Addressing the problem of 'big data' in sports: A framework for performance analysts.

Hayden G. Croft

A Thesis Submitted to Auckland University of Technology in Fulfilment of the Requirements for the Degree of.

Doctor of Philosophy (PhD)

July-2021

Primary Supervisor: Associate Prof. Kirsten Spencer

Associate Supervisor: Prof. Sam Robertson

Associate Supervisor: Prof. John Cronin

Mentor: Associate Prof. Megan Gibbons

Auckland University of Technology, Auckland, New Zealand

Abstract

An explosion of data producing technologies has changed the requirements of the modern performance analyst. It has been claimed that in 2020 approximately 50 billion 'connected devices' existed, signalling that the 'era of big data' had begun. Coaches now have access to unlimited, unstructured and potentially uncontextualized information, which could reinforce their biases. New technologies exist for collecting, organising, storing, and presenting 'large' or 'big data,' however there are limited framework for performance analysts to follow when using these tools in sport. Without a framework key aspects of tool development may be missed, and errors or poor ecological validity may eventuate. The objective of this PhD was to create a framework for performance analysts to use when building information presenting tools. The over-arching question of this thesis is: what process, or framework, will allow an analyst to present large data sets to coaches in a relevant, meaningful and ecologically valid manner?

A literature review was conducted to identify how a framework for 'large' or 'big data' may best be approached. The following areas were identified as requiring investigation: a) coach tactical behaviours during matches; b) valid data sources in sport; c) data organisation systems; and d) visualisation techniques for coaches. A systems design approach and action design research methodologies were used to guide the development of the framework.

The first stage of the framework encourages performance analysts to observe the behaviours of coaches, during matches, and summarise these identifying the most common themes. This was conducted in the sport of netball, which was used as a context for the other stages of the development in this PhD. Once these themes were identified, data sources were investigated and evaluated in Stage Two. In Stage Three, a large sample of suitable data was collected, and a cloud-based data pipeline was developed to store it in a database as information. Stage Four took this information, used it to create population normative values, for each performance indicator, and coaches guided the representation of these. Coaches were then presented, in Stage five, with the aligned information, live during netball matches, and their behaviours were observed.

At the completion of the five stages a working example of the framework, embodied in a tool, was presented to coaches and sport science practitioners and survey feedback collected. This tool was found to be ecologically valid, to the users, and future research was identified to improve these tools. The framework is proposed as valid method, for performance analysts, in creating tools that assisted coaches navigate large datasets in netball and other sports.

Table of Contents

List of Figures

List of Tables

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extend has been submitted for the award of any degree or diploma of a university or other institution of higher learning.

Hayden G Croft

9/7/2021

Co-Authored Works

The publications listed below are a result of the research conducted in fulfilment of the degree of Doctor of Philosophy.

- Chapter 2.0 Manuscript 1:

  Croft, H., Spencer, K., Robertson, S., & Cronin, J. (2021). The problem with big data and coaching: closing the gap – a review. *Submitted to Journal of Human Sport and Exercise (4th June 2021)*

  Author contributions: Croft, H (85%), Spencer, K (5%), Robertson, S (5%), & Cronin, J. (5%).

- Chapter 3.0 Manuscript 2:

  Croft, H. G., Spencer, K., & Robertson, S. (2020). What tactical and technical comments do coaches make during netball matches? A content analysis in netball. International Journal of Performance Analysis in Sport. DOI: 10.1080/24748668.2020.1846112

  Author contributions: Croft, H (85%), Spencer, K (10%), Robertson, S (5%)

- Chapter 4.0 Manuscript 3:

  Croft, H., & Spencer, K. (2021a). An evaluation of data sources in netball performance analysis. *(Under review) The Journal of Sport and Exercise Science.*

  Author contributions: Croft, H (90%), Spencer, K (10%)

- Chapter 5.0 Manuscript 4:

  Croft, H., Spencer, K., Taurua, N., & Wilton, E. (2021). Creating a live and flexible normative dataset for netball. *Frontiers in Sport and Active Living.*

  https://doi.org/10.3389/fspor.2021.743612

  Author contributions: Croft, H (85%), Spencer, K (10%), Wilton, E (3%), Taurua, N (2%)

- Chapter 6.0 Manuscript 5:

  Croft, H. G., & Spencer, K. (2021b). Coach engagement with live netball statistical information. *(Under review). Frontiers in Sport and Active Living.*

  Author contributions: Croft, H (85%), Spencer, K (10%), Robertson, S (5%)

- Chapter 7.0 Manuscript 6:

Croft, H., Spencer, K., & Bowden, M. (2021). Developing information tools for presenting coaches with large datasets. (*Under review*) Journal of Computer Science in Sport.
Author contributions: Croft, H (80%), Spencer, K (10%), Bowden, M (10%)

Presentations

Croft, H. (2020). An evaluation of data sources in netball performance analysis. *Sport and Exercise Science New Zealand Conference*, Christchurch, New Zealand November 25 - 27th.

Croft, H., & Spencer, K. (2020). An evaluation of data sources in netball performance analysis. *Otago Polytechnic, Institute of Sport, Exercise and Health, Intensive Week Symposium*, 10th November 2020, Sargood Centre.

Croft, H., Spencer, K., & Robinson, S. (2018). What tactical and technical comments do coaches make during sports matches? A qualitative analysis in netball. *World Congress of Performance Analysis of Sport 12*. Opatija, Croatia. September 19-23. 2018.
https://www.researchgate.net/publication/327667680_ISPAS-2018-final

We, the undersigned, agree to the percentage of contribution to the chapters identified above:

Supervisors

Kirsten Spencer

Sam Robertson

John Cronin

Mentor

Megan Gibbons

Collaborators

Emily Wilton

Noeline Taurua

Matt Bowden

Acknowledgements

There are many people that should be acknowledged for their contributions to this PhD. Firstly, to my supervisors who have all guided me through this process. Kirsten Spencer, you have been a positive and consistent influence throughout this PhD. I am most thankful for your flexibility to let me run off and work as a performance analyst and then refocus on this thesis after all the ups and downs of international sport. Your perspective has helped guide what can be very difficult to conduct, applied research. Sam Robinson and John Cronin, my second and third supervisors, who provided crucial guidance at the beginning and end of this thesis. It was fantastic to have input from two experienced and thorough academics. For creating a flexible and supportive workplace my Dunedin and Otago Polytechnic (OP) based mentor, Associate Professor Megan Gibbons, must be thanked. It can't be understated how important these aspects of OP have contributed to me completing this work.

Important contributions were made by Dr Emily Wilton and Dame Noeline Taurua who provided coaches perspectives to manuscript 4, and Matt Bowden for his skills in developing visualisations for the tool in manuscript 6. All participants, although anonymous, have been crucial to the learnings and findings of this work and for this reason an executive summary will be created to inform netball coaches and sport science staff of the practical applications of this work.

I would like to thank Netball New Zealand and the Silver Ferns high performance staff. Debs, Esther, Hotts, Mark, Mel, Guy, Rodney, and Noels (again). You have all had a huge impact on my thinking, around netball, throughout this PhD, thank you.

On a more personal note, I need to thank my family for pretending they understand what I do :). Even when it seemed that I probably wasn't doing my PhD anymore. The good thing about my family is they are always positive, challenging and kept my work grounded and applied. It has been a challenging few years, which have taken their toll, however it has shown me how close we all are. There was always someone at home or on the road helping when we needed it most.

Finally, and most importantly, to my wife Laura. You make a huge footprint on the world through your work, supporting people and families. I can only hope that my work has a fraction of this impact! You have two dogs, three kids, a career and a husband who travels (sometimes)…… what can I say. Your (you're) awesome :).

Finally, Emily, Oliver and Benjamin. I can come home from a big day of work and you remind me that coaching hockey, or footsal, or netball or…….. is more important than work.

Ethical approval

Ethical approval was granted by the Auckland University of Technology Ethics Committee (AUTEC) - Ethics Application

Ethical Approval

Ethical approval was grant for this PhD under the following applications:

16/267 Determining what topics or themes are important to performance for team sport coaches in real-time (Manuscript 2).

18/68 Data sources and collection in sport: (Manuscript 3). Organising, normalising and analysing different type of sports data (Manuscript 4).

21/33 Determining what statistical information is important to team sport coaches in real-time (Manuscript 5 - observation, Manuscript 6 - survey)

1.0 Introduction

1.1 Literature review

Sports performance analysis is a sub-discipline of sport science and is also a profession recognized by many sports as well as governing bodies such as the International Society of Performance Analysis of Sport (ISPAS). The role of the performance analyst is to, 'systematically observe and record athlete performance' and once this has occurred, 'facilitate and enhance feedback between coaches and athletes' (Sport and Exercise Science New Zealand, 2019). This definition is a simple explanation of the profession, however, the ways in which this occurs can depend greatly on situational and environmental factors. Hughes and Bartlett (2019) explained performance analysis as a combination of the fields of biomechanics and notational analysis, due to both fields sharing many similarities. This could be disputed as too narrow as Glazier (2017) explained other fields, including physiology and psychology, should be included to promote a more holistic approach to sports performance analysis.

Regardless of these fields the purpose of performance analysis is generally agreed upon, which is to record and process information, interpreting events and assess performance in an objective manner. This may occur in a practice or competition setting (Drust, 2010; Laird & Waters, 2008; Mooney et al., 2016; Middlemas et al., 2018;). The common methods that are used to define the field include the use of information, either visual feedback (video analysis) or objective statistics (data analysis), to improve decision making (English Institute of Sport, 2021; O'Donoghue, 2015). The subdiscipline of video analysis has been researched with investigations into timing, content and efficacy (Gil-Arias et al., 2016, 2019; Groom et al., 2005, 2011, 2012; Guadagnoli et al., 2002; Martínez et al., 2016; Middlemas & Harwood, 2018; Nicholls et al., 2018). A coach can now learn the best ways to provide visual augmented feedback to their players to enhance learning (O'Donoghue P. , 2006). Through these studies a greater understanding has been developed on the effects of video on player learning, motivation, and confidence.

The use and understanding of objective statistics have been explored in various ways with a common approach being the creation of performance indicators (PIs). Many studies have looked to define PIs and key performance indicators (KPIs) for sports to help players and coaches understand which contribute to successful performance. Examples include football (Lago-Peñas et al., 2011), netball (Bruce et al., 2018), rugby (Jones et al., 2004), basketball (García et al., 2013) and Australian football (Young et al., 2019), studies which all made recommendations to what the factors were that related to winning and losing. However as explained by Ofoghi et al (2013) there could be the potential for a 'straightforward' analysis, which looks to link a few predictor variables with a dependant variable, to lose some of the deeper context within the dataset. Complex analysis techniques, including data-

mining, need to be investigated as they can reveal underlying relationships between factors that could influence sports performance. The benefits of this approach have been demonstrated using various complex analyses in decathlon (Cox & Dunn, 2002), rugby (Vaz et al., 2010; Lamb & Croft, 2016), netball (Croft, Wilcox, and Lamb 2018) and basketball (Lamb et al, 2010). These studies created insights using various forms of cluster analysis including techniques like self-organising maps. Ofoghi et al (2013) identified four distinctive data-mining categories, these were clustering, classification, relational modelling and rule mining. All these techniques require large datasets which need filtering, format conversion, extraction, structural conversion and descriptive conversion.

How is PA used for learning across a range of sport currently

Performance analysis research is extensive and diverse across a range of sports with football being frequently studied. Various systematic reviews have summarised our understanding of the technical, tactical, physical and psychological aspects of the sport (Caicedo-Parada et al., 2020; Goes et al., 2021; Low et al., 2020; Sarmento et al., 2014, 2018; Whitehead et al., 2018). Within the technical and tactical requirements of football several studies have explored positions, coordination and structures of players and teams (Bueno et al., 2021; Goes et al., 2021; Jara et al., 2019; Ortega et al., 2016; Rein & Memmert, 2016; Rey et al., 2015). In the physical domain global positioning technologies have been used to explore the physical characteristics of players at various levels of competition (Abbott et al., 2018; Barnes et al., 2014; Bloomfield et al., 2007; Bradley et al., 2013; Jara et al., 2019; Ortega et al., 2016). Basketball also has several systematic reviews which include a performance analysis focus including passing (Maimón, Courel-Ibáñez, & Ruíz, 2020), referee decision-making (García-Santos, Gómez-Ruano, Vaquera, & Ibáñez, 2020) and physical demands on players (Petway , Freitas, Calleja-González, Medina Leal , & Alcaraz, 2020). Research into other sports, like Australian Football (Manson et al., 2021), has highlighted the benefits of targeted feedback, on performance, or modelling player actions and movements with match outcomes (Young et al., 2019; Hiscock et al., 2012).

Netball and performance analysis

Netball is a team sport, similar to basketball, in that players score points from 'shooting' a ball through a hoop, however there is no backboard, players cannot dribble the ball, and different positions are only allowed to move into specific areas of the court. Like basketball, netball leans itself to performance statistics and some commercial providers of this data exist (Champion Data, 2020; Point 9 Analytics, 2019). Netball is a sport that utilises performance analysis, with 70% of the top 10 ranked international teams employing a performance analyst within their management team during the 2019 Netball World Cup. Performance analysis research also exists in netball with many researchers looking

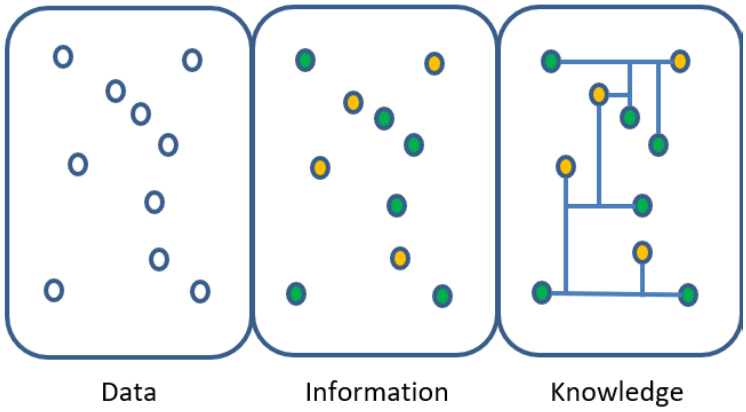
at performance indicators, modelling complex data for patterns, injury prevention and competition wide trends across time (Bruce et al., 2018; Croft, et al., 2018; Fox et al., 2013; O'Donoghue et al., 2008). Research by McLean et al (2017; 2019) has looked at both netball and football, describing the need for a domain-based approach to analysis considering 'sum of all parts' instead of traditional indicators of performance and predictors. This is also expressed by Smith and Bedford (2020) who investigated computer technologies which could automate data collection in netball. These studies both make claim to the need for more complex systems analysis of netball due to the lack of traditional performance indicator research linking predictive factors with winning.

Performance analysts in netball frequently draw on research (Groom & Cushion, 2005; O'Donoghue, 2006; Middlemas & Harwood, 2018; Mason et al., 2021) to inform the design of their video and feedback process. Due to the abundance of data and statistical information available in netball, processes and a framework are needed to help the performance analyst inform the design of information feedback tools for use by players and coaches.

The era of big data and information overload

According to Mayer-Schönberger and Cukier (2014) there has been an 'explosion' of data since the 2000's due to new technologies that enable data to be generated about almost anything. Data is defined in many ways, however this PhD will follow the following definition. 'Data are raw material of information, typically numeric' (Zins, p. 485, 2007). This concept can be described as 'dots' on a page (Figure 1) (Kaushik, 2016). Information, however, is data which is 'collected together with commentary, context or analysis so as to be meaningful to others' (Kaushik, p. 1, 2016). This would be adding colour to the dots as seen in Figure 1. Often data and information are used interchangeably, however, this does not acknowledge the relationship between the two. It can also create confusion for coaches interpreting large amounts of data as information, which has little to no meaning.

Figure 1. *A visualisation of the continuum from data to knowledge, adapted from (Kaushik, 2016).*



In sport, this saturation of information is so great that some coaches may not have the time or cognitive resources to integrate it effectively into their decision making. Due to the existence of this information, a coach could potentially select information only to reinforce their opinions, meaning that the recent 'explosion' of data may in some cases be to the detriment of coach decision making. Products such as Sportscode™ and Prozone™, enable coaches to access large sets of descriptive statistics and live video replays during matches. In football the use of big data has been systematically reviewed (Goes et al., 2021) with a framework proposed for the integration of sport science and computer science. Wearable microtechnology which generate these data have been explored across various football codes, showing that there are a range of sensor derived measures available for quantifying peak match demands (Whitehead et al., 2018). Hughes and Franks (2008) suggested that some coaches reviewed statistics to reinforce their opinions about the events they remembered within the game. Nash and Collins (2006) determined that coaches adopted this approach because decision making was a process where coaches identified solutions based on a familiar situation with a known solution from their past. If there was no known solution to a problem then a coach could use a set of cognitive rules to problem solve, however, problem solving consumes cognitive resources. Therefore, to reduce the cognitive strain, finding a way of capturing, organising, and presenting statistical information to coaches in an ecologically valid way may help. If the continuum in Figure 1 is revisited, once information is organised, contextualised, and analysed, coaches can then combine this information with their personal 'experience, intuition and expertise' (Zins, p. 485, 2007) to create knowledge.

Coach gut feel (tacit knowledge) vs scientific findings

Tacit knowledge is crucial in time constrained situations and should not be excluded from decision making. Nash and Collins (2006) gave an example of a football coach who made decisions which won games, as evidence of tacit knowledge. They argued that coaching is not an art, rather a coach draws

on information from their sport science staff, players, own observations and combine with experiences from their past to make good decisions. Netball also has examples of coaches' who have made decisions that led to the success of their team. The New Zealand netball team coach Dame Noeline Taurua decided to play two long range shooting players at the 2019 Netball World Cup, a 'dual shooting threat,' when almost all other teams played one 'tall holding shooter' with a supporting goal attack. As Nash and Collins (2006) pondered within a football context, how did she know that this strategy would lead to success and be considered as a key factor in her team's success? Her vast experience in a long coaching career led to her gaining knowledge that was effective for the given situation. Tacit knowledge is often taken for granted and is often unarticulated and abstract (Sternberg & Horvath, 1999). Hence it is important that performance information should support tacit knowledge not exclude or contradict it. Choosing the elements of the player's performance to observe and review would normally be based on a coach's previous experience with these elements. Tacit knowledge (Nash & Collins, 2006) is instinctive or intuitive which is developed and accessed at appropriate times influencing decision making. This has historically been referred to as the art of coaching, however, it is argued that this process is more scientific by nature.

The potential gap between coaches and scientists

Drust (2010) explains that performance analysis is widely valued by coaches as an important part of the feedback process. Many coaches and sport scientists work together effectively to improve athlete performance; however, Williams and Kendall (2007) identified a potential lack of understanding between some sports scientists and elite coaches of the needs of each profession due to each having different foci. There is a risk that a sports scientist investigates problems in the context of their specific discipline e.g. biomechanics or physiology, as it is easier to control variables and improve construct or content validity. Mackenzie & Cushion (2013) explain that modern performance analysis research generally utilises a reductionism approach within a positivist paradigm, where individual parts are analysed to understand the system as a whole. This type of reductionism approach could contribute to a gap in current knowledge, as many elite coaches focus on solving problems, which tend to be holistic in nature, requiring an approach that incorporates a fusion of the knowledge within the areas of sports science.

Ecological validity

Ecological validity is described as how similar an evaluation of behaviour or intervention is, when compared to realistic contextual conditions (Baxter, et al., 2015). It's focus is on the findings being relevant and within the context of the sport. It is debated that researchers' sometimes

misunderstand ecological validity (Holleman et al., 2020; Davids., 1988) when claiming their findings are representative of the 'real world'. In sports performance analysis this is described as how well information is aligned with the situational and environmental needs of a coach or player during real performances i.e., is the information provided, by a performance analyst during a sports match, able to be acted upon? The objective of the outcomes of this PhD are to provide solutions to sport practitioners, particularly coaches and performance analysts, so strong ecological validity is a crucial consideration.

Ecological validity is related to observational vs. experimental research, where observations can infer correlational understanding of phenomena. In the case of conducting research that has strong ecological validity for the coaches' information needs, observation of their behaviours or requirements is known as the most appropriate approach. Throughout this PhD an attempt has been made to gain coach insights, feedback and observations of their behaviours, to ensure strong ecological validity was maintained.

Systems design approaches

Until recently (O'Donoghue, et al., 2018) there has been a paucity in the literature on systems design for performance analysis. The convention has been for performance indicators (Hughes & Bartlett, 2002), published in controlled research studies, to be tested by the performance analyst against their team's performance. The pre-eminent article on performance indicator design and their application, Hughes and Bartlett's (2002), has over 840 citations and is deemed to be the conventional approach for sport researchers and practitioners. The evolution of information technologies, innovative data generating devices and improvements in software development, has however led to an increase in what can be measured, therefore this conventional approach needs to be revisited (McLean et al., 2017; 2019). O'Donoghue et al. (2018) explained an alternative approach, which considered five additional stages in helping a performance analyst provide ecologically relevant information to a coach (Figure 2).

The first stage was 'research and planning'. This stipulated the importance of literature to inform the purpose or question that the coach needed to be answered, whilst also including informed discussions with the coaches. Based on the 'gap' between scientists and coaches (Williams & Kendall, 2007), this PhD believed coach involvement was essential at this stage of the design process as it helped establish ecological validity for the coach as the end user.

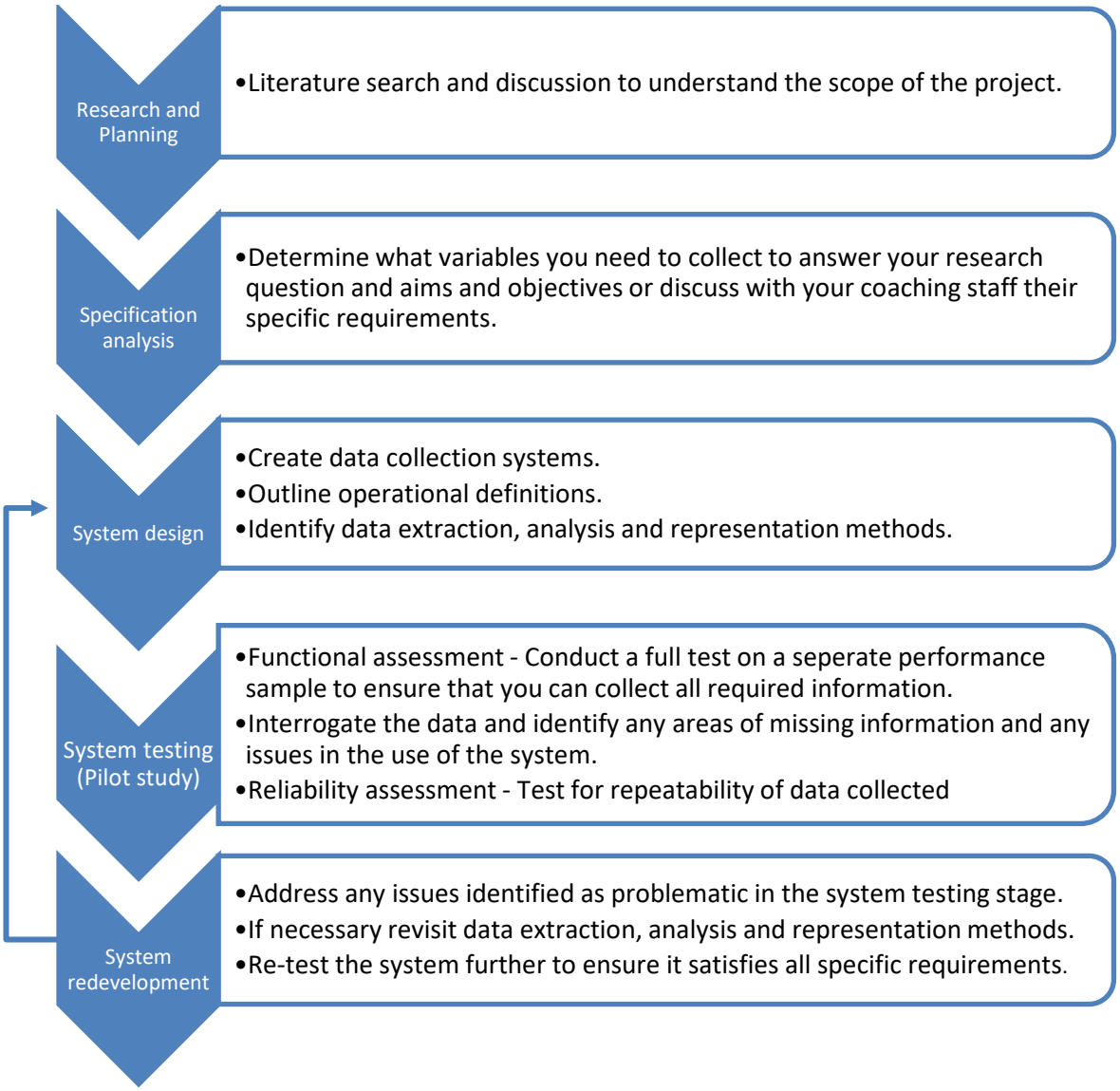
Stage two was 'specification analysis', where the variables that were needed to answer the coaches questions were defined through an investigation of literature. With the increasing sources of

data over the last 20 years (Mayer-Schönberger & Cukier, 2014) this search should not just be limited to literature but should also include information technologies available to practitioners.

