is required.

4.4.3 Practical Applications

To the authors knowledge this study is one of only very few studies to report the test/retest reliability and associated measurement error of a simple and easily administered protocol for assessing and monitoring technical performance in youth footballers over time. The generation of the typical error and smallest worthwhile change enables identification of meaningful changes in performance over time. Whilst the establishment of typical error and smallest worthwhile change is a positive step in the development of this assessment tool, interpretation of performance changes must take into consideration the measurement error associated with performance between trials. The typical error reported in this study is often greater than the SWC and medium SWC which suggests that the assessment protocol may be insufficient for detecting small changes in performance. In spite of this, the objective feedback can add value to the player appraisal process which will support the traditionally subjective opinion of coaching experience. Furthermore, the breakdown of a player's technical performance provides valuable information during the goal setting process in the formation of individual player development plans which could subsequently stimulate players to engage in focussed deliberate practice. Given the poor reliability of ball manipulation actions, shooting actions and some contextual key performance indicators, practitioners should decide whether or not these should be included in the analysis. One argument for including these technical actions is that they could be considered 'game changing actions'. For example, although actions such as attacking 1v1's and shooting ability are highly variable, is the ability to be successful in this situation valued even if success is sporadic? And is part of the development process about practicing the consistency of these actions and reducing variability?

4.5 Personal Reflection and Professional Skills Development

During this stage of my research, I was required to engage the learning of a new skill. An essential component of my newly developed assessment tools design involved notational analysis through the use of the performance analysis software SportsCode, of which I had no previous experience. In the early stages, using this software was frustrating as I was a complete novice. I felt that the software was very advanced and that I would struggle to generate the results I needed in a time efficient manner. Below are two extracts from my reflective diary which convey my feelings and anxieties at the time and how my views began to change after a period of perseverance:

> "*Had a go with SportsCode trialling coding parameters which I could potentially use for the tool. I felt that this process was over-complicated and perhaps a bit more*

complicated than it needs to be. I also had a go at trying the coding process using a hand notation system. I found that this was very labour intensive as it only allowed for one player to be analysed at a time. After using Sportscode I felt that the data output was also very messy and required a lot of work to tidy it up so that it could be worked with. I am beginning to fear that the tool may be impractical when timely assessment of many players is required. Furthermore, I am beginning the question the whole process… who cares about this data? Will it make a difference? Will the coaches/ managers value the information that is being provided? Is it worthwhile?"

"I was feeling a bit disheartened after yesterday's trial. I decided to have a play about with SportsCode again to see if I could find a way to speed up the process. I realised what the problem was… I had the video in 6 games instead of one big video file. I tried it with this method and tagged 3 players at the same time. It took me about 20 minutes and the data output was brilliant, it required no manipulation or tidy up so that it could be presented. This was a good break through. Furthermore, after I had a look at the data it gave me, I had a look back at previous data gathered a year or so ago (before the beginning of my prof doc) and the performance index method seemed a lot more stable than percentage success which I had found in the research. The variation was much lower, so I think this is the way to go. I feel a lot happier with today's breakthrough and am now excited about the process. I showed an example of the data it could provide about the players to my Head of Coaching and he also was quite excited about the potential of objective feedback that we could provide with the players. Also, we had a discussion about how we could use the tool if I can prove its validity and reliability. After reviewing the data with the Head of Coaching and Head of Youth, it provided them with some evidence which supported their subjective opinion on one of the players who they were on the fence about whilst deciding who to offer full time contracts to. Although, at this stage I recommended that the data should not be used to base decisions of this ir magnitude on, they said that it confirmed their opinion and validated it. As a result, I think the player may be offered a full-time contract to begin in the summer. This was the kind of thing I wanted the tool to be capable of, so that players who have potential aren't let go and overall, I was proud of this"

As my project progressed, I began to feel more confident with the software. The process of continually learning and watching YouTube tutorials allowed me to explore more advanced features. Learning to use this software effectively has turned out to be extremely beneficial as in my current role (at the time of writing this), I am using the concept of objective data for player and team development at a first team level where I am utilising this software during analysis to provide performance insights to coaching staff.

In reference to the research and professional aims and objectives set in Chapter 1, the professional learning outcomes obtained from this stage of my project were identified as:

- To develop research and technical skills
- To develop transferable knowledge, concepts and processes which can be utilised in other contexts

CHAPTER 5: Establishing the Criterion-Based Validity of a Newly Developed Assessment Tool for Assessing and Monitoring Technical Performance in Youth Footballers

5.1 Introduction

The technical demands of elite football performance have increased exponentially over the past decade (Barnes et al, 2014). Previous research has demonstrated that a significant correlation exists between the technical performance of a team and subsequent team success (Liu et al, 2016; Rampinini et al, 2009). Despite the availability of this information, research investigating the assessment and monitoring of technical performance in both youth and senior players is under-represented compared with other domains of research within the sport (Reilly & Gilbourne, 2003). One reason for this disparity could involve the complexity of technical performance, and the difficulty with measurement in a continuous sport that is highly dynamic and spontaneous in nature. However, recent advances in technology have provided coaches and practitioners with tools to objectively measure technical performance within the sport (Farrow & Robertson, 2017). The availability of such technology presents researchers with an opportunity to investigate this relatively un-explored area of performance and further enhance our understanding of technical performance to aid player development. Early models of skill assessment, derived in physical education literature, have provided frameworks upon which new models, utilising modern technology, can be developed (Grehigne, 1997; Oslin et al, 1998).

Within previous research, numerous methods have been designed to assess technical performance in football (Lehyr et al, 2018; McDermott et al, 2015; Waldron & Worsfold, 2010). Each assessment tool varies significantly in its methodological constructs and represents both 'closed' skill tests, performed in isolated test conditions outside of competitive performance, and 'open' skill tests, performed within a simulated variation of the competition environment. Although this research presents a significant contribution to the existing literature, several aspects of research methodology must be taken into consideration before any tool can be confidently administered within a practical setting (Robertson et al, 2014). One important aspect of research methodology which has been overlooked in most studies relates to criterion-related validity and the relationship, if any, that exists between performance in a test and in competition. For a performance test to be truly valid within a practical setting, practitioners and coaches must ensure that what they are measuring and monitoring in vivo, is correlated to the intended performance outcome, in this case competition performance. To the authors knowledge, there are limited research studies that have investigated the correlation between technical performance in a test and competition technical performance in football.

Rubajcyk and Rokita (2015) reported that young football player's performance in two isolated dribbling performance tests did not correlate well with technical performance in a 5 v 5 small-sided game scenario when judged by a panel of experts. Two possible explanations for this could involve the subjective method of performance assessment during the small-sided game scenario and the lack of specificity of the football-specific tests selected to competition performance (Bergkamp et al, 2019). The method of performance assessment utilised within the small-sided game protocol encompassed an opinion-based appraisal of performance by expert judges with a co-efficient of variation ranging between 34.12 and 37.86%. The high variability observed in this method of assessment suggests that a more reliable and objective method of performance assessment is desired. In relation to the selection of football-specific performance tests, various authors have reported that a lack of requirement for decision making, information gathering from player surroundings, and lack of opposition pressure during isolated performance tests does not reflect performance in competitive situations (Bergkamp et al, 2019; Vaeyens et al, 2008). In support, Rubajczyk and Rokita (2015), and Wilson et al (2017) reported a non-significant relationship between performance in isolated football-specific performance tests and individual match efficiency (successful skill attempts/number of attempts) in competitive 11 vs. 11 games. Rubajczyk and Rokita (2015) suggested that the lack of relationship could be due to the external focus demand during 'closed skill' performance tests compared to competition demands. Furthermore, 3-year longitudinal data from 1134 youth players on a German TID programme reported that performance in isolated passing, dribbling, ball control and shooting tests did not successfully discriminate between adult performance level (Leyhr et al., 2018). Although a multitude of factors contribute to the 'elite' adult performance levels that lie outside the facet of technical performance, they concluded that their selected test battery may not correlate well with competition performance and hence, may be a poor predictor of performance. Information generated from these studies provide valuable insight into performance and highlights the need for the development of an assessment protocol that better reflects competition performance. It is likely that the absence of external stimuli during 'closed test' protocols plays a critical role in the predictive value of these tests to competition performance, which can be complex and chaotic in nature.

In top-level professional football, technical performance within competition can now be monitored thanks to advancements in modern technology (Liu et al, 2016). Longitudinal technical performance tracking within competition could provide a feasible method of assessment which will empower key stakeholders with valuable insight into individual and team performance. However, despite the availability of this technology and service provision in the professional game, at present, no such service is available within youth football due to financial and operational constraints. Consequently, to maximise player development, it is of critical importance that an assessment tool is developed that provides meaningful information and has a strong predictive value towards competition performance. Due to the substantial variation in technical performance during competitive match performance (Liu et al, 2016), and the associated constraints highlighted within academy football for match data collection, training performance could provide a feasible method for assessing technical performance. However, before application of such an assessment within a training context, work must be done to explore the relationship (if any) that exists between training and competition performance and to ascertain if training performance carries any predictive value towards performance in a competitive situation. The use of small-sided games has become a popular training modality within football (Sarmento et al, 2018). It has been extensively documented that small-sided games can successfully replicate technical, physical and tactical competition demands (Dellal et al, 2012). Within existing literature, it has been suggested that the use of small-sided games may be a suitable method of assessing technical performance due to their similar demands compared with competition performance (Bennett et al, 2017), however, further investigation into their measurement potential is required. Therefore, the aim of this study is to explore the relationship between technical performance observed in a small-sided game assessment protocol and in competition. It was hypothesized that a relationship exists between technical performance in the TEEM protocol and competition performance.

5.2 Methods

5.2.1 Experimental approach to the problem

The purpose of this study was to explore the relationship between technical performance in a small-sided game technical performance assessment protocol (TEEM) and technical performance in competition. This study intends to establish criterion-based validity, which represents one stage in the development of new systematic observation tool for assessing technical performance in youth football (Brewer & Jones, 2002; Robertson et al, 2017). Technical performance in the TEEM protocol was assessed using the methods outlined in Chapter 3.

5.2.2 Video Recording Data

Permission to access the video recording data of a Scottish Premier League Football Academy was requested. As part of the club's development programme, TEEM's are performed bi-annually to provide objective feedback to all academy players. All technical assessments and competitive matches are video recorded by the club's academy performance analyst and stored on the club's internal database. Data from one technical assessment and match footage from 3 competitive matches were used from 21 academy players (aged 17 ± 1.6 years). Test/retest reliability of the TEEM was established in chapter 4 and therefore, due to time constraints, the decision was made to use only 1 technical assessment for analysis. Only players who had completed the technical assessment and participated in 3 full 90-minute competitive matches were included in the analysis.

Due to the variability observed within performance in competitive matches (Liu et al, 2016), McKenzie and Cushion (2013) recommend that a minimum of 3 matches are required for analysis to account for the contextual variation in performance. Examples of these contextual factors which play an influential role in competition performance include opposition quality, team tactics and environmental conditions. In line with academy procedure, all academy players and parents are required to sign an informed consent document which ensures that all players and parents give permission to be photographed and filmed within academy training and matches.

5.2.3 Coding Procedure

The coding procedure and operational definitions for football actions used for analysis in both the technical assessment and the competitive matches are outlined in chapter 3. Acceptable test / retest reliability and inter / intra-observer reliability for the TEEM was previously established (chapter 4). The same code window (Hudl Sportscode, Elite Version 10) was used for the technical assessment and match analysis to ensure continuity between performance measures and observer training was completed prior to the analysis procedure to ensure observer competence.

5.2.4 Statistical Analysis

Frequency data was generated from the analysis software and was used for subsequent analysis. The average number of actions performed over the 3 competitive matches was used. The technical performance frequency data from the TEEM and competitive matches were compared. The frequency of technical actions was selected for comparison between the assessment protocol and the competitive game scenario for two reasons: 1) to ensure that all key performance indicators were being recorded in both scenarios and 2) to compare any differences in player behaviour in the two different contexts. For example, does a player play with more freedom during the assessment protocol in a training environment compared with a competitive situation where there is more pressure to perform which will consequently this will affect their habitual behaviour. In addition, the 'Skill Efficiency Performance Index's' (SEPI, *see chapter 3)* in the technical assessment and in the competitive matches were compared. Furthermore, to assess differences in the frequency of actions between performance in the TEEM and competitive matches, the number of actions performed per minute was calculated. This was done by dividing the number of actions attempted by the total playing time (in minutes) of each scenario (36 min for TEEM and 90 min for competitive matches). All data were analysed using the SPSS statistical analysis software package (v.16; SPSS Inc., Chicago, IL). If no successful or unsuccessful actions were recorded in the 4 main SEPIs in either the assessment protocol or over the 3 competitive matches, data were excluded from the analysis. This resulted in a sample size of only 14 in the shooting SEPI. Actions were observed across all other SEPIs. Data were tested for normality using the Shapiro-Wilk normality test ($P > 0.05$). Results of the normality indicated that the frequency data was not normally distributed and therefore selection of a non-parametric test for subsequent analysis was required. In contrast, the Shapiro-Wilk normality test revealed that the SEPI data followed a normal distribution. For the frequency data, a Spearman's Rank Correlation test was used and for the SEPI data, a Pearson's Correlation test was used to investigate the relationship between performance in the technical assessment protocol and performance in a competitive match. The strength of the correlation was interpreted using the guidelines recommended by Evans (1996) where r = 0.00-0.19 - very weak; 0.20-0.39 - weak; 0.40-0.59 - moderate; 0.60-0.79 - strong and 0.80-1.00 - very strong.

5.3 Results

5.3.1 Frequency Data

A Spearman Rank correlation coefficient analysis between the number of key performance indicator actions performed in the technical assessment protocol and over an average of 3 competitive matches was conducted (figure 9). Results of the analysis identified a moderate correlation (r = 0.42 ± 0.17; *range = 0.10 - 0.71*) between the frequency of actions performed in the assessment protocol and the average frequency of actions performed over 3 competitive matches in youth footballers. The frequency of actions performed in the technical assessments, the average frequency over 3 competitive games, the associated correlation coefficients and the number of actions performed per minute can be seen in table 8. The frequency of actions performed per minute was greater in TEEM protocol compared with the frequency of actions performed over the 3 competitive matches (0.22 actions per minute vs. 0.06 actions per minute respectively) despite the greater length of time spent in competitive matches compared with the assessment protocol (90 vs. 36 min respectively). The percentage of actions attempted under pressure was greater during competitive matches compared with the TEEM (51.2 vs. 46.5% respectively)

5.3.2 Skill Efficiency Performance Index (SEPI) Data

The SEPI's for each of the 4 key performance indicators (1st Touch, Ball Manipulation, Passing and Shooting) in the TEEM and the average over 3 competitive matches can be seen in table 9. Pearson correlation analysis revealed a trivial to strong correlation (figure 10) between technical performance in the TEEM and the average technical performance over 3 competitive matches (r = 0.05 - 0.73). The strength of the correlation in each of the individual SEPIs between the TEEM and the average over 3 competitive matches can be seen in table 8. In 3 of the 4 SEPIs, performance was higher in the TEEM compared with the average over 3 competitive matches.

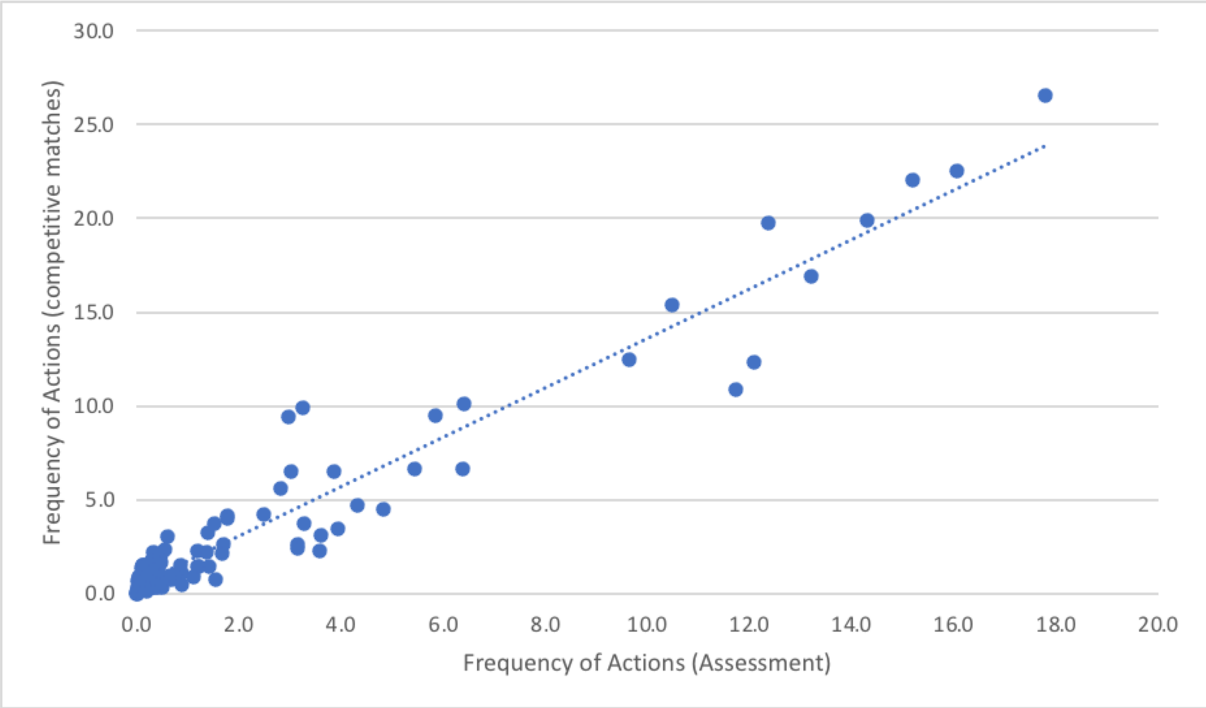

Figure 9. Frequency of actions performed in each of the key performance indicators in the TEEM (*n* = 1) compared with actions performed during competitive matches (*n* = 3) (*r* =0.42 ± 0.17).