Stage three was 'system design', where the collection of data, operational definitions, data extraction, analysis and representation were defined. This is a significant and complex process when working with large volumes of data, where automation of the system needs to be implemented, (O'Donoghue and Holmes, 2015).

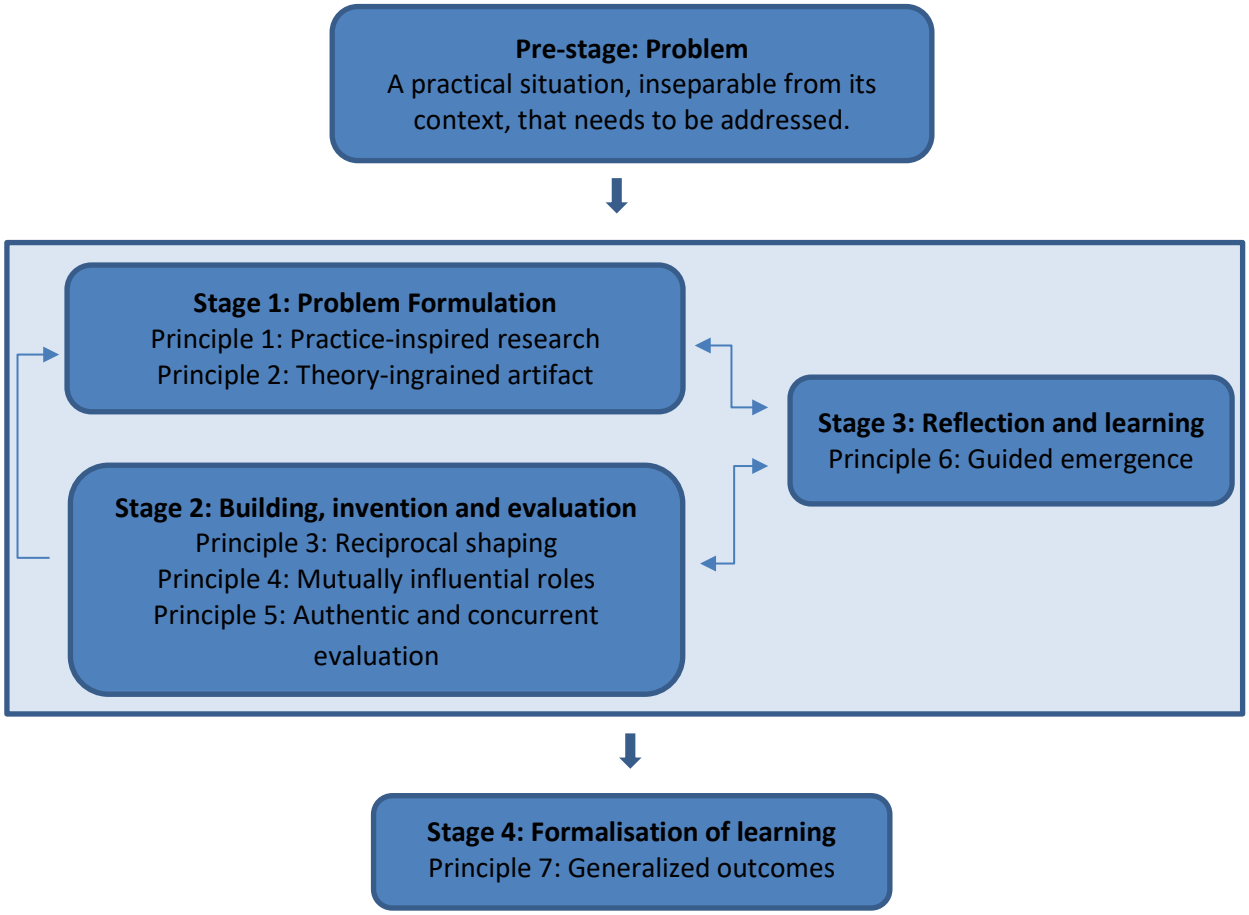
Stage four and five were 'system testing' and 'system redevelopment' where functional and reliability issues were assessed and remedied. This is a very important stage of system development, however, it was debated that this whole process should be cyclical and return to the first stage of coach feedback and interaction with the information provided. This cyclical process would enhance ecological validity for the coach as the end user.

Figure 2. *Stages of systems development, adapted from O'Donoghue et al., (2018)*



An Action Design Research (ADR) approach has been implemented in sports performance analysis research (Connell & Spencer, 2020) to develop a player effectiveness analysis system (PEAS). This framework was developed incorporating a multidisciplinary approach to performance analysis (Glazier, 2010) and demonstrated good ecological validity (Baxter et al., 2015). This methodology created an artifact and then used an iterative process to improve the design of the PEAS. Petersson and Lundberg (2016) described four stages of ADR (Figure 3) as: 1) problem formulation; 2) building, intervention, and evaluation; 3) reflection and learning; and, 4) formalisation of learning.

Figure 3. *The four stages and principles of Action Design Research, adapted from Sein, et al, (2011).*



This approach shares some similarities to the Systems Development approach described by O'Donoghue et al. (2018) where the process follows five stages (Figure 2). These are: 1) research and planning, 2) specification analysis, 3) systems design, 4) system testing, and 5) system redevelopment. There is also an iterative process where, upon completion of stage 5, stage 3 is revisited. The Systems Development approach was designed for both developing research studies or for a practitioner to build and analyse workflow and may include video and data. However, the Systems Development approach is not specifically designed for automating the management and utilisation of big data. Big or large data requires automation of integrating multidisciplinary datasets and hence needs a more specialised approach.

Based on the rationale above, including the need for greater ecological validity of information for coaches and the problems associated with big data, a tool needs to be developed for performance analysts that assists them in developing a tool for coaches to use to inform decision making pre-, during and post-match. This tool needs to attempt to negate coach bias and provide a holistic approach to information analysis while retaining high congruency with the coach and players

understanding of the game. The format of this PhD looked to build upon the system development process outlined by O'Donoghue et al. (2018) and used elements of the ADR approach.

1.2 Rationale and Significance of the Study

The current issues identified within the sport literature and practice are;

1. A lack of research to understand the best methods for presenting statistical information to coaches and or players.
2. A need for this research in the context of netball for performance analysts.
3. The issue of 'big data' and increased data access for coaches assisting with confirmation bias.
4. A lack of understanding between coaches and scientists' requirements necessitating greater ecological validity.
5. A need to allow for a coach's tacit knowledge to be complimented by research information.

These five issues suggest the need for a framework to be developed for performance analysts to assist in the design of their workflows and processes for presenting data and information to coaches. Within the research context O'Donoghue et al. (2018) have described a set of five stages for systems design research which could be followed in designing a framework.

1.3 Research Question

Performance analysts who work with coaches in netball, and other sports, in the 'era of big data', should attempt to address the five issues stated above. This therefore creates the need for a framework to be developed for performance analysts to use when designing their data and information statistics workflows. Previous frameworks in other fields, such as data mining (Ofoghi et al., 2013) or 'the coaching process' (Cushion, 2007) are well established, researched, and are applied within coaching. However, such a framework does not currently exist for the performance analyst when working in an applied setting, with large data sets and presenting the information in an ecologically valid format to coaches.

Therefore, the overarching research question of this PhD is; 'what process, or framework, will allow an analyst to present large data sets to coaches in a relevant, meaningful, and ecologically valid manner to assist coach decision making?'

1.4 Aim of this PhD

The aim of this PhD is to create a framework that performance analysts can use to guide coaches objectively through large data sets to answer questions related to strategic performance. This will be achieved in five stages guided by a systems design approach:

1. Determine what themes are important to performance for coaches in real-time.
2. Understand what data is available for the themes identified and make recommendations therewith.
3. Develop a way to capture and organize this data for easy recall and consumption.
4. Contextualize and normalize this data for answering coaches questions.
5. Understand coach's behaviours when using a data presenting tool.

It is expected the findings of each stage will benefit performance analysts, in designing information tools, not only working within netball but also other sports.

1.5 Overall Design of the Thesis

This PhD is presented in Format Two thesis (manuscript format) also known as 'by publication', which requires the student to submit for publication a series of papers that link together within a larger body of work. All papers/chapters within this thesis will be published, however, at the time of examination must be either 'in preparation for submission to a peer-reviewed journal or submitted to a peer-reviewed journal or published or in-press or accepted by a peer-reviewed journal' (Parthiban, p. 1, 2020) . The status of publication will be indicated throughout the document for each manuscript.

Due to this thesis being a manuscript format there was repetition throughout. This was necessary as the manuscripts needed to be published, in isolation from the rest of the PhD and still be understandable to the reader.

Through this process, this PhD will explore the question of whether the large data sets that are already available within sport can be captured, organised, and presented to coaches in an understandable and ecologically valid way. It will mostly employ a mixed methods approach. These will include observations of participants in their natural sporting environments, surveys, literature-based design and an overarching systems design methodology similar to that explained by O'Donoghue et al. (2018). A post-positivism paradigm will frame this research, as it acknowledges that knowledge is a combination of theory and practice, where different methods can be utilised (Henderson, 2011; Ryan, 2006). The aim of this qualitative descriptive (Sandelowski, 2000) approach is to develop an understanding of what is being investigated while maintaining strong ecological validity. This should ensure that the framework and tool developed will be beneficial to coaches and solve a problem that currently exists.

This post-positivism mixed-methods PhD consists of four progressive stages that develop a framework for analysts and coaches, assisting them to navigate complex 'big data' in a way that is ecologically valid. Initially a literature review manuscript outlines previous research in fields related to each of the

stages. This is followed by a series of research manuscripts (Figure 4) with prelude sections between each chapter to create coherency between manuscripts. Finally, a discussion, conclusion and future research directions section will bring together the findings from each study and provide recommendations for coaches, performance analysts and researchers.

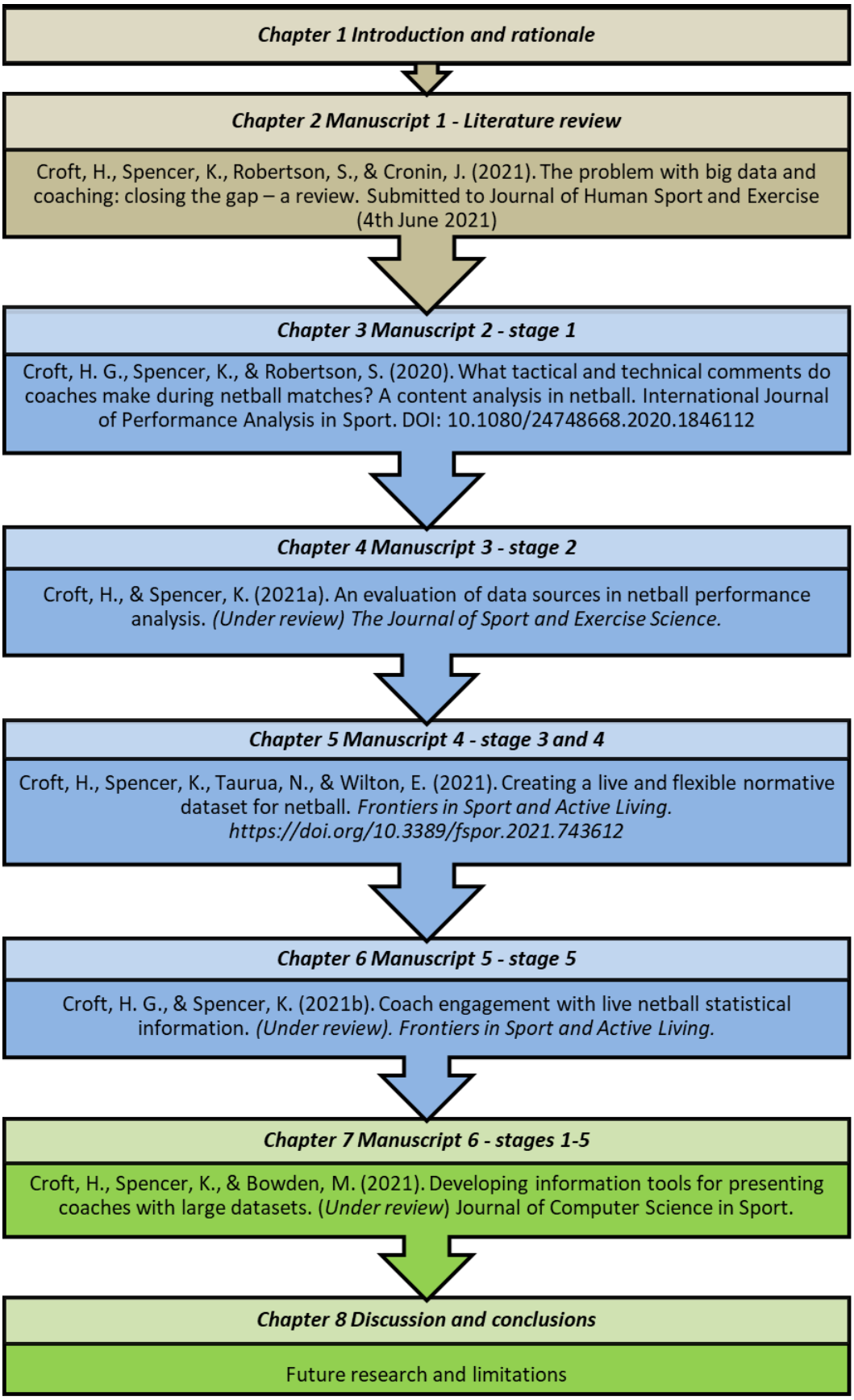
1.6 Objectives of this Research

The mixed methods approach used throughout this PhD allowed the objectives of each stage of the design to be achieved. These objectives were:

1. To understand the variables considered important by coaches and players in court and field sports as they analyse the tactical and technical aspects of a game and what elements are important to their decision making (chapter 2; manuscript 2 - stage 1).
2. To identify the current sources of data that exist within netball and determine which of these data sources would assist a coach or player when making tactical decisions (manuscript 3 - stage 2).
3. To develop a data-organisation technique that allows for data to be captured and organised so it can be presented to coaches for decision making (chapter 3; manuscript 4 - stage 3).
4. To normalise and contextualise this data to give it meaning and make it understandable to a coach (chapter 3; manuscript 4 - stage 4).
5. To understand coach's behaviours when using information presentation tools for decision making in matches (chapter 4; manuscript 5 - stage 5).
6. Finally, this research will develop and implement a pilot tool that demonstrates the key elements developed from stages 1-5. It will apply the principles of a framework that can be utilised by performance analysts when preparing information feedback interfaces for coaches and players that can be used in real-time (chapter 5; manuscript 6).

The structure of this PhD can be observed in Figure 4. The brown stage represents the rationale for a framework, the blue stage represents the systems design methodology, and the green demonstrates the resulting framework and learnings from this research.

Figure 4. *Overview of the thesis structure to determine a framework for guiding coaches through complex data in sport.*



Prelude Chapter 2.

Chapter 2: manuscript 1, 'The problem with big data and coaching: closing the gap – a review' was a review of literature from areas that related to the five stages discussed in Figure 2, informed by the systems design approach. This chapter looked to evaluate literature that was applied in nature, attempting to maintain ecological validity, and was a starting point for informing the design of each stage of the proposed framework. The ongoing review of literature, in each of the subsequent manuscripts (2-6), was an attempt to ensure there was solid theoretical underpinnings to all stages throughout this PhD.

The sections throughout manuscript 1 are presented in the same order as the stages presented in the proposed framework.

- Stage 1. What themes are important to performance for coaches in real-time?
  Chapter 2. What information is used by coaches when making decisions during a match?
- Stage 2. What data is available for the themes identified in stage 1?
  Chapter 3. Data sources in sport.
- Stage 3 and 4. Capture, organize, contextualize and normalize this data in a database.
  Chapter 4. Organising and analysing different type of sports data.
- Stage 5. Present visualisations to coaches and understand how they engage with it.

The purpose of chapter 2 was to identify relevant literature which informed the design of each stage of the framework proposed in this PhD. Literature was reviewed appropriate to each of the stages identified from systems design methodology, and was summarised with recommendations for each of the chapters .

Chapter 2.0 (manuscript 1): The problem with big data and coaching: Closing the gap – a review.

This chapter comprises the following manuscript (1) submitted to Journal of Human Sport and Exercise (4th June 2021)

Croft, H., Spencer, K., Robertson, S., & Cronin, J. (2021). The problem with big data and coaching: closing the gap – a review. *Submitted to Journal of Human Sport and Exercise (4th June 2021)*

Author contributions: Croft, H (85%), Spencer, K (5%), Robertson, S (5%), & Cronin, J. (5%).

2.1: Introduction

According to Mayer-Schönberger and Cukier (2014) there has been an "explosion" of data since the 2000's due to new technologies that enable data to be generated about almost anything. In sport this saturation of data is so great that coaches do not have the time or cognitive resources to integrate it effectively into their decision making, rather like all humans they have a propensity to use it to support their beliefs (Mercier & Sperber, 2011). This phenomenon is known as conformational bias and leads to decision making based on samples of the evidence, potentially taken out of context, which unfairly represent the situation. Conformational bias has been explained as "creationists see fossils as evidence of God, evolutionary biologists see fossils as evidence of evolution" (McNerney, 2011). This means that the recent growth in data availability may be to the detriment of coach decision making as easier access to information will allow greater conformational bias.

This is further enabled by products like Sportscode™ OPTA Sports Data™ and Prozone™, which provide coaches with access to large sets of descriptive statistics and live video replays during matches. Hughes and Bartlett (2008) explained that most coaches reviewed statistics to reinforce their opinions about the event they remember within a game. Nash and Collins (2006) explain that this is due to problem solving being avoided as it requires large amounts of cognitive resources. This means that basing a decision on previous experiences is preferable, however may not be the best decision.

With the increased availability of data and coaches' propensity to utilize information poorly, finding a way of capturing, organising and presenting this information back to coaches in an ecologically valid way needs to be identified. The aim of this paper is to explore a process that will guide coaches objectively, limiting bias's, through large data sets to help them answer questions related to strategic performance.

2.2: Methodological approach to sourcing literature

Paré and Kitsiou (2017) explain that there are nine approaches to conducting literature reviews, with scoping reviews providing effective for emergent topics (Daudt, van Mossel & Scott. 2013). This paper will utilise a scoping review approach as it will enable and understanding of the extent, range and nature of research activities in fields related to data analytics in sport. With scoping reviews inclusion and exclusion criteria need to be established to help identify research that aligns with the question. This was achieved in this manuscript by reviewing literature under the following topics which align with the stages of O'Donoghue et al's (2018) systems design framework: 1) What information is used by coaches when making decisions during a match? 2) Data sources and collection in sport, 3) Organising and analysing different type of sports data, 4) Visualizing and feeding back information to coaches. The final part of this paper will identify the gaps in the literature and then make recommendations on what future research is needed.

Four online sources were utilised for this study (Sport Discus, Google Scholar, Taylor and Francis and ProQuest) and the search criteria used for inclusion in this review were; coach discussions, sport information technology, normative data sport, visualising sport data. Abstracts were reviewed for relevance to each of the four review areas and those deemed appropriate where reviewed in full text.

2.3: What information is used by coaches when making decisions during a match?

To ensure data is ecologically valid, and therefore relevant to a coach, it must be determined what coaches analyse and communicate to players during a game. This type of systematic observation has been utilised by researchers in volleyball (Zetou et al., 2011) for the behaviours of coaches. Findings showed that tactical instruction (17.4%) was the most common, followed by general instructions (15.9%) and technical instruction (12.4%), with encouragement, motivation, other comments and demonstration making up the rest of their behaviours. Bloom et al., (1999) and Horton, Baker and Deakin (2005) highlighted the importance of tactical instruction also. These results are hard to apply as they focused on broader themes and did not disclose the specific details about the topics within each of these themes.

This leaves a gap in the literature around the specific themes and topics communicated by coaches within the tactical element of team sport. Until this research is conducted alternative methods, explained below, would be to allow coaches to select topics they are most interested in

knowing about. This could introduce an element of conformational bias, however if the topics are broad in their focus this could somewhat mitigate the bias.

2.4: Data sources in sport

There is an opportunity to expand the incorporation of data into coach and player decision-making however there are various challenges that exist in doing this. Liebermann et al. (2002) conducted a review of similar scope to this paper, however their focus was directed towards information technology that provides feedback about skilled or technical, not tactical performance. While it provided important information, in an applied manner, for coaches and practitioners technology has evolved since 2002. Additionally, Liebermann et al. (2002) did not consider a significant attribute of coaching during a game (Zetou, et al. 2011) tactical information. With the rapid advance of information technology an update is required.

Information is prevalent in sport, from sensory devices such as GPS, accelerometers and video cameras to observations by the coach, analysts, players or fans. These data generators provide direct streams of data to a computer or person and measure many aspects of performance depending on their application. Some devices require humans to interpret the data, like notational analysis software, while others automatically collect data and calculate information based on criterion, like heart rate monitors. Regardless of the device, the data produced comes in a variety of formats and scales requiring normalisation before it can be compared. Below are examples of different categories of data generators that commonly exist in team sports.

2.4.1 Ubiquitous Computing Systems (UCS)

Ubiquitous computing is potentially one of the biggest sources of information within a sporting environment and could be one of the most important for this research. The term ubiquitous computing systems refers to the ever-growing process of embedding computers, usually in the form of micro-processors, in everyday objects to measure and communicate information about it. As described by Chi (2008) accelerometers, gyroscopes, microphones and cameras are becoming commonplace in sports research and lend themselves as useful for measuring real time training and competition performance. What is not acknowledged by Chi is that the information generated by these devices is rarely combined in a holistic way to provide greater insights.

2.4.2 Notational Analysis Systems

Notational analysis systems do allow for a broader collection of information, based on observed behaviours or events, and integrate well with video. They can be custom built, with relative ease, to

measure anything that can be observed during a match. O'Donoghue and Holmes (2015) outlined a range of commercially available systems that allowed observed information to be digitised and organised for quantitative analysis. These fit into two categories, those which are used for assessing the technique of short-term events and those for broader whole or part game events. For tracking events and analysing performance live, during a team sport match, notational analysis programs like Sportscode™, Dartfish™, Focus™ and Nacsport™ are used to cross-tabulate variables with the aim of understanding performance in a reductive manner. The problem with these systems, not considered by O'Donoghue and Holmes (2015), is that unlike ubiquitous computing systems they cannot be automated and require a human operator.

2.4.3 Publicly Available Information

There are many other sources of information available for analysts and coaches that could be incorporated into their analysis of performance. In the modern era of technology advancements mobile devices allow sports fans to express their opinions via social media or via direct communication with friends. Although there is no way to assess the accuracy or validity of these opinions, they do carry levels of information that could be useful if collected from vast quantities. Currently technologies like Topsy™ allow anyone to track the frequency that keywords appear in conversations on Twitter. Topsy has been used in research focused on understanding medical networks and how these utilise micro-blogging sites for engagement with the wider community (Mishori et al., 2014). Technologies like Topsy™ use algorithms, described as data aggregators, which search all Twitter™ posts much like a Google™ search using text recognition techniques. Although somewhat simplistic it demonstrates the potential for extracting information from the public domain.

2.5: Organising and analysing different type of sports data

2.5.1: Databases in Sport

Conventional methods for managing data generally include the use of a database. This combined with other data processing tools appears to be the most obvious solution with some literature having focused on their design specifically for sport science data. Vincent et al. (2009) overviewed the design and implementation of databases in sport with specific focus given to their use as a performance analysis tool. They summarized the most effective uses of databases, data management and content management systems as a solution for removing coaches' reliance on "gut feel" or intuitive thinking about performance. The authors claimed that databases could be more than a number storing system, and focused on the use of databases for managing and organizing video both locally and online. Once data is collected and organised in a database the next step is to identify how to analyse it and produce

information for the coach. With very large datasets the challenge is to know what information to select.

2.5.2: Reductive vs data-mining approaches

O'Donoghue (2015) explained the reductive approach as being an effective process for making vast quantities of sport performance information manageable and useful to coaches and players. He does not however, acknowledge that important information could be lost during this process, nor that important performance patterns are complex by their multi-disciplinary nature. This reductive approach is essentially the practice of selecting information that is perceived as important whilst removing noise, or information that has no direct relationship with performance. The development of performance indicators (PI), variables that represents important aspects of performance, assists in this reductive approach. Much of the sports performance research published in recent years has focused on the development and application of PI's (Kraak et al., 2016; Hughes & Bartlett., 2002) for individual sports, however this reductive approach could lack sensitivity and ecological validity as it does not take into account the complexity that exists around individual players, coaches and situational circumstances. This is a challenge that faces all reductive approaches.