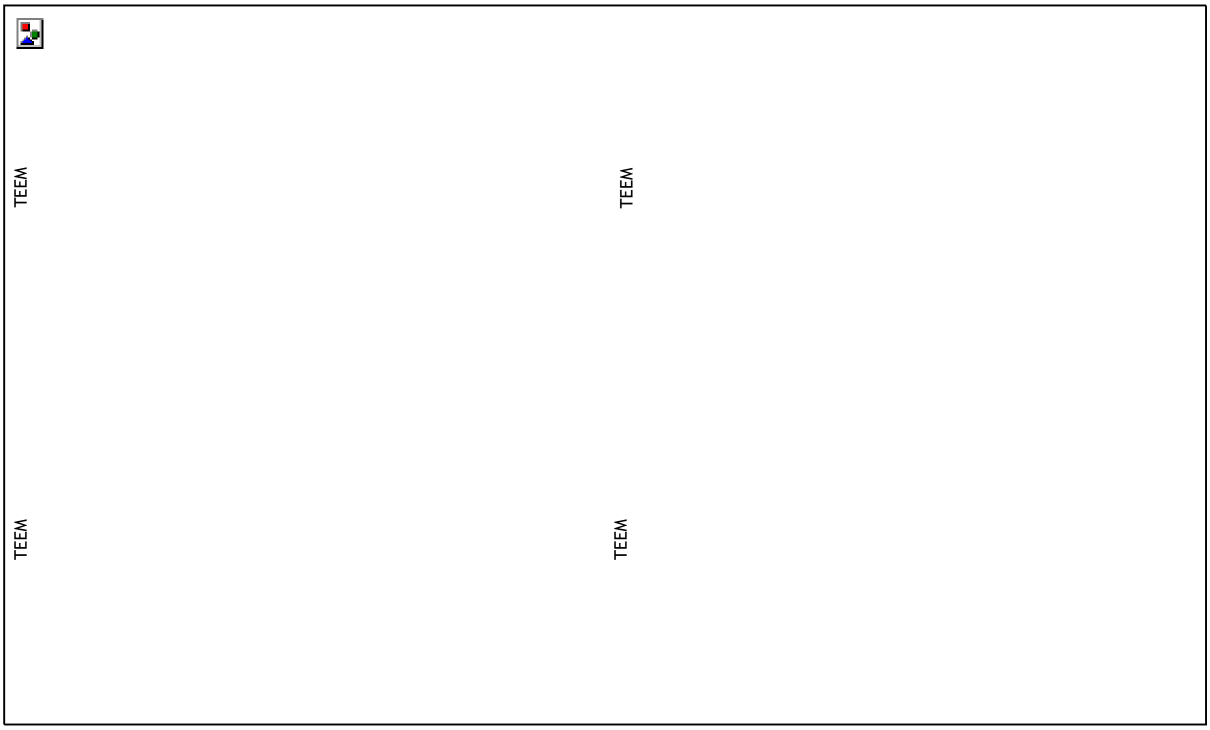

Figure 10. The individual SEPI's observed in the TEEM compared with the average SEPI's observed over 3 competitive matches (*r = Pearson Correlation Coefficient)*

Table 8. Differences in the frequency of actions performed in the TEEM vs. the average of 3 competitive matches and correlation coefficients. Significant correlation coefficients are represented by * *($p < 0.05$)*

Key Performance Indicator	Assessment Mean (±SD)	Competitive Matches Mean (± SD)	Correlation Coefficient	Assessment (actions/ min)	Competitive Matches (actions/ min)
1st Touch Successful	**26.6 ±7.7**	**17.8 ± 7.1**	**0.64 ***	**0.74**	**0.20**
1st Touch Unsuccessful	**1.4 ± 1.4**	**1.4 ± 1.0**	**0.22**	**0.04**	**0.02**
Forward	16.1 ± 6.0	11.2 ± 4.7	0.50 *	0.45	0.12
Side	5.0 ± 2.3	4.7 ± 2.1	0.17	0.14	0.05
Back	6.8 ± 3.7	3.4 ± 2.9	0.71 *	0.19	0.04
ll Manipulation Successful	**4.2 ± 3.0**	**1.8 ± 1.8**	**0.55 ***	**0.12**	**0.02**
Manipulation Unsuccessful	**1.0 ± 1.1**	**0.8 ± 1.0**	**0.52 ***	**0.03**	**0.01**
1v1	1.3 ± 1.6	0.3 ± 0.7	0.61 *	0.04	<0.00
Skill/ Move	3.6 ± 2.9	1.8 ± 1.6	0.61 *	0.10	0.02
Drive	1.3 ± 3.1	0.4 ± 0.7	0.34	0.04	<0.00
Passing Successful	**22.5 ± 5.5**	**16.1 ± 5.9**	**0.57 ***	**0.63**	**0.18**
Passing Unsuccessful	**4.5 ± 2.4**	**4.8 ± 2.4**	**0.26**	**0.13**	**0.05**
Forward	12.9 ± 5.1	9.8 ± 4.6	0.42	0.36	0.10
Side	7.6 ± 2.7	7.0 ± 2.7	0.42	0.21	0.07
Back	6.6 ± 3.2	4.1 ± 2.3	0.32	0.18	0.02
Short	25.8 ± 5.6	18.4 ± 6.5	0.32	0.72	0.20
Long	1.2 ± 1.2	2.4 ± 1.9	0.48 *	0.03	0.03
Penetrating Passes	4.0 ± 3.6	2.9 ± 2.3	0.58 *	0.11	0.03
Shooting successful	**1.5 ± 2.0**	**0.1 ± 0.3**	**0.60 ***	**0.04**	**<0.00**
Shooting Unsuccessful	**3.1 ± 2.2**	**0.6 ± 0.8**	**0.27**	**0.08**	**0.01**
On Target	2.8 ± 2.9	0.4 ± 0.6	0.57 *	0.08	<0.00
Off Target	1.8 ± 1.5	0.3 ± 0.6	0.30	0.05	<0.00
Smash	2.8 ± 2.7	0.5 ± 0.6	0.36	0.08	<0.00
Finish	1.8 ± 1.6	0.3 ± 0.6	0.35	0.09	<0.00
Side Foot	1.2 ± 1.1	0.1 ± 0.3	0.51 *	0.03	<0.000
Instep	3.2 ± 3.1	0.6 ± 0.8	0.31	0.09	0.01
1st Time	1.6 ± 1.6	0.3 ± 0.5	0.10	0.04	<0.00
Not 1st Time	2.9 ± 3.3	0.5 ± 0.6	0.58 *	0.08	0.01
Decision Making					
Best	25.9 ± 8.9	16.8 ± 6.0	0.61 *	0.72	0.18
Ok	7.0 ±3.6	3.6 ± 2.2	0.14	0.19	0.04
Poor	4.0 ± 2.5	3.8 ± 2.0	0.28	0.11	0.04
Pressure					
Tight	30.1 ± 10.8	22.2 ± 10.7	0.66 *	0.84	0.25
Loose	34.6 ± 9.0	21.2 ± 9.5	0.39	0.96	0.24
% of Actions Under Pressure	46.5	51.2	-	-	-
Unforced Errors	**2.5 ± 1.7**	**2.2 ± 1.3**	**0.17**	**0.07**	**0.02**

5.4 Discussion

5.4.1 General Discussion

This study, to the authors knowledge, was one of very few to explore the relationship between technical performance in a small-sided game-based assessment protocol (TEEM) and performance during competitive matches. This study attempted to establish the criterion-based validity of the TEEM and support the use of small-sided games as a means of assessing technical performance in youth football (Bennet et al, 2017). This study reports a moderate correlation between the frequency of actions performed in the TEEM and the average number of actions performed across 3 competitive matches and a moderate correlation between the SEPI score achieved in the TEEM and the average SEPI score achieved across 3 competitive matches. The wide range of correlation coefficients observed in both the frequency data and the separated SEPI data in this current study perhaps highlights the complexities in assessing human behaviour in a random and chaotic environment.

One factor that may contribute to the varying strengths of correlation coefficients observed in this study is the low prevalence of occurrences in some of the actions/behaviours within both the TEEM and in a competition environment. For example, if there is a consistently low number of actions recorded across all participants in both the TEEM and in the competition environment (in particular the contextual information accompanying the shooting and ball manipulation technical actions), then consequently, this will result in a strong correlation between the two measures due to the systematic process of rank ordering which is applied during the statistical correlation test.

Table 9. Table 9 shows the descriptive statistics for the SEPI score in the TEEM and the average SEPI score over the 3 competitive matches. *SD = standard deviation*

SEPI	Assessment Mean (± SD)	Competitive Matches Mean (± SD)
1st Touch	75.7 ± 4.6	70.1 ± 4.6
Ball Manipulation	55.8 ± 4.7	52.0 ± 3.0
Passing	68.9 ± 5.6	63.3 ± 5.3
Shooting	46.9 ± 3.3	48.3 ± 1.33

Furthermore, in technical actions/behaviours where frequency counts are low, it is difficult to draw meaningful conclusions about how performance in one situation relates to another if there is limited information (low frequencies) on the proficiency of that technical action. This was the reason for the inclusion of the SEPI. As a result, a misleading interpretation of the true strength of the relationship is provided. To accommodate for this phenomenon during analysis of frequency data, it is of critical importance that each measure records an adequate number of occurrences to ensure that true inferences about the strength of the relationship can be postulated (Hughes et al, 2001). A similar study conducted by Cobb et al (2018) reported similar limitations when designing a newly developed small-sided game-based protocol whereby specific technical and tactical events with low a frequency of occurrences reduced the reliability of their tool. The process of technical actions/ behaviours stabilising over time is known as 'normative profiling' and is predicated on the fact that actions which occur with a higher frequency are associated with less 'noise' and therefore any analysis can be conducted with more precision (O'Donoghue, 2005). Although, it is evident that the TEEM in this study allowed for greater repetition of actions compared with a competitive match (table 8), some technical actions require substantially more data to be collected throughout subsequent observations in the TEEM and/ or competitive matches for performance to 'stabilise'.

In the current study, the global SEPI scores were selected as a performance measure to account for the low prevalence of certain technical actions. Memmert and Harvey (2008) recommended the use of a performance index over the use of percentage success to reduce the variation observed in scores due to low frequency counts. For example, if a player performs one shot on goal and this shot is successful, the corresponding score in the percentage success model would be 100%. If the same player were to attempt 10 shots, it is unlikely that the percentage success (in this case 100%) will remain constant. By adopting the SEPI, the corresponding value for the same situation would be 52.4. As the player achieves more successful shots, the SEPI would increase in smaller, gradual steps. As the player achieves more successful shots, the concomitant and steady rise in performance score also reflects our confidence in true performance levels. These smaller fluctuations in performance score are also associated with lower variation. When the SEPI scores are compared in the TEEM compared with competitive match play, the correlation coefficients reveal a moderate to strong correlation 3 of the 4 SEPIs (1st Touch, ball manipulation and passing) and a very weak correlation in shooting. One explanation for the very weak correlation in shooting could relate again to the low frequency of occurrences in match play compared with the TEEM (table 8). In football, there is a consistent trend in a greater number of unsuccessful shots compared with successful shots, as scoring goals could be considered the most difficult part of the game. With more opportunities to shoot during the TEEM, it is perhaps expected that the shooting SEPI is lower (due to more unsuccessful shots) compared with competitive matches where there is a significantly lower number of shots per match. With the exception of the shooting SEPI, greater stability is offered in the global SEPI scores when compared with the contextual measures accompanying the SEPIs (e.g., direction of pass) as measured by frequency counts that present varying strengths of correlation to match performance.

To the authors knowledge, there is limited existing research investigating the relationship between performance in a small-sided game-based scenario and competition performance. The investigation of this relationship represents a critical step in establishing the criterion-based validity of any assessment tool designed for measuring technical performance (Bergkamp et al, 2018). Bergkamp et al (2020) investigated this relationship and employed a similar research design to this study. They reported moderate correlations in 6 of the 9 selected performance indicators and trivial/small correlations in the remaining 3. Whilst the key performance indicators selected for analysis were different between the study by Bergkamp et al (2020) and the current study, Bergkamp et al (2020) reported similar correlation coefficients for forward passes (0.38 vs. 0.42 respectively) and a small difference in shots on target (0.38 vs. 0.57 respectively). However, two main significant differences exist between the current study and the study by Bergkamp et al (2020). Firstly, the number of key performance indicators selected within the current study was more than 3 times that in the study by Bergkamp et al (2020). In reflection, the contextual information surrounding each action was implemented in this study to try and capture player behaviour and playing habits, however, the low frequency counts in some of the performance indicators may only add to further 'noise' and a lower correlation to match performance instead of adding value. It may be the case that a lower number of specific key performance indicators, as adopted in the study by Bergkamp et al (2020), may be adequate to provide the necessary value. Secondly, the relative playing area per player was significantly different in both studies. In the current study, the pitch size ($100m^2$ / player) was selected to replicate game demands as close as possible following the recommendations by Frauda et al (2013). When using the same extrapolation of pitch sizes as suggested by Frauda et al (2013), the relative playing area per player in the study by Bergkamp et al (2020) would be $184m^2$ / player which may inadvertently affect player behaviour due to different playing constraints such as the space to play longer passes. Perhaps future investigations should apply the key performance indicators (with accompanying contextual information) adopted in the current study to ascertain if the larger pitch will observe changes in player behaviour.

Results of the current study demonstrate a trivial to strong correlation between technical performance in the TEEM and during competition. However, closer examination of the individual key performance indicators reveals a wide range of variation in the correlation coefficients. Although not all technical actions/behaviours performed in a small-sided game environment are strongly correlated to match performance, this study contributes a step towards understanding how technical performance in a training scenario transfers to the criterion performance. It is of the authors opinion that although small-sided games present a practical and viable option for assessing technical performance with moderate task-representativeness, they fail to replicate the behavioural technical demands and intensity of competition entirely. This difference in playing intensity is perhaps supported by the higher percentage of actions attempted under pressure observed during competitive match play compared with the TEEM (table 8).

5.4.2 Limitations

Despite the varying strengths of correlation between individual technical performance measures in the TEEM and competition performance observed in the current study, three main limitations were identified which could have improved the correlation coefficient further. Firstly, the collection of more data would have given the technical actions/behaviours a higher frequency count which in turn would have given them the opportunity to 'stabilise'. Realistically, this data collection process would have extended beyond three competitive matches and would also encompass a higher number of observations in the TEEM. In the study by Bergkamp et al (2020), players participated in between 11-17 small-sided games in comparison to just 1 in the current study. Furthermore, Hughes et al (2004) suggest in their field hockey example that it may take as many as 30 matches for the number of shots key performance indicator to stabilise. Secondly, participants in the video footage were senior academy players (aged 15+), which limits the generalisability of the findings to younger players, as it is possible that younger players elicit greater variation in performance due lower skill levels. Lastly, the effect of growth and maturation was not accounted for during the TEEM or competition performance. Although it is possible to group players according to their biological age rather than their chronological age, it is impossible to control for the opposition team during competition without the application of a performance strategy detailing specific rules and regulations from the governing body. Therefore, it is possible that later maturing players display lower performance levels when competing against earlier maturing players which does not reflect their true performance capabilities.

5.4.3 Practical Applications

Findings from this study suggest that the TEEM may offer a practical and feasible method for assessing and monitoring technical proficiency as measured by the global SEPIs discussed throughout this thesis. The author recommends that when implementing the TEEM protocol, measures should also be taken to ensure playing intensity is as high as possible to better reflect competition performance. These measures could include playing for individual points (Unnithan et al, 2012) or including the assessment as part of a (bi)annual player report. However, the limitations of the TEEM should be considered when assessing the transfer of player behaviours and playing habits due to the high level of variation in strengths of correlation between training and competition. If individual player learning through the identification of specific behaviours is to be assessed and monitored, then the author suggests that this should be conducted within competition performance. Furthermore, the author recommends that the data collection process should be performed on a regular basis and should be extended over a longitudinal period. Lastly, data collected from the TEEM should not be used on its own, for the benefits of player appraisal and selection/deselection but should be part of a holistic process which is supported by additional methods such as match performance and subjective coach opinion (Seighartsleitner et al, 2019). Future research should longitudinally investigate the TEEM's ability to detect change over time which would further enhance the tools scientific robustness and advocate its use within an applied setting.

5.4.4 Conclusion

This study demonstrated that technical performance in the TEEM protocol showed a moderate correlation to competition performance over 3 matches. Results of this study established the criterion-based validity of the measurement tool and provided one important step in the process of developing a new systematic observation tool for assessing and monitoring technical performance in youth footballers.

5.5 Personal Reflection and Professional Skills Development

A new quality assurance measure created during the Scottish Football Association's restructuring and introduction of an increased level of standards within professional academies involved the requirement to issue each individual player with an annual evaluation and personal development plan. This stage of my research project came at an ideal time and would allow me to disseminate the information I had gathered to coaching staff, players and parents. To this point, I was apprehensive about using the data for this purpose until I had established the necessary validity measurement properties and was confident about the methodological robustness of the tool. Therefore, by establishing the criterion-based validity in this study around the same time as the individual player reports were due to be issued, I felt this would be a perfect time to present the results to coaches, players and parents. Furthermore, I felt that any critique of the process offered by any of the key stakeholders would help me develop my tool further.

After a conversation with my Academy Manager and Head of Coaching, we agreed that I would take responsibility for providing objective data about technical performance based on the results of my assessment of technical performance using my tool rather than the subjective opinion of the coaches which often incites contrasting opinions from other coaches and parents. Following this conversation, I issued reports to every academy player detailing the information I had generated from my assessment protocol. In reflection, this process allowed me to target several of the research and professional aims and objectives outlined in chapter 1:

- I was required to explain the process in front of all academy parents and coaches in presentation format (around 100 people)
- Secondly, I had to compile the reports and present them in a format which I thought would be most appropriate

Whilst presenting the process by which the data was collected and presented (to around 100 people), I felt comfortable and relaxed which was unexpected as normally I am very nervous when presenting. Overall, I felt the event went well and that this was a positive step towards developing my ability to communicate via public speaking. After issuing reports to all academy players, the feedback I received was that the reports were 'too scientific' and the parents were struggling to understand them, never mind the kids! After this feedback, I realised they were correct and decided to try another approach. Consequently, I came up with the idea of issuing each player with a 'Talent Card' where the data from the technical and physical assessments were combined and presented in a format which resembled the player profiles in the video game FIFA or in the once popular Match Attax cards. An example of these cards can be seen in appendix. I was proud of this idea as the feedback I received was excellent and got the players engaged in a competitive way where they would compare their cards with teammates and try and improve their 'attributes' for the next time. This was exactly the impact I wanted from the whole process and I am hoping that this will positively impact my club going forwards through promotion of individual deliberate practice out with academy training and giving players ownership of their development.

Professional learning outcomes achieved during this stage of my project were:

- To improve my communication skills (specifically public speaking) and dissemination of information for various audiences
- To practically apply knowledge and insight formulated from the process into my club's curriculum to aid player development
- To engage in regular public speaking to disseminate research findings

CHAPTER 6: The Longitudinal Monitoring of Technical Performance in Youth Footballers and the Influence of Biological Maturity

5.1 Introduction

The process of TID has become a highly scrutinised topic in the past decade (Nesti & Sulley, 2014). The world's elite football clubs direct substantial resources into developing academy systems to uncover and develop football talent who possess the competencies required to compete at the elite level. In tandem with this investment, there has been an increased research focus on talent development in attempts to guide the process utilising evidence-based practice (Ali, 2000). Traditionally, in many cases, the process of player appraisal and evaluation has been predicated on subjective opinion of qualified and experienced coaching staff (Larkin & Reeves, 2018). Whilst this information is valuable, recent research has demonstrated that a multidimensional approach to identification and development is more effective than subjective coach opinion alone (Sieghartsleitner et al, 2019). Within this multidimensional approach, the longitudinal assessment of technical performance and regular player appraisal is an important facet. Up until recently, most research designs have been cross-sectional in nature and have investigated measurement property parameters such as validity and reliability (Cobb et al, 2018; Garcia-Lopez et al, 2013; Unnithan et al, 2012). However, recent research has utilised longitudinal study designs whereby football specific motor performance has been tracked throughout a players' developmental years (Huijen et al, 2013; Honer & Votteler, 2016; Lehyr et al, 2018; Zuber et al, 2016). Results of these studies demonstrated that football specific motor performance, as assessed by 'closed skill' performance tests, significantly improved with age. Furthermore, they demonstrated that football specific performance tests were powerful enough to discriminate between 'elite' and 'non-elite' performers where performance level was determined by subjective opinion or by selection/de-selection for professional and/or national teams. Gullich (2014) reported that only 33% of the players selected for the U16 German Youth National Squad were retained at U19 despite 'surviving' numerous stages of the selection procedure, which encompassed multiple 'closed skill' performance tests that began at the U11 age level. This evidence suggests that the test battery implemented during the selection and development procedure failed to identify 66% of the U19 German Youth National Squad which in turn, questions its utility and prognostic value.

One evident limitation identified within the results of previous research studies was that performance in the 'closed skill', football-specific performance tests did not correlate well with senior performance level attained, thus questioning their utility within talent development. One possible explanation for this could be the lack of specificity of the 'closed skill' performance tests and the resulting low prognostic value for predicting competition performance (Bergkamp et al, 2019). Bergkamp et al (2019), classified performance tests based on an explanation from selection psychology known as the 'fidelity' scale. Tests with 'low fidelity' "have relatively little overlap with the criterion in terms of the behaviour the player should show and the context in which the player must perform" (Bergkamp et al, 2019, p 1327). An example of a 'low fidelity' test is the isolated assessment of physical capabilities for evaluating performance or even predicting success. Assessment of physical capabilities represents only one facet of a multi-factorial and complex process which therefore limits its utility as an isolated measure (Martinez-Santos et al, 2016). A 'high fidelity' test refers to a situation "when the predictor becomes more similar to the criterion in terms of behaviour, task and contextual constraints" (Bergkamp et al, 2019 p 1237). The authors state that 'closed skill' tests lack the contextual constraints and requirements for information processing and decision making associated with competition performance, thus limiting their practical application. Therefore, the design and implementation of assessment protocols which better reflect competition performance should be a priority in reference to TID programmes. Results from previous chapters in this thesis have established content validity (chapter 3), reliability and measurement error (chapter 4) and criterion-based validity through investigating the correlation between small-sided game performance and competition performance (chapter 5). The research carried out up to this point has provided an argument for the potential inclusion of this TEEM into an academy's player appraisal, given the limitations identified throughout this thesis are appropriately considered. The final stage of tool development involves the establishment of tool sensitivity and responsiveness (Robertson et al, 2017). As this assessment protocols are intended to monitor longitudinal player development, and not just to provide a one-off cross-sectional assessment, this final step represents a critical stage in tool validation.