Middlemas, S. (2014) discussed an alternative, non-statistical, reductive approach based on coach and player requirements. This required reliance on tacit knowledge gained by coaches and players from their experiences and is heavily biased towards their past as previously explained. It does however create greater buy in and is more closely aligned to what the coaches and players are trying to achieve. The integration of this approach to help contextualize the data being processed might somehow improve coach and player understanding, however this requires further investigation.

There are several techniques that can be implemented for dimension reduction including, principal component analysis, linear discriminant analysis and canonical correlation analysis (Shaw & Jebara, 2009). Further investigation is required into these selection techniques, as they possess strengths and weaknesses depending on the data sets and applications. For example, O'Donoghue and Holmes (2014) described principal component analysis as ideal for identifying critical subsets to analyse in a live setting where not a lot of information is available, yet this can be troublesome in that it produces a non-unit and non-norm scale for interpreting scores.

2.5.3: Artificial neural networks

A promising solution, which could be applied to this study is offered by Bartlett (2006) who summarised the potential for artificial intelligence (AI) to advance sports biomechanics with its handling of large data sets. These AI technologies could provide similar benefits to the field of sports

performance analysis, especially if the emergence of new sources of information can be captured and organised. In support of this is Silva et al. (2007) who applied artificial neural networks (ANN) to information extracted about swimmers anthropometry, conditioning status, technique efficiency and other functional evaluations e.g. hydrodynamics. This showed the strength of ANN's in dealing with multi-disciplinary data. They found that the ANN was able to predict performance to within 0.8% of real swim performance times. Obviously, the information that the ANNs are provided is crucial to their accuracy. These AI techniques show great potential for not only sports biomechanics but also sport performance analysis (Croft, Lamb, & Middlemas, 2015; Lamb & Croft, 2016), however the identification of the correct input information, or variables, requires a structured process.

2.6 Visualizing and feeding back information to coaches.

Once a tool is developed that converts data into meaningful information, the way this information is presented back to coaches and players is crucial. It must be easily understandable, relate to their objectives and questions as well as being ecologically valid. As sports people rely heavily on visual information, especially in team sport, it seems appropriate that investigation into visual representations of data and information be conducted. Friendly and Denis (2001) described data visualization as a form of descriptive statistics. This classification is somewhat inaccurate as data organisation and visualisation techniques like self-organising maps (SOMs), a form of ANN's, provide an alternative to statistics and provide visual representations of the data in a low dimensional way, avoiding the need for a reductive statistical approach. Friendly and Denis (2001) did however, correctly describe data visualisations as the visual representation of abstract data in order to reinforce human cognition. The types, modes and timing of visual representations are all factors that need to be explored.

2.6.1: Visual analytics

Thomas and Cook (2005) explained the field of visual analytics as "the science of analytical reasoning facilitated by interactive visual interfaces (p.4)" with these interfaces enabling the user to understand the nature of the data sets. These data sets can be very large, diverse and incomplete yet with appropriate visualisation methods humans can understand and make judgements about them. Thomas and Cook's (2005) statement supports the concept that data visualisation is not a form of statistics. The benefits also relate well to sport, with player's, analysts and coaches processing large amounts of information during a game, in an attempt to make accurate decisions and understand what is happening in a given situation. Given this reliance on the visual system in sports it would make sense that analysis of data, by coaches and players, collected from the environment should also be

visual. Other authors support the use of visual analytics; Keim et al. (2006) stated that the strength of visual analytics is that it "combines the advantages of machines with strengths of humans" (p.10) implying that this could be an idea tool for coaches and players. However, for ecological validity to exist visualisations would need high congruency to what coaches and players see during a game and according to Thomas and Cook (2005) visual analytics must allow hypothesis and scenario-based reasoning to occur. This insight is useful when developing the coach interface as hypothesis and scenario-based reasoning would work in an interface that has an "expert systems" design as outlined by Lapham and Bartlett (1995). For feeding back information to the coaches there have been attempts to produce visual analytics which are coach friendly with glyph-based visualisations being one of these. Legg et al. (2012) designed a glyph-based visualisation technique for live coach feedback during rugby matches. This technique utilised an arrangement of icons to create an intuitive representation that is claimed to be easily understood by coaches. Direct endorsement by an international rugby team's analyst was provided as evidence of its effectiveness as well as a student focus group supporting its ease of use. Greater and more thorough evaluation of this technique is required as it provides an interesting approach to visualising information.

2.7 Future research

2.7.1 Coach focused variable selection.

Direct observation has been recognized as a valid research methodology (Erickson, 2009). Conducting an analysis of the coaches verbal communication (theme and topics) during live sports matches would provide a robust and valid source of the information about what coaches communicate to the field or court during a game. An inductive content analysis process (Neuendorf, 2002) is a method that can extract themes and topics from audio recordings made of the coaches' conversations during matches.

2.7.2 Identifying valid data sources in sport.

A review of existing technologies that can provide information about the strategic and tactical elements of a court and or field sports matches needs to be conducted. To do this Palmius's (2007) recommendations of comparing information systems with a set of important operational and measurable criteria needs to be implemented. A checklist needs to formulated and include important factors, such as: relevance of the data to the topics coaches discuss, suitability to a real-time setting and accessibility.

2.7.3 Creating a data organisation system.

Before data sources can be used they need to be organised to allow fast access to information. Vincent et al. (2009) have reviewed the design and implementation of databases in sport and discussed the benefits of automated vs. manual data entry. Advanced methods for storing data have been explained by Cheolkon and Joongkyu (2008), while Lapham and Bartlett (1995) described expert systems as a method for organising data which "simulates the actions of an expert". Bartlett (2006) summarised the potential for artificial intelligence (AI) for analysing large data sets. These various findings need to be tested for their ability to find relationships and group data on the data sources relevant to the coaches' perceptions.

2.7.4 Visualisation techniques which allow understanding.

Thomas and Cook (2005) explained the field of visual analytics as "the science of analytical reasoning facilitated by interactive visual interfaces (p.4)" while Keim et al. (2006) stated that the strength of visual analytics is that it "combines the advantages of machines with strengths of humans". Visual analytics could provide a simplistic way for coaches to understand data which represents the answers to their questions. With sample data sets, further work is needed to assess which techniques are most relevant and usable in a real-time setting.

2.8 Conclusions

This review identifies gaps in the literature which require further investigation and a guide to how a practitioner might select, organise and visualise large multivariate data sets. Future research needs to look to implement some of the following steps: a) coach tactical behaviours during matches, b) identify valid data sources in sport, c) create a data organisation system, and d) Implement visualisation techniques which allow understanding of complex relationships by coaches during a game. Future research is described below which may help solve these problems and bridge the gap between coaches and scientists.

Prelude Chapter 3.

The previous chapter, 'The problem with big data and coaching: closing the gap - a review' reviewed applied literature that was relevant to the five stages of a systems design approach, and the development of a framework for performance analysts. Recommendations for future research proposed the use of inductive content analysis to identify coaches behaviours and needs during netball matches. It also identified categories of information technologies that could be considered when producing information about performance. An overview of data organisation and contextualisation was given, followed by data visualisation techniques for managing 'large' or 'big data'. When applied to the development of a framework, investigation into the content of coach communications was the Stage One, as seen in Figure 5.

Figure 5.

Stage One of the proposed framework; "evaluate the coaches behaviours and needs during a live netball match".

To develop a deeper understanding of Stage One (Figure 5), of the proposed framework, the "research and planning" phase of a systems design approach, included the use of literature. An observation of coaches behaviours also needed to be conducted. Manuscript 2 utilised inductive content analysis of coaches spoken communications during live netball matches, as a form of direct observation.

The purpose therefore of chapter 3 was to create categories of spoken themes, that could be used by performance analyst to identify information systems with ecological validity. These spoken themes gave an indication about the types of information that coaches might find useful during a game. Using this approach was hoped to better align data with what coaches are interested in when making decisions.

Chapter 3.0 What tactical and technical comments do coaches make during sports matches? A content analysis in netball.

This chapter comprises the following manuscript (2) published in The International Journal of Performance Analysis in Sport, 2020.

Croft, H. G., Spencer, K., & Robertson, S. (2020). What tactical and technical comments do coaches make during netball matches? A content analysis in netball. International Journal of Performance Analysis in Sport. DOI: 10.1080/24748668.2020.1846112

Author contributions: Croft, H (85%), Spencer, K (10%), Robertson, S (5%)

3.1 Introduction

3.1.1 Coach bias

Applied performance analysis most commonly appears in the form of collecting match footage and statistics. Researchers have suggested that coaches tend to track and review statistics that reinforce their opinions about the game, or memorable events within it, which in turn leads to highly subjective decision making (Hughes & Franks, 2008). The decision-making process adopted by expert coaches has been summarised by Nash and Collins (2006) as a process where individuals will identify a solution based on a familiar situation with a known solution from their past. If there is no known solution then the individual will use a set of cognitive rules to problem solve, however problem solving is generally avoided as it ties up cognitive resources. These familiar situations and solutions construct a coach's tacit knowledge and can be viewed as being intuitive due to their subconscious nature. Franks and Miller (1991) studied football coaches' observational skill and utilised a video training technique to improve their recall of events in a match. One of their findings was that coaches were incapable of recalling more than 40% of the pertinent events in a match. This finding, coupled with Hughes and Franks (2008) description of biased decision-making, reinforces the need for structured collection of information about performance in sport. If information is to be collected and presented to coaches during a sports match, then it must relate to the way they interpret the performance but also be robust and objective.

3.1.2 The growth of data and information

In sport the recent 'explosion' of information (Mayer-Schönberger and Cukier, 2013) has meant that coaches may not have the time or cognitive resources to integrate this information effectively into their decision-making. However, due to the existence of this large volume of information coaches are

easily able to selectively reinforce their previous opinions, meaning that this data may in some cases be to the detriment of developing coach decision making. Therefore, finding a way of capturing, organising and presenting the large volume of information available back to coaches in an ecologically valid way needs to be identified so it can be used by the coaches to enhance their decisions.

3.1.3 The gap between science and coaching

It has been found that coaches believe that sports science research does not currently meet their needs and that there is a need for better dissemination of sport science findings (Williams & Kendall, 2007). This highlights the requirement for future research into performance, including that which organises and evaluates large sources of information, to have a problem-solving focus and an applicability to the coach. Specifically, visualisations, data summarisations or statistics need to have a high congruency with coaching feedback. This may be somewhat difficult for many researchers as it requires an in-depth understanding of the context (constraints, experience, and competition) and environment in which the coach and players are working. Prior to providing performance information to a coach, a researcher needs to develop a process which incorporates an understanding of the coach context within the evaluation. This would require the researcher to observe the environment and the coach's actions within it.

William and Kendall (2007) identified that the "gap" or lack of understanding between sports researchers and elite coaches may be due to the two professions having different foci. Sports researchers tend to investigate problems in the context of their specific discipline e.g. biomechanics. Coaches however, focus on solving problems for their players, which can be holistic in nature, requiring an approach that incorporates many different areas of sports science research. For example, if a player is having difficulty shooting a netball goal, a coach will work through the technical, physiological and psychological elements of this skill to determine where improvement may be made. Choosing the elements of the netball player's performance to observe and review would normally be based on a coach's previous experience with these elements, also known as tacit knowledge. Tacit knowledge has been explained (Nash and Collins, 2006) as instinctive or intuitive knowledge which is developed and accessed at appropriate times influencing decision making. This has historically been referred to as the art of coaching, however it is argued that this process has underlying scientific constructs.

3.1.4 Coach behaviours during a match

To address this disconnect between coaches and researchers the first step is therefore to determine what coaches analyse and communicate verbally to players during a game. Once this information has

been captured, relevant information can be provided to enhance the effectiveness of the coach decision-making process. Zetou et al. (2011) successfully captured and categorised the behaviours of volleyball coaches through systematic observation. Their finding showed that tactical instruction (17.4%) was the most common behaviour followed by general instructions (15.9%) and technical instruction (12.4%), with encouragement, motivation, other comments and demonstration making up the rest of their behaviours. The importance of tactical instruction was also highlighted by Bloom et al., (1999) and Horton, Baker and Deakin (2005). While these studies are insightful, they do not disclose the specific details about the topics within each of these themes, therefore making the results hard to apply generically. Research is required to investigate the specific themes and topics communicated by coaches within the technical and tactical elements of evasive team sport.

3.1.5 Grounded theory and content analysis

Grounded theory has been described (Weed, 2009) as a "total methodology" for investigating the nature of the social world. Understanding the nature of coaches, communications during a match, may provide insights into the types of information that would assist their decision-making. Content Analysis is a methodology that sits within grounded theory and allows the researcher to study and understand the nature of spoken language, written text, interviews, pictures, and other communication media (Krippendorff, 2004; Neuendorf, 2002), within either quantitative or qualitative frameworks. Qualitative research has many strengths that make it more suitable for describing coach behaviour and developing theories which quantitative research does not due to its focus on testing hypotheses. Strauss and Corbin (1998: p. 137) suggest that "an interplay between induction and deduction" is a requirement of effective grounded theory research. This means that a model is developed from the communications data (deductive) and then applied to the rest of the data (inductive) and then the model is continually revised and evolved.

3.2 Methods

Direct observation is recognized as a valid research methodology for understanding the nature of human behaviour (Erickson, 2009). This study conducted an analysis of coach's verbal communication during live sports matches. An inductive content analysis process (Neuendorf, 2002) with an independent researcher, acting as a 'critical friend', extracted themes and topics from audio recordings made of the coaches' conversations during matches. These themes were analysed and the most common identified. The aim of this study is to determine the specific themes of tactical and technical communication, thus behaviour analysis tools such as the Coaching Behaviour Recording Form (CBRF; Tharp & Gallimore, 1976) or the Coaching Analysis Instrument (CAI; More & Franks, 1996),

were not used as these focus on more generalized behaviours.

The Auckland University of Technology Ethics Committee (AUTEC) granted ethical approval for this study under application number 16/267.

3.2.1 Participants

Six female participants were recruited for this study. They were either head or assistant coaches from New Zealand, ANZ Premiership™ netball teams. The coach's conversations were recorded live during competition games within the competitive season. Recordings began when the games started i.e. on the umpire's whistle and will end at the final whistle. All audio recordings were processed and anonymised by a 3rd party, prior to coding, to protect the identity of the participants. This study received ethical approval by the AUT University Ethics Committee and participants' informed consent was obtained.

3.2.2 Equipment

The participants placed a Sony™ (Tokyo, Japan) Stereo Digital Voice Recorder in their pocket which was connected to a Polsen™ (New York, USA) Omnidirectional Lavalier Microphone, attached to a convenient location on the collar of their clothing. This captured audio was transcribed and the transcriptions were analysed with MAXQDA™ Qualitative Analysis software. The thematic results were tabled and qualitatively analysed with Microsoft Excel™.

3.2.3 Procedures

Each participant wore the audio recording device during one ANZ Premiership netball game and all spoken dialogue between themselves and members of the management team as well as players was be recorded. At the end of the game, the audio recordings were sent directly to an independent transcription service, allowing the participants to remain anonymous from the primary researcher. The transcribers only included the participant's comments in the transcriptions and did not include comments of the people they were speaking with.

3.2.4 Data Analysis

The transcribed, anonymised audio recordings were processed with MAXQDA™ Qualitative Analysis software. An inductive content analysis approach was undertaken by the lead author, and themes were organised into hierarchical model, containing raw data themes, lower order themes, higher order themes and general dimensions with frequencies. A progressive coding technique was implemented with description and analysis of meaning conducted (Braun. & Clarke., 2006). The lead authors experience, of working in elite netball for 7 years, was seen as a strength of this study as

interpretation and analysis of the comments could be a limiting factor. Self-correcting verification strategies were used to improve 'trustworthiness' (Holt, & Sparkes., 2001) and second author acted also as a 'critical friend' on two occasions to also optimize 'trustworthiness' and minimize bias (Smith, & Sparkes, 2006). During this process, it was identified that the technical and tactical comments were intertwined and that these categories were best combined into one model.

3.3 Results

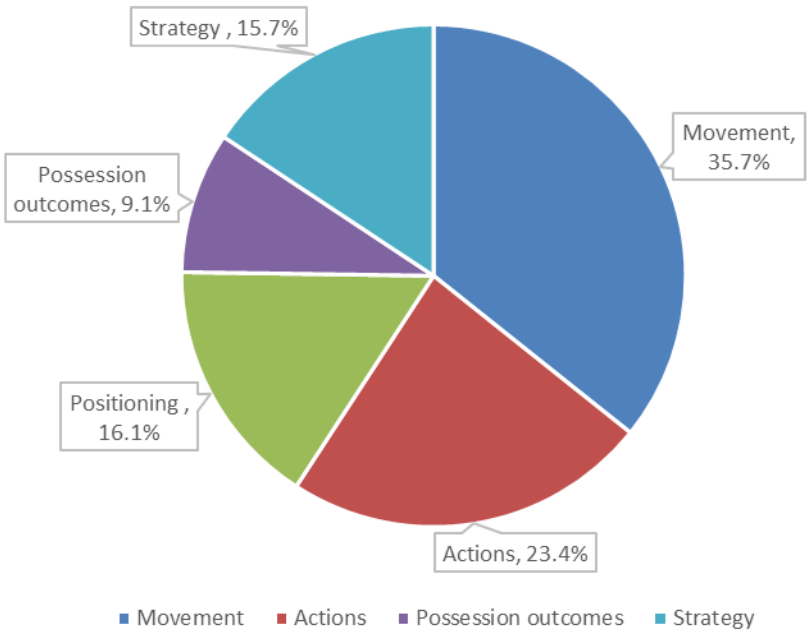
3.3.1 General Dimensions

All results presented here are representative of all six participants and were organised using the process described above, from the raw data themes.

As seen in figure 6, five general dimensions were discovered from the raw data, with 'movement' (38.4%) the most common, followed by 'actions' (23.9%), 'positioning' (17.3%), 'possession outcomes' (10.6%) and finally 'strategy' (9.8%). A total of 747 raw data themes, or comments, were coded from the transcriptions to identify these dimensions with 267 of these within the 'movement' dimension, 166 within the 'actions', 120 in 'positioning', 74 in 'possession outcomes' and 68 in the 'strategy' general dimension.

Figure 6.

The percentage of comments made by netball coaches, during matches, within the five general dimensions.



There were some interesting differences within the higher (2nd) order themes for the coach's comments. 'Movement spatial' was the most frequent category of comments with 151 raw data themes (20.2%) being identified. This was closely followed by 'attacking actions' with 132 comments (17.7%). 'Player positioning' - 83 comments (11.1%), 'movement intensity' - 69 comments (9.2%) and 'defensive strategy' - 57 comments (7.6%) were all moderately observed. 'Team positioning' (5.0%) and 'gain/loss' (4.1%) where also present to a lesser extent.

Table 1.

Five technical/tactical general dimensions and themes for each identified from the content analysis of six netball coaches comments.

General dimensions	Frequency	2nd order themes (higher)	Frequency	%
Movement	267	Movement timing	20	2.7
		Movement intensity	69	9.2
		Movement spatial	151	20.2
Positioning	120	Opposition movements	27	3.6
		Team positioning	37	5.0
		Player positioning	83	11.1
Actions	175	Opposition actions	17	2.3
		Defending actions	26	3.5
		Attacking actions	132	17.7
Possession outcomes	68	Centre pass	10	1.3
		Gains/Losses	31	4.1
		Scoring	27	3.6
Strategy	117		6	0.8
		Attack strategy	16	2.1
		Defensive strategy	57	7.6
		Errors	5	0.7
		Vision	16	2.1
		Niggle	17	2.3

3.3.2 Movement dimension

The largest dimension, 'movement', contained four higher order themes, as seen in table 1, which included 'movement timing', 'movement intensity', 'movement spatial' and 'opposition movements'.

Table 2.

Model of the general dimension "movement" linked to its higher, lower and raw data themes.

General dimension	2nd order themes (higher)	1st order themes (lower)	#	Raw data themes (examples)
Movement (267)	Movement timing	Timing of movement	5 1 5	*Good girl [Player 5]. That was good movement.* *Um first phase, you're timing that really well over the line.*
		Driving (run)	2 4	*Drive into the ball, meet it.*
	Movement intensity	Explosive movement	3 2	*And if you're going to start that rush, make sure you rush really hard and rush them again.*
		Footwork	1 3	*Feet, feet, feet, feet.*
	Movement spatial	Create space	1 8	*get deeper on them so those, those real up to go back so we can get some space.*
		Limited space	1	*we'll either haven't got enough room to do anything or we just going back and using the space. We shouldn't need to.*
		Connection with team	9	*Get more connection between you two and just, let's go*
		Movement in circle	3 1	*roll off and come back to the post to go again, alright?, and that will give*
		Move with opposition	2 9	*Stay on her [Player 7].*
		Where to move	4 4	*Come to the top.*
		Changing direction	1 1	*That's the girl, big change of direction.*
		Angles	8	*Oh that angle's a bit narrow.*
	Opposition movements		2 7	*Yeah but you've gotta look at [name 12] to see if she's stuck on that drive or if she's creating space 'cause sometimes when she goes, [name 14]'s holding onto her here and she can't get past and then you come down there and then [name 9] and this person,*

3.3.2.1 Movement spatial

'Movement spatial' was the most common high order theme being 20.8% of all comments. 'Where to move' was the most common (n=44) theme within this. This was generally in the form of instruction

from the coach who was on the bench, courtside communicating directly to the player on court; “Get in the middle” (Coach 2), however there were a lot of longer comments during breaks in the game as illustrated here:

> Yeah so you either need to come, it depends what the centre’s doing, but you need to come to a different space and then work off it for a square or you need to drive and come back and let the ball go to the square, so then it can go to you on the other side but you need to drive deep to come back because this one’s gonna be hunting for it, alright? (Coach 4)

These comments demonstrate both simple on-court instruction (Coach 2) and complex long discussion that occur during a break in the match (Coach 4). These comments were very explicit and tended not to question the players about where they thought they should move. At times, as shown in the comment by Coach 4 above, there was a lot of information for the players to absorb.

‘Movement in circle’ and ‘move with the opposition’ were also very prominent in the coach’s comments and at times, these could have been categorised with the ‘where to move’ themes. The coaches tended to comment a lot about the multiple actions within the circle and highlighted that movement in the circle involved multiple players who needed to be considered. This is seen in this comment: “It’s a great high take. When you’re high in the circle, look for [Player 19] coming in and then running that baseline.” (Coach 2)

3.3.2.2 Movement intensity

There were several comments on ‘movement intensity’ (9.2%) under the themes of ‘driving (run)’ (n=24) and ‘explosive movement’ (n=32) which showed the coaches did discuss the more physical aspects of the game. Most of these comments were short and from courtside in the form of direct instruction as seen here, “Good work. Nice driving, keep strong” (Coach 6), “Drive her hard and then pop back in” (Coach 1) and “That’s it [Player 4], turn quickly” (Coach 2). Although movement intensity was not as frequent as movement spatial it still was common across most coaches observed.

There were times when comments about ‘explosive movement’ were mixed with other technical comments, which made the categories again somewhat mixed as seen here “Feet, keep moving, move, move, move again. Keep the movement, angle in, angle feet, angle in, sight your base.” (Coach 6).

3.3.3 Actions dimension

The second largest dimension ‘actions’ only had three higher order themes; ‘opposition actions’, ‘defending actions’ and ‘attacking actions’. ‘Attacking actions’ (n=132) was the largest higher order theme identified in this study, as seen in table 3.

Table 3.

Model of the general dimension 'actions' linked to its higher, lower and raw data themes.

General dimension	2nd order themes (higher)	1st order themes (lower)	#	Raw data themes (examples)
Actions (166)	Opposition actions	Traits	9	*what she's, what does is she floats in and jumps in at the last minute with her bum, like she jumps in, she takes quite a bit of space.*
		Slow opposition pass	1	*stop them from keeping the ball moving, the minute they stagnate up and hold the ball for two seconds, they don't know what to do with it, alright?*
		Predicting opposition actions	7	*because otherwise [player 9] will come.*
	Defending actions	Defending a shot	7	*your jump is effective, it's intimidating, yeh, keep aggressive on the ball*
		Defending	15	*Yep. It is, it is. So come in, play that front strip game strong, need lots of work from you on D and we actually have to go for those intercepts coming through.*
		Jumping	4	*Yes, great jump [player 21].*
	Attacking actions	Shooting	36	*It's just got a back spin on it at the last minute.*
		Passing	46	*That was a little just soft pass, I wasn't sure it was gonna make it.*
		Feeding (pass into circle)	25	*Short, short, good girl.*
		Receiving a pass	25	*Drive into the ball, meet it.*

3.3.3.1 Attacking actions

The second most common higher order theme was 'attacking actions' (17.7%) with all 4 lower order themes within it well represented; 'shooting' (n=36), 'passing'(n=46), 'feeding'(n=25) and 'receiving a pass' (n=25). The other lower order themes in this dimension where diverse however not frequent. Most comments in this theme where short, explicit and from court-side during the game.

'Passing' was a very frequent, short and explicit theme with "We need to make sure we keep the ball moving, just keep the ball moving, as you know it's a shitty short pass, if it's a flat ball in front, just keep the ball moving, alright?" (Coach 4) the longest of these, during a break in the game. Most passing comments followed the characteristics of these quotes: "Pass" (Coach 3), "Give it to her." (Coach 2) and "Come on, give it, give it, give it, give it." (Coach 6). 'Receiving a pass' was also common with a

few longer comments in breaks of the game instructing players to technical improvements i.e. “So we just need that, we need actively catching a ball ‘cause they’re gonna chop the ball out of your hands alright so make sure you actively catch it.” (Coach 4). Many shorter instructions from court-side occurred with a strong focus on actively catching the ball i.e. “Meet the ball.” (Coach 6), “Drive into the ball, meet it.” (Coach 6) and “Reoffer [Player 3]” (Coach 2).

Shooting was also often referred to with many of these comments taking the form of praise and encouragement “keep your shot going, good girl, get in there” (Coach 5) and “Don’t step in [name 7], come on, keep...” (Coach 6). Longer comments were more statistically focused and generally occurred during breaks in play or between coached: “Score [Player 4]. Forty-eight, twenty-five. Great shooting. A hundred percent in that quarter. Eight from eight.” (Coach 2) and “Yeah. There we go. Ninety-seven and a half percent.” (Coach 2).

Feeding was also a common comment and closely related to comments about ‘movement in circle’. Unlike ‘passing’ and ‘receiving a pass’, comments about feeding, which is a form of pass, where slightly longer and more instructional. Good examples of this are: “Once you get in there, we probably need to go back to a little bit of circle edge and use each other but yeah.” (Coach 2), “ There’s room to feed low ball...low, just right down on, yeah at their feet.” (Coach 6) and:

> And they did zero long feeds in that one [Player 21]. They did. Yeah. So nothing, they’re not doing any feeds from off that circle. Probably because [Player 11]’s coming out. Yeah. But if they are out, don’t worry about them. (Coach 2)

3.3.4 Positioning dimension

The third largest dimension ‘positioning’ only contained two higher order themes, these were ‘team positioning’ and ‘player positioning’. ‘Player positioning’ was the largest of the two.

Table 4.

Model of the general dimension ‘positioning’ linked to its higher, lower and raw data themes.

General dimension	2nd order themes (higher)	1st order themes (lower)	#	Raw data themes (examples)
Positioning (120)	Team positioning	Court position	29	*On her [name 1]. Yes, [name 2]. Middle, middle. Oh gee. Middle [name 1].*
		Hold width	8	*so we need to keep that width on*
	Player positioning	Hold the opposition	25	*Keep her there [name 28].*
		Body positioning	42	*Take the front, take the front.*
		Body contact	8	*we need someone to, somebody hit somebody early, not physically, if it is physically, it doesn’t matter, just*

			slow them down, get a bit of instruction, just slow that blow down because they're getting it out and they're scoring it so that's huge so we need everyone
	Squeezing	1	*That happened in like slow mo. That was quite funny. No, it's 'cos she, she squeezed it real hard and it*
	Marking an opponent	7	*Oh we're getting a bit separate and then we've got a shooter cutting back, we're missing her. That's about it.*

3.3.4.1 Player positioning

Within the 'positioning' dimension there were only two higher order themes, however 'player positioning' (11.1%) was extremely common in the context of all comments. All coaches were represented in this theme and most comments were short and explicit, for example "Back, backs, back, backs on [Player 21]" (Coach 1) or "protect the space on the shoulder 'cause they're just trying to take her shoulder out." (Coach 4). There were examples of comments that were longer and strategic in focus like:

> …but we're just not moving enough. Sometimes when we're coming in, we're coming in on the back of that player who's got the front feed, even if we got to the side, then we're opening our mate up on the other side. (Coach 6)

3.3.5 Possession outcomes dimension

One of the smaller general dimensions was 'possession outcomes', which contained the three higher order themes 'centre pass', 'gains/losses' and 'scoring'. 'Gains/losses' and 'scoring' where the largest themes which when combined accounting for a large number of the overall comments in this study.

Table 5.

Model of the general dimension 'possession outcomes' linked to its higher, lower and raw data themes.

General dimension	2nd order themes (higher)	1st order themes (lower)	#	Raw data themes (examples)
Possession outcomes (74)	Centre pass		10	*Um [name 32], on that centre pass, she's just looking for you the whole time.*
	Gains/Losses	Tips or turnovers	13	*Yeah. She, I think she got one tip off it. Early on.*
		Gain possession	11	*we've only five though so we wanna get that up a little bit, yep.*

	Intercept attempts	2	*ok on defence, as a team, really great defence coming through the court. WA and GA can you have a go at those cross courts 'cause we can get those.*
	Rebounds defensive	5	*They're down to 47 seconds, I hope she gets the bloody rebound off it. We're going into extra time mate.*
		4	*So two things there, one we need to pick up our scoring rate on that and anything we get, we score*
Scoring	Scoring from turn-over	9	*Nine in nine. Turnover to score's forty-four percent.*
	Shooting outcome	10	*Hundy. Ninety-three percent. Eighty-nine.*
	Scoring from centre-pass	4	*Yeah. So they're four from nine*

3.3.5.1 Gains/Losses and Scoring

The 'possession' dimension was less frequent in the coaches comments, however when the higher order themes of 'gains/losses' (4.1%) and 'scoring' (3.6%) are combined they still contributed significantly. Comments within 'gains/losses' generally focused on 'gaining possession' or 'tips or turnovers' with discussions many times occurring between coaches about the frequency of gains or turnovers, e.g. "What's the um…? How many gains have we got? Yeah. What's our turnover to score look like?" (Coach 1). The higher order theme 'scoring' also followed this pattern with comments like "Nine in nine. Turnover to score's forty-four percent." (Coach 2).

3.3.6 Strategy dimension

Although the 'strategy dimension' was smallest, within its five higher order themes, 'defensive strategy' was one the largest in the whole study. One of the defining characteristics of this dimension was the length of the comments. When 'strategy' was discussed, the coaches spoke at length regardless of the whether they were speaking to players or the other coach.

Table 6.

Model of the general dimension 'strategy' linked to its higher, lower and raw data themes.

General dimension	2nd order themes (higher)	1st order themes (lower)	#	Raw data themes (examples)

Strategy (68)			6	*So yeah. It, it, I guess it reinforces that our game plan's working. Or it tells us that there's somewhere we need to pick up the lack on.*
	Attack strategy	Balance / variety strategy	7	*Um I think just when we need that variety on, let's try to get back to that middle space. Yep. And hit that top. Yeah.*
		Centre Pass movement	9	*So on that centre pass, so [name 24] if you're starting off the line, they're letting [name 7] get her space there and then they're blocking you and then you're coming down in here um with the GD dropping back and you're both caught on that side.*
	Defensive strategy	Attack opp	8	*Track and attack bub*
		Contesting centre pass	13	*Good timing except we've gotta shut down that first, yeh, let's try coming over in a wedge rather than a flat 'cause when we get together, we only need one of us to step forward and they get the slide, she's sliding on us, just sliding through the side, angle on a wedge so we've both got our back and our feet up*
		Team defence	3	*So whole team effort on defence, everyone defending from the start 'cause any turnover ball we can generate is gonna give us more opportunity to score so when we're on attack*
		Circle defence	21	*Behind, yeah, but we're not setting up right in the circle either to even have an opportunity and we're not pushing them wide enough outside the circle.*
		Screening	3	*Yep. In a way it's kind of good, 'cos if she goes then you can just pop. Yeah. Yeah. She's, she's screening and coming to your inside.*
		Hold 1 on 1	9	*It's almost like she has to just stay one on one and while she's got her in that position, she keeps here there. ___*
	Errors		5	*We've just gotta not lose the ball down here.*
	Vision		11	*What's on their shoelaces. Always shoelaces, always.*
		Look for space	5	*Space, space, space, space, space. Ooh.*
	Niggle		3	*Yep, you're doing some really good little niggly stuff and racking with her so just keep that going*
		Drawing a penalty	1	*Even if it's a little pop and then hopefully, I'm hoping we can draw the contact from it.*
		Confuse Opp	12	*um, let's see if we can upset [name 9] a little bit in this quarter so*
		Holding	1	*She's holding with her arm every time isn't she?*

3.3.6.1. Defensive strategy

'Defensive strategy' (7.6%) was one of the most regularly commented categories across all coaches with the majority of this focusing on 'circle defence' (n=21) and 'contesting centre pass' (n=13). Comments tended to be very long and complex, with multiple parts:

> Second phase, [name 17]'s going really hard on the first, you just creep creep creep, they'll look long, [name 18] is looking in quite quickly and going on and attack, yeah, so [name 17] come up hard on hers and they've, a first, she's gone first base centre pass then it's coming into [name 18] here.

Some comments where short and directed to players on-court, "Post [name 17], post, stay stay stay [name 19]" (Coach 6), however these were not as common as in other general dimensions. One of the limiting factors of the comments in this dimension was that due to their complex nature they tended to cross over with other dimensions i.e. the above quote on circle defence contains actions, movements, positioning. However, the overall theme of the comment is one of strategy.

3.4 Discussion

This study has identified seven higher order themes that netball coaches frequently discuss or make comment on during matches. This comes from a total of 18 higher order themes sitting within 5 general dimensions discovered during an inductive content analysis. The next stage of this study is to put these seven most common higher order themes into context through the use of existing literature. A review of literature found that five key studies identified PI's in netball. Across each of these there is agreement with the findings of this study, however no study was broad enough to include all seven of the common higher order themes.

3.4.1 Comparison to Previous Research

A comparison (Table 7) of existing research demonstrates the context of the findings of this study. O'Donoghue, Mayes, Edwards, & Garland (2008) , Bruce, Brooks, & Woods, (2018)and Croft, Willcox, & Lamb, (2018) focused on the tactical and technical aspects netball. Studies by Hewit, Cronin, Button & Hume (2011) and Paget, Spencer, & Kilding, (2015) focused primarily on the physical elements of the game while there were no studies that identified PI's focused on measuring 'team positioning' or 'player positioning'. This shows there are gaps in the netball literature that require further investigation. It is acknowledged that these studies may have not intended to assess all aspects of performance.

Tables 7.

Alignment between the high order themes and existing netball literature.

Second order themes (higher)	%	O'Donoghue et al. (2008)	Croft et al. (2018)	Bruce et al. (2018)	Hewit et al. (2011)	Paget et al. (2015) - umpires
Movement timing	2.7				Y	Y
**Movement intensity*	*9.2*				*Y*	*Y*
**Movement spatial*	*20.2*				*Y*	*Y*
Opposition movements	3.7					
**Team positioning*	*5.0*					
**Player positioning*	*11.1*					
Opposition actions	2.3		Y			
Defending actions	3.5		Y			
**Attacking actions*	*17.7*		*Y*	*Y*		
Centre pass	1.3	Y	Y			
**Gains/Losses*	*4.1*	*Y*	*Y*	*Y*	*Y*	
Scoring	3.6	Y	Y	Y		
Strategy (generic)	0.8					
Attack strategy	2.1		Y			
**Defensive strategy*	*7.6*		*Y*			
Errors	0.7		Y			
Vision	2.1				Y	
Niggle	2.3					

** Indicates the most common higher order themes.*

O'Donoghue, Mayes, Edwards, & Garland (2008) identified 13 performance indicator (PI), with their operational definitions, in netball. These were; number of centre passes, number of interceptions, number of sidelines, number of backlines, number of toss ups, number of def rebounds, number of penalty / frees, number of turnovers, number of missed shots, number of goals, % centre passes to goal, % turnovers to goal, % shots scored. These 13 PI's align with the seven most common higher order themes of Attacking Actions and Gains/Losses which is approximately 22.5% of the tactical conversations by coaches in this study.

More recent research (Bruce, Brooks, & Woods, Team and seasonal performance indicator, 2018) has utilised non-metric multidimensional scaling (nMDS) to identify PI's, which is a technique used in ecology for identifying similarity between items. The authors identified; Attempts, Shoot Percentage, Goal Assists, Rebounds, Centre Pass Receives, Intercepts, Deflections, Penalties and Turnovers as PI's of importance, however they could not relate these to synchronous league-wide evolution of the way the game is played. These PI's align with those of O'Donoghue, Mayes, Edwards, & Garland (2008) and fit within the same higher order themes resulting in 22.5% similarity.

The Croft, Willcox, & Lamb, (2018) paper shares again the same common higher order themes as Bruce, Brooks, & Woods, (2018) and O'Donoghue, Mayes, Edwards, & Garland, (2008) however they did include Defensive Strategy also. This equates to approximately 30.3% similarity with the coaches common higher order themes in this study. Croft, Willcox, & Lamb, (2018) did however cover elements of Defensive Strategy, being the only paper reviewed to do so.

Hewit, Cronin, Button, & Hume, (2011) surveyed 52 netball coaches and associated staff finding that; "1)fast and sharp change of direction movements inclusive of rapid decelerations and explosive accelerations; 2) aerial changes of direction (i.e. the ability to turn fully in the air prior to landing); 3) single-leg jumping ability; 4) awareness of the ball, teammate and opponents; and 5) interception timing and accuracy" where considered the most important aspects of netball player development. This study agrees with these findings however only in the Movement Intensity, Movement Spatial and Gains/Losses second order themes. This equates to approximately 34.6% of the conversations observed by coaches during netball games.

Spencer, Natasha, Farley, & Kilding, (2019) focused primarily on the physical elements of the game for umpires in netball. This included the physical, physiological and movement characteristics of umpires during match with triaxial accelerometers, heart rate monitors and video coding to classify movement. These measures or PI's are useful but also isolated to a small proportion, 29.4% for the 'movement intensity' and 'movement spatial', of the higher order themes found in the results of this study.

3.4.2 Future Research

The use of content analysis is a common technique for understanding spoken, transcribe language, however there are limitation to the technique in that interpretation can create errors is categorising the comments. There are also many times where the comments could have been placed in multiple categories within the model. Future research needs to focus on alternatives to content analysis that could provide a more robust approach. Artificial intelligence and computer algorithms could provide a solution as methods like Natural Language Processing have been used successfully in other fields (Peifeng et al., 2007).

There also needs to be further research in performance analysis that looks at multi-variant or systems based, non-linear, analysis of a wide range of performance indicators. As shown in this paper many studies focus on just one aspect of performance, however performance is complex and influenced holistically by many factors that are fall in the physical, psychological, technical and tactical domains. This study provides a good starting point for the identification of performance indicators as its findings are aligned to the end user, the coach, which makes it highly ecologically valid.

Finally, this study has found that there is a lack of research identifying PI's in the areas of 'team positioning' or 'player positioning', which are areas that coaches commonly discuss during netball matches. Research into these areas would help inform coach discussion during matches and help identify measurable assessment for a more evidence-based coaching.

3.5 Conclusion

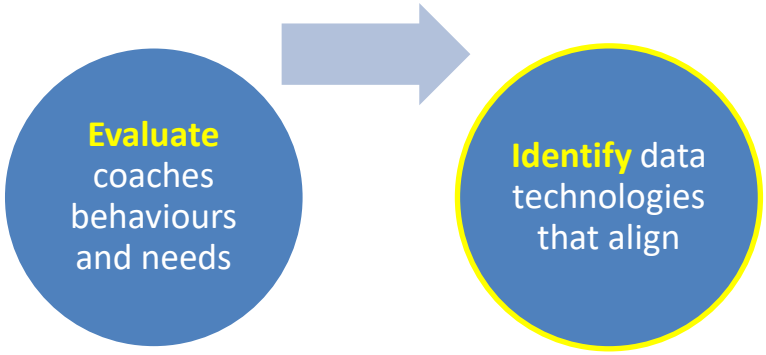
Netball coaches discuss a wide range of tactical themes during matches, however the literature only partially supports these with studies generally focusing on either physical PI's, like 'movement intensity' and 'movement spatial', or on 'attacking actions', 'gain/loss' of possession or more general 'defensive strategy'. Coaches also discuss 'team positioning' and 'player positioning' however there is a lack of literature that looks at these in a tactical context. Considering 16.1% of all coach dialog, in this study, was found to be focused on these areas, further research and work needs to be completed.

Prelude Chapter 4.

The previous chapter identified themes that a group of coaches discussed during a netball match. These included the general dimensions, movement, actions, positioning, possession outcomes and strategy. There were also several higher order themes which fit within each of these dimensions and gave a range of tactical areas where information could be useful to a coach. These findings provided insights in the context of netball, a highly tactical evasive team sport. This approach has now been shown as a valid method for understanding coaches behaviour during matches and added clarity to how Stage One (Figure 7) of the framework could be applied.

Now that tactical themes have been Identified, the next stage in presenting large datasets to coaches requires research into information technologies that provide data. This would create a second stage to the framework in Figure 7. Research evaluating information technologies in sport is not widespread. Information technologies are constantly evolving; therefore these limited studies lose their relevance quickly, and need to be repeated more frequently. As there was very little information published about netball specific information technology, there was a need for this to be conducted.

Figure 7. *The second stage of a framework to guide the use of data for coach decision making.*



If data, and therefore information, can be accessed for each of the tactical themes identified, then it could be aligned in a feedback tool for a coach to use during a game. For this to be achieved the next step would be to conduct a review of what technologies and data sources currently exist in netball that are aligned to the coaches tactical themes (chapter 4).

The purpose of chapter 4 was to identify information technologies that align with the spoken themes of coaches and then to assess each against set criteria. These criteria needed to establish whether the data were accessible, valid, and compatible.

Chapter 4.0 An Evaluation of Data Sources in Netball Performance Analysis

This chapter comprises the following manuscript (3) under revision with The Journal of Sport and Exercise Science.

Croft, H., & Spencer, K. (2021). An Evaluation of Data Sources in Netball Performance Analysis. *(Under review) The Journal of Sport and Exercise Science.*

Author contributions: Croft, H (90%), Spencer, K (10%)

4.1 Introduction

The importance of feedback has been acknowledged in literature for not only learning but also for its psychological impact upon players (Groom & Cushion, 2005). The balance of positive and negative feedback is an important consideration; however, the relevance or ecological validity of the information also plays an important role. Ecological validity is a term that has been debated in literature (Holleman et al, 2020; Davids, K., 1988) and is sometimes misunderstood by researchers who claim their research is representative of the "real world". The definition used in this manuscript is the realism with which a design matches the user's real work context (Hartson & Pyla, 2012). The provision of data to coaches during sports matches should consider if the information is ecologically valid to the coach.

The importance of recording events and presenting summarised information to coaches has been demonstrated (Nicholls & Worsfold, 2016) as essential for accurate decision making. However limited research has attempted to align this information with actual coaches' behaviours. It is argued that many data producing and presenting technologies do not align with coaches planning and approach to a sports match. Research conducted by Croft et al. (2020) provides insight into what coaches discuss during netball matches. The themes identified from this study were used as criteria for the selection of data for coaches during a match, creating a more ecologically valid approach to data feedback during sports matches.

Complex dynamic systems have been highlighted as a future direction for sports performance research (Reed & Hughes, 2006), however for these systems to be useable, large volumes of data are required (Rein & Memmert, 2016). Combining this approach in a way that is useful and meaningful to a coach requires not only large amounts of organised data, but also data that is aligned with a coach's view of the game.

This paper outlines the types of data that currently exist in sport and the various systems and devices that produce it. It identifies the many opportunities to expand the incorporation of data as well as the challenges of integrating data into the analysis process to potentially improve coach and player decision-making. Liebermann et al. (2002) conducted a review of similar scope, however their focus was directed towards information technology that provides feedback about skilled or technical, not tactical performance. While it provided important information for coaches and practitioners an update is required, due to the rapid advances in information technology, with a focus on the largest attribute of coaching during a game (Zetou et al, 2011) tactical information.

4.1 Sources of Data in Sport

Data sources and types are in abundance in high performance sports. Information technologies that produce these data can be expensive with annual subscriptions for some team sport global positioning systems (GPS) or computerised notational analysis software costing between $5,000-20,000 (NZD). There are many other information systems that should be considered, some which have no cost and others being very expensive. This section gives examples of different categories of data generating information technologies and then overviews how each produces and presents this data and in what format.

Several categories of information sources have been identified by (Croft H. , Spencer, Robinson, & Cronin, 2021). These included ubiquitous computing systems (UCS), notational analysis systems and public domain data sources. UCS are essentially those where sensors and microprocessors are embedded into everyday objects and then data is gathered about the sensory characteristics of the object (Chi, 2008). Examples of these include heartrate monitors and GPS embedded in wearable devices like watches or accelerometers, gyroscopes, microphones and cameras. These devices are data collectors and do not produce information until algorithms or software interpret their measures against set criteria.

Computerised video analysis software, as outlined by (O'Donoghue & Holmes, 2015), either take video images and automatically produce information about participant kinematics or events within the footage, or allow an operator to generate objective and subjective information manually. These software share many attributes with notational analysis software which allow the operator to generate information through observation during a sports match. Some like Hudl-Sportscode™, Dartfish™, Focus™ and Nacsport™ allow for automation of feedback via dashboards, from this "coded" data.

Another important category of information, that could be of use to performance analysts and coaches, is public domain information. This can exist as algorithms, webpages, or application

programming interfaces (APIs) which are shared freely to the public for analysis and research. APIs can be directly connected to as a data stream, while algorithms like Keyhole™ have been used to create information about social media language and the now defunct Topsy™ have been used to assess medical networks (Mishori et al, 2014). Along with sport league APIs, possibly the most useful public domain information to performance analysts and coaches is that which can be "scraped".

Data Scraping is also a common method for capturing publicly available information on the internet. This information commonly appears in the form of tables. Tables that list performances, or inputs, for a team or individual can be easily copied and pasted into a CSV (comma-separated values) file. This information, if aligned with identifying information, for example Match ID or Player ID, can be very useful. Target Internet. (2020) outlines the various forms of data scraping including website scraping, where a user can copy information into a locally stored spreadsheet or CSV. This process can be automated (Provost, 2016) using a "get data" function in Microsoft Excel™ or other similar product, which creates a live connection to the webpage where the data is stored. This approach is useful in that it gives access to large amounts of historical data stored on the internet.

4.2 Methodology

Based on the study by Croft et al. (2020), 18 higher order themes were identified for tactical discussions by coaches during live netball matches (Table 8). This paper will identify data that could be used to create information about these themes during a live netball match. A review of existing literature and public domain information will be conducted to evaluate what data types currently exist.

Using search phrases and terms which included the general dimensions from Table 8, netball research that collected quantitative data was collated via a Google Scholar™ web search. Quantitative data was required as this is suitable for inclusion in an automated database in real-time during netball matches. From within the methodologies of each paper, information technologies that produced data related to the higher or lower order themes, as identified by Croft et al. (2020), were evaluated against inclusion criteria.

Table 8.

Content analysis model developed by Croft et al. (2020) from coach conversations during netball matches.

General dimensions - Level 1	2nd order themes (higher) - Level 2	#	%	1st order themes (lower) - Level 3	#
Movement (267)	Movement timing	20	2.7%		5

Positioning (120)	Movement intensity	69	9.2%	Timing of movement	15
				Driving (run)	24
				Explosive movement	32
				Footwork	13
	Movement spatial	151	20.2%	Create space	18
				Limited space	1
				Connection with team	9
				Movement in circle	31
				Move with opposition	29
				Where to move	44
				Changing direction	11
				Angles	8
	Opposition movements	27	3.6%		27
	Team positioning	37	5.0%	Court position	29
				Hold width	8
	Player positioning	83	11.1%	Hold the opposition	25
				Body positioning	42
				Body contact	8
				Squeezing	1
Actions (175)	Opposition actions	17	2.3%	Marking an opponent	7
				Traits	9
				Slow opposition pass	1
				Predicting opposition actions	7
	Defending actions	26	3.5%	Defending a shot	7
				Defending	15
				Jumping	4
	Attacking actions	132	17.7%	Shooting	36
				Passing	46
				Feeding (pass into circle)	25
				Receiving a pass	25
Possession outcomes (68)	Centre pass	10	1.3%		10
	Gains/Losses	31	4.1%	Tips or turnovers	13
				Gain possession	11
				Intercept attempts	2
				Rebounds defensive	5
					4
	Scoring	27	3.6%	Scoring from turn-over	9
				Shooting outcome	10
				Scoring from centre-pass	4
Strategy (117)		6	0.8%		6
	Attack strategy	16	2.1%	Balance/variety strategy	7
				Centre Pass movement	9
				Attack opposition	8
	Defensive strategy	57	7.6%	Contesting centre pass	13
				Team defence	3
				Circle edge defence	21
				Screening	3
				Hold 1 on 1	9
	Errors	5	0.7%		5
	Vision	16	2.1%		11

				Look for space	5
					3
	Niggle	17	2.3%	Drawing a penalty	1
				Confuse opposition	12
				Holding	1
		747	100.0%		747

4.2.1 Criterion for inclusion

To assess different data sources for their suitability for use by coaches during netball matches a set of inclusion criterion must be established. Palmius. (2007) described the task of comparing information systems as one which requires a set of operational and measurable criteria, which are thought to be important. The follow operational criteria were used to assess the suitability of these information technologies:

- Ecological validity - relevance of information to the topics the coaches discussed in Croft et al. (2020). This was assessed at all three levels, general dimensions, higher (2^{nd}) order themes and lower (1^{st}) order themes, to identify coverage and specificity.
- Face validity – does the information produced by the technology appear to be useful to the coach or players? For example, when they see the information or data, can they act upon it?
- Compatibility: data type and format - whether the data can be directly compared to that from other sources and the way it is organised.
- Compatibility: speed and accessibility - the speed with which the data can be captured and stored as well as whether the data can be accessed easily from its current source.