During an individual players' development years, it is anticipated that technical performance could change either positively or negatively due to a multitude of potential factors (Goto et al, 2018). One important factor which may impact technical performance is growth and maturation (Melan et al, 2010). During adolescence, an individual undergoes one specific period of accelerated growth, known as Peak Height Velocity (PHV), whereby the rate of growth increases exponentially (Balyi & Hamilton, 2004). The age at which this occurs exhibits inter-individual variation, however it usually occurs sooner in females than in their male counterparts. Following PHV, a rapid improvement in physical attributes is evident, such as speed, agility and explosive power (Fransen et al, 2017). A consistent issue highlighted within the TID research is a phenomenon known as the maturation selection bias (Helsen et al, 2005). The maturation selection bias is a pertinent issue within European academies and occurs when early maturing players are favoured over later maturing players due to their superiority in physical performance as a result of a more advanced stage of maturation (Vandendriessche et al, 2012). However, despite its well documented impact on physical performance, comparatively little research has investigated the potential impact of PHV on technical performance.

Moreira et al (2017) observed a significant positive correlation between biological maturity and involvements with the ball, successful passes, number of passes, game effectiveness and number of tackles in Brazilian academy youth football players. Their findings suggest that PHV may influence technical performance in a similar manner to physical performance. In addition, Huijgen et al (2010) demonstrated that technical performance in an isolated, 'closed skill' dribbling performance test exhibited greater improvements in the U16-U17 age group and the lowest rate of improvement during the U14-U16 age group. They suggested that this pattern could be a direct result of 'adolescent awkwardness', which is defined as "a disruption of motor control during peak height velocity" (Huijgen et al, 2010 p 696). In contrast, Vandendriessche et al (2012) reported no significant differences between early and late maturing players in football specific, 'closed-skill' dribbling performance tests. One possible reason for this disparity could be that Vandendriessche et al (2012) elected to recruit a homogenous group of talented footballers who were selected for the Belgium National Youth Squad. With this consideration, it is possible that the assessment protocol utilised in this investigation lacked the sensitivity to discriminate between such a highly skilled and homogenous group. To further our understanding, future research should explore the impact of maturation on technical performance during longitudinal player development. This should allow coaches and practitioners to better predict the quality of our young players and aid in the decision making and development process.

Therefore, the aims of this study are as follows: (1) to establish the current technical assessment protocols sensitivity to change and responsiveness; (2) to explore the impact (if any) of maturation on technical performance; and (3) to establish the assessments discriminative validity and ability to discriminate between players of different maturational stages. The author hypothesises that the assessment protocol would be sensitive enough to detect large improvements in technical performance over a twelve-month period but may not be sensitive enough to identify small changes over a short period of time given the measurement error established in chapter 4. Furthermore, I anticipate a significantly better performance in players with more advanced biological maturation compared with players at an earlier stage of maturational development.

5.2 Methods

5.2.1 Video Recording Data and Growth and Maturation Data

Permission to access the video recording and growth and maturation data of a Scottish Premier

League Youth Academy was requested. All data had been collected by youth academy staff in accordance with academy procedures and adhered to policies on data protection and sharing of that data. Written gatekeeper consent was obtained following the issuing of a research project information sheet prior to the analysis of any data. The gatekeeper was made aware that the academy was free to withdraw any data from the project at any time without reason. After permission was granted, video recording data of technical assessments (previous chapters) and growth and maturation data were collated for 30 academy players (age 13 ± 2.6 years).

5.2.2 Experimental Approach to The Problem

This longitudinal study design measured technical performance at 3 time points over a 12-month period. Technical assessment data were collected at 0, 6- and 12-month intervals from August 2018 to August 2019. Growth and maturation data were collected every 3 months as part of the academy's development programme. At each time point, both technical assessment data, and growth and maturation data were collected. Following data collection, the interaction effect between technical performance, time and maturation was explored.

5.2.3 Measures

Technical assessments from the video recording data were performed by the academy performance analyst. The academy analyst was qualified to degree level, had 3+ years as an analyst within an academy setting and was familiar with the academy technical assessment procedures. Details of the assessment protocol can be viewed in chapter 3 of this thesis. The TEEM consisted of 6 x 6-minute games in a 6 v 6 format with goalkeepers. Teams were randomly reconfigured after each 6-minute game to prevent bias in team selection. To facilitate a competitive environment, each player competed individually for a total point score at the end of the assessment procedure. Each player was awarded 2 points for a win, 1 point for a draw, 0 points for a loss and 1 point for every goal their team scored. All games were recorded for subsequent notational analysis (see chapter 3, section 3.3.2 and 3.3.3). In addition to this, a position-specific comparison of the number of actions performed during the assessment protocol was conducted to identify any differences in player behaviour which may provide additional information on what attributes are commonly performed based on player position.

Growth and maturation data were collected in 3-month intervals in accordance with the academy programme. All data was collected by the academy sports scientist who had 5+ years' experience in growth and maturation data collection. Player height, seated height and weight was collected in accordance with the protocol developed by Mirwald et al (2002). This method has been previously validated and is commonly used within research and practice (Rommers et al, 2019).

Following data collection, data was entered into an online growth utility predictor (www.usak.ca/kin-growthutility/phv_ui.php) to generate a prediction of how far the individual is from their Age of Peak Height Velocity (APHV) in years and a Predicted Adult Height (PAH). This information was then used to group players based on their stage of maturation (pre PHV = more than 0.6 years before PHV; mid PHV = between -0.6 and 0.6 years from PHV and post PHV = more than 0.6 years after PHV).

5.2.4 Statistical Analysis

All data were analysed using SPSS (v.16; SPSS Inc., Chicago, IL). In accordance with recommendations for a within-subjects, repeated measures study design, a multilevel regression model was designed for data analysis (Kreft & De Leeuw, 1998). As a preliminary step in the construction of the model, various models were explored using the Schwarz's Bayesian Criterion (BIC) parameter to find the best balance between model complexity and fit (Vellejo et al, 2014). After model exploration, it was concluded that incorporating a Random-Slope did not further improve the model based on the BIC parameter, however, inclusion of the Random-Slope allowed for determination of explained variance. Therefore, a Random-Intercept with Random-Slope mixed model consisting of two levels: repeated measures or time (level 1) and individual players (level 2) was selected to meet the requirements of the data sets within-subject, longitudinal structure which encompassed multiple measurement points over time, within the same individual. Each separate SEPI was used as a dependent variable within the model. Time was used as the repeated measures independent variable. The Scaled Identity repeated covariance type was selected, which assumed that measurements from each time point had the same variation but were independent of each other. Scaled identity was selected as this covariance type was considered the most appropriate after the previous establishment of SEPI typical error in Chapter 2. Factors included in the model were time and stage of maturation, both of which represented categorical data. Fixed effects included in the model were time and stage of maturation. Stage of maturation was included as a random effect and a random Intercept was included, and the individual players were selected as a grouping variable to determine individual variation around the mean intercept. An intraclass correlation coefficient was calculated to determine how much variance was attributed to stage of maturation. This was calculated using the following formula: repeated measures variance + within-player variance/within-player variance * 100. Statistical significance of technical performance (SEPI) predictors was determined using a significance value of $p < 0.05$. Confidence intervals (95%) were presented in estimates of fixed effects to identify numerical differences between SEPI score and the 3 time points and between SEPI score and stages of maturation. The frequency of actions comparison between playing positions was analysed using a one-way analysis of variance (ANOVA). The frequency of actions was selected instead of the SEPIs for the position-specific comparison to investigate the any position-specific differences in the number of occurrences in the individual technical actions. Statistical significance was determined using a significance value of $p < 0.05$. A Tukey post-hoc test was used to determine significance between groups.

5.3 Results

5.3.1 SEPI and Maturation Data

The descriptive statistics for SEPI scores across the 3 measurement time points and the 3 player classifications based on stage of biological maturation are shown in table 10. Overall, all groups experienced an increase in performance in the 1st touch and passing SEPI's from point 1 to point 3, however, only the change in performance from time point 2 to time point 3 in the 1st touch SEPI reached statistical significance. Results of the multilevel regression model are shown in table 11. Results of the multilevel regression model also revealed that stage of biological maturation was a significant predictor of technical performance in the 1st touch and passing SEPI's but not for ball manipulation and shooting performance. Players classified as mid-PHV demonstrated the largest average improvement in first touch and passing compared with pre- and post-PHV players from measurement point 1 to point 3 (4% vs. 1.56 and 2.48% respectively). Furthermore, mid-PHV players exhibited greater improvements in ball manipulation compared with pre- and post-PHV players from measurement point 1 to point 3 (3.58% vs. 0.40 and 1.42% respectively). Pre-, mid- and post-PHV players experienced a decrease in shooting performance from point 1 to point 3. Post-PHV players achieved higher scores than pre-PHV players in first touch and passing across all measurement points and higher scores than mid-PHV players in 2 out of 3 measurement points for first touch and across all measurement points for passing.

The stage of maturation variance (table 11) provides an estimate of how much variation in SEPI score can be explained by stage of maturation. The intraclass correlation coefficients show that an average of 23.6% of the variation in SEPI scores can attributed to stage of maturation (*range 12.77-31.42 %*). Stage of maturation significantly contributed to the prediction of players performance in the first touch and passing SEPIs but no other measure of technical performance ($p < 0.05$). Figure 11 provides graphical representation of trends in the SEPI scores over time. Figure 10(a) and (c) show a positive linear trend in the 1st touch and passing SEPIs over a 12-month period. Figure 10 (b) shows no change in ball manipulation performance over the 12-month monitoring period whilst figure 10 (d) shows a significant decrease in shooting performance. Figures 10 (e) and (g) show that players at the most advanced stage of biological maturation (post-PHV) demonstrated significantly higher levels of technical proficiency in the 1st touch and passing SEPI's, suggesting

Table 10. Descriptive statistics for the SEPI scores over the 3 measurement points in relation to stage of maturation. The change in score represents score at measurement point 3 - score at measurement point 1

SEPI	TIME POINT 1	TIME POINT 2	TIME POINT 3	ΔSEPI (1 TO 3)
Pre-PHV				
1st Touch	71.89 ± 1.01	72.50 ± 5.08	73.69 ± 4.96	1.8
Ball Manipulation	54.34 ± 4.60	56.54 ± 5.26	54.56 ± 7.73	0.32
Passing	64.52 ± 7.57	65.58 ± 6.51	64.99 ± 7.17	0.47
Shooting	45.73 ± 4.61	44.70 ± 4.98	43.95 ± 5.72	-1.78
Mid-PHV				
1st Touch	74.93 ± 2.8	70.7 ± 8.51	77.19 ± 3.14	2.26
Ball Manipulation	53.03 ±4.89	54.47 ± 4.81	55.0 ± 6.53	1.97
Passing	65.73 ± 4.85	65.12 ± 6.8	69.25 ± 4.71	3.52
Shooting	47.10 ±3.46	43.82 ± 3.26	41.55 ± 6.35	-5.55
Post-PHV				
1st Touch	73.76 ± 8.11	77.45 ±5.18	77.94 ± 4.13	4.18
Ball Manipulation	53.98 ± 2.33	53.24 ± 5.14	54.76 ± 4.69	0.78
Passing	71.05 ± 4.59	70.50 ± 7.22	70.71 ± 8.16	-0.34
Shooting	46.37 ± 3.18	45.77 ± 4.51	44.96 ± 4.76	-1.41
Combined Groups				
1st Touch	72.51 ± 4.76	72.4 ± 6.41	75.03 ± 5.41	2.52
Ball Manipulation	54.04 ± 4.20	55.35 ± 4.96	54.41 ± 6.20	0.37
Passing	65.61 ± 6.87	65.78 ± 6.56	67.28 ± 7.44	1.67
Shooting	46.06 ± 4.07	44.42 ± 4.32	43.84 ± 5.40	-2.22

improvement in these parameters with age, whereas figure 10 (f) shows better ball manipulation in players at earlier stages of maturation compared with players at later stages of maturation. Figure 10 (h) shows no significant change in shooting performance across all stages of maturation.

Table 11. Results of the multilevel level regression analysis. The intercept value represents the model intercept for the post-PHV group and measurement point 3. ICC = intraclass correlation coefficient; SE = standard error; CI = confidence interval; Estimate = regression coefficient (N.B. asterisk [*] represents a significant difference from the intercept)

		1ST TOUCH			BALL MANIPULATION			PASSING			SHOOTING		
	Parameter	**Estimate**	**SE**	**95% CI Lower Upper**	**Estimate**	**SE**	**95% CI Lower Upper**	**Estimate**	**SE**	**95% CI Lower Upper**	**Estimate**	**SE**	**95% CI Lower Upper**
Fixed Effects	*Time Intercept*	78.46	1.38	75.7- 81.2	54.27	1.44	51.3-57.2	71.08	1.81	67.5- 74.7	44.13	1.27	41.5- 46.7
	Time Point 3	0	0	-	0	0	-	0	0	-	0	0	-
	Time Point 2	-2.4 *	0.99	-0.43—0.41	0.66	1.22	-1.79 - 3.11	-0.66	1.9	-3.24- 1.92	0.88	0.99	-1.10- 2.88
	Time Point 1	-1.87	1.04	-3.95-0.21	-0.93	-0.73	-3.46- 1.59	-0.72	1.36	-3.44- 1.98	2.44*	1.03	0.36 - 4.51
	Post-PHV	0	0	-	0	0	-	0	0	-	0	0	-
	Mid-PHV	-3.63 *	1.51	-6.63- -0.63	0.16	1.73	-3.28- 3.62	-3.81	1.97	-7.73- 2.89	-0.77	1.46	-3.68- 2.14
	Pre-PHV	-4.12 *	1.71	-7.54—0.69	0.87	1.74	-2.62-4.37	-5.08 *	2.25	-9.58- -0.58	-3.71	1.56	-3.51- 2.76
Random Effects	*Repeated Measures Variance*	13.18	2.66	8 .87- 19.57	21.09	3.99	1 4.55- 30.58	23.04	4.52	15.68-33.85	13.61	2.62	9.32- 19.86
	Stage of Maturation Variance	6.04	2.15	3.00- 12.13	3.09	1.80	0.98- 9.69	9.67	3.39	4.87- 19.23	3.56	1.5	1.54 - 8.22
	ICC for Stage of Maturation Predictor (%)	31.42			12.77			29.56			20.73		

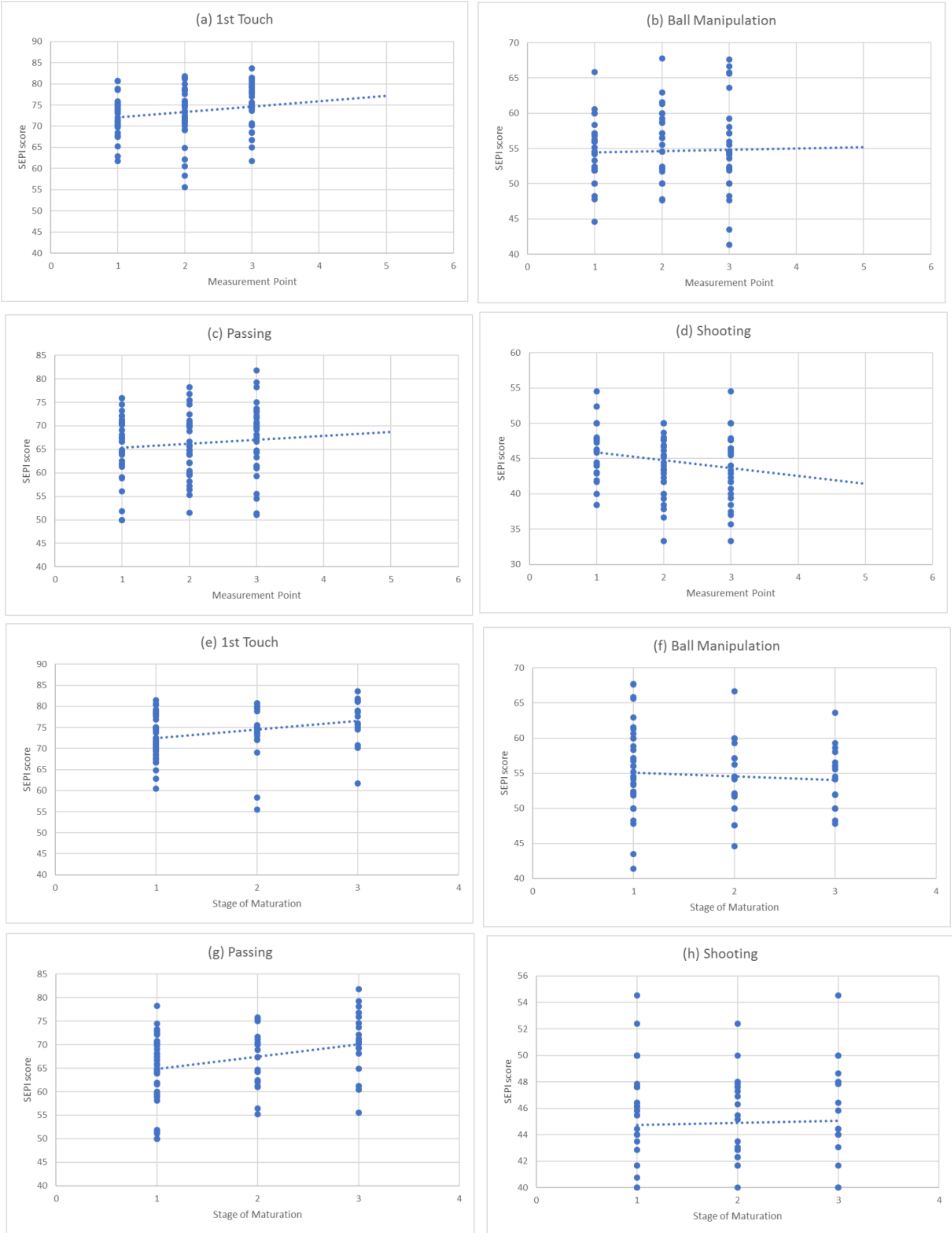

Figure 10. Figures (a)-(d) represent the SEPI score trends over the 12-month monitoring period. A linear forecast trendline was applied to the data. Figures (e)-(h) represent the differences in SEPI scores in relation to stage of maturation. The stage of maturation groups (pre-, mid- and post-PHV) were dummy coded into the corresponding values: 1 = pre-PHV; 2= mid- PHV; 3 = post- PHV

5.3.2 Position-Specific Frequency Data

The results of the one-way ANOVA for the frequency of attempted actions across the 3 playing position are shown in figure 12. Midfielders attempted a significantly higher number of first touches than attackers. Midfielders also attempted a higher number of ball manipulations than both defenders and attackers. Attackers attempted significantly less passes than both defenders and midfielders but significantly more shots than defenders. As expected, attackers attempted the highest number of shots per assessment.