For each factor, operational ratings were created which identified the suitability of the information for live decision-making. Upon completion the information systems were categorized from poor to excellent and recommendations made about their consideration for use as a live analysis tool.

4.3 Results

As can be seen in Table 9, information systems used by Vanwanseele et al. (2014) were rated poor for relevance to skill learning and also poor for compatibility due to the fact that 3-dimensional motion analysis cannot be conducted live during sports matches. Hetherington et al. (2009)'s 2-dimension video analysis was also rated as poor for compatibility as 2D digitisation required the kinematic feature to be perpendicular to the camera to avoid parallax error. This is virtually impossible in a live sporting context due to the players moving freely around the court. The information captured by Hetherington et al. (2009) was rated good for its relevance to skill learning, in that it would be information a player or coach could act upon.

The observed frequency data from Hale & O'Donoghue. (2007) would be valuable to coaches and players for improving decision making around a number of lower order themes, however due to slow motion footage being required to digitise and rate these events it would be very hard to conduct live during a match and therefore received a rating of poor-moderate. A small sample of these events could however be captured by a skilled operator.

The heart-rate and accelerometery data captured by Chandler et al. (2014) could be made available in a live setting, however extraction of this from the commercial software would require further investigation. Rated perceived exertion (RPE) could be monitored in the breaks between quarters, however would detract from the discussion between players and coaches at these times, meaning that the compatibility of these systems is only moderate. The information gained however would have good relevance to a number of lower order themes and therefore could be good feedback.

Higgins et al. (2009) use of GPS for collecting performance information shows good validity for 4 of the 5 general dimensions with a range of information that would also be relevant for feedback on skill execution. As professional netball is conducted indoors conventional GPS does not work however there are some systems that allow for an indoor mode. Further investigation with specific brand and models would need to be conducted and access to this data live for integration with other data sources would also need to be looked at, meaning a good-moderation rating was given.

The observed notational analysis approaches adopted by Davidson & Trewartha. (2008) as well as the Point 9 Analytics™ and Champion Data™ websites were rated as excellent and good-excellent respectively, for their validity to skilled feedback. There were several higher and lower order themes aligned with this data. The notational analysis software used by Davidson & Trewartha. (2008) can be customised to capture any observed event and can also live export data as it is being captured, meaning it was rated excellent. The other two company data sets would require an API to allow integration of this data live with other sources, however the Champion Data™ data is published in real-time online for observation.

Table 9.

Information systems and their suitability against criteria identified.

Study	Ecological validity: Dimensions and order themes	Face validity: Usefulness	Compatibility: data type and format	Compatibility: speed and availability
Hetherington, S., King, S., Visentin, D., & Bird, M. L. (2009). A kinematic and kinetic case study of a netball	Poor *(General dimensions:* Movement	Good	Poor *2D coordinates, 25fps (time series)*	Poor

shoulder pass. *International Journal of Exercise Science, 2*, 243-253.	*Lower 1st order themes:* Movement timing Explosive movement			
Chandler P, Pinder S, Curran J, et al. (2014). Physical demands of training and competition in collegiate netball players. *J Strength Cond Res. 28(10):*2732–7.	Moderate *(General dimensions:* Movement *Lower 1st order themes:* Driving (run) Explosive movement Changing direction Angles	Good	Good *Heart rate, RPE, accelerometer player load*	Moderate
Hale SL, O'Donoghue P. (2007). Addressing turning and direction changes when using the Bloomfield Movement Classification. *Int J Perform Anal Sport, 7*:84–89	Moderate *(General dimensions:* Movement Actions *Lower 1st order themes:* Changing direction Angles Driving (run) Explosive movement Footwork Jumping	Good	Good *Observed frequencies, duration, percentage time (slow motion video)*	Poor-moderate
Vanwanseele B, Stuelcken M, Greene A, Smith R. (2014).The effect of external ankle support on knee and ankle joint movement and loading in netball players. *J Sci Med Sport, 17*:511-515	Moderate *(General dimensions:* Movement *Lower 1st order themes:* Footwork Timing of movement Angles	Poor	Moderate *3D motion analysis coordinates - range of motion, force plate – ground reactive force (time-series)*	Poor
Higgins T, Naughton GA, Burgess D. (2009). Effects of wearing compression garments on physiological and performance measures in a simulated game-specific circuit for netball. *J Sci Med Sport, 12(1)*:223–6.	Good *(General dimensions:* Movement Positioning Actions Strategy *Lower 1st order themes:* Driving (run) Movement in circle Court position Hold width	Good	Moderate *GPS – distance, velocity, (time series)* *Note: Indoor version required*	Moderate-good

	Centre Pass movement			
Davidson, A and Trewartha, G. (2008). Understanding the physical demands of netball: A time-motion investigation. *Int J Perform Anal Sport, 8*: 1–17.	Moderate *(General dimensions:* Movement *Higher 2nd order themes:* Movement timing Movement intensity Movement spatial	Excellent	Excellent *Notational analysis – coded events live XML (live) or .csv files (live)*	Excellent
Champion Data. (2019, December 10). *ANZ Premiership*. Retrieved from Champion Data: https://mc.championdata.com/anz_premiership/	Excellent *(General dimensions:* Positioning Actions Possession outcomes Strategy *Lower 1st order themes:* Team defence Circle edge defence *Higher 2nd order themes:* Team positioning Player positioning Opposition actions Defending actions Attacking actions Centre pass Gains/Losses Scoring Errors	Good-Excellent	Excellent *Website data and CSV data, XMLs also available*	Good-Excellent
Point 9 Analytics. (2019, December 10). *2019 ANZ preseason*. Retrieved from Point 9 Analytics: www.point9analytics.co.nz/2019-anz-preseason/	Good-Excellent *(General dimensions:* Positioning Actions Possession outcomes *Lower 1st order themes:* Team defence Circle edge defence *Higher 2nd order themes:* Team positioning Player positioning Opposition actions Defending actions Attacking actions Centre pass Gains/Losses	Good-Excellent	Good-Excellent *Website data and CSV data*	Good

Scoring
Errors

4.3.1 Themes without information technology data

All five general dimensions identified by Croft et al. (2020); movement, positioning, actions, strategy, and possession outcomes, were represented in the review of data sources in this paper. However, for meaningful insights to be gained, data explaining the performance of lower and higher order themes needed to be identified. The results of the online website and literature review showed that only 13 lower (1^{st}) order themes were matched in Table 9 with data creating information technologies, whereas 22 lower order themes were not. Information at this level is quite specific and could be explained by the broader higher (2^{nd}) order themes, where 12 were identified in the online website and literature review while five were not.

The main areas that this review did not identify information technology data for were opposition movement, attack strategy, defence strategy, vision and niggle. For the attack and defence strategy higher order themes that were not represented, the lower order themes of centre pass movement, circle edge defence and team defence, where represented with information. This highlights that finding information about opposition movements, vision and niggle were not available, however as vision and niggle are quite subjective this may not be achievable. Opposition movement is an area where further investigation, and information technology development, would be beneficial.

4.4 Discussion

The results of this paper (Table 9) show there are two categories of data that are available to provide information to coaches about the performances in a netball game. The first is those used in research for data collection purposes including 2 and 3-dimentional video digitisation software, GPS, accelerometers, and heartrate monitors (Table 9). These are generally difficult to access in real time, in a format that can be uploaded to a generic databased, however there is potential with further work to develop solutions for each.

The second category are those that are designed for performance analysts or spectators to be used in real-time. These include notational analysis software like Hudl-SportsCode™, Nacsport™, and Dartfish™ which allow the coding of events by a trained operator. Public domain webpages which update throughout games also use this data collection method and can be data or web scraped either manually or via a software product like Microsoft-Excel™. Based on the review conducted in this paper the second category of data sources is better suited to collection and importation into a cloud-based database. This finding agrees with many studies and websites that describe the design of performance

analysis systems (O'Donoghue & Holmes, 2015; Arastey, 2020; What is sports performance analysis, 2020).

This paper has also identified that current information technologies do not provide sufficient information around opposition movement. This could be due to the intrusive nature of wearable ubiquitous tracking devices, like GPS, where an opponent is very unlikely to wear a device or share data for an oppositions benefit. Automatic video tracking systems do exist in sports like basketball; however these are expensive and require significant infrastructure within a venue to operate successfully. Products that track players from a single vantage point are available (Veo Cam™), however the algorithms that are trained to track the movements of players still need further development.

The timeframe in which this data can be included in an analysis system will determine its suitability for use in understanding the current state or the performance or its use in providing historical context to a performance. Both XML files and data scrape CVS's can be used to produce a historical or normative dataset, and for also assessing the current state of performance.

4.5 Conclusion/Recommendations

It is recommended that operator controlled notational analysis software, like Hudl-SportsCode™, Nacsport™ or DartFish™, are the best suited to providing live information and data to coaches and players during netball matches. This is due to the excellent face and ecological validity that can be provided, as well as the data types, formats, speed of access and accessibility. Connecting to a data provider who use this technology, like Champion Data™, to produce netball information, is also an excellent option. These providers also have large historical databases that could be utilised to provide context to live performances.

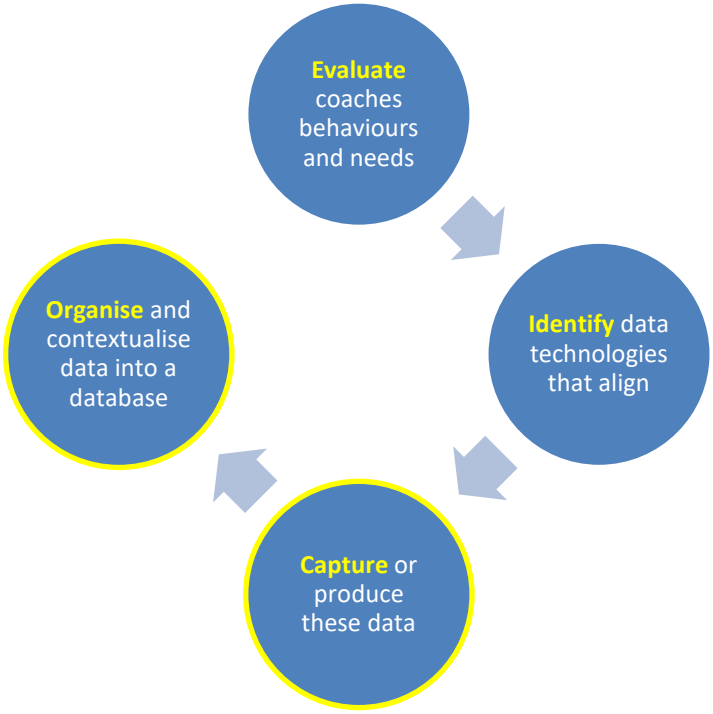
Historical datasets give context to the information being analysed and normative data tables are one example of this. Normative data tables are used in various disciplines (Fox, O'Malley, & Blake, 2014) yet their use is limited in sports performance analysis. Convention has been to research key performance indicators (KPIs) and performance indicators (PIs) in isolation with inferential statistics. These PIs are then generally used as targets for players. Further research is required to utilise historical datasets in producing normative data. This normative data could then provide context to a range of variables during a match, as to whether they are relatively high or low for a performance for a player is a specific position. This will be the focus for the next part of this ongoing research.

Prelude Chapter 5

The previous chapter attempted to linked themes from chapter 3, to information systems that produced data relating to these themes. These two chapters informed our understanding of the first two stages of an overarching framework (Figure 8). Chapter 4 found that the public domain Champion Data™ website and Hudl-Sportscode™ software was aligned with the themes that coaches discussed in chapter 3. These sources also met the requirements of a live coaching tool, with strong compatibility and accessibility. The next step in the systems design approach was to capture or produce data in real-time, as depicted in Figure 8. Due to the size and complexity of this stage, it is presented as two parts within the framework in Figure 8.

Figure 8.

A pipeline framework for presenting data to coaches during sports matches.



To build upon the first two stages of the framework the next stage of a systems design approach was 'system design' where the data collection, data extraction, analysis and representation was defined. Automation was a crucial part of this complex process, especially when working with large volumes of data. Once captured, the data needed to be organised, contextualised and held within a database ready for consumption and visualisation within a modern visual analytics tool. These two distinctive steps are those that skilled performance analysts may have already be proficient in, however as

changes to database technologies has been rapid and web-based systems have automatically process data, via algorithms, within ‘pipelines’. This meant that performance analysts may need to continue their education and learn to programme these.

Therefore the purpose of chapter 5 is to design a platform for capturing, processing and organising data so it can be presented as a normative data table. These normative data tables will provide context for players about their performance against historical performances at the same level of competition and position played. Expert coach feedback will be used to guide how this data is best visualized and to maintain ecological validity.

Chapter 5.0 Creating a live and flexible normative dataset for netball.

This chapter comprises the following manuscript (4) which was published in *Frontiers in Sport and Active Living.*

Croft, H., Spencer, K., Taurua, N., & Wilton, E. (2021). Creating a live and flexible normative dataset for netball. *Frontiers in Sport and Active Living. https://doi.org/10.3389/fspor.2021.743612*

Author contributions: Croft, H (85%), Spencer, K (10%), Wilton, E (3%), Taurua, N (2%)

5.1 Introduction

Previous research (Croft et al., 2020) has identified tactical themes that netball coaches discuss during matches and subsequently data from sports performance analysis tools that align with these themes has been evaluated. Public domain data providers like Champion Data™ or notational analysis software like Hudl-Sportscode, allow access to this data via either APIs (in full) or live coding of events, respectively. The twin benefits of this is that the data is generated in real-time and immediately presented to the coach courtside, enabling the coach to reflect-in-action enhancing professional practice and learning (Edwards S., 2017). Therefore, it is essential that the data is presented in a way that is understandable and has strong face validity to assist coach decision making during a match. Historical data and information may provide context to this live data, specifically by indicating whether current-performance is above or below relative past-performance i.e. in comparison to normative data.

5.1.1 Normative data

The importance of normative data (O'Connor, 1990) to aid the establishment of standards of excellence for physicians is standard procedure in medical research. Sport physiology research has developed and utilised normative data, predominately for movement screening (Fox, O'Malley, & Blake, 2014) and fitness assessment (Tomkinson, et al., 2018), anthropometry (BMI, skinfolds) and physical conditioning (submaximal VO2max).

Performance analysis research (O'Donoghue P. , 2005) has developed normative profiles, although these studies have presented normative data tables that generally are fixed and only represent performances from a set period for a specific population.

There are currently no standardised normative data sets present in netball literature. Recent research (Bruce, Brooks, & Woods, 2018) has identified performance indicators that differ between

competitions, or those that are related to winning. However, this research does not provide coaches with usable measures that give context to an individual's performance. Thus, the development of normative data tables for netball would enable the player or the coach to understand if their current performance, for example goal volume, was relatively high, moderate or low. This fixed normative dataset could also be used as a tool in player development pathways to assist with identifying areas of strength and/or weakness in a player, and for rehabilitation targets if recovering from injury. However, comparisons to a fixed data set may need to be viewed with caution, as performance trends and strategies are dynamic constructs. The solution is to create a dynamic normative data table from an automated database that continually updates as performances are added.

5.1.2 Databases and data management

A database combined with other data processing tools appears to be the most obvious solution with some literature having focused on their design specifically for sport science data. Vincent et al. (2009) overviewed the design and implementation of databases in sport with specific focus given to their use as a performance analysis tool. They summarized the most effective use of databases, data management and content management systems is as an answer to the coaches' subjective reliance on "gut feel" or intuitive thinking about performance. The complex nature of navigating databases and creating simple interfaces for coaches that allows them to easily interpret, coordinate and implement the information within should be considered.

In their simplest form, a database can be a spreadsheet much like those built in Microsoft Excel™. In contrast, they can also be very complex with products available that allow capturing, storage, retrieval, interpretation, reporting, and dissemination of information, for example Fusion- Sports$_{inc}$ Smartabase™ athlete management system. Good database design is crucial for usability, however if the data entered is not accurate or valid then they can produce false information. Two possible data entry methods (Vincent et al., 2009) are:

1. An automated computer algorithm which enables fast inputting as data is created without checks for accuracy.
2. Manual inputting which can be slower but more accurate.

This statement is somewhat generalised as it could be argued that a well-designed, automated process, with validity, reliability, and accuracy checks (O'Donoghue, 2015) could be more accurate than manual inputting which is susceptible to human error. As this study will look to automate as much of the process possible, checks for accuracy need to be integrated via validation conditions. Well-designed databases or data management systems should avoid unnecessary redundancy and duplication of information. These recommendations are based on a relational model for databases

(Codd, 1970) that involves organizing data into tables of rows and columns.

5.1.3 Sources of data and variable selection

The themes identified by Croft et al. (2020) were utilized in this study and focused on movement, positioning, actions, possession outcomes and strategy, with layers of lower and higher order themes. (Croft & Spencer, 2021) identified both publicly available information (Champion Data™) and coded information produced using a notational analysis software (Hudl-Sportscode™, Dartfish™, Nacsport™) were best suited for use in a dynamic database tool. This was due to them having good face and ecological validity, could be accessed, and organised easily and as publicly available information like Champion Data™ have several years of historical data available, information can be scraped manually or sourced via an XML file. This research has aligned existing data sources to the tactical themes' coaches discuss during games. By incorporating expert coach feedback into the design of data visualizations or data tables ecological validity should be maintained across the design process.

The purpose of this study is to design a platform for capturing, processing and organising data so it can be presented as a normative data table. These normative data tables will provide context for players about their performance against historical performances at the same level of competition and position played. Expert coach feedback will be used to guide how this data is best visualized and to maintain ecological validity.

5.2 Materials and Methods

5.2.1 Equipment

A personal computer device connected to high-speed internet was used to upload the XMLs and CSV file type to an AWS S3 - Amazon Simple Storage Service (a cloud storage service). Event driven pipelines within the AWS Amazon Website Service suite, processed the data, integrating and transforming it ready for consumption in an analysis software tool Tabeau™. Tableau™ was connected to this database and a normative data table was built to summarise means, standard deviations and percentiles for each level of competition and position.

5.2.2 Data source

Data was sourced from coded XML files produced in Hudl-Sportscode™ and a public domain website (Champion Data, 2020). Data was collected between 2014-2020 from; International netball matches, ANZ Premiership (formally Championship), Suncorp Super Netball, Beko National Netball League,

New Zealand (NZ) U23 matches, NZ U19 matches and NZ Secondary Schools competitions. Approximately 20,320 rows of individual player data were collated.

5.2.3 Data acquisition

Both manually scraped and coded XML data sources were collected and loaded into the relational database. The database accepted this data in both raw XML format, which required processing via an algorithm to extract the variables in a format useable within the database. The database was also designed to accept scraped data within a CSV file (Comma-separated values). This data was organised into a single paged matrix format, with one performance per player, per row and variables organised into columns. These samples were tested for completeness and input errors.

5.2.4 Data preparation and cleaning

As scraped and coded data required large repetition of inputting as well as dimensions being added manually, such as competition, team and year, data cleaning was required to ensure an accurate dataset. This was achieved in two ways, firstly visually observing each row of data for missing, misaligned and obviously incorrect numbers. Secondly the data was graphed and checked for averages greater than zero in categories where a player could not have achieved a frequency, i.e. a Centre cannot attempt a goal. Data that was identified as incorrect was either corrected or removed from the dataset. Most errors appeared to be incorrectly inputted player position.

5.2.5 Upload and consumption

Once the dataset was complete the XMLs and .csv tables were imported into the AWS S3 - Amazon™ Simple Storage Service cloud storage and an event driven pipeline automatically transformed the data ready to be consumed by the Tableau™ dashboard. This process of transformation and consumption takes a few seconds depending on internet connections and dataset size. A Tableau™ dashboard was designed to allow the dataset to be filtered by competition level, including international, national (ANZ Premiership and Beko), Australian national, NZ under 23yrs, NZ under 19yrs and secondary schools notational tournament. A filter was also built to allow selection of the year of the performance and the teams competing. This allows the user the ability to select more specific populations if required.

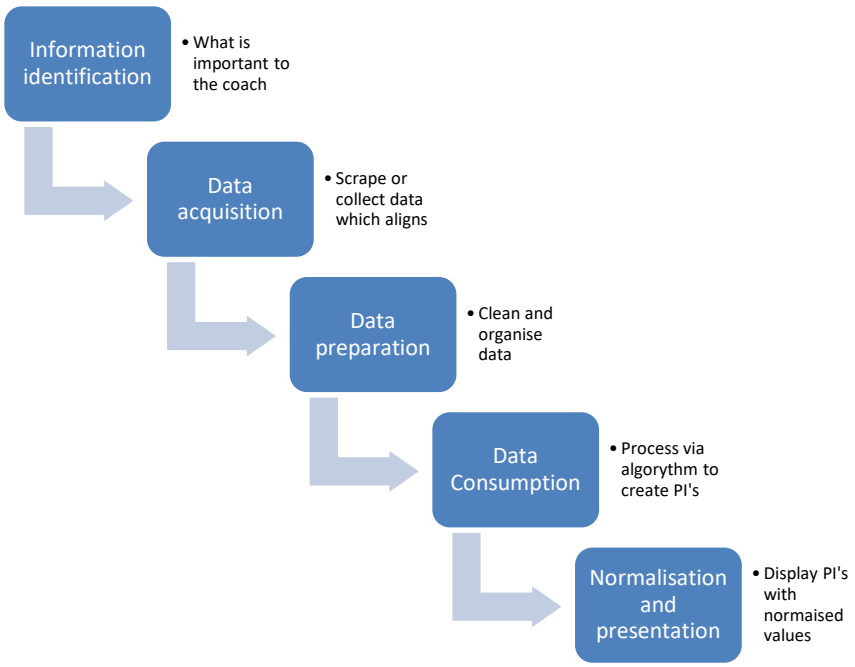
5.2.6 Normalization and presentation

Once in the Tableau™ dashboard further transformation of the data was completed. The performance measures were normalised into per 60 mins played. This was done so that all performances could be

compared across competitions, regardless of the time on court. Sixty minutes was chosen as this is the normal length to an international netball match. For presentation on the normative table, percentiles and medians were used as (Hopkins, 1997) when the data is not of a normal distribution, i.e. there are a lot of performances at the larger end of the scale, an average would not accurately represent the data set. Twenty fifth, 50th and 75th percentiles were chosen for this study, however 90th and 95th percentiles are easily generated using the software's features.

Figure 9.

Pipeline for automated data management for display in visualization tools.



5.3 Results

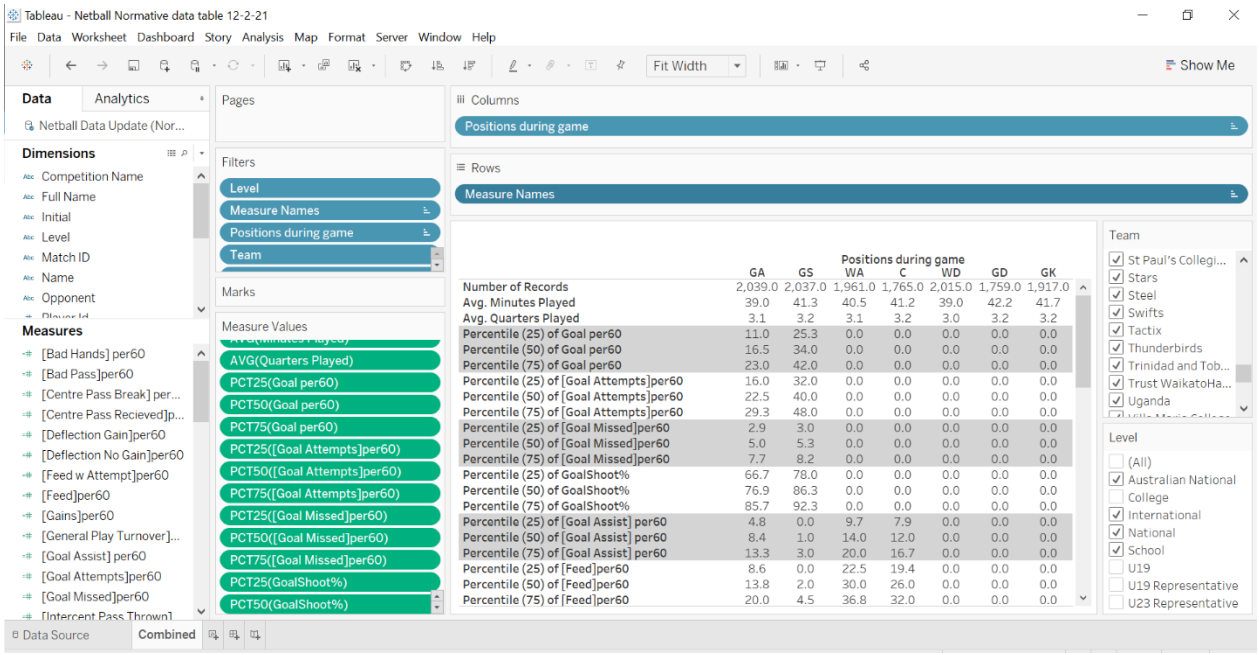
The normative data table (Figure 10) presented here is divided into vertical layers, or clustered rows, for each level of competition and then into columns for each playing position. This allows easy navigation by the user and comparisons for each position. Positions that are similar in their roles, i.e. Goal Shot and Goal Attack, are ordered across the columns for ease when comparison also.

The number of samples (n=1) in each competition level is also listed to allow the user to understand the size of the population they are comparing against. This population size automatically varies dependent on additional matches being uploaded into the cloud database or the filtering option that is selected i.e. years, months, days. The normative data table presented below displays all the data, excluding those with missing data, that has been loaded into the database. The time scale for these matches occurred between the dates of 01/01/2014 and 01/07/2020.

The Dimensions under the Data tab (far top left) can be used to create filters which appear one the left and in the boxes on the far right. These dimensions include; Competition Name, Level, Measure Names, Positions During Game, Team, and Year. However, for this table only Team, Position During Game and Level have been used. In the list of Measures on the left side, a new calculated variable was created that represented the given measure as a value scaled to represent 60 minutes of play. This normalised all values to enable direct comparison, and they were then presented in the table as 25^{th}, 50^{th} (Median) and 75^{th} percentiles or interquartile ranges (Hopkins, 1997) . These can be easily adjusted to represent other percentiles i.e. 5^{th}, 10^{th}, 90^{th} and 95^{th}.

Figure 10.

Layout of Tableau™ reporting interface with filters (right side) for competition, level, and year present.



For reporting in this paper, this table was exported into Microsoft Excel™ and is presented below in Table 10. Table 10 shows, after cleaning and filtering that there were 2039 records, or player performances, for Goal Attacks (GA), 2037 records for Goal Shoot (GS), 1961 for Wing Attack (WA), 1765 for Centre (C), 2015 for Wing Defence (WD), 1759 for Goal Defence (GD) and 1917 for Goal Keep (GK). Performances where a player shifted between 2 or more positions during a game were excluded.

Some of the results presented in Table 10 include, GS's taking more shots at goal (34 vs 16.5 per 60 mins) than GA's and also taking these at a higher goal shooting percentage (86.3 vs 76.9 %). WA's were the most frequent with Feeds (30 vs 26 per 60 mins) and Goal Assist (14 vs 12 per 60 mins). They also had higher Centre Pass Receives than any other position including GA's (19.5 vs 16.7 per 60 mins). As the GA has a role receiving centre passes, feeding the circle and shooting goals, it is logical that they have lower values for these then the more specialized positions of WA and GS. The C position also has a broader role in both attack and defence and therefore has lower values than the WA for feeds (26 vs 30 per 60 mins) and goal assists (12 vs 14 per 60 mins). In comparison to the defensive midcourt position WD for gains (1.0 vs 1.3 per 60 mins). The C position does however have the highest frequency of pick up (1.7 vs 1.1 per 60 mins) which could be due to them being allowed to cover more area of the court than any other position.

In the defensive area of the court, we see a similar relationship between the GK and the GD

as to that of the GS and GA in the attacking end. For defensive measures Intercepts (2.0 vs 1.5 per 60 mins), Deflections Gain (1.9 vs 1.5 per 60 mins) and Rebound (1.5 vs 1.0 per 60 mins), the GK has consistently higher values. This may be due to the GK having more constrained court movement than the GD and not being available to receive a centre pass or feeding the circle.

Finally, an interesting finding for the number of general play turnovers (non-shooting errors) per position, is that the more attacking a player is i.e., GS, the higher the number is (5.0 per 60 mins). This may be because the majority of their role tends to be with possession of the ball, trying to score or support scoring, and therefore there is a greater opportunity for them to concede more turnovers than defensive areas of the court.

Table 10.

Combined normative data of international, national, and school level netball performance indicators.

Position	GA	GS	WA	C	WD	GD	GK
Number of Records	2,039.0	2,037.0	1,961.0	1,765.0	2,015.0	1,759.0	1,917.0
Avg. Minutes Played	39.0	41.3	40.5	41.2	39.0	42.2	41.7
Avg. Quarters Played	3.1	3.2	3.1	3.2	3.0	3.2	3.2
Percentile (25)	11.0	25.3	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Goals] per60	16.5	34.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	23.0	42.0	0.0	0.0	0.0	0.0	0.0
Percentile (25)	16.0	32.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Goal Attempts]	22.5	40.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	29.3	48.0	0.0	0.0	0.0	0.0	0.0
Percentile (25)	2.9	3.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Goal Missed] per60	5.0	5.3	0.0	0.0	0.0	0.0	0.0
Percentile (75)	7.7	8.2	0.0	0.0	0.0	0.0	0.0
Percentile (25)	66.7	78.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of Goal Shoot %	76.9	86.3	0.0	0.0	0.0	0.0	0.0
Percentile (75)	85.7	92.3	0.0	0.0	0.0	0.0	0.0
Percentile (25)	4.8	0.0	9.7	7.9	0.0	0.0	0.0
Percentile (50) of [Goal Assist] per60	8.4	1.0	14.0	12.0	0.0	0.0	0.0
Percentile (75)	13.3	3.0	20.0	16.7	0.0	0.0	0.0
Percentile (25)	8.6	0.0	22.5	19.4	0.0	0.0	0.0
Percentile (50) of [Feed] per60	13.8	2.0	30.0	26.0	0.0	0.0	0.0
Percentile (75)	20.0	4.5	36.8	32.0	0.0	0.0	0.0
Percentile (25)	7.5	0.0	15.0	13.0	0.0	0.0	0.0
Percentile (50) of [Feed w Attempt] per60	12.0	1.4	21.0	18.0	0.0	0.0	0.0
Percentile (75)	18.0	3.9	27.0	22.2	0.0	0.0	0.0
Percentile (25)	12.0	0.0	14.0	0.0	2.0	2.0	0.0
Percentile (50) of [Centre Pass Received] per60	16.7	0.0	19.5	0.0	4.2	4.0	0.0
Percentile (75)	21.7	0.0	24.9	0.0	7.5	7.2	0.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	1.9	2.9

Percentile (50) of [Gains] per60	0.0	0.0	0.0	1.0	1.3	3.5	5.0
Percentile (75)	1.0	0.0	1.0	2.0	2.9	5.9	7.5
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Intercept] per60	0.0	0.0	0.0	0.0	1.0	1.5	2.0
Percentile (75)	0.0	0.0	0.0	1.5	2.0	3.0	3.9
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Deflection Gain] per60	0.0	0.0	0.0	0.0	0.0	1.5	1.9
Percentile (75)	1.0	0.0	0.0	1.8	2.0	3.9	4.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	1.0	1.4
Percentile (50) of [Deflection No Gain] per60	0.0	0.0	0.0	1.0	1.5	2.9	3.0
Percentile (75)	1.0	0.0	1.0	2.0	3.0	4.0	5.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Pick Up] per60	1.0	0.0	1.1	1.7	1.1	1.1	1.0
Percentile (75)	2.0	1.4	2.6	3.0	2.9	2.6	2.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Rebound] per60	0.0	2.0	0.0	0.0	0.0	1.0	1.5
Percentile (75)	1.5	3.9	0.0	0.0	0.0	2.0	3.0
Percentile (25)	0.0	0.0	1.0	3.6	4.5	6.0	7.0
Percentile (50) of [Penalty Contact] per60	2.1	1.7	2.9	6.0	7.3	9.0	10.0
Percentile (75)	4.4	3.0	5.0	9.0	10.9	12.0	14.5
Percentile (25)	0.0	0.0	0.0	0.0	0.0	1.5	2.0
Percentile (50) of [Penalty Obstruction] per60	0.0	0.0	0.0	1.1	1.5	3.0	4.0
Percentile (75)	1.3	1.0	1.2	3.0	3.0	5.9	6.7
Percentile (25)	3.0	3.0	2.0	1.5	0.0	0.0	0.0
Percentile (50) of [General Play Turnover] per60	5.0	4.5	4.0	3.0	1.0	1.0	0.0
Percentile (75)	7.5	7.0	6.0	5.2	2.7	2.3	1.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Bad Pass] per60	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	1.5	0.0	1.9	1.5	0.0	0.0	0.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Intercept Pass Thrown] per60	1.0	0.0	1.3	1.2	0.0	0.0	0.0
Percentile (75)	2.0	0.0	2.9	2.6	1.1	1.0	0.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Centre Pass Break] per60	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Bad Hands] per60	0.0	1.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	1.3	2.0	1.1	0.0	0.0	0.0	0.0
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50) of [Offside] per60	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	0.0	0.0	0.0	1.0	1.0	0.0	0.0

5.3.1 Case study: Coach visualization feedback

After consultation with a netball coach, Table 11, Figure 11, 12, 13 and 14 were constructed using filters and categories in the Tableau™ software. These plots and table allow visualisation of the position units, i.e. "shooters", within different levels of competition. Shooters are the positions GS and GA, "midcourt" are the position WA, C, and WD, while "defenders" are the positions GD and GK. Feedback was given by the coach about how important the need for visualisations that were easy to understand and how easy this would be to explain to players and other coaches.

As presented in Table 11, the defender unit players in international and elite club are consistently lower in the number of "Gains" (3.9 vs 4.5 per 60mins), "Intercepts" (1.3 vs 1.9 per 60 mins), Deflection Gain (1.1 vs 2.4 per 60 mins), and Pickups (1.0 vs 1.9 per 60 mins) than school and often national league. The coach expressed the desire for a table of this type to be converted into a more visual representation of the data. Box and whisker plots are a visualization method for looking at the distribution of population data.

Some notable observations from box and whisker plot visualization (Figure 11) of the normative data include median, or 50th percentile, "goals missed" being lower at international level when compared to school (4 vs 6 per 60 mins), "Goal Shoot %" better at higher levels (85.3 vs 77.8%). Box and whisker plots were created for General Play Attack measures with midcourt and shooters separated (Figure 12). This was because their roles for these measures influence the values that might be expected, e.g. Shooters will do less feeding as they also need to be available to receive the feed in-circle. Midcourt players "Centre Pass Receives" were higher in internationals than schools (6 vs 3.5 per 60 mins), and also for "Goal Assists" (9.6 vs 5.9 per 60 mins). Additionally, "Feeds" and "Feeds with Attempt" where consistently higher for international and elite club level then school. This implies that there are more pre-shot actions at higher levels, yet only a small increase in "Goals" (24 vs 22.5 per 60 mins). This could indicate that there is better execution of these skills at international and elite club levels, which is reflected in the Defensive players data. The coach felt that these plots where not easy to understand and concept of quartiles was one, they needed more education around.

In Figure 13, Distribution Plots were used to display error measures with colours, shaded from red for less desirable performance, to green for preferable performance, shading the different 25th, 50th and 75th percentile boundaries. Although not distinctly measurable there appears to be an overall visual trend for general play turnovers (GPT) to increase as performance went down the levels i.e. from international down to schools. It was however clear that shooters had high GPTs then midcourters.

Table 11.

Comparison between various level of netballs normative data for Possession Gains measures.

	Possession Gains											
	Defenders				Mid-courters				Shooters			
	International	Elite Club	National League	Schools	International	Elite Club	National League	Schools	International	Elite Club	National League	Schools
	Gains (per 60 mins)											
Percentile (25)	1.9	2.0	2.6	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50)	3.9	4.0	5.0	4.5	0.0	1.0	1.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	6.0	6.7	7.7	7.5	2.0	2.0	2.0	2.0	0.0	0.0	0.0	0.0
	Intercept (per 60 mins)											
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50)	1.3	1.9	2.0	1.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	3.0	3.0	3.9	4.4	1.2	1.0	1.9	1.5	0.0	0.0	0.0	0.0
	Deflection Gain (per 60 mins)											
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50)	1.1	1.1	2.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	3.0	3.2	5.0	4.5	1.0	1.0	2.0	2.0	0.0	0.0	0.0	0.0
	Deflection No Gain (per 60 mins)											
Percentile (25)	1.0	1.9	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50)	2.6	3.0	3.0	3.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (75)	4.0	5.0	4.8	4.5	1.9	2.0	2.0	3.0	0.0	0.0	0.0	0.0
	Pick Up (per 60 mins)											
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50)	1.0	1.0	1.3	1.5	1.2	1.3	1.4	1.5	0.0	1.0	0.0	0.0
Percentile (75)	2.0	2.0	2.6	2.9	2.6	3.0	3.0	3.0	1.7	2.0	1.5	1.5
	Rebound (per 60 mins)											
Percentile (25)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Percentile (50)	1.0	1.0	1.9	1.5	0.0	0.0	0.0	0.0	0.0	1.0	1.4	1.5
Percentile (75)	2.0	2.1	3.2	3.0	0.0	0.0	0.0	0.0	2.0	2.8	3.1	3.0

Figure 11.

Box and Whisker Plots of various level of netballs normative data for Goal Scoring measures.

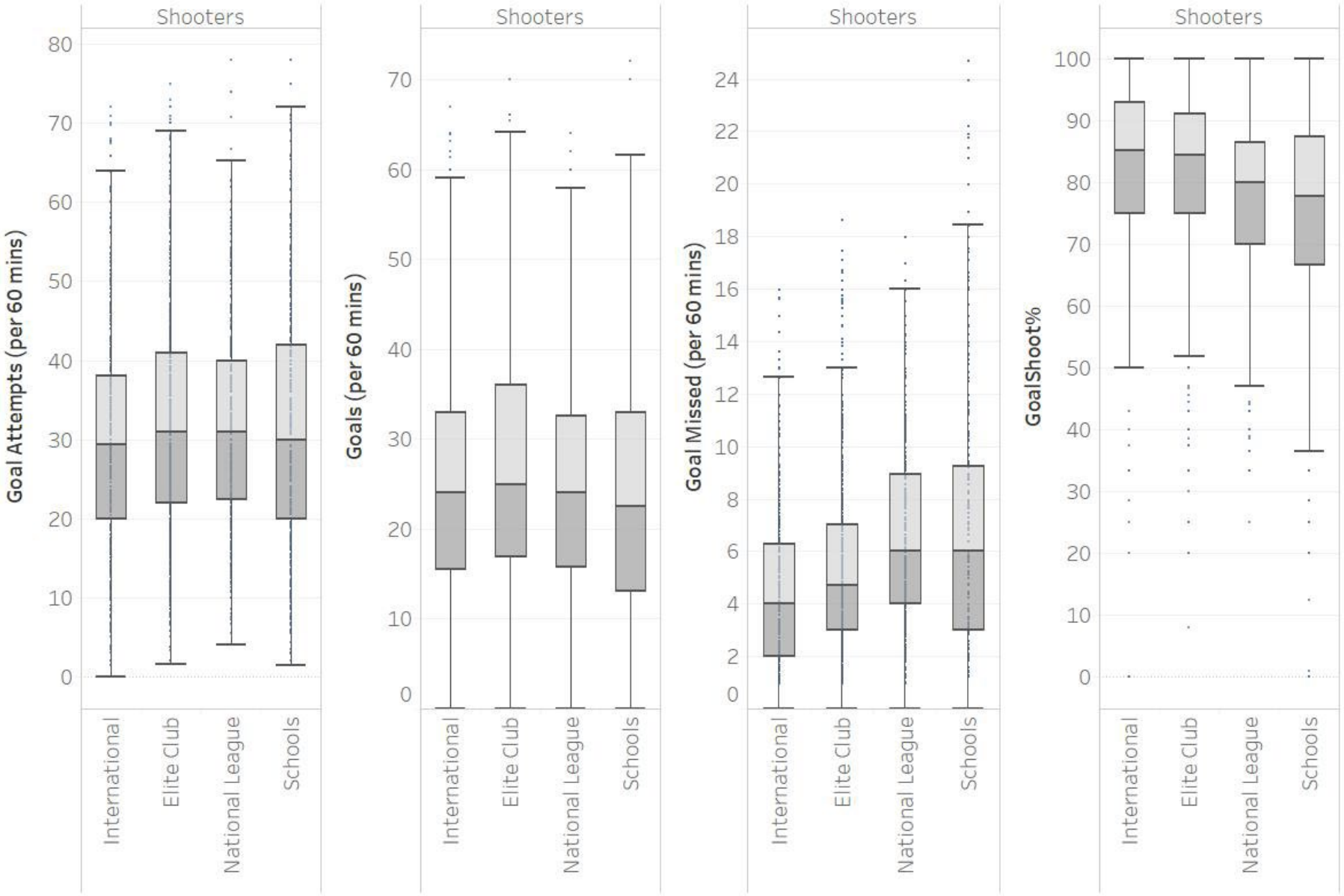


Figure 12.

Box and Whisker Plots of various level of netballs normative data for General Play Attack measures.

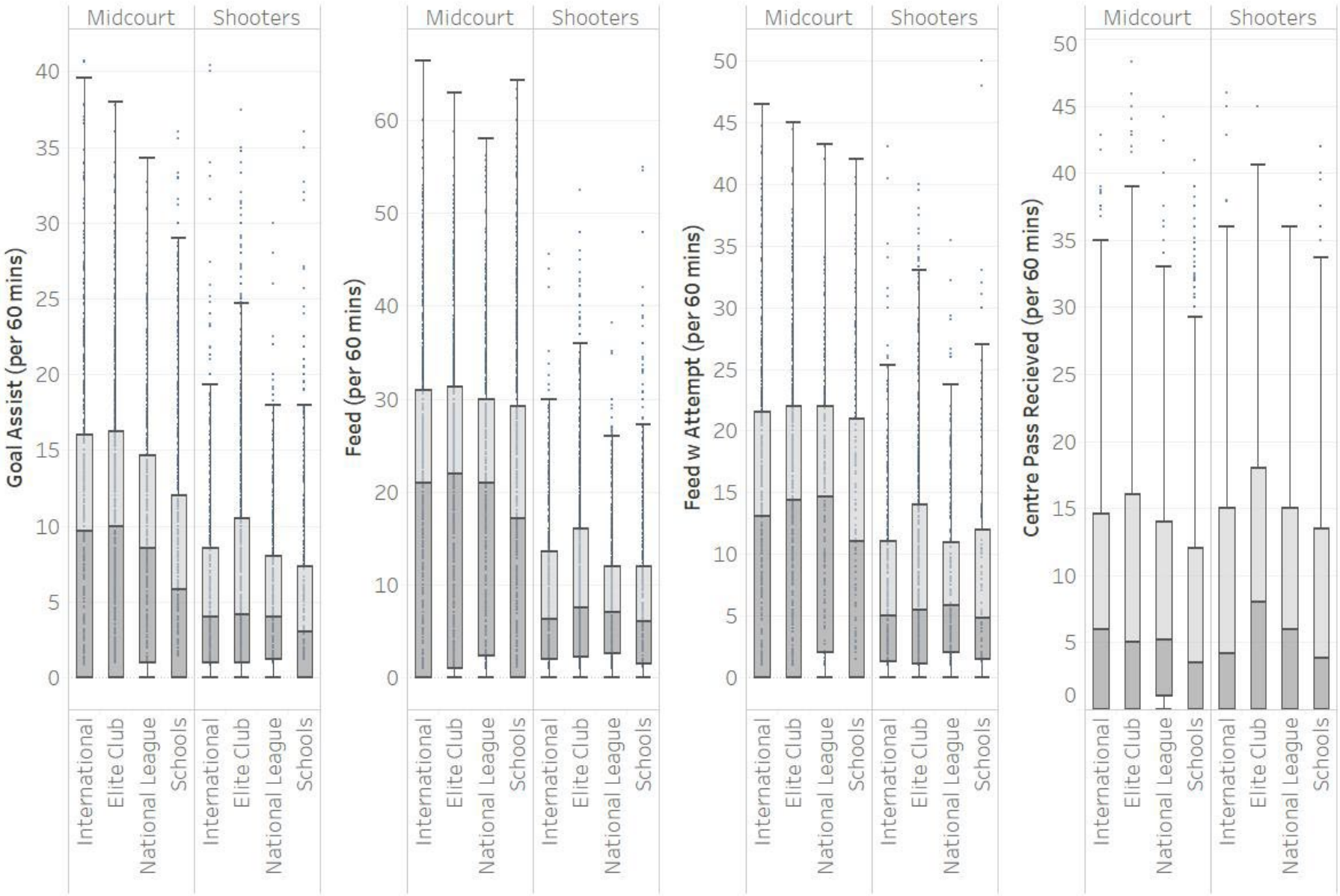
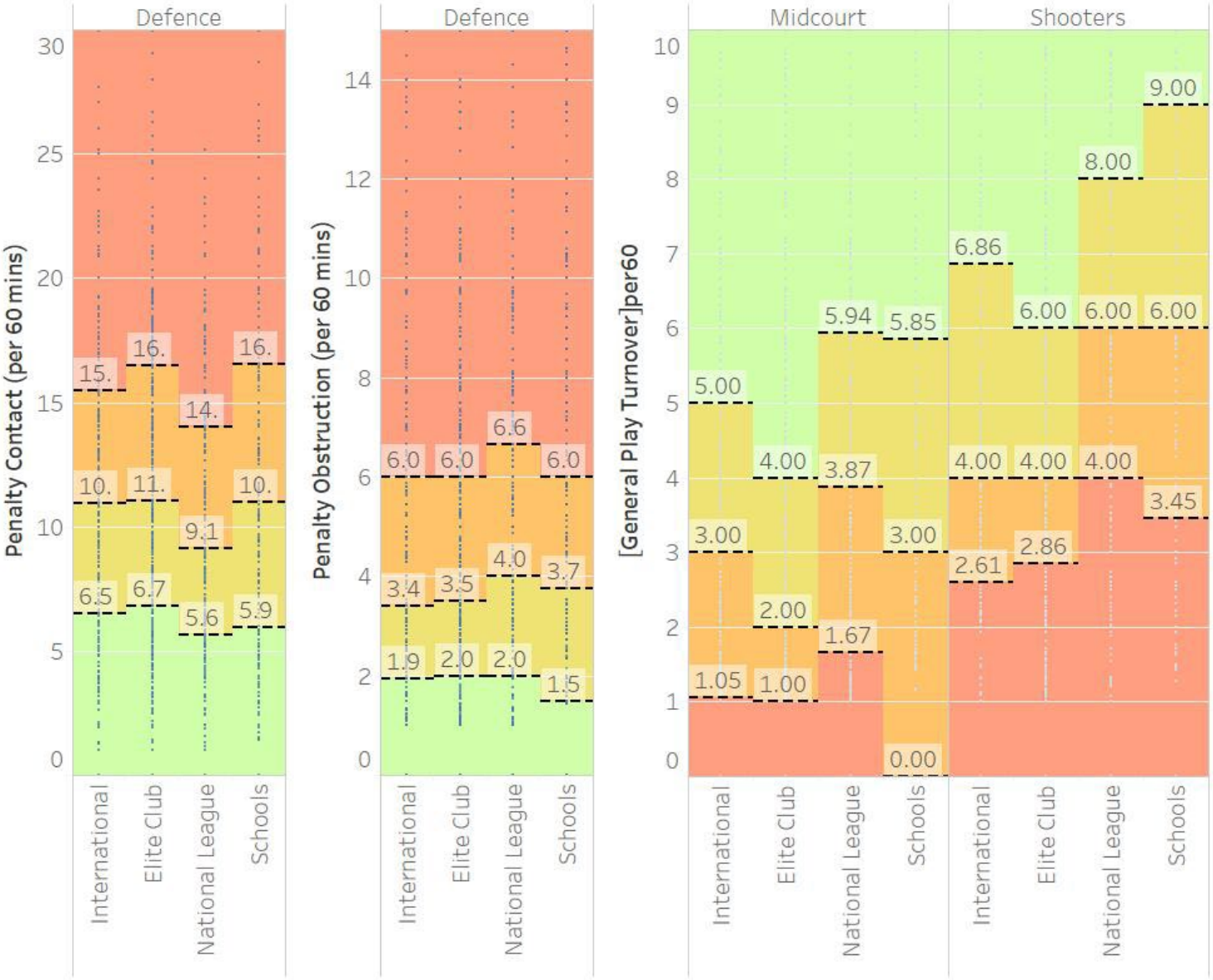


Figure 13.