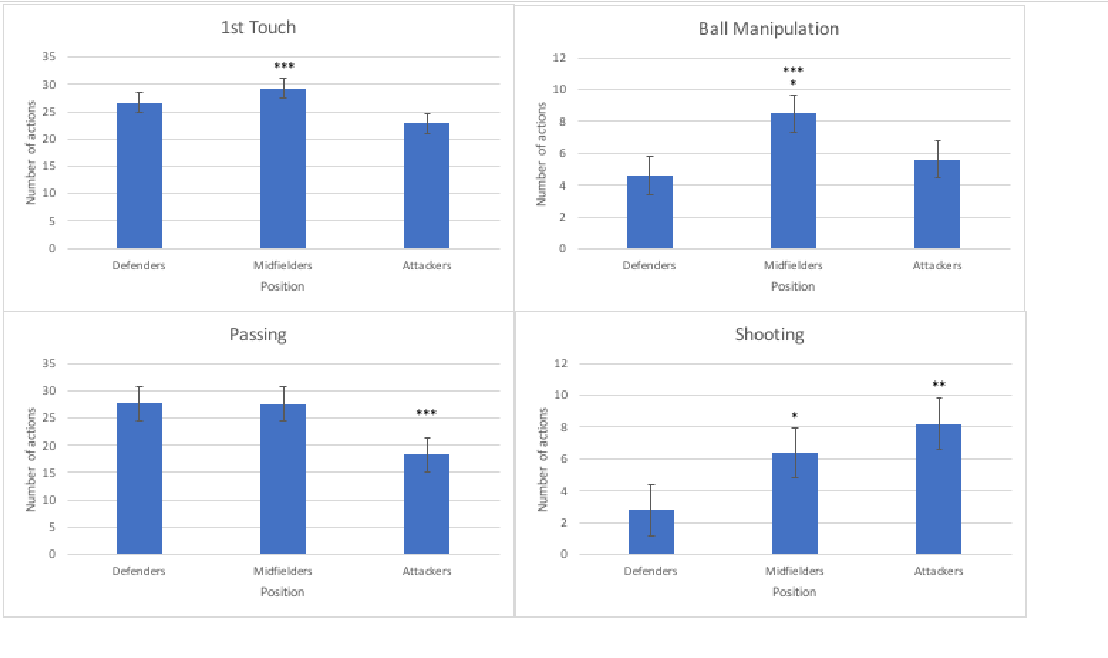

Figure 12. Results from One-way ANONVA for the frequency of actions attempted per assessment in relation to player playing position. Defenders included central defenders and full backs. Midfielders included central midfielders. Attackers included forwards and wide midfielders. Values are expressed as ± standard deviation. * represents a significant difference between defender and midfielders, ** represents a significant difference between defenders and attackers and *** represents a significant difference between midfielders and attackers as determined by the Tukey post-hoc test.

5.4 Discussion

5.4.1 General Discussion

This research aimed to investigate the responsiveness and sensitivity to change in a newly developed protocol for assessing and monitoring technical performance in youth footballers. Additionally, the investigation aimed to explore the influence of individual stage of growth and maturation on technical performance and any position-specific differences in player behaviour. Results revealed that the assessment protocol lacked the sensitivity to detect statistically significant changes in technical performance over a 12-month monitoring period. However, despite the lack of significance of time as a predictor of technical performance, a positive linear trend towards improvement of first touch and passing performance was observed over the 12-month period (figure 11). Descriptive statistics demonstrated that both first touch and passing showed an improving trend towards the smallest worthwhile change (SWC) + typical error (TE) established in chapter 4 (3.59 and 4.98 respectively). In many circumstances, the difference in age between post- and pre-PHV players exceeded the length of the monitoring period (12 months). In studies investigating longitudinal monitoring of technical performance in youth football players, all adopted a substantially longer monitoring period (3+ years) than the one in the present study (Honer et al, 2016, Huijgen et al, 2010, Lehyr et al, 2018, Zuber et al, 2016). Results in this study demonstrate that the assessment protocol successfully discriminated between post-PHV players and mid- and pre-PHV players in the 1st touch and passing SEPI's. Since biological maturation is a product of time, these trends suggest that a monitoring period, comprised of multiple measurement points expanding over a minimum of the average age difference observed between pre-PHV and post-PHV players, may present alternative observations. With this theory in mind, it is possible that an extended monitoring period may be required to detect a significant and meaningful change. This is supported by findings from chapter 4 in this thesis which suggested that the TEEM lacked the sensitivity to detect small changes in performance. In further support of this, Lehyr et al (2018) reported that across a 3-year longitudinal study, an average improvement of 3.05 and 6.97% per year in dribbling and ball control performance was observed. This was in comparison with a total improvement in performance from measurement point 1 to point 4 (3 years) of 8.93 and 19.09% in dribbling and ball control respectively, measured in isolated, 'closed skill' tests. These findings suggest that more substantial improvements in performance are observed over a longer monitoring period and that the results reported in the current study (improvements of 3.47, 0.68, 2.54 and -4.81% in 1st touch, passing, ball manipulation and shooting respectively) could potentially show a greater improvement over a similar time period. It may be the case that only certain SEPIs (e.g., 1st touch and passing) may be suitable for assessing performance over time, however, further research which extends over a longer monitoring period is required.

Two possible explanations are offered for the lack of statistically significant improvements in performance observed over a 12-month period: 1) the assessment protocol lacks the sensitivity to detect change due to the variation associated with technical performance in a dynamic situation and/ or 2) the 12-month period was insufficient for any significant adaptations in performance to occur. However, it should be noted, that although time was not a significant predictor of performance, analysis revealed that the individual players stage of maturation was identified as a significant predictor of performance in the first touch and passing SEPIs. The results demonstrate that mid- and post-PHV players exhibited significantly higher SEPI scores compared with pre- PHV players and post-PHV achieved higher scores than mid-PHV players across two measurement points for first touch and across all measurement points for passing. These results are consistent with findings reported by Moreira et al (2017) where players more advanced in their stage of maturation achieved more involvements with the ball, completed more passes and more successful passes compared with players at an earlier stage of maturation in a small-sided game-based assessment. Furthermore, Saward et al (2019) monitored a selection of elite youth footballers from 3 elite level English Academies between the age of 11-18 years across 2 competitive seasons with the aim of tracking the development of match skills. Saward et al (2019) utilised computerised notational analysis through the video recordings of 1-10 competitive matches per player. The 'match skills' selected for analysis were successful passes, shots on target, dribbles, crosses, tackles, blocks, clearances and interceptions. The authors implemented a specifically designed regression model to explore the effect of playing status (retained or released), maturity status and playing position on the development of match skills. Saward et al (2019) reported that the model predicted an increase in passing proficiency over time and was significantly influenced by maturity status, with early maturing players performing significantly more actions than their later maturing counterparts. These results are consistent with the current study. In contrast, Malina et al (2005) reported no significant differences in passing accuracy and dribbling performance in youth football players aged between 13 and 15 with differing stages of biological maturity. However, it should be noted that technical performance in Malina et al (2005) utilised 'closed skill' performance tests which lack the associated spatial-perceptual cognitive demands required during competitive match performance. This important factor should be taken into consideration during test selection as the increase in testosterone levels associated with maturation have been shown to influence spatial cognition and visuospatial ability (Aleman et al, 2004; Cherrier et al, 2001), both of which play an important role in skill execution during match-related performance (Vaeyens et al, 2008). One interesting observation from the current study was the highest average rate of improvement in first touch and passing performance occurred in the mid-PHV group, when an acceleration of testosterone production occurs (Delemarre-van de Waal et al, 2001), compared with pre- and post-PHV groups. The pattern of development within the mid-PHV group suggests that growth and maturation may stimulate a positive improvement on match-related motor skill performance due to the rapid improvement of spatial ability associated with testosterone production during maturation (Davison & Susman, 2001), however, future research employing larger sample sizes is required to investigate this further.

Another interesting finding from the current results was the lack of improvement in the ball manipulation and shooting SEPIs across all age groups and measurement points. Two explanations for this lack of change in ball manipulation and shooting scores are provided. Firstly, the descriptive statistics in table 10 reveal that ball manipulations were significantly higher in midfielders and attackers compared with defenders, which is perhaps expected given the positional technical demands. In line with position-specific technical demands, shooting was significantly higher for attackers than for midfielders and defenders. These results suggest that ball manipulation and shooting are individual and position-specific skills. This obvious assertion that certain aspects of technical performance are specific to playing position is supported by research Sarward et al (2019) who reported a significantly higher number of dribbles and shots on target performed by attacking players compared with defensive players during competitive match play in elite youth footballers. The findings in both the current study and in the study by Sarward et al (2019) suggest that specific technical actions such as dribbling and shooting are 'specialist' skills that may not require a high level of proficiency across all positions where as technical actions such as first touch and passing could be considered basic rudimentary skills that are essential across all playing positions. The content of team-based sessions within an academy setting is predominantly focussed on possession-based activities resulting in considerably more time spent practicing passing and receiving than any other skill which perhaps explains why there is a trend towards improvement for 1st touch and passing but not for ball manipulation and shooting (Partington & Cushion, 2013). It may be the case that players have limited opportunities to practice the skill of pressurised ball manipulation and shooting with the required time investment for improvement during team-based sessions, where a specific coaching curriculum is often commonplace within academy settings, however, this coaching content will vary considerably from club to club. Therefore, it could be speculated that any improvement in ball manipulation or shooting is achieved through individual practice away from organised club sessions and this time investment has been shown to be individual to the specific player and culture (Ford et al, 2012). Furthermore, these skills are thought of as exceptional skills which may separate elite players from non-elite players (Gai et al, 2019; Konefal et al, 2019; Yi et al, 2019). Therefore, it is unlikely that every player has the technical competency to be well-skilled in these situations. Secondly, the occurrence of ball manipulation and shooting actions were significantly lower than the number of first touch and passing actions during the small-sided game assessment protocol (figure 11). Consequently, it is likely that due to a low number of observed actions, the ball manipulation and shooting SEPIs failed to stabilise which resulted in too much 'noise' to detect any meaningful change in performance. To increase the number of actions performed and stabilise the performance outcomes within the TEEM protocol, it is proposed that future research could also focus on the development of additional, skill-specific, small-sided based games for assessing and monitoring ball manipulation and shooting performance. This future research should also focus on the feasibility of its application within an academy environment.

5.4.2 Limitations

Despite some interesting results generated from the current investigation, three main limitations were identified., Firstly, the small sample size used within this longitudinal study increases the probability of type II error. Furthermore, in a longitudinal study design, it is recommended that the starting pool of participants is substantially large to accommodate the inevitable drop out of participants over multiple measurement points due to de-selection, injury or unavailability. This is especially relevant within an academy setting where a high turnover of players is prevalent. Secondly, 12 months, in the context of long-term player development, represents a very small snapshot of developmental change. For example, it is unknown how much improvement will take place during a 12-month period and it is unlikely that dramatic changes in performance will be observed. Therefore, if the monitoring period was extended to at least a 3-year period (Honer et al, 2016; Leyhr et al, 2018; Zuber et al, 2016), it could be possible that based on the trend observed in the current study, significant changes in performance would be observed. It is likely that development of technical performance is a slow, long-term process and is only susceptible to small, year-on-year improvements. If exponential improvements in performance per year were possible and easy to achieve, then subsequently, there would be a likely concomitant increase in the production of elite level players. Lastly, in this study, training content and practice hours were not controlled or recorded due to a lack of recording procedures within the participating academy. The current study assumed that players would improve in a linear fashion, whereas development is likely non-linear. Furthermore, the extent to which an individual player improves is likely dependant on the number of individual practice time investment in non-organised football activities away from the academy setting such as school football and non-organised play with friends (Gullich et al, 2017).

5.4.3 Conclusions

In conclusion, the TEEM protocol investigated in the current study failed to detect any significant change in technical performance over 3 measurement points during a 12-month monitoring period. Despite this lack of significant change over time, a linear trend towards improvement in the first touch and passing SEPIs was observed. In addition, the assessment protocol successfully identified individual player stage of maturation as a predictor of technical performance in the first touch and passing SEPIs. These results suggest that the 1st touch and passing SEPIs could be used to track longitudinal technical performance in an environment which is more representative of competition performance compared with isolated 'closed skill' assessment protocols.

5.4.4 Practical Applications and Future Directions

Results of this study suggest that the TEEM protocol utilised in the current investigation can be used as part of a holistic and multifaceted TID programme. This assessment protocol can be used to detect within-player changes in 1st touch and passing performance and to discriminate between players of differing stages of maturation. As biological maturation is a product of time, it is likely that an extended monitoring period will uncover larger changes in technical performance. Furthermore, the information generated from this assessment procedure can be used as a motivational tool to drive individual development based on continuous player appraisal and feedback. Future research should aim to implement this monitoring tool over an extended time period of at least 3 years to provide conclusive evidence of test sensitivity. The assessment tool in the current study successfully discriminated between more biologically mature players compared with their less mature counterparts. Further research investigating the impact of growth and maturation on match-related technical performance is warranted which could subsequently inform early years coaching practice and content with the aim of maximising technical performance prior to age at peak height velocity.

5. 5 Personal Reflection and Professional Skills Development

As this research study developed, I was engaged in interesting conversations with my colleagues about the influence that biological maturation would have on the results of the assessments. For example, it was clear from simply watching the assessments take place that the more 'physically developed' players experienced more involvements with the ball and dominated over the less 'physically mature' players. This raised questions about test validity and whether, these results would provide a true reflection on performance. Below is an extract from my reflective diary which explains my rationale:

> *"Was thinking about why there might be some variation in performance indices with some players but not with others. When considering the players who showed the most variation in performance, the main KPI was ball manipulation. Also, the players who showed the most variation, or even a decrease in performance between two trials 6 months apart, were the smaller players. I was thinking that, ball manipulation e.g. taking players on 1 v 1 and dribbling past players, requires an element of physical performance such as speed and power. If the players who were good at ball manipulation at one stage but are less physically developed than the players in the same group 6 months later then it is logical to assume that their 'lesser' physical ability could contribute to the decrease*

in performance. For example, some players could have begun their high velocity growth spurt when others have not. This could drastically affect performance. Perhaps a way to control for this problem is to carry out the assessment protocol in bio-banding categories rather than chronological age categories. However, I've already started collecting data using chronological age groups and therefore the data may be affected.

Also, it may be worth looking at including biological maturity as a variable when building a regression model in your longitudinal study. However, I'm not sure if I'll have enough data to have enough statistical power"

As maturation data was available as part of the academy's programme, I decided to include this in my analysis. Regardless of the results presented previously in this chapter, we decided as an academy that the impact of biological maturity was a pertinent issue that should be addressed within our curriculum. The result of this was the introduction of 'bio-banding' nights whereby players were grouped according to their biological age for the first two sessions of the week as opposed to their chronological age group. Within these groupings, further groupings were created using 'ability bands' based on the results of the technical assessments. The aim of this process was to create situations where the best players were competing against each other as this fell in line with our academy philosophy that we aim to prioritise the development of the individual player over the team. This was run as a 6-month trial at the end of which, the change in technical performance would be re-assessed. Unfortunately, due to unforeseen circumstances, this was not able to be completed. I feel this pilot project again highlighted the impact that my professional doctorate project has had on my organisation and I consider these to be positive additions to our curriculum. This was ultimately one of the main aims prior to commencement of the programme and I am proud that I have been able to achieve this at some level.

Lastly, following completion of my final study, I decided that I needed to practice communicating my findings to a wider audience and thus took the decision to organise a CPD event open to the general public which would provide the opportunity to present. I invited additional speakers to contribute to the event but unfortunately due to an unprecedented global pandemic, the event was unable to take place. Alternatively, I decided to embrace modern technology and arrange the event in the format of an online webinar. In the end, I was able to reach a wider audience with over 100 attendees which greatly surpassed the number we would have been able to accommodate at the venue I had previously arranged. The online event itself ran smoothly and attracted interest. The speakers delivered their content very well, however, I was unhappy with the delivery of my own. I felt extremely nervous which was unlike the situation previously mentioned in chapter 4 where I felt comfortable and relaxed. I think the thought of knowing there were a number of highly respected practitioners/ researchers/ coaches in the audience made me more nervous. I also found speaking into a camera more nerve racking than presenting to a group. My nervousness was exacerbated when my slide notes were lost due to the 'share screen' feature which had to be enabled to deliver your presentation. From this experience I have highlighted that I need more practice in presentation skills, especially in front of an academic audience. I feel that this is one aim/ objective from the research and professional aims and objectives outlined in chapter 1 that will require ongoing and regular practice throughout my professional career. However, despite my somewhat negative personal appraisal, my supervisor received some positive feedback from another academic who attended the webinar. The feedback detailed that "he enjoyed my presentation and thought it was well presented with a clear message and implications". In summary, I believe this final stage of my project facilitated my development in the following areas (chapter 1):

- To improve my communication (specifically public speaking) and dissemination of information for various audiences
- To practically apply knowledge and insights formulated from the process into my organisations curriculum to aid player development
- To engage in regular public speaking to disseminate research findings

CHAPTER 7: Synthesis

7.1 Overarching Discussion

Throughout this thesis, the main objective has been to establish the necessary measurement properties of a newly developed tool for assessing and monitoring technical performance in youth footballers. Recently, authors have inferred that this is a process that has lacked the methodological scrutiny required for the successful validation of newly developed test protocols (Robertson, 2013). Furthermore, the work presented in this thesis has addressed some key, and often overlooked methodological issues identified within previous assessment of technical performance (Bergkamp et al, 2019; Koopmann et al, 2020). Firstly, by establishing necessary measurement properties such as the smallest worthwhile change (SWC) and typical error (TE), practitioners and coaches can make more informed inferences about true changes in performance. This in turn provides more in-depth information for data interpretation and dissemination and makes the TEEM a viable tool for monitoring player progression/regression in a long-term player development programme. Secondly, by establishing criterion-based validity (chapter 5), a measure of task-representativeness is demonstrated. This relationship between performance in a training-based environment and competition performance suggests that the TEEM developed in this thesis represents a more task-specific measure of performance than 'closed-skill' based assessment protocols, which are common within TID programmes. The use of isolated or 'closed-skill' assessment protocols have previously been reported to lack transfer to performance (Bergkamp et al, 2019), therefore the research presented in this thesis offers a contribution to the existing research in the field of technical performance assessment in youth football. It must be noted however, that the TEEM also exhibits its own limitations in terms of capturing all behavioural playing habits which are expressed during competition. Lastly, the research presented in chapter 6 offers a contribution to the longitudinal research existing in this topic area. Koopmann et al (2020) reported that 75% of all studies investigating the assessment of technical skills in talented youth athletes adopted a cross-sectional study design. This over-representation of cross-sectional research limits the application of many current assessment tools for longitudinally monitoring player development. In summary the TEEM developed within this thesis provides a promising method of assessing and monitoring an important aspect of performance which has been suggested to heavily influence the outcome of a match and ultimately the success of a team (Filetti et al, 2017; Rampinini et al, 2009).

Recent literature and systematic reviews have highlighted the requirement for the development of assessment tools that: 1) closely represent the criterion task (i.e., competition performance); 2) are valid and reliable; and 3) are easily applicable within a practical setting (Bergkamp et al, 2019; Koopmann et al, 2020; Robertson, 2013). Authors concur that methods which bear a closer resemblance to competition performance show promising signs for application within a practical environment, especially with the development and availability of modern technology. Furthermore, the importance of technical skills and their role within TID has been demonstrated within previous research. For example, Koopmann et al (2020) reported that 93% of studies investigating the assessment and monitoring of technical performance conveyed positive discriminatory, explanatory and/or predictive benefits. However, to the author's knowledge, research investigating the longitudinal development of technical performance utilising an assessment tool which is 'competition based' or 'high-fidelity' (Bergkamp et al, 2019) in design is scarce. It is the author's opinion, that studies in this thesis will provide a welcome contribution to the body of existing research and offer a potential solution to some of the methodological issues and challenges around the establishment of appropriate measurement properties in newly developed assessment tools. Whilst researchers, practitioners and coaches recognise the need for developing more robust assessment protocols, further work is required to optimise practical application and dissemination of the information generated.

7.2 Research Findings and Practical Applications

On completion of this thesis, 4 main research findings can be offered as a contribution to the existing research in the field of assessing and monitoring technical performance in youth footballers. This thesis aimed to establish the necessary measurement properties of a newly developed observation tool that can be practically applied within an academy environment. The first main finding of this thesis is that a tools content validity can be systematically and scientifically established through the qualitative Delphi data collection method. Whilst the Delphi method offers a logical process for collecting and filtering subjective data, the context in which the tool is being designed for should be taken into consideration. With this in mind, one major factor which will heavily influence the tools generalisability, is the size of the study sample recruited to participate. For example, if the sample is too low, as is probably the case in chapter 3 of this thesis, then researchers run the risk of failing to capture potentially important information. Alternatively, the more participants who contribute to the first round of questionnaires, the more information you will have to filter down and consequently, there is less risk of important information being overlooked. However, if the participant pool is too large then this could be associated with a higher risk of low compliance rates and increased difficulty for consensus to be reached due to disagreement between participants. As the TEEM designed within this thesis was developed for use within one professional academy, the author decided that in order to keep the tool contextually relevant, only a selection of the most highly qualified and experienced coaches from that specific club was chosen. In summary, the Delphi method facilitated the establishment of content validity for the development of this new assessment tool, however, due to the selective sampling of participants used in chapter 3, the tools generalisability is limited. The Delphi method utilised within chapter 3 represents a robust a scientific process that can be applied to develop contextually relevant tools.

The second main finding from the research conducted within this thesis is that the newly developed TEEM demonstrates good reliability across 2 different measurement points, and between and within-observer. To the author's knowledge, the study conducted in chapter 4 was the first study to present typical error and smallest worthwhile change measures to deepen our understanding of the variation associated with performance between trials and interpretations of 'true' changes in performance. Although the magnitude of this typical error is open to interpretation, providing a measure of this can be seen as a positive step towards improving our understanding. However, it should be noted that the range of typical error associated with the individual key performance indicators (0.72 – 8.24 AU) suggests that some technical actions/behaviours are more stable than others (O'Donoghue, 2005). One explanation for this could be related to the subjective nature of certain key performance indicators such as opposition pressure and decision making. Although operational definitions were designed to reduce variability during the coding process, these indicators still remain open to interpretation by the individual observer/coder, thus limiting their reliability. In addition, the low number of frequencies recorded for specific contextual aspects of the global SEPI may also explain the strength of the reliability measures. For example, if the frequency counts are consistently low across all assessments (e.g., first time shots), then the resultant typical error would be low. Therefore, the author recommends that before implementing the TEEM, the most stable (and also valuable) key performance indicators must be selected for analysis. After a sufficient data collection period extended over time, other key performance indicators may 'stabilise' and it is at this point that these can be included in any player appraisal process. This in turn, will provide a more reliable and accurate reflection of performance.

The third main finding from this research process relates to the discussion in the previous paragraph. Chapter 5 of this study reported a moderate relationship between performance in the TEEM and the average performance across 3 competitive matches. Similar to the results reported in chapter 4, a wide range of correlation coefficients were observed within the individual technical actions/behaviours. The main source of this variation in strength of correlations was primarily related again to the contextual aspects of each of the global SEPIs. The author suggests that the contextual elements of the SEPIs (such as direction of 1st touch and type of ball manipulation) are associated with too much 'noise' for use following only one assessment. Due to a combination of low frequencies and different competition demands compared with a training environment such as opposition quality, team tactics, opposition pressure and intensity of play (Liu et al, 2016), the author recommends that only the most stable and valuable measures of technical performance are selected if choosing to conduct assessments occasionally (e.g., bi-annually). If the TEEM assessments are conducted regularly throughout a season and enough data is generated to allow all technical actions/behaviours to stabilise then this would also be recommended. However, this then becomes a question of practicality and whether or not this volume of data collection is feasible and fits within a development curriculum. In summary, the author suggests that the TEEM is a valid and reliable tool for assessing skill proficiency in certain aspects of technical performance, as long as measures are taken to ensure competition intensity and measurement error is acknowledged. However, the results presented in chapter 5 suggest that the contextual/behavioural aspects of technical performance do not correlate well between performance in the TEEM and competition performance and therefore, behaviour (or learning) should be assessed within competitive match-play.

The last main finding reported from chapter 6 in this thesis is the TEEM's ability to detect changes between players of different maturational statuses. Results of this study describe the potential effects of maturation on technical performance in youth footballers with post-PHV players performing significantly better than mid- and pre-PHV players in 2 out of 4 SEPIs (1st touch and passing). Unfortunately, time was not identified as a significant predictor of technical performance following the fitting of the regression model. However, this may have been explained by the length of the monitoring period adopted in chapter 6 (12 months). In the context of long-term player development, 12 months perhaps only represents only a small window in a player's development, and, to the authors knowledge, little research has investigated how much change in technical performance can be expected. It should be noted however, that a positive upward trend was observed when a regression line was projected over subsequent measurement points. In the limited number of previous studies investigating changes in technical performance over time, a minimum monitoring period of at least two competitive seasons was employed (Honer et al 2016, Huijgen et al 2013, Lehyr et al 2018, Sarward et al, 2019). Initially, it was intended that the monitoring period employed in chapter 6 of this thesis was longer, however, participant drop out due to deselection and injury resulted in a significant reduction in participant number and consequently reduced statistical power over 4 measurement points (18 months). Therefore, a monitoring period of only 12 months was possible. Until research further investigating the TEEMs ability to detect changes in performance over time has explored the influence of a longer monitoring period (minimum of 2 years but ideally longer), the author cannot confidently conclude that the tool is sensitive enough to detect such a change. However, one interesting observation is the significant difference between maturational stages in the 1st touch and passing SEPIs. It is logical to hypothesize that since age is a product of time, time would emerge as a significant predictor over a longer monitoring period.

In summary, the results presented throughout this thesis demonstrate the difficulty and complexity of monitoring technical performance in football. Due to the random and unpredictable nature of the game, technical performance is associated with substantial variation between observations. However, it should be noted that 1st touch and passing performance demonstrate the most stability across multiple observations and are associated with the least variation in performance. This is likely explained by the higher number of occurrences recorded compared with other technical actions and contextual descriptors of these actions such as types of ball manipulations and penetrating passes. One possible solution to the low frequency of actions problem is a longitudinal data collection process. The collection of more data facilitates the stabilisation of technical actions and subsequently will provide a more complete and accurate assessment of performance. The TEEM developed in this study offers one possibility for assessing skill proficiency which can be easily applied within an academy environment. Although the reliability and validity of certain technical actions was poor, practitioners should decide whether or not their inclusion is still warranted given their potential influence in important moments of the game. For example, although game changing actions such as attacking 1v1's and shooting ability are variable, is the ability to be successful in this situation valued even if success is sporadic? It could also be argued one important objective of player development is to reduce this variation in skill success and to increase the number, efficiency and consistency of technical actions. Therefore, for its ease of application, feasibility and ability to discriminate between players of stages of maturation, the TEEM could be a welcome addition to a club's assessment protocol. However, as previously mentioned, for assessing and monitoring player learning and behaviour, the author recommends assessment within a competitive situation.

7.3 Future Directions

As evidenced by the work conducted within this thesis, the use of the TEEM or any other small-sided game-based protocol has a place in the monitoring of technical performance in youth footballers. However, these types of assessment protocol have associated limitations, primarily in assessing the transfer of player behavioural learning and playing habits. The use of objective data in football has increased exponentially over the previous decade (Rein & Memmert, 2016). With the development of modern technology and its increasing accessibility and affordability, the use of data analytics at academy level has become an emerging prospect for TID (De Silva et al, 2018). The provision of objective data at a youth development level can add significant value to the selection and development process which will support traditional, subjective coach evaluation and facilitate the decision-making process. Whilst research in this field is still at an early stage, future work should focus on data collection and the benchmarking of performance levels across academies, leagues and countries. The sharing of this information could provide valuable feedback to clubs and organisations regarding current player standard. Following this benchmarking process, an objective evaluation of player quality can be carried out and appropriate interventions introduced if necessary. A practical example of this process would be the benchmarking and evaluation of technical performance in young (< U13) footballers. Should the results reveal lower technical performance characteristics at this level compared with other clubs/ organisations/ countries, an intervention can be implemented (such as the introduction of home-based, individual skill development programmes) to supplement an academy learning curriculum. Furthermore, the effectiveness of this academy curriculum or individual coach methodology can be evaluated longitudinally through extended monitoring periods. Evaluation of the effectiveness of different coaching methodologies (e.g., repetitive technique practices vs. small-sided game-based practices) would be welcomed within the coaching science community. Lastly, further research is required to confirm the effects of biological maturation on technical performance observed in the results the study in chapter 6 of this thesis and previously reported in others (Sward et al, 2019). Although research in this area is scarce, future investigation would provide valuable information regarding the long-term development of technical performance in youth footballers. Whilst the future directions highlighted in this paragraph present excellent opportunities for further research, it is imperative that a consistent and validated methodological approach is accepted by practitioners and coaches which in turn will ensure validity and reliability of data sources.

7.4 Limitations

Whilst the use of this TEEM offers a promising solution to objective performance assessment in youth football, 3 main limitations should be taken into consideration during practical application. Firstly, the content of this observation tool was established based on the subjective opinion of coaches employed by one club. Therefore, the generalisability of the tool is limited and may not represent the philosophies and views of the greater football community. Furthermore, there may exist some inconsistencies in the definitions of technical actions between clubs/organisations. Any investigations which extend across multiple clubs/organisations should ensure consistent, and detailed operational definitions are established and appropriate observer training is completed prior to data collection. This limitation manifests from the qualitative research design adopted in Chapter 3 and specifically, with the use of the Delphi method. Secondly, due to the subjective nature of the coding process during analysis, there will always exist an element of human error when interpreting technical actions. Therefore, it is essential that measurement error is established (chapter 3) and taken into consideration. As previously mentioned, detailed operational definitions and adequate observer training will likely reduce measurement error, however, it is impossible to eradicate it almost entirely without the use and application of advanced technological techniques remain in their infancy at this current time. Lastly, whilst the requirement for longitudinal research is highly necessary, this method of research design has associated limitations and complexities. Longitudinal research requires a large participant pool and extends over a significant period of time due to the high possibility of missing data associated with selection/deselection and injury. Therefore, this type of investigation is difficult to implement within one organisation and will likely require cooperation from multiple organisations or be driven by a governing body. For example, a nationwide project ran by a countries footballing association which is aimed at collecting this objective data to be stored centrally. Such a project would provide valuable benchmark data for talent identification/ development at both a club and national level.

7.5 Meta-Reflection

When approaching the end of my professional doctorate journey, an important personal step for me was taking some time to look back on the process and reflect on what I have achieved from both a professional and academic perspective. An obvious personal achievement is the completion of this thesis. The magnitude of work completed during this thesis is the largest I have completed in my lifetime which in itself is a great personal achievement, especially when simultaneously working in a highly demanding full-time role within professional football. This achievement proves to myself that I am capable of handling multiple, large scale projects.

On a deeper level, this professional doctorate process has facilitated my development as a practitioner in various respects. When considering the professional aims outlined in chapter 1, the different stages on this process has provided me with unique and specific opportunities to achieve these which are detailed throughout this thesis in the reflective accounts written in the various chapters. The main professional learning outcomes which I feel have been achieved (or not achieved) are summarised in the remainder of this section. Firstly, this research project has initiated the learning of new technical skills which are highly relevant in day-to-day practice. The main examples of these newly developed technical skills are user training with the advanced performance analysis software Hudl SportsCode and a deeper understanding of research and complex data analysis concepts such as multi-level regression. Furthermore, I feel my ability to efficiently search for relevant academic literature has improved significantly. The value of learning these new skills is highlighted in the fact that at the time of writing this, I have been offered new position at my organisation as a first team data analyst. Without the learning that has taken place during this professional doctorate process, this change of direction in my career would never have occurred. Secondly, the opportunities provided for me to succinctly and clearly explain my research process and findings to various audiences has helped me develop my dissemination skills and creativity in reporting. An example of this is highlighted in appendix 2 which shows the '*talent card'* idea which was used to improve player/parent engagement in the feedback process. However, with perceived improvements in reporting skills through written reports, I feel I have not made the same progress in my public speaking ability, especially when addressing academic audiences. This conclusion is reached simply from a personal feeling during speaking in which I still feel very nervous and anxious whilst delivering my content. I feel this is an area of professional development which will continue beyond the completion of the professional doctorate process. Lastly, the writing of this section of my thesis perhaps highlights my last professional learning outcome from the process as a whole. I now appreciate the value of reflective practice and I have been able to develop a personal process which has enabled me to reflective on my professional practice. This in turn has provided me with an

DESIGN → APPLY → ASSSESS → REFINE

internal framework which I can use for professional and personal development. This framework sounds simple and obvious, but it helps me conceptualise ideas and is summarised in the following 4 words:

This professional doctorate journey began with the presenting of a research question by my academy manager: *which is the best way to assess and monitor the technical performance of our young players so we can provide them with valuable feedback to facilitate their development?* Ultimately this was my primary aim and I wanted to make a significant contribution to my organisation. I feel that the research carried out and the knowledge obtained from completing this process has allowed me to successfully achieve this aim and is perhaps the one I am most proud of. The impact of my research and the knowledge I have acquired is reflected in 3 main key additions to our academy programme. Firstly, the bi-annual implementation of the TEEM developed in this thesis into our academy programme, highlights the potential perceived value of the tool to my organisation. Secondly, the introduction of specific, detailed and objective individual player development plans has provided players and parents with valuable feedback on current performance levels and future directions for development. I feel that this process has added a 'growth mindset' culture within our academy environment by promoting individual practice away from our structured sessions. Lastly, players were supported and encouraged to perform this individual practice away from structured sessions by the introduction of an intervention based on the overarching results of the assessment process. Following the reporting of results, an online app-based technical programme was offered to each academy player whereby they could follow a structured individual skills-based programme from their mobile phones or tablets to promote deliberate practice. Although in the early stages, I believe this will improve the skill levels, or at least improve the mindset, of our players in the long-term which I consider a commendable achievement and was the ultimate goal at the beginning of this professional doctorate process.

REFERENCES

Aleman, A., Bronk, E., Kessels, R.P., Koppeschaar, H.P. & van Honk, J. (2004). A Single Administration of Testosterone Improves Visuospatial Ability in Young Women, Psychoneuroendocrinology, 29 (5): 612 – 617

Ali, A. (2011). Measuring Soccer Skill Performance: A Review. Scandinavian Journal of Medicine & Science in Sports, 21 (2): 170-183

Ali, A., Williams, C., Hulse, M., Struwick, A., Reddin, J., Howartg, 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

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

Atkinson, G. & Nevill, A.M. (1998). Statistical Methods for Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. Sports Medicine, 26 (4): 217-238

Badin, O. Smith, M., Conte, D. & Coutts, A. (2016). Mental Fatigue Impairs Technical Performance in Small-Sided Soccer Games. International Journal of Sports Physiology and Performance, 11(8): DOI 10.1123/ijspp.2015-0710

Balyi, I., & Hamilton, A. (2004). Long-Term Athlete Development: Trainability in children and adolescents. Windows of opportunity. Optimal trainability. Victoria, BC: National Coaching Institute British Columbia & Advanced Training and Performance Ltd.

Barnes, C., Archer, D.T., Hogg, B., Bush, M. & Bradley, P.S. (2014) The Evolution of Physical and Technical Performance Parameters in the English Premier League. International Journal of Sports Medicine, 35 (13): 1095 – 1100

Barnett, L.M., van Beurden, E., Morgan, P.J., Brooks, L.O. & Beard, J.R. (2010). Gender Differences in Motor Skill Proficiency From Childhood to Adolescence: A Longitudinal Study. Research Quarterly for Exercise and Sport, 81 (2): 162 – 170

Bennett, K.J.M., Novak, A.R., Pluss, M.A., Coutts, A.J. & Fransen, J. (2019). Assessing the Validity of a Video-Based Decision-Making Assessment for Talent Identification in Youth Soccer. Journal of Science and Medicine in Sport, 22 (4): 729 - 734

Bennett, K.J.M., Novak, A.R., Pluss, M.A., Stevens, C.J., Coutts, A.J. & Fransen, J. (2017). The Use of Small Sided Games to Assess Skill Proficiency in Youth Soccer Players: A Talent Identification Tool. Science and Medicine in Football, 2 (3): 231 - 236

Bennett, K.J.M., Vaeyens, R. & Fransen, J. (2018). Creating a Framework for TID in Emerging Football Nations, Science and Medicine in Football, 3 (1): 36 - 42

Bergkamp, T. L. G., den Hartligh, R. J. R., Frencken, W. G.P., Niessen, S. M. & Meijer, R. R. (2020). The Validity of Small-Sided Games in Predicting 11 vs 11 Soccer Game Performance. PLoS ONE, 15 (9): e0239448

Bergkamp, T.L.G., Niessen, A.S.M., den Hartigh, R.J.R., Frencken, W.G.P & Meijer, R.R. (2019). Methodological Issues in Soccer Talent Identification Research. Sports Medicine, 49 (9): 1317-1335

Borges, P.H., Rechenchosky, L., Depra, P.P., Vaz Ronque, E.R., Greco, P.J., Menegassi, V.M. & Rinaldi, W. (2017). Impact of Aerobic Power, Strength of Lower Limbs and Speed on Technical Skills in Young Soccer Players. Journal of Exercise Physiology, 20 (1): 221 – 230

Boud, D. (1999). Avoiding the Traps: Seeking Good Practice in the Use of Self- Assessment and Reflection in Professional Courses. Social Work Education, 18(2): 121-133

Bradley, B., Johnson, D., Hill, M., McGee, D., Kana-ah, A. Sharpin, C., Sharp, P., Kelly, A., Cumming, S.P. & Malina, R.M. (2019). Bio-banding in Academy Football: Player's Perceptions of a Maturity Matched Tournament. Annals of Human Biology, DOI: 10.1080/03014460.2019.1640284

Brandes, M. & Elvers, S. (2017). Elite Youth Soccer Players' Physiological Responses, Time-Motion Characteristics and Game Performance in 4 vs. 4 Small-Sided Games: The Influence of Coach Feedback. Journal of Strength and Conditioning Research, 31 (10): 2652-2658

Brewer, C.J. & Jones, R.L. (2002). A Five-Stage Process for Establishing Contextually Valid Systematic Observation Instruments: The Case of Rugby Union. The Sport Psychologist, 16 (1): 138-159

Carling, C., Williams, A. M., & Reilly, T. (2005). The Handbook of Soccer Match Analysis. Abingdon, UK: Routledge.