Distribution Plots of various level of netballs normative data for General Play Attack measures.



Finally, the coach communicated that simple table with a benchmark number be produced (Figure 14). The concept of the 75th percentile, or upper quartile, was agreed upon as a benchmark by the coach, as is it represented an attainable number for players if they improved. This is somewhat subjective, however aligned with the coach's beliefs giving a level of ecological validity.

Figure 14.

Example coaches normative table showing 75th percentile values for values for performance indicators.

		Goal	Goal Attempts	Goal Missed	GoalShoot%	Goal Assist	Feed	Feed w Attempt	Centre Pass Recieved	Rebound	General Play Turnover
GA	International	25	31	6	91	13	19	16	20	1	7
	Elite Club	23	28	7	86	15	22	19	23	1	7
	National League	22	29	9	80	12	18	16	22	2	8
	Schools	23	31	9	85	11	18	17	20	2	9

5.4 Discussion

5.4.1 Workflow and Automation

This paper's purpose was to develop and explain an automated process for presenting and manipulating normative data for netball matches across various populations. It was also proposed that this process should allow filtering of data dimensions (independent variables) and measures (dependent variables). As a result, two data tables (Table 10 and Table 11) were presented to not only demonstrate the functionality of the process, but also to provide insights into some of the characteristics of the population data. The process followed several stages including data acquisition, data preparation and cleaning, and upload and consumption.

The data acquisition stage of this process did use some automation via the XML file upload, however, to create a large historical data set manual data scraping was required. Future data acquisition into this process could however be fully automated.

5.4.2 Normative data

The results of Table 10 showed percentiles, or quartile ranges, for 22 performance indicators identified by Croft and Spencer (2021) as being highly valid to coach and player decision making. Table 10's results help describe the characteristics of each position with some being more specialized. Examples of this were GS's and GK players having the highest values for shooting and defensive turnovers, respectively. GA, WA, WD, GD were represented in a greater number of measures; i.e. Goal Assists, Feeds, Feeds w Attempt, Centre Pass Receives etc., then other positions, meaning they had

involvement in more aspects of the game. The C position was the most generalized with representation in most measures except those within the circle. This representation tended to be as a low to moderate frequency for all measures.

With the normative data measures presented a coach could evaluate their team's performance for strengths and weaknesses by looking at performances that are either above or below the median (50% percentile). This could assist in decision making around areas for improvement, or strengths that could be played to.

5.4.3 Case study

To demonstrate the flexibility of the tool and to assess its face (Middleton, 2020) and ecological validity (Ransdell, 1993) (Davids, 1988), consultation was undertaken with an elite netball coach to determine what groupings of positions, levels of competition and variables would be useful when applying this tool to coaching. Table 2 was constructed to demonstrate the differences on normative values for international, elite club, national club, and school levels. Player positions were combined into "units" so that players could understand the broader requirements of different areas of the court.

With Table 11, a coach or selector could consider player pathways and relative strengths of developing players. An example of this could that if a GA player have a goal shooting volume similar to that of a GS, they may be able to change positions and still meet the requirements of the more specialize GS position.

Some difference between levels were observed in the case study (Table 11). These described the higher levels of competition (international and elite club) as having better goal accuracy, more pre-circle movements and less turnovers conceded. This could imply that as players develop and move to the higher levels, they tend to be more accurate with their ball handling skill, play faster with more actions and have better ball retention.

6.0 Conclusion

The overall objectives of this paper were to not only produce a live tool that could display data to coaches during matches but also demonstrate this with a normative dataset, that presents useful contextualisation to netball coaches and players about their performance statistics. Interesting insights into the normal values for netball player performances were found, with differences between specialised and more generalised positions observed. Finally, through consultation with expert coaches, a simple visual representation of benchmark statistics was developed. This simple table (Figure 14) was seen to have the best ecological validity for the coach.

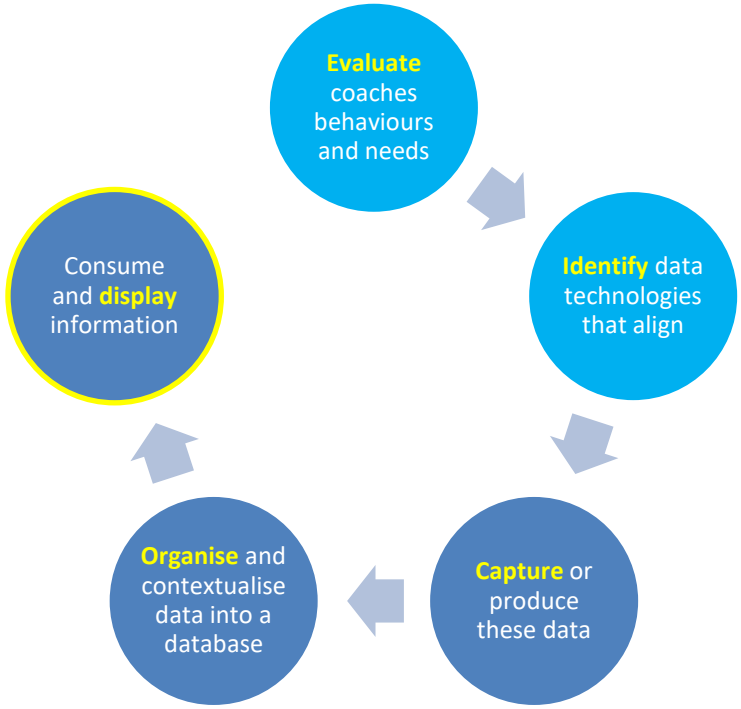
Prelude Chapter 6

The previous chapter explored and developed an approach for collecting, organising and normalising XML and CSV data files, from computerised analysis tools. These tools included Hudl-Sportscode™ and data scraped from public domain websites stored in CSV files. This data was stored in an AWS™ bucket and algorithms transformed it within data pipeline and it was then stored in a cloud database. The database was connected to data visualisation tool Tableau™ where it was displayed as normative data for coaches to evaluate in a case study.

The case study was an opportunity to gain feedback from two expert coaches centred on the usability, understandability, and overall ecological validity of various data presentation formats. The feedback from the coaches' showed their preference was for simple numerical representations of the netball population data. Specifically using the 75th percentile or upper quartile value as targets or benchmarks for players to aim, for each performance indicator. These developments contributed to the final phase, 'system design', of the systems design approach, where the automation of data collection, data extraction, analysis and representation was defined.

Figure 15.

A framework for presenting data to coaches during (dark blue) sports matches (phase five).



The final stage of the systems design framework was 'system redevelopment', where functionality and reliability issues were assessed and remedied. Chapter 6 observed coaches engaging

with a statistical tool, which incorporated the information sources identified in chapter 4. Chapter 6 adopted a similar methodology to chapter 3, 'descriptive observation', as this allowed comparison between findings.

The purpose of chapter 6 was to observe coaches in-game behaviours, while using a statistics tool, i.e., recording which information they access, and for how long. It gave insights into the function and design that was required. Once understood, this information was used in the final design of a tool, that helped guide coaches through 'large' or 'big data'.

Chapter 6.0 Coach engagement with live netball statistical information.

This chapter comprises the following manuscript (5) in preparation for submission to The Journal of Performance Analysis in Sport.

Croft, H. G., & Spencer, K. (2021b). Coach engagement with live netball statistical information. *(In preparation for submission) International Journal of Performance Analysis in Sport.*

Author contributions: Croft, H (85%), Spencer, K (10%), Robertson, S (5%)

6.1 Introduction

Netball is a dynamic, fast paced, team sport where coaches and players are required to make decisions quickly during a match. Live video replays are possible via modern streaming technologies, however due to the continuous and high-speed nature of the game, coaches have limited time to look at a video replay device. Statistical data is utilised by coaches as it provides insights into the current state of the game and can be used to support decision making. Public domain websites (Champion Data, 2020) broadcast live statistical information about the state of netball match in many international, elite club, national and school competitions. The live information that this website provides includes scoring, team statistics, individual player performance, opposition player performance, team's performance comparisons, shooting performance, top 5 player rankings, score sequences, and player interchange. These categories align with many of the tactical themes used by coaches as observed in Croft et al. (2020) research into coach discussions.

6.1.1 Literature Review

Coach behaviour is not well understood, and due to this lends itself to descriptive research (Cushion et al., 2012). Croft et al. (2020) explored the spoken behaviours of coaches, identifying themes around tactics and strategies. Subsequent work has looked to align and contextualise information with these themes (Croft & Spencer, 2021). When this statistical information is presented to coaches it then becomes necessary to observe their interactions with it (Edwards, 2017) to help understand its value and therefore validity.

Observational field research is an effective tool for understanding behaviour (Brown, 2020) as it provides a depth of information. This gives observation strong ecological validity and is a good method for understanding coach behaviour and subsequently what information may be important to them during a sports match. Previous research (Cushion et al., 2012) has developed tools for

instrumented coach observation. The use of computers and software, e.g., Hudl-Sportcode™, have enabled more complex analysis of behaviours and prove to have better reliability and validity than pencil and paper notational analysis.

Technology has been described as both a source of knowledge and a resource for coaches, however it is critically unexplored in its understanding of coach learning (Cushion & Townsend, 2018). Observations of coaches' interactions with information technology has the potential to add to the evidence base and improve our understanding of the role it can play. Aligning these behaviours with the spoken tactical themes identified in previous research (Croft et al., 2020; Zetou et al, 2011) will give insights into whether a coaches spoken communications have good congruency with decision making.

Recent research has looked at the use of statistical information by coaches in providing feedback to athletes and players (Nicholls et al., 2018; Middlemas et al., 2017). However, there is a lack of research considering coaches interactions with the information during a sports match. This is important for our understanding of what and how information should be presented to netball coaches, during matches, on a data visualisation platform so that the coaches use the information to inform their decisions.

The purpose of this study was to describe a small group of netball coaches use and perceptions of statistical information during game-time, that is aligned with their communication of themes within the game which are important to their decision making (Croft et al., 2020). This will be achieved through computerised observation of elite coach interactions with statistics during live netball matches.

6.2 Methodology

6.2.1 Participants

Six elite netball coaches, who currently coach in either the National Netball League (NNL) or ANZ Premiership (ANZP) where recruited voluntarily via email. Participants approached the researchers to participate in the study after see information about it via their high-performance manager. Their involvement was kept confidential from all parties. The inclusion criterion of coaching at NNL or ANZP was set, as it aligned with the participants from the Croft et al. (2020) study and allowed comparison.

6.2.2 Equipment

Each coach was provided with an Apple iPad™ device, which was connected to the internet via a local high speed Wi-Fi network. The iPad was configured to show a live statistics website (https://mc.championdata.com/anz_premiership/index.html) on the right 2/3rds of the screen and

the devices internal camera on the left 1/3rd of the screen. The camera's perspective was of the coach enabling identification of when they looked at the screen and for how long. The website had several pages of statistics, which align with those identified as important. A screen recording was taken of the of the iPad™, identifying all interactions with the device and the participant's visual gaze.

Hudl-Sportscode™ notational analysis software was used to code the website pages, and therefore category of statistic, that coaches looked at during a netball match, and the amount of time looked at each. Table 12 shows the statistics screen name with what information was displayed and an overall description for the page.

Table 12.

Content of statistics screens available to coaches, for navigation, during their live netball match.

Statistics screen name	Information displayed	Description
Scoring	Possession source to goal frequency and success Score worm graphic Last score and misses Lead and scoring streak	Statistics that provide information about trends and current state of scoring events
Team Stats	Time in Possession Gains, Intercept, Deflection Resulting in a Gain Penalty – Contact & Obstruction Possession Changes General Play Turnover Shooting zones and success	Combine actions of the team that are either attacking or defensive and reflect losing or maintaining possession of the ball
Player Grid	Player time in the game Frequency of action variables in the match, including: Goal, goal attempts, goal shoot%, goal assist, feed, feed w attempt, gains, intercept, intercept pass thrown, deflection gain, deflection no gain, rebound, centre pass received, pick up, penalty contact, penalty obstruction, general play turnover, bad pass, bad hands, offside, centre pass break	Individual actions for own team that are either positive, scoring, possession gaining, or errors that lead to losses of possession of penalties.
Opposition Player Grid	Same as player grid but for opposition players	Individual actions for opposition team that are either positive, scoring, possession gaining, or errors that lead to losses of possession of penalties.
Both Teams Player Grids	Goal, goal attempts, goal shoot%, feed, rebound, centre pass received, penalty, general play turnover	An abbreviated player grid allowing for comparison of both teams' individual players
Goal Shooter	Current player goal shoots, attempts and successful%	A view of the shooting performances of the 4 players on

		court at the time of viewing, for both teams.
Top 5	Individual player counts, for: Feed, goal assists, centre pass received, penalty, general play turnover	A view showing the 5 players on court for highest counts for 5 individual actions.
Score Flow	Shooting attempt outcome, ordered by game time, and current score.	A temporal list of scoring and non-scoring attempts for both teams
I/Change	Individual player substitutions, positions, and time	A temporal record of player substitutions with position and game time these are made

6.2.3 Procedures

As the coaches watch the live NNL™ or ANZP™ netball match, they sat with the iPad™ statistics device on their lap, and were instructed to look at the screen when information about the game was of interest to them. Before the match began the 'iPad's' 'screen record' feature was activated so that all information present on screen was recorded into an mp4 format video file. This meant that the following information was captured; the website subpages selected, the zooming they did on the device, what information they looked at, and via the web camera screen display their visual gaze frequency and length.

6.2.4 Data analysis

An inductive data analysis approach (Elo & Kyngäs, 2008; Neuendorf, 2002) was adopted, where open coding, creation of categories and abstraction was implemented. This approach aligns this study with that conducted by Croft, Spencer & Robertson (2020), conducted on the same population of participants, allowing for comparison of results. Once the video data was captured the files were loaded individually into Hudl-Sportcode™, where each of the coaches gazes where coded for length and group of statistics viewed. This enabled data to be recorded for the frequency, type and duration of statistics viewed.

6.3 Results

As observed in Table 13, coaches spent the most time looking at 'Team Stats' information, in both total duration and frequency respectively (167.8s & 30.5 per match). However, the amount of time they looked at this information, per view, was one of the shortest (5.5s per view) of all categories, implying that they glanced often. This was followed by the 'Scoring' (133.3s per match) and 'Both Teams' (127.8s per match), however 'Scoring' looked at less often (16 per match) for longer (8.4s vs. 6.6s per view). Additionally, the 'Player Grid' and 'Opposition Player Grid' screens was looked at for the longest duration (11s and 10.6s per view) however only 10.8 and 3 times per match, per coach.

This could be due to the large volume on information available in this category. 'Opposition Player Grid', 'Goal Shooter', 'Top 5', 'Score Flow' and 'I/Change' were looked at the least (31.9s, 10.8s, 33.0s, 19.0s, 3.9s per match) respectively for comparatively low frequencies.

Overall coaches tended to look at the statistics screen for 7.8s at a time, averaging 71.9 views per game with a total viewing time of 646.7s per game. This equates to approximately 11% of the game time.

Table 13.

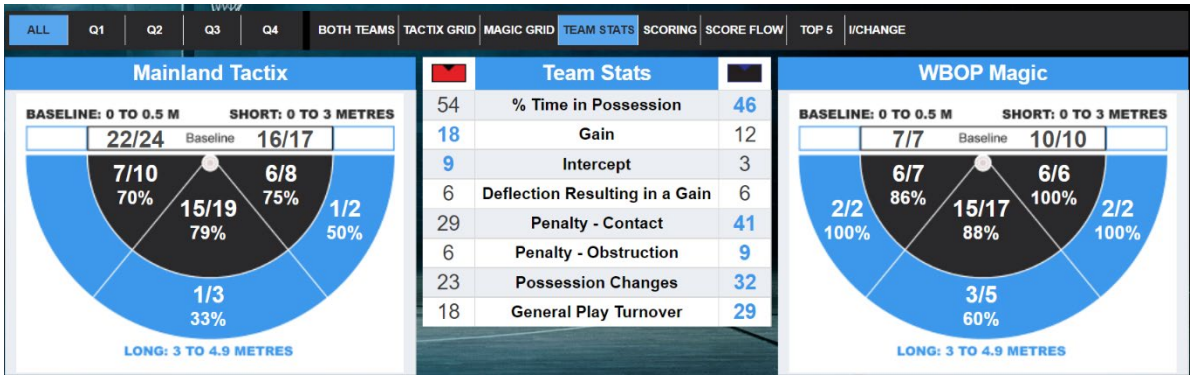
Summary of the statistical information viewed, for frequency and duration, by coaches during live netball matches.

Statistical categories	Time viewing (s)		Frequency viewing		Time per view (s)
	per coach	total	per coach	total	per coach
Scoring	133.3	799.9	16.0	96	8.3
Team Stats	167.8	1007	30.5	183	5.5
Player Grid	119.2	715	10.8	65	11.0
Opposition Player Grid	31.9	191.6	3	18	10.6
Both Teams	127.8	766.6	19.5	117	6.6
Goal Shooter	10.8	64.8	1.7	10	6.5
Top 5	33.0	198.1	3.3	20	9.9
Score Flow	19.0	114	4.5	27	4.2
I/Change	3.9	23.3	0.5	3	6.6
Average	71.9s		10.0		7.8s
Total	646.7s		89.8		

The most significant of these results is that coaches favoured the 'Team Stats' screen which is a screen that (Figure 16) presents measures of possession gains and losses, in the statistics; 'Time in Possession', 'Gains', 'Intercept', 'Deflection Resulting in a Gain', 'Penalty – Contact & Obstruction', 'Possession Changes', and 'General Play Turnover'. It does also display shooting zones for each team on both sides, however it was observed that this centre 'Team Stats' block was zoomed into for most of the time.

Figure 16.

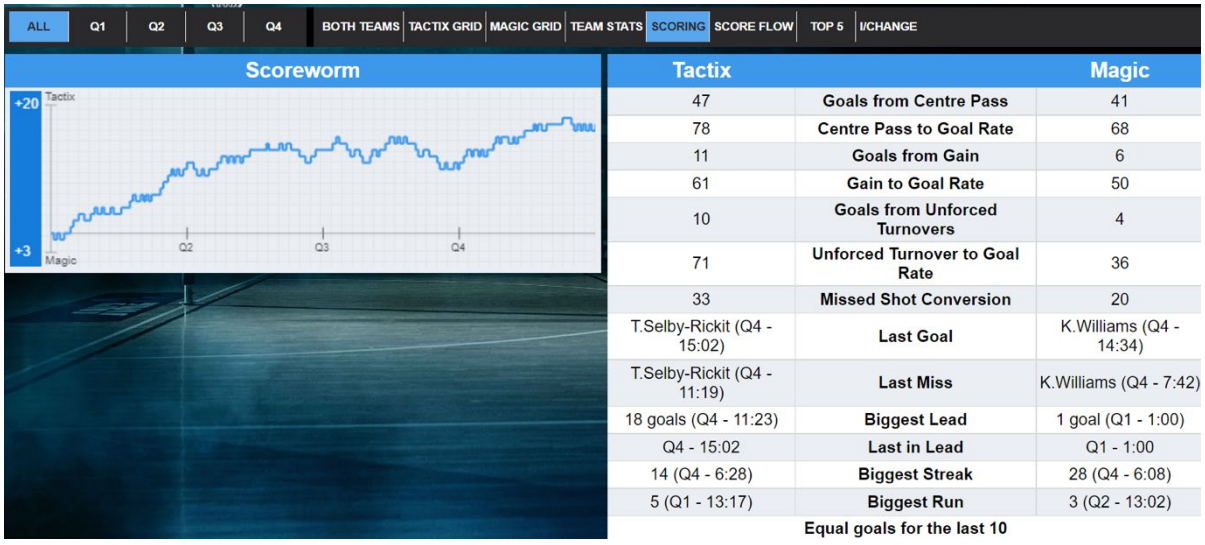
Screen capture of the Team Stats category that coaches viewed information off during live netball games.



Another category that the coaches tended to focus on was 'Scoring' (Figure 17). The 'Scoring' category shared similarities to the most popular category 'Team Stats' in that it displayed team outcome information related to success of the teams in scoring goals from 'Centre Passes', 'Gains' and 'Unforced Turnovers'. This category focused on measures more directly related to goal scoring, than the 'Team Stats', and it provided a visualisation of points difference between teams across the match time.

Figure 17.

Screen capture of the Scoring category that coaches viewed information off during live netball games.

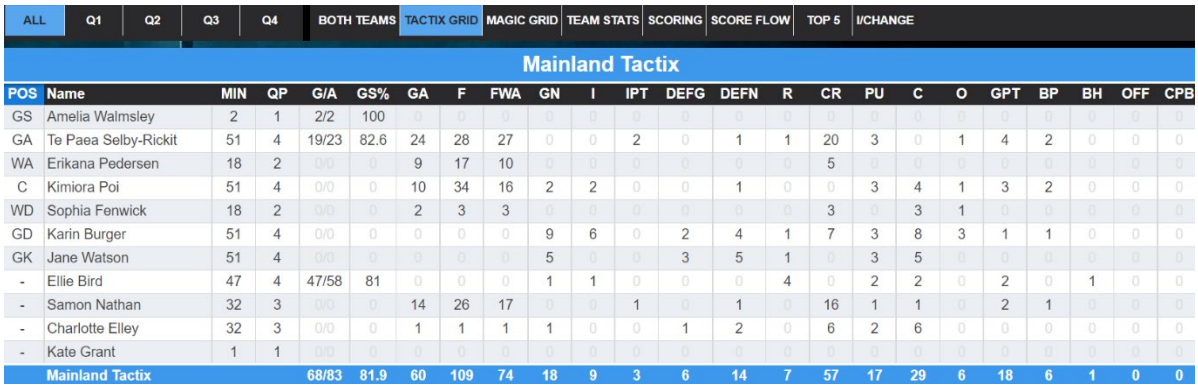


The other categories of information most engaged with by coaches were the 'Player Grid' and 'Both Teams' grids (Figure 18 and Figure 19). Both categories display similar information about the individual performances of players in the coach's team and then also a comparison of individual players for both teams. If it is considered that these categories are different representations of the same group of individual statistics, the total time coaches viewed these is larger (139.7s and 116.2s per game) then the single largest category (190.1s per game) 'Team Stats'. It was observed that some

coaches moved between the 'Team Stats' or 'Scoring' categories, and the individual player stats categories, 'Player, Opposition and Both Team' grids, which may imply there are different levels of information between them i.e., the 'Team Stats' and 'Scoring' statistics are made up of combined individual statistics.

Figure 18.

Screen capture of the Teams category that coaches viewed information off during live netball games.



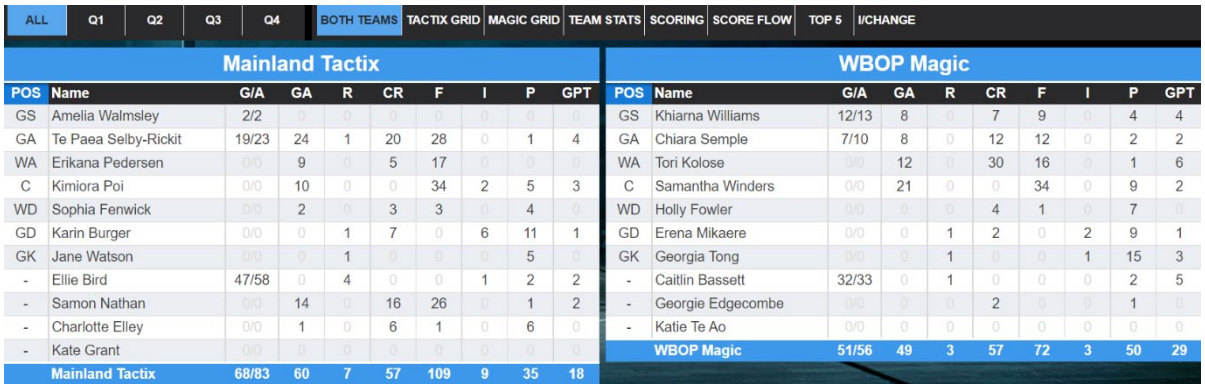
ALL | Q1 | Q2 | Q3 | Q4 | BOTH TEAMS | TACTIX GRID | MAGIC GRID | TEAM STATS | SCORING | SCORE FLOW | TOP 5 | I/CHANGE

Mainland Tactix

POS	Name	MIN	QP	G/A	GS%	GA	F	FWA	GN	I	IPT	DEFG	DEFN	R	CR	PU	C	O	GPT	BP	BH	OFF	CPB
GS	Amelia Walmsley	2	1	2/2	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GA	Te Paea Selby-Rickit	51	4	19/23	82.6	24	28	27	0	0	2	0	1	1	20	3	0	1	4	2	0	0	0
WA	Erikana Pedersen	18	2	0/0	0	9	17	10	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0
C	Kimiora Poi	51	4	0/0	0	10	34	16	2	2	0	0	1	0	0	3	4	1	3	2	0	0	0
WD	Sophia Fenwick	18	2	0/0	0	2	3	3	0	0	0	0	0	0	3	0	3	1	0	0	0	0	0
GD	Karin Burger	51	4	0/0	0	0	0	0	9	6	0	2	4	1	7	3	8	3	1	1	0	0	0
GK	Jane Watson	51	4	0/0	0	0	0	0	5	0	0	3	5	1	0	3	5	0	0	0	0	0	0
-	Ellie Bird	47	4	47/58	81	0	0	0	1	1	0	0	0	4	0	2	2	0	2	0	1	0	0
-	Samon Nathan	32	3	0/0	0	14	26	17	0	0	1	0	1	0	16	1	1	0	2	1	0	0	0
-	Charlotte Elley	32	3	0/0	0	1	1	1	1	0	0	1	2	0	6	2	6	0	0	0	0	0	0
-	Kate Grant	1	1	0/0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Mainland Tactix			**68/83**	**81.9**	**60**	**109**	**74**	**18**	**9**	**3**	**6**	**14**	**7**	**57**	**17**	**29**	**6**	**18**	**6**	**1**	**0**	**0**

Figure 19.