Castillo, D., Los Arcos, A. & Martinez-Santos, R. (2018). Aerobic Endurance Performance Does Not Determine the Professional Career of Elite Youth Soccer Players. The Journal of Sports Medicine and Physical Fitness, 58 (4): 392-398

Cherrier, M.M., Asthana, S., Peskind, E. & Tendresse, S. (2001). Testosterone Supplementation Improves Spatial and Verbal Memory in Healthy Older Men, Neurology, 57 (1): 80 - 88

Clemente, F.M., Sarmento, H., Costa, I.T., Enes, A.R. & Lima, R. (2019). Variability of Technical Actions During Small-Sided Games in Young Soccer Players. Journal of Human Kinetics, 69: 201 - 212

Clemente, F.M., Wong, D.P. & Martins, F.M.L. (2014). Acute Effects of the Number of Players and Scoring Method on Physiological, Physical and Technical Performance in Small-Sided Soccer Games. Research in Sports Medicine, 22 (5): 380-397

Cobb, N.M., Unnithan, V. & McRobert, A.P. (2018). The Validity, Objectivity and Reliability of a Soccer-Specific Behaviour Measurement Tool. Science and Medicine in Football, 2 (3): 196-202

Coelho-e-Silva, M., Figuiredo, A., Simoes, F. & Seabra, A. Discriminatin of U-14 Soccer Players by Level and Position. International Journal of Sports Medicine, 31 (11): 790-796

Cohen, J. (1998). Statistical Power Analysis for the Behavioural Sciences. Routledge

Côté, J., Baker, J., & Abernethy, B. (2007). Practice and Play in the Development of Sport Expertise. Handbook of Sport Psychology. John Wiley & Sons, Inc

Davison, K.K. & Susman, E. J. (2001). Are Hormone Levels and Cognitive Ability Related During Early Adolescence? International Journal of Behavioural Development, 25 (5): 416 - 428

Delemarre-van de Waal, H.A., van-Coerverden, S.C. & Rotteveel, J. & de Gruyter, W. (2001). Hormonal Determinants of Pubertal Growth, Journal of Pediatric Endocrinology & Metabolism, 14 (Suppl 6): 1521 - 1526

Dellal, A., Chamari, K., Wong, D.P., Ahmaidi, S., Keller, D., Barros, R., Bisicotti. & Carling, C. (2011). Comparison of Physical and Technical Performancein European Soccer Match-Play: FA Premier League and La Liga. European Journal of Sports Science, 11(1): 51-59

Dellal, A., Owen, A., Wong, D.P., Krustrup, P., van Exsel, M. & Mallo, J. (2012). Technical and Physical Demands of Small vs. Large Sided Games in relation to Playing Position in Elite Soccer. Human Movement Science, 31 (6): 957-969

Deprez, D.N., Fransen, J., Lenoir, M., Philippaerts, R.M. & Vaeyens, R. (2015). A Retrospective Study on Anthropometrical, Physical Fitness and Motor Coordination Characteristics that Influence Dropout, Contract Status and First Team Playing Time in high-Level Soccer Players Aged Eight to Eighteen Years. Journal of Strength and Conditioning Research, 29 (6): 1692 – 1704

De Silva, D. V.S.X., Caine, M., Skinner, J. & Dogan, S. (2018). Player Tracking as a Tool for Physical Performance Management in Football: A Case Study from Chelsea Football Club Academy. Sports, 6 (4): 130

Diamond, I.R., Grant, R.C., Feldman, B.M., Pencharz, P.B., Ling, S.C., Moore, A.M. & Wales, P.W. (2014). Defining Consensus: A Systematic Review Recommends Methodologic Criteria for Reporting of Delphi Studies. Journal of Clinical Epidemiology, 67)4): 401-409

Di Salvo, V., Gregson, W., Atkinson, G., Tordoff, P. & Drust, B. (2009). Analysis of High Intensity Activity in Premier League Soccer. International Journal of Sports Medicine, 30 (3): 205 - 212

Drust, B. (2019). Applied Science and Soccer: A Personal Perspective on the Past, Present and Future of a Discipline. Sports Performance & Science Reports, 56: v1

Eubank, B.H., Mohtadi, N.G., Lafeve, M.R., Wiley, J.P., Bois, A.J., Boorman, R.S. & Sheps, D.M. (2016) Using the Modified Delphi Method to establish Clinically Consensus for the Diagnosis and Treatment of Patients with Rotator Cuff Pathology. BMC Medical Research Methodology, 15 (56):

European Club Association (ECA). (2019). Youth Academy Study. https://www.ecaeurope.com/media/4545/eca-youth-academy-study-condensed-version-2019.pdf

Evans, J. D. (1996). Straightforward Statistics for the Behavioral Sciences. Pacific Grove, CA: Brooks/Cole Publishing.

Farrow, D. & Robertson, S. (2017). Development of a Skill Acquisition Periodisation Framework for High-Performance Sport. Sports Medicine, 47 (6): 1043 - 1054

Fawver, B. & Hodges, N.J. (2017). Using the 'Expert Performance Approach' as a Framework for Examining and Enhancing Skill Learning: Improving Understanding of How Experts Learn. Frontline Learning Research, 5 (3): 139 - 154

Fernandez-Navarro, J., Frauda, L., Zubillaga, A. & McRobert, A.P. (2018). Influence of Contextual Variables on Styles of Play in Soccer. International Journal of Performance Analysis in Sport, 18 (3): 423 – 436

Fédération Internationale de Football Association (FIFA). (2018). 2018 FIFA World Cup Russia: Global Broadcast and Audience Summary. https://resources.fifa.com/image/upload/2018-fifa-world-cup-russia-global-broadcast-and-audience-executive-summary.pdf?cloudid=njqsntrvdvqv8ho1dag5

Figuiredo, A.J., Goncalves, C.E., Coelho-e-Silva, M.J. & Malina, R. (2009). Youth Soccer Players, 11-14 Years: Maturity, Size, Function, Skill and Goal Orientation. Annals of Human Biology, 36 (1): 60-73

Filetti, C., Ruscello, B., D'Ottavio, S. & Fanelli, V. (2017). A Study of Relationships Among Technical, Tactical, Physical Parameters and Final Outcomes in Elite Soccer Matches Analyzed by a Semiautomatic Video Tracking System. Perceptual and Motor Skills, 124 (3): 601-620

Ford, P., Couglan, E. & Williams, M. (2009). The Expert-Performance Approach as a Framework for Understanding and Enhancing Coaching Performance, Expertise and Learning. International Journal of Sports Science and Coaching, 4 (3): 451- 463

Ford, P., Carling, C., Garces, M. & Marques, M. (2012). The Developmentak Activities of Elite Soccer Players Aged Under-16 Years from Brazil, England, France, Ghana, Mexico, Portugal and Sweden. Journal of Sports Sciences, 30 (15): 1652-1663

Forsman, H., Grasten, A., Blomqvist, M., Davids, K., Liukkonen, J. & Konttinen, N. (2016). Development of Perceived Competence, Tactical Skills, Motivation, Technical Skills and Speed and Agility in Young Soccer Players. Journal of Sports Sciences, 34 (14): 1311-1318

Fransen, J., Woods, C.T., Bennett, K. & French-Collier, N. (2017). Modelling Age-Related Changes in Motor Competence and Physical Fitness in High-Level Youth Soccer Players: Implications for TID. Science and Medicine in Football, 1 (3): 203 - 208

Frauda, L., Zubillaga, A, Caro, O., Fernandez-Garcia, A.I. Ruiz-Ruiz, C. & Tenga, A. (2013). Designing Small-Sided Games for Training Tactical Aspects in Soccer: Extrapolating Pitch Sizes from Full-Size Professional Matches. Journal of Sports Sciences, 31 (6): 573-581

Fulton, J., Kuit, E., Sanders, G. & Smith, P. (2012). The Role of the Professional Doctorate in Developing Professional Practice. Journal of Nursing Management, 20 (1): 130 – 139

Gai, Y., Leicht, A., Lago, C. & Gomez, M. (2019). Physcial and Technical Differences Between Domestic and Foreign Soccer Payers According to Playing Positions in the China Super League. Research in Sports Medicine, 27 (3): 314-325

Garcia-Lopez, L.M., Gonzalez-Vilora, S., Gutirrez, D. & Serra, J. (2013). Development and Validation of the Game Performance Evaluation Tool (GPET) in Soccer. Euroamericana Magazine of Sports Sciences, 2 (1): 89-99

Gonzalez-Rodenas, J., Lopez-Bonida, I., Aranda-Malaves, R., Tudela, A., Sanz-Ramirez, E. & Aranda, R. (2020). Technical, Tactical and Spatial Indicators Related to Goal Scoring in European Elite Soccer. Journal of Human Sport & Exercise, 15 (1): 186-201

Gonzalez-Villora, S., Serra-Olivares, J., Pastor-Vicedo, J.C. & Teoldo da Costa, I. (2015). Review of the Tactical Evaluation Tools for Youth Players, Assessing the Tactics in Team Sports: Football. Springerplus, 4(1): 663-680

Goto, H., Morris, J. & Nevill, M. (2018). Influence of Biological Maturity on the Match Performance of 8- to 16-Year-Old Male Youth Soccer Players. The Journal of Strength and conditioning Research, 33 (11): 1

Grehaigne, J. (1997). Performance Assessment in Team Sports. Journal of Teaching in Physical Education, 16: 500-516

Gullich, A., Kovar, P., Zart, S. & Reimann, A. (2017). Sport Activities Differentiating Match-Play Improvement in Elite Youth Footballers – A 2-Year Longitudinal Study. Journal of Sports Sciences, 35 (3): 207-215

Hasson, F., Keeney, S. & McKenna, H. (2000). Research Guidelines for the Delphi Survey Technique. Journal of Advances Nursing, 32 (4): 1008-1015

Helsen, W.F., Winckel, Jan.V. & Williams, M. (2005). The Relative Age Effect in Youth Soccer Across Europe. Journal of Sports Sciences, 23 (6): 629-636

Hill, B.A. & Sotiriadou, P. (2016). Coach Decision-Making and the Relative Age Effect on Talent Selection in Football. European Sport Management Quarterly, 16 (3): 292-315

Honer, O. & Votteler, A. (2016). Prognostic relevance of motor talent predictors in early adolescence: A group – and individual- based evaluation considering different levels of achievement in youth football. Journal of Sports Sciences, 34 (24): 2269-2278

Hopkins, W. (2004). How to Interpret Changes in an Athletic Performance Test. *Sports Science,* (8): 1-7

Hughes, M., Evans, S. & Wells, J. (2004). Establishing Normative Profiles in Performance Analysis. International Journal of Performance Analysis in Sport, (1): 1 – 26

Hughes, M. & Bartlett, R.M. (2002). The Use of Performance Indicators in Performance Analysis. Journal of Sports Science, 20 (6): 739-754

Huijen, B., Elferink-Gemser, M.T. & Ali, A. (2013). Soccer Skill Development in Talented Players. International Journal of Sports Medicine, 34 (6): 720-726

Huijgen, B.C.H., Elfferink-Gemser, M.T., Post, W. & Visscher, C. (2010). Development Of Dribbling In Talented Youth Soccer Players Aged 12-19 Years: A Longitudinal Study. Journal of Sports Sciences, 28 (7): 689-698

James, N. (2006). Notational Analysis in Soccer: Past, Present and Future. International Journal of Performance Analysis in Sport, 6 (2): 67-81

James, N., Mellalieu, S.D. & Jones, N.M.P. (2005). The Development of Position-Specific Performance Indicators in Professional Rugby Union. Journal of Sports Sciences, 25 (1): 63-72

Joihansson, A. & Fahlen, J. (2017). Simply the Best, Better Than All the Rest? Validity Issues in Selections in Elite Sport. International Journal of Sports Science & Coaching, 12 (4): 470-480

Kelly, D.M. & Drust, B. (2009). The Effect of Pitch Dimensions on Heart Rate Responses and Technical Demands of Small Sided Soccer Games in Elite Players. Journal of Science and Medicine in Sport, 12(4): 475-479

Kirkendall, D.T. (2020). Evolution of Soccer as a Research Topic. Progress in Cardiovascular Diseases, 63 (6): 723-729

Kokstejn, J. & Musalek, M. (2019). The Relationship Between Fundamental Motor Skills and Game-Specific Skills in Elite Young Soccer Players. Journal of Physical Education and Sport, 19 (S1): 249 - 254

Kokstejn, J., Musalek, M., Wolanski, P., Murawska-Cialowicz, E. & Statsny, P. (2019). Fundamental Motor Skills Mediate the Relationship Between Physical Fitness and Soccer-Specific Motor Skills in Young Soccer Players. Frontiers in Physiology, 10: 596

Konefal, M., Chmura, P., Zajac, T., Chmura, J., Kowalczuk, E. & Andrzejewski, M. (2019). Evolution of Technical Activity in Various Playing Positions in Relation to Match Outcomes. Biology of Sport, 36 (2): 181 - 189

Koopmann, T., Faber, I., Baker, J. & Schorer, J. (2020). Assessing Technical Skills in Talented Youth Athletes: A Systematic Review. Sports Medicine, 50 (12): 1593-1611

KPMG Sport Advisory Practice. (2020). The European Champions Report 2020. https://www.footballbenchmark.com/home

Kreft, I. & De Leeuw, L. (1998). Introducing Multilevel Modelling, London: Sage.

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): e0175716

Leyhr, D., Kelava, A., Raabe, J. & Honer, O. (2018). Longitudinal Motor Performance Development In Early Adolescents And Its Relationship To Adult Success: An 8 Year Prospective Study Of Highly Talented Soccer Players. PLoS ONE, 13 (5): e0196324

Link, D., Lang, S. & Seidenschwarz, P. (2016). Real Time Quantification of Dangerousity in Football Using Spatiotemporal Tracking Data. PLoS ONE, 11 (2): e0168768

Liu, H., Gomez, M., Goncalves, B. & Sampaio, J. (2016). Technical Performance and Match-to-Match Variability in Elite Football Teams. Journal of Sports Sciences, 34 (6): 509-518

Lorrains, M., Ball, K. & MacMahon, C. (2013). Performance Analysis for Decision Making in Team Sports. International Journal of Performance Analysis in Sport, 13 (1): 110-119

Malina, R.M., Cumming, S.P., Kontos, A.P., Eisenmann, J.C., Ribeiro, B. & Aroso, J. (2005). Maturity-associated Variation in Sport-specific Skills of Youth Soccer Players Aged 13-15. Journal of Sports Sciences, 23 (5): 515-522

Martinez-Santos, R., Castillo, D. & Los Arcos, A. (2016). Sprint and Jump Performance Do Not Determine the Promotion to Professional Elite Soccer in Spain, 1994-2012. Journal of Sports Sciences, 34 (24): 2279-2285

McDermott, G., Burnett, A.F. & Robertson, S.J. (2015). Reliability and Validity of the Loughborough Soccer Passing Test in Adolescent Males: Implications for Talent Identification. International Journal of Sports Science & Coaching, 10 (2+3): 515-527

McKenzie, R. & Cushion, C. (2013). Performance Analysis in Football: A Critical review and Implications for Future Research, Journal of Sports Sciences, 31 (6): 639-676.

Melan, C., Cronlin, J., Oliver, J. & Hughes, M. (2010). Talent Identification in Soccer: The Role of Maturity Status on Physical, Physiological and Technical Characteristics. International Journal of Sports Science & Coaching, 5 (4): 571- 595

Memmert, D. & Harvey, S. (2008). The Game Performance Assessment Instrument (GPAI): Some Concerns and Solutions for Further Development. Journal of Teaching in Physical Education, 27: 220-240

Mills, A., Butt, J., Maynard, I. & Harwood, C. (2012). Identifying Factors Perceived to Influence the Development of Elite Youth Football Academy Players. Journal of Sports Sciences, 30(15): 1593-1604

Mirwald, R.L., Baxter-Jones, A.D.G., Bailey, D.A. & Beunen, G.P. (2002). An Assessment of Maturity from Anthropometric Measurements, Medicine and Science in Sport and Exercise, 34 (4): 689 - 694

Moalla, W., Fessi, M.S., Makni, E., Dellal, A., Filetti, C., Di Salvo, V. & Chamari, K. (2018). Association of Physical and Technical Activites with Partial Match Status in a Soccer Professional Team. Journal of Strength and Conditioning Research, 32 (6): 1708 – 1714

Mor, D. & Christian, V. (1979). The Development of a Skill Test Battery to Measure General Soccer Ability. The North Carolina Alliance for Athletics, Health, Physical Education, Recreation, 15 (1): 30-39

Moreira, A., Massa, M., Thiengo, C.R., Lopes, R.A.R., Lima, M.R., Vaeyens, R., Barbosa, W.P. & Aoki, M.S. (2017). Is the Technical Performance of Young Soccer Players Influenced by Hormonal Status, Sexual Maturity, Anthropometric Profile, and Physical Performance? (2017). Biology of Sport, 34 (4): 305-311

Nesti, M. & Sulley, C. (2014). Youth Development in Football: Lessons from the Worlds Best Academies, London: Routledge, Taylor & Francis Group

Nevill, A., Atkinson, G. & Hughes, M. (2008). Twenty-Five Years of Sports Performance Research in the Journal of Sports Sciences. Journal of Sports Sciences, 26 (4): 413 - 426

Nicholls, S.B. & Worsfold, P.R. (2016). The observational analysis of elite coaches within youth soccer: the importance of performance analysis. International Journal of Sports Science and Coaching, 11 (6): 825-831

O'Donoghue, P. (2005). Normative Profiles of Sports Performance. International Journal of Performance Analysis in Sport, 5 (1): 104 - 119

Oslin, J.L., Mitchell, S.A. & Griffin, L.L. (1998). The Game Performance Assessment Instrument (GPAI): Development and Preliminary Validation. Journal of Teaching in Physical Education, 17: 231-343

Ostojic, S.M., Castagna, C., Calleja-Gonzalez, J., Jukic, I., Idrizovic, K. & Stojanivic, M. (2014). The Biological Age of 14-year-old Boys and Success in Adult Soccer: Do Early Maturers Predominate in the Top-level Game? Research in Sports Medicine, 22 (4): 398-407

Owen, A.L., Wong, D.P., Paul, D. & Dellal, A. (2013). Physical and Technical Comparisons Between Various-Sided Games Within Professional Soccer. International Journal of Sports Medicine, 35 (3): 286-292

Partington, M. & Cushion, C. (2013). An Investigation of the Practice Activities and Coaching Behaviour of Professional Top-Level Youth Soccer Coaches. Scandinavian Journal of Medicine & Science in Sports, 23 (3): 374 - 382

Philippaerts, R.M., Vaeyens, R., Janssens, M, van Renterghem, B., Matthys, D., Craen, R., Bourgois, J., Vrijens, J., Beunen, G. & Malina, R.M. (2006). The Relationship Between Peak Height Velocity and Physical Performance in Youth Soccer Players. Journal of Sports Sciences, 24 (3): 221 - 231

Phillips, E., Davids, K., Renshaw, I. & Portus, M. (2010). Expert Performance in Sport and the Dynamics of Talent Development. Sports Medicine, 40 (4): 271-283

Praca, G.M., Soares, V.V., Alves da Silva Matias, C.J., Teoldo da Costa, I. & Greco, P.J. (2015). Relationship between tactical and technical performance in youth soccer players. Revista Brasileira de Cineantropometria Desempenho Humano, 17 (2): 134-144

Rampinini, E., Impellizzeri, F.M., Castagna, C., Coutts, A.J. & Wisloff, U. (2009). Technical Performance During Soccer Matches of the Italian Serie A League: Effect of Fatigue and Competitive Level. Journal of Science and Medicine in Sport, 12(3): 227-233

Reeves, M.J., Littlewood, M.A., McRobert, A.P. & Roberts, S.J. (2018). The Nature and Function of Talent Identification in Junior-Elite Football in English Category One Academies. Soccer & Society, 19 (8): 1122 - 1134

Reilly, T. & Gilbourne, D. (2003). Science and Football: A Review of Applied Research in the Football Codes. Journal of Sports Sciences, 21 (5): 693-705

Reilly, T., Williams, A.M., Nevill, A. & Franks, A. (2000). A Multidisciplinary Approach to Talent Identification in Soccer. Journal of Sports Sciences, 18: 695 – 702

Rein, R. & Memmert, D. (2016). Big Data and Tactical Analysis in Elite Soccer: Future Challenges and Opportunities for Sports Science. SpringerPlus, 5 (1): 1410 – 1424

Roberts, S. J., Rudd, J.R. & Reeves, M.J. (2020).. Efficacy of Using Non-linear Pedagogy to Support Attacking Players' Individual Learning Objectives in Elite-Youth Football: A Randomised Cross-over Trial. Journal of Sports Sciences, 38 (11-12): 1454-1464

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, 3 (2):

Robertson, S.J., Burnett, A.F. & Cochrane, J. (2014). Tests Examining Skill Outcomes in Sport: A Systematic Review of Measurement Properties and Feasibility. Sports Medicine, 44 (4): 501-518

Rommers, N., Mostaert, M., Goossens, L., Vaeyens, R., Witvrouw, E., Lenoir, M. & D;Hondt, E. (2019). Age and Maturity Related Differences in Motor Coordination Among Male Elite Youth Soccer Players. Journal of Sports Sciences, 37 (2): 196-203

Rubajczyk, K. & Rokita, A. (2015). Relationships Between Results of Soccer-Specific Skill Tests and Game-Related Soccer Skill Assessment in Young Players Aged 12 and 15 Years. TRENDS in Sport Sciences, 4 (22): 197-206

Sarmento, H., Clemente, F.M., Harper, L.D., da Costa, I.T., Owen, A. & Figueiredo, A.J. (2018). Small Sided Games in Soccer – A Systematic Review. International Journal of Performance Analysis in Sport, 18 (5): 693 - 749

Sarward, C., Morris, J.G., Nevill, M.E. & Sunderland, C. (2019). The Effect of Playing Status, Maturity Status and Playing Position on the Development of Match Skills in Elite Youth Football Players Aged 11-11 Years: A Mixed-Longitudinal Study, European Journal of Sport Science, 19(3): 315 - 326

Sieghartsleitner, R., Zuber, C., Zibung, M. & Conzelmann, A. (2019). Science or Coaches' Eye? – Both! Beneficial Collaboration of Multidimensional Measurements and Coach Assessments for Efficient Talent Selection in Elite Youth Football. Journal of Sports Science and Medicine, 18 (1): 32 - 43

Strudwick, T. (2016). Soccer Science. Champaign, IL: Human Kinetics.

Swinton, P.A., Hemmingway, B.S., Saunders, B., Gualano, B. & Dolan, E. (2018). A Statistical Framework to Interpret Individual Response to Intervention: Paving the Way for Personalized Nutrition and Exercise Prescription. Frontiers in Nutrition, 5 (41).

Tenga, A., Holme, I., Ronglan, L.T. & Bahr, R. (2010). Effect of Playing Tactics on Achieving Score-Box Possessions in a Random Series of Team Possessions from Norwegian Professional Soccer Matches. Journal of Sports Sciences, 28 (3): 245-255

Thomas, C., Fellingham, G. & Vehrs, P. (2009). Development Of A Notational Analysis System For Selected Soccer Skills Of A Womens College Team. Measurement in Physical Education and Exercise Science, 13 (2): 108-121

Towlson, C., Cobley, S., Midgley, A.W., Gerrett, A., Parkin, G. & Lovell, R. (2017). Relative Age, Maturation and Physical Biases on Position Allocation in Elite-Youth Soccer. International Journal of Sports Medicine, 38 (3): 201-209

Unnithan, V., White, J., Georgiou, A., Iga, J. & Drust, B. (2012). Talent Identification in Youth Soccer. Journal of Sports Sciences, 30(15): 1719-1726

Vaeyens, R., Lenoir, M., Williams, A.M. & Philippaerts, R.M. (2008). TID Programmes in Sport Current Models and Future Directions. Sports Medicine, 38(9): 703-714

Vaeyens, R., Lenoir, M., Williams, M.A. & Phillippaerts, R.M. (2008). TID Programmes in Sport: Current Models and Future Directions. Sports Medicine, 38 (9): 703-714

Valenete dos Santos, J., Coelho e Silva, M.J., Simoes, F., Figueiredo, A.J., Leite, N., Elferink-Gemser, M.T., Malina, R.M. & Sherar, L. (2012). Modelling Developmental Changes in Functional Capacities and Soccer-Specific Skill in Male Players Aged 11-17, Pediatric Exercise Science, 24 (4): 603 - 621

Vallejo, G., Tuero-Herrero, E., Nunez, J.C. & Rosario, P. (2014). Performance Evaluation of Recent Information Criteria for Selecting Multilevel Models in Behavioural and Social Sciences, International Journal of Clinical and Health Psychology, 14 (1): 48 - 57

Van Maarseveen, M.J.J., Oudejans, R.R.D. & Savelsbergh, G.J.P. (2017). System For Notational Analysis In Small-Sided Soccer Games. International Journal of Sports Science and Coaching, 12 (2): 194-206

Vandendriessche, J.B., Vaeyens, R., Vandrorpe, B., Lenoir, M., Lefevre, J. & Philippaerts, R.M. (2012). Biological Maturation, Morphology, Fitness and Motor Coordination as Part of a Selection Strategy in the Search for International Youth Soccer Players (age 15-16 years). Journal of Sports Sciences, 30 (15): 1695 - 1703

Verburgh, L., Scherder, E.J.A., van Lange, P.A.M. & Oosterlaan, J. (2016). The Key To Success In Elite Athletes? Explicit And Implicit Motor Learning In Youth Elite And Non-Elite Soccer Players. Journal of Sports Sciences, 34 (18): 1782 – 1790

Visser, J., Geuze, R.H. & Kalverboer, A.F. (1998). The Relationship Between Physical Growth, the Level of Activity and the Development of Motor Skills in Adolescence: Differences Between Children with DCD and Controls. Human Movement Science, 17 (4): 573-608

Waldron, M., Harding, J., Gray, A. & Barrett, S. (2020). A New Foot-mounted Intertial Measuremrnt System in Soccer: Reliability and Comparison to Global Positionig Systems for Velocity Measurements During Team Sport Actions. Journal of Human Kinetics, DOI: 10.2478

Waldron, M. & Worsfold, P. (2010). Differences in the Game Specific Skills of Elite and Sub-Elite Youth Football Players: Implications for Talent Identification. International Journal of Performance Analysis in Sport, 10(1): 9-24

Williams, A.M. (2000). Perceptual Skill in Soccer: Implications for TID. Journal of Sports Sciences, 18 (9): 737-750

Wilson, R., David, G.K., Murphy, S.C., Angiletta, M.J., Niehaus, A.C., Hunter, A.H. & Smith, M.D. (2017). Skill Not Athleticism Predicts Individual Variation In Match Performance Of Soccer Players. Proceedings Of The Royal Society of London B, 284 (1868): 20170953

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 TID: Utilising Evolutionary Theory in Sports Science. Journal of Sports Sciences, 34(21): 2074-2086

Yi, Q., Gomez-Ruano, M., Liu, H., Zhang, S., Gao, B., Wunderlich, F. & Memmert, D. Evaluation of the Technical Performance of Football Players in the UEFA Champions League. International Journal of Environmental Research and Public Health, 17 (2): 604 - 616

Yi, Q., Groom, R., Dai, C., Liu, H. & Ruano, M.A.G. (2019). Differences in Technical Performance of Players From 'The Big Five' European Football Leagues in the UEFA Champions League. Frontiers in Psychology, 10: 2738

Zuber, C., Zibung, M. & Conzelmann, A. (2016). Holistic Patterns As An Instrument For Predicting The Performance Of Promising Young Soccer Players – A 3-Years Longitudinal Study. Frontiers in Psychology, 7: 1088

APPENDICES

Appendix 1 Training Plan

Development of any research project requires careful consideration of both the internal and external factors which could contribute to the success or failure of the project. Therefore, the creation of a detailed research plan which identifies both these internal and external determinants is a critical component in the initial stages of research development. This research plan will consist of 3 sections which will include a professional self-assessment in order to assess my current abilities as a researcher and as a practitioner; a literature review for rationale and brief methodology of proposed research ideas and a logistical training plan to establish a realistic time frame for completion.

PROFESSIONAL SELF-ASSESSMENT

In current professional practice, continuous professional development (CPD) has become an integral necessity imposed by many governing bodies. CPD embodies various methods which contribute to the development of personal professional practice and such methods include: further learning through knowledge attainment; critical reflection of practice and self-assessment. Boud (1995) defines self-assessment as *"the involvement of students in identifying standards and/ or criteria to apply their work and making judgements about the extent to which they have met these criteria and standards"*. Self-assessment enables students and professionals to truthfully assess their current ability in relation to the professional field in which they aspire to practice in. This process remains important as it facilitates the development of a goal-orientated development framework to which an individual can set goals and strategize methods for achieving these goals. The remaining part of this section will highlight various strengths and weaknesses identified by the self-assessment process which in turn will provide a framework for goal-orientated professional development in conjunction with the completion of this research project.

RESEARCHER COMPETENCY PROFILE

To begin the self-assessment of my capabilities as a researcher, I utilised the Vitae Researcher Development Framework (2009) which is a globally recognised tool validated by the research community. **Figure 1** provides a summary of my analysis of current researcher competencies. From the pre-determined categories set by the Vitae Researcher Development Framework, I identified that my main strengths lay in my subject knowledge and information seeking as depicted in **figure 1**. During my education and professional experience to date, I feel that I have developed strong subject knowledge which is evidenced by successful completion of both my Undergraduate Degree and my Master's Degree. Furthermore, the practical experience I have gained through developing and implementing a new sports science programme within my organisation has required me to explore many research areas in order to establish best practice which has consequently enhanced both my subject knowledge and information seeking skill set. I am familiar with current and previous research in my topic area but also in other areas which impact my daily practice and I have now become familiar with authors and researchers who have made significant contributions to the field. Furthermore, I pride myself on taking time to keep up-to-date with current literature by carrying out regular searches in online databases in order to optimise my professional practice. By utilising this information seeking skill set, it enabled me to stimulate a thought process which helped me develop ideas relating to the development of a research project designed to overcome an issue identified within my organisation. Investigating the subject area further fuelled my curiosity and enthusiasm which in turn contributed further to the construction of a research idea.

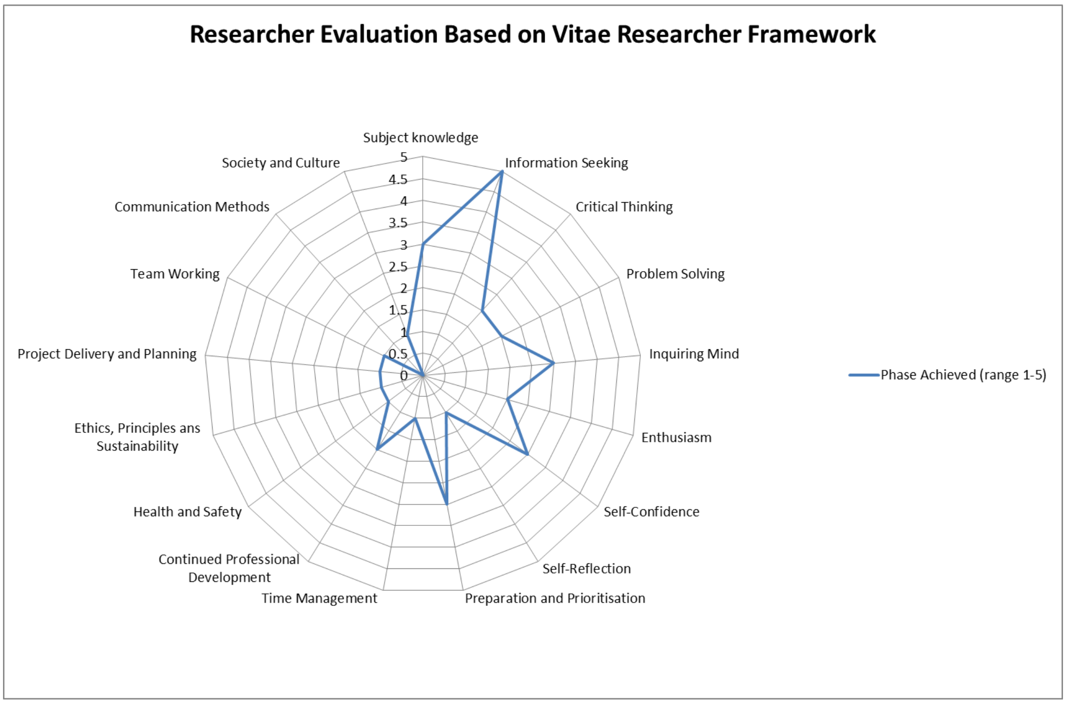

Figure 1. Evaluation of my researcher competencies identified by Vitae Researcher Framework

Two competencies highlighted from the Vitae Framework that I feel are important areas in my future professional development are communication methods and self-reflection. Throughout my education, I feel one of my main areas of weakness has manifested in my oral communication skills when delivering information to large groups of peers. I feel that I am competent in communicating on a one-to-one basis or in groups of colleagues of whom I know well, however, I feel less confident when delivering presentations or ideas in an academic setting. Evidence of this is highlighted by poor results from assignments obtained during my undergraduate and post graduate study that involved academic presentations. Therefore, I aim to improve my confidence in my basic oral communication skills by continuing to seek any opportunity to deliver presentations in order develop my confidence in uncomfortable positions. Furthermore, I will explore any relevant literature which will provide me with any techniques, such as projecting positive body language, which in turn will help me develop my confidence in these situations. Progress in my development will be measured intrinsically through personal satisfaction of performance and also from feedback from colleagues. In addition, a long term aim is to develop my skills to a standard whereby I am able to successfully deliver content and ideas at conferences or in academic settings. I feel that the process of completing this research project will provide me with an excellent opportunity to develop my oral communication skills as sharing project developments and ideas in an academic setting will be required. Furthermore, completing the viva process in a pressure situation will also provide invaluable presentation experience.

The second area of development which I feel warrants attention is my ability to engage in critical reflection of practice. Within professional practice, critical reflection has become and essential and valued requirement and to date, I have experienced little exposure to this method of personal development. I feel that my development as a practitioner will benefit substantially from becoming compliant in critical reflection. Evidence of this is supplied simply by possessing no record of previous reflective practice and being unaware of critically reflective techniques. I feel that the Professional Doctorate programme will allow me to address this area as one of the course modules requires me to provide a documented reflective account of the research process. This will compel me to develop my reflective techniques and refine an individualised process. By analysing my researcher competencies, it is apparent that I could have expanded on numerous areas which I feel warrant development. It should be noted however, that as a first stage researcher, I expect that my competencies will lie within the initial stages of development and that for the purposes of this self-assessment, I have highlighted the competencies which I feel are most important to me as a practitioner. On completion of this self-appraisal process, I recognise that there are other areas which warrant development, however, I feel that these areas are not necessarily directly linked to this current project and are more relevant perhaps to a future job role. In spite of this, I plan to re-visit these researcher competencies following the conclusion of this current project to further my skill set as an applied practitioner.

PROFESSIONAL COMPETENCY PROFILE

As a sports science practitioner, it is logical to aspire to be recognised and accepted as an accredited member of my governing body, which in this case, is the British Association of Sport and Exercise Scientists (BASES). To assess my competencies as a practitioner, I applied the BASES accreditation competency profile to my own professional practice and evaluated my competency using a scale of 1-4. Results from my evaluation are presented in **figure 2**.

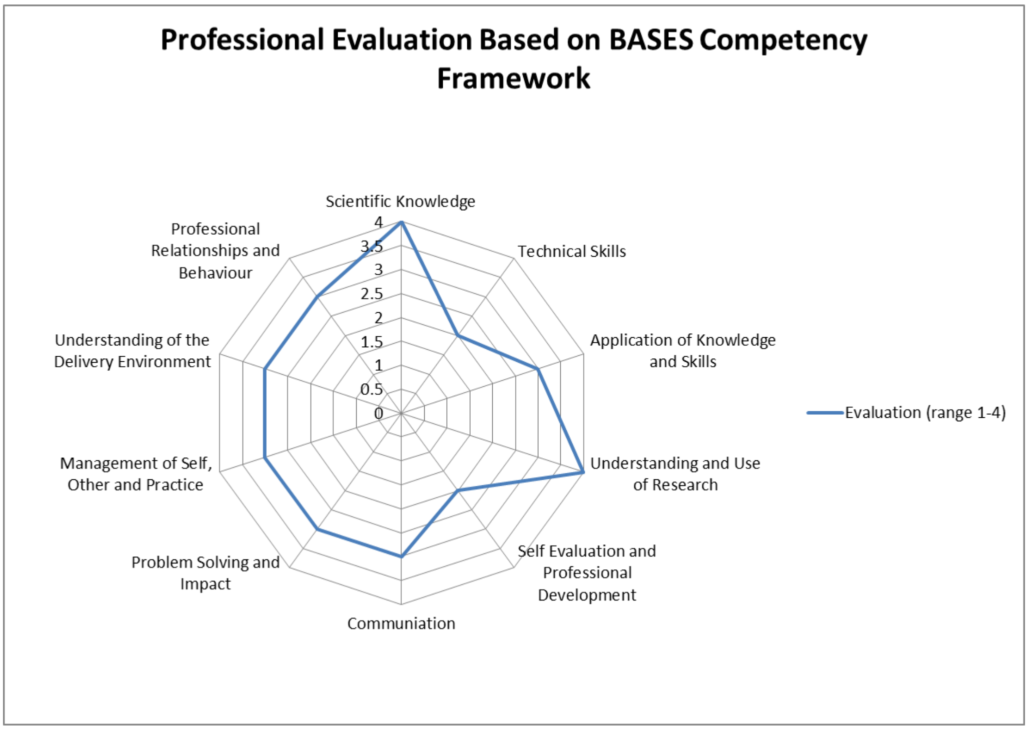

Figure 2. Evaluation of my professional practitioner competencies determined by the BASES accreditation competency profile.

Competencies that I feel warrant further elaboration for self-development purposes are technical skills and problem solving and impact. From my education and work experience to date, I have had limited exposure to laboratory techniques and various technology now commonplace in the sports science field which is a direct result of financial limitations and availability of equipment and technology (such as GPS) within my organisation. Although I feel I possess the required knowledge to effectively utilise such technology, I feel a lack of hands-on experience in day-to-day practice inhibits my professional skill set in this area. Lastly, I feel that in order to become a recognised practitioner in my field, it is important to me personally and for my organisation that I begin to contribute to the development of relevant research projects. However, to date, I acknowledge that I have not made any significant contributions to the research in my field and, in order to develop this key competency further, I plan to contribute to the scientific research in my chosen field by means of accepted publication and to continue to contribute to the literature beyond the completion of my Professional Doctorate study.

BEHAVIOURAL CHARACTERISTICS

To assess my behavioural characteristics, I utilised the Behavioural Profile Report developed by PDA International (2004), the results of which are depicted in **figure 3**. The Behavioural Profile Report is an internationally recognised method of assessing behavioural characteristics and tendencies which can influence how an individual operates in professional practice and is used by many organisations worldwide.

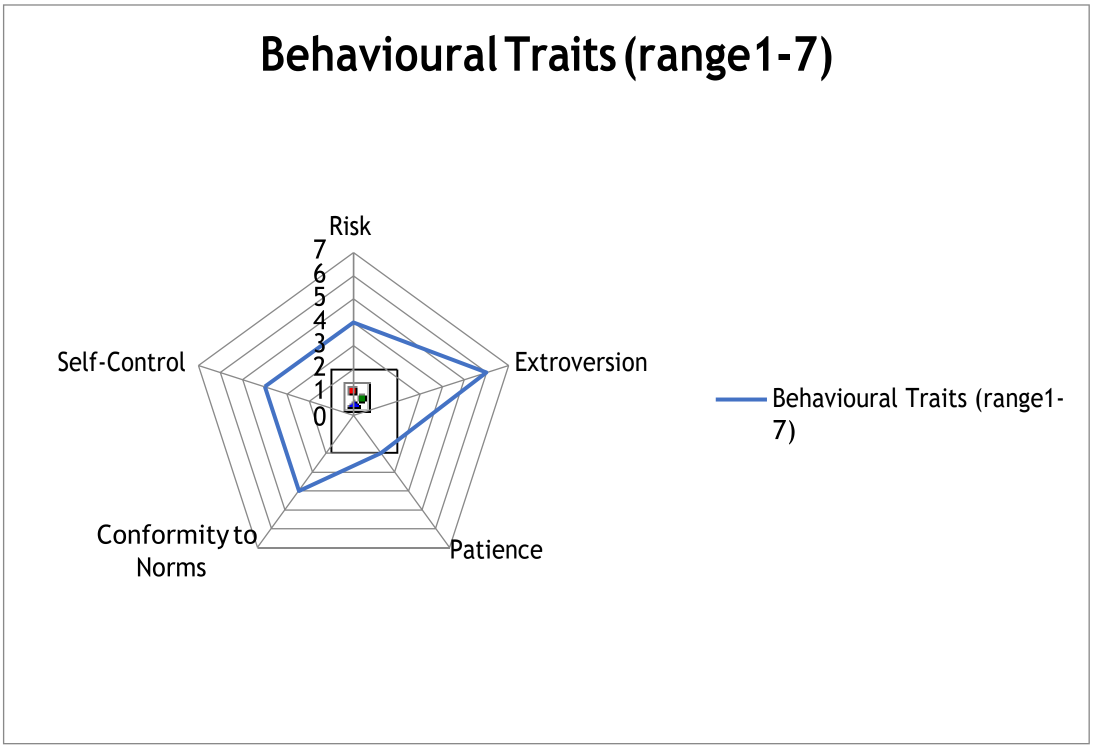

Figure 3. Behavioural traits as determined by the Behavioural Profile Report (PDA International, 2004).

From the report, the two main areas which have fallen towards the 'intense' side of the scale are extroversion and patience. The report accurately revealed that I am a sociable person and enjoy working as part of a team. I feel that this is an important personality trait in my current professional role as the success of my project relies on using participants from groups that have tight training schedules and it is important that, in some circumstances, I am able to successfully convince team coaches to sacrifice some of their training time for the purpose of data collection. This means that the ability to build effective inter-personal relationships becomes an important factor. Furthermore, the report further suggests that "I am eager to please" and require regular appraisal to consolidate performance. In my current job role, no formal procedures are in place which provides feedback of professional performance such as formal appraisals or annual reviews. I feel that this situation makes it difficult to assess my performance as a practitioner and that in order to ensure on-going progression; this is a vital process which must be addressed. Furthermore, as I am a sole member of staff in my department, I lack a support network and mentor to interact with which can be useful in circumstances where important decisions have to be made. Following this, the results of the Behavioural Profile Report also identified that I am an impatient person and have a desire to obtain results quickly. The report stipulates that I am proactive and like to get things started quickly which in turn can lead to trying to manage too many projects at the one time. Implications for impatience also lie in the fact that, although I am very evidence based, it is possible that I may make impulsive decisions which could consequently affect the effectiveness of my decisions.

In summary, it is clear to me that in order to progress as a practitioner; I must seek to install methods of regular formal professional appraisals from superiors in order to ensure continuous quality service provision. Furthermore, I should try to establish a support network through which I can present ideas and receive critique or justification for any decisions that are required to be made. Lastly, I should think carefully about which projects to take on and slow down my decision making process by taking time to contemplate and critically reflect. On closer observation of the other personality traits outlined in **figure 3**, there doesn't appear to be any other potential issues that warrant further expansion in terms of impacting the success of my research project.

LITERATURE REVIEW

INTRODUCTION

In modern football over the years, steady improvements in physical and technical performance have resulted in the evolution of our game making it more exciting than ever before (Barnes *et al.*, 2014). Clubs throughout the world invest millions on developing the best academy structures and research centres, which are invariably built around a clubs individual playing philosophy, to develop the next star of the future that will hopefully progress onto first team and international level. However, in the case of talent development, it is still unlikely that we possess all the necessary objective data to clearly define the specific attributes a young aspiring footballer requires to successfully compete at the highest level. By being able to combine subjective data obtained from many years of quality coach practice with objective data attained from scientific investigation, we could better equip our clubs with valuable insights into talent development which in turn would reduce the element of chance and improve the prospects of player progression through scientific process in order to aid the highly complex and multi-factorial process. Therefore, it would seem logical to assume that the more objective data we have at our disposal, the better the chances will be of improving the process of developing a high-quality future generation of players.

To date, sports science research has had a strong focus on improving physical performance and injury resistance through sub-topics such as strength and conditioning research, training load optimisation and athlete monitoring, however, one important aspect of performance which is perhaps often overlooked in the literature is technical performance and how this can be best optimised throughout a player's development (Reilly & Gilbourne, 2003). Perhaps one of the reasons why this is the case, is, that measuring and tracking the technical performance of young players through their developmental years and into their professional careers is difficult to implement with no fully accepted 'gold standard' method of assessing performance. Furthermore, it is important to understand why measuring technical performance is important and the value of the data that the process can provide. Relatively recently, Barnes *et al.* (2014) reported, through the largest collection of match data to date utilising 22,846 player observations from players in the English Premier League, that over 7 consecutive seasons, we have seen a steady improvement in players technical ability with players performing over 40% more passes, receiving 17% more passes and having an 8% better pass completion rate in the most recent season compared with the first season of data collection. This study alone highlights the importance of technical ability in modern football. In further support of this analysis, Rampinini et al. (2009) and Dellal et al (2011) also demonstrated the importance of technical performance and the impact of superior technical performance has on success across 3 top European domestic leagues. By taking into consideration the data presented here, it would seem pertinent and imperative that we are able to assess and monitor the technical ability of our aspiring players. In order to achieve this, we must first recognise the need for the development of a valid and reliable assessment protocol which is specific to the aspects of performance we want to measure. Previous research studies have adopted a magnitude of various assessment protocols with varying degrees of success and limitations. In order to further develop a robust assessment protocol which can be easily applied in a practical setting to aid in the continuous development of our players, we must first explore the existing literature to gain an insight into the research area.

'Closed Skill' Assessment Protocols

Several authors to date have attempted to pioneer the development of various tests that assess aspects of technical performance in youth footballers. A wealth of previous research exploring the development of assessment protocols for measuring technical performance has focussed primarily on assessing individual aspects of technical performance in an isolated skills assessment situation (Ali et al., 2007; Borges et al., 2017; Huijgen et al., 2013; Keller at al., 2016; McDermott et al., 2015; Praca et al., 2015; Reilly et al., 2000; Rubajczyk & Rokita, 2015 & Wilson et al., 2016). Two of the most frequently adopted assessment protocols of this nature include the Loughborough Passing Test (Ali et al., 2007) and the General Soccer Ability Skill Test Battery (Mor & Christian, 1979). Both tests require participants to successfully complete a football specific task in the shortest time possible. For example, the Loughborough Passing Test requires the participants to complete 16 passes against pre-determined targets in the quickest time possible. Results from studies adopting protocols such as these have shown that the tests are capable of successfully distinguishing between elite and non-elite players (McDermott et al., 2015; Huijgen et al., 2013 & Keller et al., 2016).

One major limitation with the implementation of 'closed skill' tests to assess technical performance lies within the factors relating to the ecological validity of the test and how it correlates to actual competitive game performance. The process of football performance has been identified as a series of actions which comprise of communication (information gathering from the surrounding environment), decision making (what to do) and action (execution of a skill) (Praca et al., 2015). This process requires both an internal and external focus, so therefore, it could be argued that tests isolating only one aspect of technical performance that do not require a great deal of external focus and in turn do not represent competitive match play scenarios. In support of this, Rubajczyk and Rokita (2015) reported in a study of young football players in Poland that performance in 'closed skill' football specific tests correlated poorly with performance in a 5v5 small sided game related test (0.325-0.452). However, it must be highlighted that, in this particular study, performance in the small sided game test was measured using a subjective opinion of 5 'expert' coaches resulting in an intra-observer coefficient of variation ranging from 34-37% and therefore the reliability of this method of assessment should be taken into consideration. In addition to this, Praca et al (2015) also state that "good technical performance requires stable structures that are able to elicit similar responses in similar contexts as well as flexible structures that are able to permit the execution of a technique based on the person-environment task relationship i.e. in a game situation", providing support for the development of an objective assessment protocol which takes into consideration the demands of competitive match play.

'Game Related' Skill Assessment Protocols

The concept of assessing technical performance in competitive match play situations has been previously investigated by a number a research teams and has resulted in the development of various methodologies. Previous research has attempted to develop a model which takes into consideration the element of external focus that is required during match play, many of which have evolved from a physical education perspective as opposed to elite sport (Garcia-Lopez et al., 2013; Grehaigne, 1997; Nadeau et al., 2008; Oslin et al., 1998 & Waldron& Worsfold, 2010). One of the earliest and perhaps most innovative attempts, derived from physical education teaching literature, to develop a game-based skills assessment model was conducted by Oslin et al (1998) and was named the Game Performance Assessment Instrument (GPAI). This model involved participants competing in a small-sided game format, during which, observers would observe each individual performer for a 10-minute period and record efficient skill executions, inefficient skill executions, effective decisions made and ineffective decisions made using a hand notation system. Results from the initial development studies demonstrated that the test was reliable for skill executions across 3 sports (0.971, 0.844 and 0.850 for soccer, basketball and volleyball respectively). Furthermore, the study reported that the assessment was able to discriminate between 'high performers' and 'low performers' with effect sizes ranging from 1.58-1.93 across all 3 sports and also inter-observer reliability scores that ranged between 0.73-0.97. However, although this original research article was innovative, various limitations have been identified and in a follow-up article by Memmert and Harvey (2008), the authors raised some concerns about the GPAI and also offered some suggestions on how to rectify them. One issue raised by Memmert and Harvey (2008) was the subjectivity of the coding system used to discriminate between efficient and inefficient skill executions. This came to fruition following the perhaps over-simplistic framework by which the observers assessed performance (**table 1**). Furthermore, there are several other factors that should also be taken into consideration which will be briefly discussed: 1) as the player is assessed in real-time,

1. Decisions Made	Player chooses to pass to an open teammate Player chooses to shoot when appropriate
2. Skill Execution	Reception – Control of pass and set up of ball Passing – Ball reaches target Shooting – Ball stays below head height and is on target

3. Support	The player appeared to attempt to support the ball carrier by being in/ moving to an appropriate position to receive a pass

Table 1. Criteria for appropriate/ efficient action rating (Oslin et al., 1998):

it could be possible that some actions may be missed by observers; 2) If there was any doubt about the coding of one particular action and the observer hesitates in any way, this could also affect the outcome of the selected code and finally 3) By observing a 10 minute snapshot of performance, it could be possible that the player is going through a 'quiet spell' where the number of technical actions performed is reduced which may not provide an accurate representation of true performance.

Following on from the work by Oslin et al (1998), Garcia-Lopez et al (2013) also developed a new novel approach to assessing technical and tactical performance solely for the purpose of football. The main difference between the two research projects was that Garcia-Lopez further developed the game based model by introducing a video analysis component to the protocol by which small-sided game situations were recorded and subsequently re-visited for coding and analysis. However, attention is drawn to some of the limitations relating to this research which should be considered: 1) no test/ re-test measure was taken to ensure reliability of the assessment protocol 2) validity is only measured by investigating the relationship between performance and level of expertise and is perhaps more suited for talent identification purposes rather than talent development and finally 3) the framework describes detailed descriptions of coding events for technical and tactical parameters which include both on the ball and off the ball actions; however, it could be suggested that the framework has now perhaps become over-complicated and labour intensive which questions the study's practicality and justification for its inclusion in an overall test battery - thereby further supporting the need for a practically applicable and easy-to-administer technical assessment protocol which takes into consideration the demands of competitive match play.

Previous research teams have adopted technical ability assessment procedures based around the foundations of a small sided game format, during which, performance is assessed. This approach makes sense given the number of ball interactions during small-sided games compared with competitive matches (Kelly & Drust, 2009; Barnes et al, 2014) thus allowing players to perform a high number of technical actions in a shorter time frame making analysis more time-efficient. In contrast to this, Waldron and Worsfold (2010) designed and implemented an inventive research protocol which aimed at assessing differences in game specific skills between elite and sub-elite youth football players during competitive match play. However, the applicability of this method of assessment for monitoring the continuous development of young players within an academy setting would appear to be one of its main limitations. By employing this methodology, sport scientists/ analysts are required to analyse data for every player during full or partial duration competitive matches meaning that the process becomes very labour intensive and time consuming with a high number players demanding analysis. Furthermore, continuous monitoring of players is required to track progress in development and therefore ongoing analysis throughout the playing season would further contribute to the labour and time demands.

In summary of the selected previous research studies presented in this literature review, several consistent limitations appear to emerge from the methodology. Firstly, there appears to be a wide variety of game formats and conditions during which performance is measured. Such variation in conditions could consequently affect technical performance as changes in available space, and in turn, time on the ball may affect the decision-making process and resulting execution of technical actions. Secondly, with the exception of the 'closed skill' tests, no research study has reported reliability statistics meaning that no methodology can be considered as being 'gold standard' due to uncertainty around reliability measures. Thirdly, all previous research studies to date have focussed their efforts around discriminating between 'elite' and 'sub-elite' players which provides a useful tool for the purposes of talent identification and in turn a strong construct validity measure, however, to my knowledge no study has reported any methodology's sensitivity to change over a longitudinal period which is a crucial factor in talent development (Vaeyens et al, 2008). Lastly, the aforementioned studies apply a multitude of various coding systems that range from perhaps over-simplistic, to over-complicated for analysis of technical performance which in turn may influence the practical applicability of the assessment protocol.

Despite these limitations, it can be agreed that utilising a small- or large-sided game format containing a video analysis component provides compelling objective data about the technical abilities of youth soccer players. It is therefore the opinion of the author that a consistent methodology with simpler practical applications would provide a valuable tool for assessing a critically important aspect of performance. With this in mind, the aim of this research project is to remodel the existing research in order to develop an easily applicable, valid and reliable technical ability assessment protocol which is capable of tracking improvements in performance over a period of time for the purposes of talent development in elite youth footballers. Following on from this paragraph, a brief outline of the proposed studies which will construct this research project will be presented.

Study 1

Development of Assessment Tool – Content Validity and Initial Instrument Design

Experimental Approach to the Problem

This study will investigate coaches' perception towards which attributes are deemed most important for elite technical performance. Using a small group of expert coaches, the DELPHI technique (Dalkey & Helmer, 1963) will be applied to explore personal opinion and subsequently, data will be consolidated and filtered down until all coaches agree on specific content to be included in assessment protocol which in turn will determine the framework. The framework will also attempt to incorporate a tactical element which will assess decision making, a key determinant in the success of technical actions (Garcia-Lopez et al, 2013). Following this process, the assessment tool will be created by combining the newly acquired content data and a previous research design which utilises a small sided game protocol for the purpose of talent identification (Unnithan et al, 2012).

Study 2

Development of Assessment Tool – Reliability

Experimental Approach to the Problem

This study will determine the reliability and reproducibility of the newly developed assessment instrument. The assessment protocol will be administered to a sample of youth football players within a professional academy. Within 7 days, the assessment protocol will be administered again and a reliability statistic determined. The assessment protocol will be analysed retrospectively using videos obtained during the procedure and analysis will be performed by the same observer to provide details of intra-observer reliability.

Study 3

Using the Assessment Tool – Face Validity

Experimental Approach to the Problem

This applied project will investigate the extent to which the assessment tool is sensitive enough to detect changes in technical performance for the purpose of talent development. The project will employ a longitudinal intervention study which will involve analysis of assessment results over a competitive season. The assessment will be administered to sample of players at 3 time points throughout a competitive season during which players will carry out their habitual training routine in line with the clubs coaching strategy. Following data collection, results will be analysed to determine any statistical differences between the assessment protocols. Furthermore, effect sizes will be calculated to determine the magnitude of change. Results of this study should demonstrate a significant improvement between assessment 2 and assessment 1 and also between assessment 3 and both assessments 2 and 1.

Study 4

Using the Assessment Tool – Face Validity

Experimental Approach to the Problem

This study will further improve on the protocols validity by investigating the relationship between the test and real-life competitive match play. Four competitive matches will be recorded and subsequently analysed from a sample of youth footballers within a professional academy. The same coding and scoring system used in the skills assessment will be used to provide an objective analysis of technical performance in competitive match play. A correlation analysis will then be applied to determine any relationship between performance in the technical ability assessment as determined by objective results generated by the assessment instrument and competitive match play.

LOGISTICAL RESEARCH PLAN

The final section of this research plan will provide a SWOT analysis and graphical depiction of proposed time frames for completion. The SWOT analysis will highlight any potential internal or external concerns which may pose a threat to the success of this project. The SWOT analysis can be seen in **table 2.**

INTERNAL		EXTERNAL	
Strengths	**W**eaknesses	**O**pportunities	**T**hreats
1. Part of day-to-day practice so is easy to collect a large data set (initial studies) 2. The complexity of the project is relatively low and manageable meaning it should be easily implemented	4.Failure to scientifically validate the tool which will consequently impact the success/ possibility of completing the remaining research studies	7. In my job role, I have full control over the SS side of the programme and therefore I can implement the investigation freely 8. The research hopefully has the potential to have a	10. Termination of employment 11. Fitting aspects of the research into an already busy schedule i.e. first team data collection, training intervention study

3.Follows a logical progression from initial development of the tool into practical application which will intend to make it highly relevant in a practical setting	5. May be overly ambitious and time consuming 6. Requires the use of techniques I am unfamiliar with (video analysis)	positive impact on my organisation and other organisations in Scotland 9. Will benefit my club in the long term and will provide a powerful tool to aid our decision making process as a multi-disciplinary team	12. Variance in day-to-day plans and schedules e.g. re-scheduling of training to accommodate games 13. Changes to academy structure in Scotland i.e. Winter Seasons

Table 2. A SWOT analysis identifying internal strengths and weaknesses of the research idea and external opportunities and threats which have the potential to influence the success of the project.

One of the potentially limiting factors observed from the SWOT analysis which presents itself in both the internal (weaknesses) and external (threats) sections of the analysis relate to the issue of time management (5 and 11). One method of optimising time management is through the use of a GANT chart (**figure 5**) which provides visual and schematic representation of how the various components of the project will be structured in order to ensure timely completion.

The proposed timescales presented in **figure 5** are constructed around the current structure of the Youth Academy Programme and senior competitive season in Scotland. It should be noted however, that due to a potential re-organisation of the Youth Academy Programme, it is possible that the competitive season for academy players in this country may be moved to the winter months, meaning that the GANT chart presented in **figure 5** will have to be re-configured to accommodate for this process. In spite of this, I am confident that there will be no issues regarding data collection.

The final item to be discussed following the presentation of the **table 1** and **figure 5** relates to project 2. As depicted in **figure 5**, it can be seen that the data collection period for this segment of the project is spread almost entirely throughout the 3 year data collection period. The reason for this, is because due to the fact the project requires data collection from our 1st team players, I will need to seek permission from the first team manager to carry out the data collection and the time he grants me permission to do so will be completely at his discretion. Due to the nature of the technical ability assessment tool (small sided game), I do not foresee any issues gaining permission to carry out the assessment as the team regularly uses small sided game protocols as part of their habitual training regime.

In conclusion, this section of the research plan, which considers logistical deliberations, is a vital component of the research project process as factors such as time management and organisation can ultimately determine its success. I am confident that I have considered all the internal and external factors which may pose a threat to the completion of my project and furthermore, I feel have I have established realistic time scales which can be adhered to in order to successfully and simultaneously complete my project within my day-to-day practice.

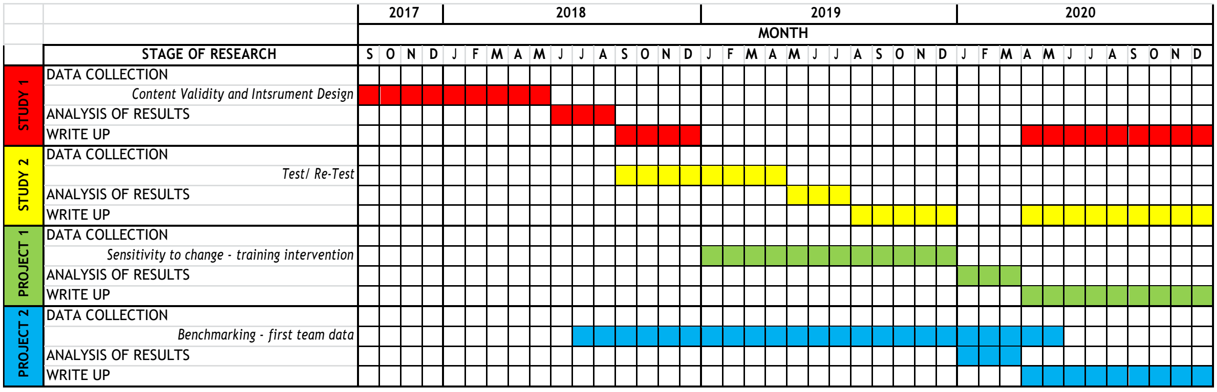

Figure 5. A GANT chart outlining a basic, realistic time plan for completion of the research project.

Appendix 2 Practical Applications: Talent Card