Screen capture of the Both Teams category that coaches viewed information off during live netball games.



ALL | Q1 | Q2 | Q3 | Q4 | BOTH TEAMS | TACTIX GRID | MAGIC GRID | TEAM STATS | SCORING | SCORE FLOW | TOP 5 | I/CHANGE

Mainland Tactix

POS	Name	G/A	GA	R	CR	F	I	P	GPT
GS	Amelia Walmsley	2/2	0	0	0	0	0	0	0
GA	Te Paea Selby-Rickit	19/23	24	1	20	28	0	1	4
WA	Erikana Pedersen	0/0	9	0	5	17	0	0	0
C	Kimiora Poi	0/0	10	0	0	34	2	5	3
WD	Sophia Fenwick	0/0	2	0	3	3	0	4	0
GD	Karin Burger	0/0	0	1	7	0	6	11	1
GK	Jane Watson	0/0	0	1	0	0	0	5	0
-	Ellie Bird	47/58	0	4	0	0	1	2	2
-	Samon Nathan	0/0	14	0	16	26	0	1	2
-	Charlotte Elley	0/0	1	0	6	1	0	6	0
-	Kate Grant	0/0	0	0	0	0	0	0	0
	Mainland Tactix	**68/83**	**60**	**7**	**57**	**109**	**9**	**35**	**18**

WBOP Magic

POS	Name	G/A	GA	R	CR	F	I	P	GPT
GS	Khiarna Williams	12/13	8	0	7	9	0	4	4
GA	Chiara Semple	7/10	8	0	12	12	0	2	2
WA	Tori Kolose	0/0	12	0	30	16	0	1	6
C	Samantha Winders	0/0	21	0	0	34	0	9	2
WD	Holly Fowler	0/0	0	0	4	1	0	7	0
GD	Erena Mikaere	0/0	0	1	2	0	2	9	1
GK	Georgia Tong	0/0	0	1	0	0	1	15	3
-	Caitlin Bassett	32/33	0	1	0	0	0	2	5
-	Georgie Edgecombe	0/0	0	0	2	0	0	1	0
-	Katie Te Ao	0/0	0	0	0	0	0	0	0
	WBOP Magic	**51/56**	**49**	**3**	**57**	**72**	**3**	**50**	**29**

Figures 20 and 21 demonstrate that in the most categories of statistics, coaches viewed the three individual player categories in the most consistent manner. These were 'Player Grid', Opposition Player Grid', and 'Both Teams', where in all 3 categories, at least 5 of the 6 coaches demonstrated similar behaviours.

For the team-based statistical categories 'Scoring' and 'Team Stats', there was extremely high engagement in both duration and frequency for some coaches, however many other coaches had low or no engagement. This may reflect that, of the 6 coaches there, were differing levels of understanding of this type of statistic and therefore a difference in perception of their value.

Figure 20.

Column chart demonstrating the individual variation, in frequency, between participants engagement with categories of statistics.

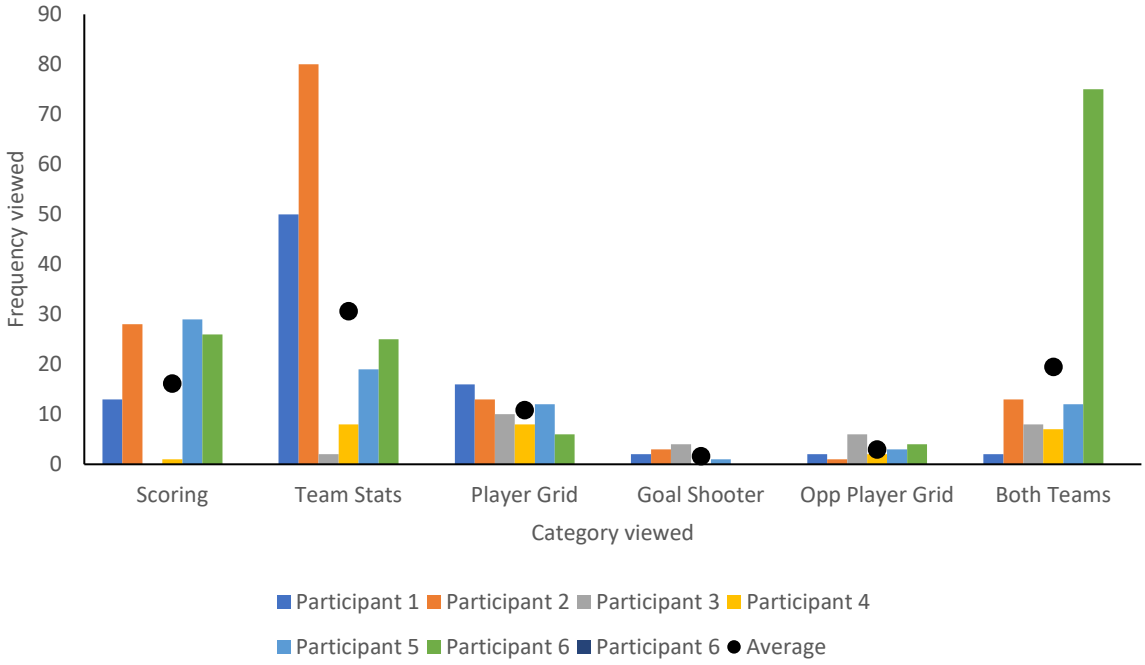
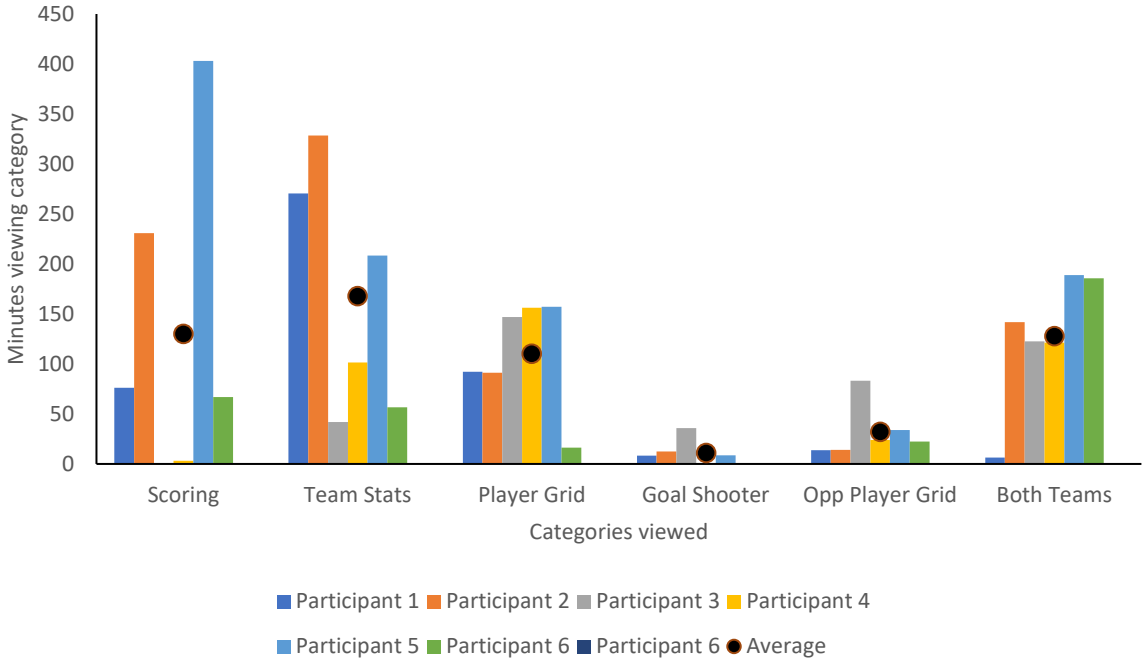


Figure 21.

Column chart demonstrating the individual variation, in duration, between participants engagement with categories of statistics.



Not only was the frequency of engagement insightful, but also the patterns of movement

between categories or screens. Many coaches tended to stay on a screen such as 'Team Stats' or 'Scoring', glancing frequently, then moving to another category before returning to this screen. Table 14 show the most popular sequence of categories that all coaches moved between were from 'Scoring' to 'Team Stats', 'Player Grid' to 'Team Stats' and from 'Team Stats' back to 'Player Grid'.

Table 14.

Frequency of movements between statistical categories, for all coach, during live netball matches

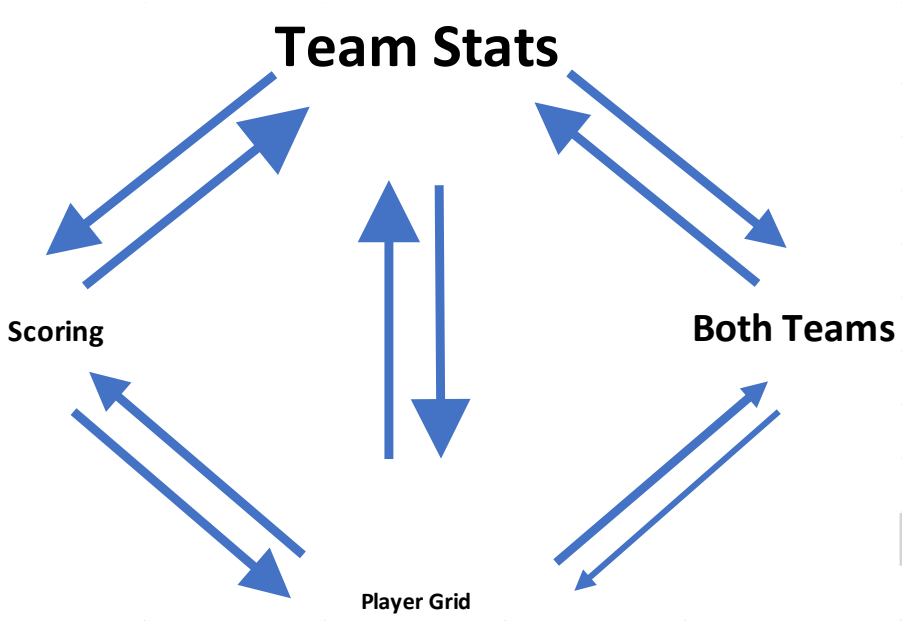
Start screen		End screen	Frequency
Scoring	-->	Team Stats	26
Player Grid	-->	Team Stats	24
Team Stats	-->	Player Grid	23
Team Stats	-->	Scoring	20
Both Teams	-->	Team Stats	20
Scoring	-->	Player Grid	17
Team Stats	-->	Both Teams	15
Player Grid	-->	Scoring	15
Player Grid	-->	Both Teams	10
Both Teams	-->	Player Grid	7

To better visualise these movements, a network graph (Figure 22) displays the categories with arrows and font size, weighted for frequency and movement between categories, in the direction the coach moved. This shows that all arrows pointing towards 'Team Stats' are heavily weighted, implying that a return to this category after being in another, was the most common movement. Movement away from this category was also high, which could be due to it being the most common that coaches looked at.

Font size was set at 10 percent of the category frequency displayed in Table 13, while arrow thickness (point size) was set at 25 percent of the frequency of the combination displayed in Table 14.

Figure 22.

Network graph of the coach movements between statistical categories, weighted for font size and arrow thickness to scale.


Team Stats
Scoring
Both Teams
Player Grid

6.4 Discussion

Observing coach interactions (Brown, 2020) with statistical information, during netball matches has provided useful insights into their behaviours. Through computerised analysis (Cushion et al., 2012), frequencies, and sequences of engagement, with categories of information has shown, netball coaches tend to prefer 'Team Stats' however this is somewhat inconsistent between participants.

Coaches engaged more consistently with individual player statistics ('Player Grid', 'Opposition Player Grid' and 'Both Teams'), which make up the 'Team Stats' and 'Scoring'. This could be explained in the complexity of the statistical information, with the 'Team Stats' and 'Scoring' categories being constructed from the combined actions of players and possession outcomes. Nash and Collins (2006) described coaches as avoiding the use of too many cognitive resources for unfamiliar situations and tending to rely on knowledge from a similar situation. This could explain why we see coaches favour more familiar or simplistic statistics during games, as there are already high cognitive demands on them.

The analysis of movement between categories provides some interesting insights into coaches use of this information. The most common movements, as demonstrated in Table 14 and Figure 22, were from 'Team Stats' and other categories, often returning back to 'Team Stats'. One of the most common of these patterns was to move from 'Team Stats', a group of combined team statistics, to a category that represented individual player performance, i.e., 'Player Grid'. The 'Team Stats' are a combination on individual player actions, or events, and this movement may indicate that coaches navigate layers of statistics to "dig down" for causes of the measures. This layer of information follows a similar framework as those used in grounded theory (Weed, 2009) and subsequently content analysis (Krippendorff, 2004; Neuendorf, 2002), where dimensions are comprised of higher order themes and then lower order themes.

Probably the most interesting and informative findings are those produced when the results of this study were compared with those themes identified by Croft et al. (2020). It was found that 58.3% of these themes were represented in the categories viewed by the coaches in this study. Coincidentally the largest category, 'Team Stats', was the most representative of the spoken themes with 46.6% similarity. Although a qualitative representation of frequency, this reinforces the concept that there is a relationship between spoken behaviour and engagement in statistics. This justifies the assumption that spoken behaviours are indicative of information coaches require.

As seen in Table 15, the individual player statistics, 'Player Grid' (37%) and 'Opposition Player Grid' (39.3%) also had strong alignment with the spoken themes from Croft et al. (2020), with comparative player themes aligning with 'Both Teams Player Grid' (26.1%). This also indicates a relationship between what coaches discuss and view statistically.

The ‘Scoring’ (24.9%) category of statistics have weaker alignment with the spoken themes of coaches. This may represent that they view it often (16 views per match, at 8.3s each) because it is an outcome measure, yet do not need to discuss it to interpret its meaning, as it is easy to understand.

Table 15.

A comparison of the spoken themes identified by Croft et al. (2020) and the coach engagement with statistical categories.

(Croft, Spencer, & Robertson, 2020) Spoken behaviour	Statistics screen categories	Alignment % of spoken themes	Description
Scoring Attacking actions Defending actions	Scoring	24.9	Statistics that provide information about trends and current state of scoring events
Player positioning Team positioning Attacking actions Defending actions Errors Attack strategy Defensive strategy	Team Stats	46.6	Combine actions of the team that are either attacking or defensive and reflect losing or maintaining possession of the ball
Centre pass Gains/Losses Attacking actions Defending actions Errors Attack strategy Defensive strategy	Player Grid	37	Individual actions for own team that are either positive, scoring, possession gaining, or errors that lead to losses of possession of penalties.
Opposition actions Centre pass Attacking actions Defending actions Gains/Losses Errors Attack strategy Defensive strategy	Opposition Player Grid	39.3	Individual actions for opposition team that are either positive, scoring, possession gaining, or errors that lead to losses of possession of penalties.
Opposition actions Centre pass Gains/Losses Attacking actions Errors	Both Teams Player Grids	26.1	An abbreviated player grid allowing for comparison of both teams individual players
Errors Attacking actions	Goal Shooter	18.4	A view of the shooting performances of the 4 players on court at the time of viewing, for both teams.
Attacking actions Centre pass Errors	Top 5	18.4	A view showing the 5 players on court for highest counts for 5 individual actions.

Attacking actions	Score Flow	17.7	A temporal list of scoring and non-scoring attempts for both teams

A limitation of this study is that the statistics screen presented to the coaches did not present information about the players movements, which is a category identified by Croft et al. (2020) as of interest to coaches during games. Approximately 35.8% of the themes categorised in that study related to 'movement timing', 'movement intensity', 'movement spatial', and 'opposition movements'. This would require information technology that was either worn by the players (Chandler et al., 2014) or an automated video tracking systems that could produce these metrics (Barris & Button, 2008). With the recent advent of commercial systems (Glory League, 2021) that can be installed in an arena and automate the collection of video information this could become a possibility in the future. These systems would need to provide an application programming interface (API) for this integration to occur.

6.5 Conclusion

There are several findings and insights discovered in this manuscript that can help with the design of information systems for coaches during live netball games. These findings could also provide insights for other sports also. Firstly, computerised observation of coach's behaviours is a valuable method for understanding coach's requirements. This approach provides feedback on the ecological validity of statistical tools and can help with the design and alignment of these tools to coach's needs. Performance analysts should also use simplistic statistical measures to avoid cognitive demand on coaches. As coaches already face high cognitive loads during matches, which a complex and multidimensional in nature. A layered organisation of statistical information should also be implemented to allow the coaches easy movement between 'layers' information. The limitations of these results to be acknowledged. There were only six participants, and they were observed in a single sport, netball. Future research and product development needs to include the measure of 'movement' based statistics, as this is an area which coaches discuss often and will help inform decision making during matches. Products do exist however the ability to integrate this with other performance information is limited.

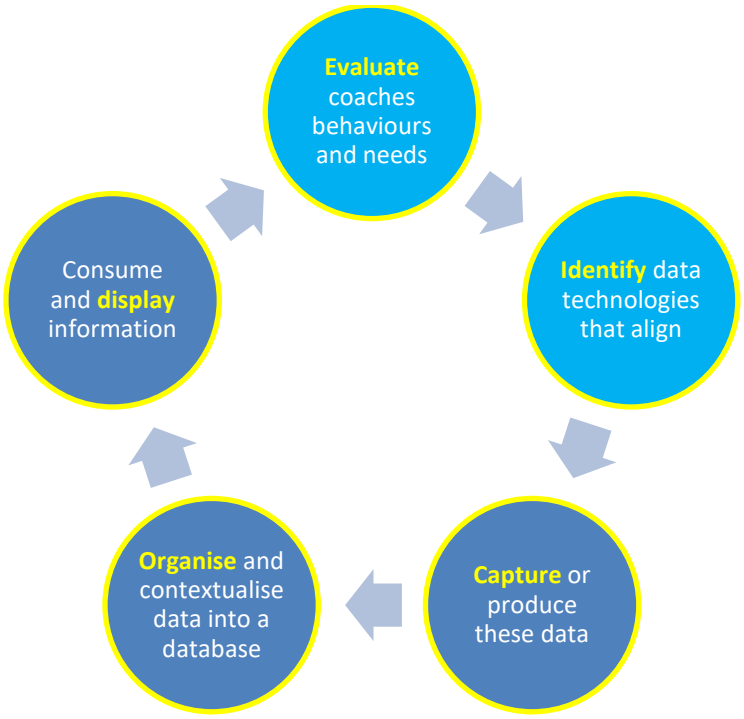
Prelude Chapter 7.

Chapter 6 identified that some coaches may utilise statistics information for substantial amounts of time during netball matches. The research findings suggested that the design of the statistics tool could incorporate team statistics and player statistics in a layered manner. This would allow the coaches to navigate through layers of complexity. The 'Team Stats' was the most common category viewed by the coaches and they tended to move between this category and others frequently. When the categories viewed were compared to the themes determined in chapter 3 there was some alignment with 58.3% similarity. This suggested that coaches verbal themes could be aligned with areas that coaches may be inclined to seek further information.

The layering of information and preferred statistical categories, observed in chapter 6, needed to be integrated into a tool to demonstrate how these findings can be acted upon. To understand if ecological validity had been maintained feedback from users also needed to be assessed.

Figure 23.

A complete framework for the development of tools to navigate large datasets.



Chapter 7 demonstrated how the proposed framework (Figure 23) could be used to develop a tool for coaches. It used the results from the previous five manuscripts to develop a web-based information tool. The usefulness of the tool to enhance decision-making was evaluated by six elite netball coaches and sport science staff for the five dimensions identified in chapter 3. They

subsequently provided feedback on their likeliness to use the tool in the future, during netball matches, providing an assessment of its ecological validity.

The purpose of chapter 7 was to assess a validity of a tool developed from the framework proposed in Figure 23. This tool incorporated each of the stages of the framework and the learnings identified in the previous 5 chapters. It was test for ecological validity via a net promoters score survey participated in by 6 high performance practitioners.

Chapter 7.0 Developing information tools for presenting coaches with large datasets.

This chapter comprises the following manuscript (6) in preparation for submission to Journal of Computer Science in Sport.

Croft, H., Spencer, K., & Bowden, M. (2021). Developing information tools for presenting coaches with large datasets. (*In preparation for submission*) Journal of Computer Science and Sport.

Author contributions: Croft, H (80%), Spencer, K (10%), Bowden, M (10%)

7.1 Introduction

There is a growing body of literature rationalising and describing the opportunities and risks associated with big data in sport (Vermeulen & Yadavalli, 2018; Passfield & Hopker, 2017). There are more data producing technologies in existence than any other time in history, with claims of 20.8 to 50 billion connected devices existing world-wide in 2020 (Aeris, 2021). Large data repositories are now held in the 'cloud' presenting easy access to these datasets for those with ways of using them. Recommendations are made on how to conduct retrospective research of these datasets to improve knowledge of sport and exercise sciences (Passfield & Hopker, 2017). However, warnings are given around data validity, data security and protection, as well as athlete autonomy, as risks to the reputability of research (Vermeulen & Yadavalli, 2018).

Data is described as "just random dots on a page" with no pattern (Vermeulen & Yadavalli, 2018). Data becomes information by adding colours to these dots, however until connections are made between the coloured dots no knowledge exists. Only once knowledge exists, can a coach or athlete act upon it. Beyond this, insights and wisdom are built over time and are the traits of expert coaches. Kaushik, (2016) describes a hierarchy where data is at the lowest level, and wisdom the highest, with information, knowledge and insights in between. This infers that more needs to be done the just collecting and assessing data. For coaches to develop insights and wisdom from data, information and knowledge needs to be developed for them.

The way big data is used can be guided by Kitchin's (2015) description of the three V's. They are Volume: the size of data sets, Velocity: how quickly the data arrives and can be accessed after the event and, Variety: how diverse the data is in its type and organisation. These characteristics need to be considered when designing tools to access and utilise insights from large data sources.

The development of a framework, for sports performance analysts to follow when developing data presentation tools to coaches, is a desired requirement in the era of 'big data'.

Several approaches have been utilised in the design of tools for both research and applied performance analysis (Connell & Spencer, 2020; O'Donoghue et al., 2018). No current framework exists for developing tools that present data and information to coaches in sport. The objective of this study was to demonstrate the development of one such tool through implementing the proposed framework set out in Figure 24. Each stage of the framework was created reflecting on research in netball. Once constructed, the tool was then evaluated by elite netball coaches and sport science staff to determine the level of ecological validity achieved.

Sources of data

An evaluation of data sources (Croft & Spencer, 2021), that align with coaches spoken behaviours (Croft, Spencer, & Robertson, 2020b) identified the XML files from the Champion Data™ statistics website provided data about many of these themes. Scaped public domain data organised on comma-separated values (CSV) files and XML files coded in Hudl-Sportscode™ were also identified as suitable. An automated programming interface (API) was not available, from Champion Data™ at the time of this tool development, however this was expected to be available in the future. For this reason, the tool in this manuscript was not able to be presented live during games. The XML and CSV files were uploaded into an Amazon AWS™ bucket and a pipeline prepared the data and stored it in a cloud database ready for consumption by a visualisation tool.

Data visualisation

Lowe & Matthee (2020) claimed that data visualisation tools should be designed to complement human cognitive and visual capabilities. They should be fast, automated, and intuitive. Their review of 31 research articles found six categories of requirements. These were: data reduction, dimensionality reduction, fast retrieval of results, interactivity, scalability, and user assistance. These factors should be considered when selecting a platform to visualise information. Orad (2019) rated the Microsoft™ data visualisation tool highest for 'completeness of vision' and 'ability to execute' from a group of 21 analytics and business intelligence platforms. For this study PowerBI™ was selected to display information to coaches.

Information contextualisation

Contextualisation was explore by Croft H. , Spencer, Taurua, & Wilton, (2021b) in netball data with a single representative percentile value being favoured by coaches. This was due to its simplicity and understandability for themselves and players. In the design of a tool that represents the

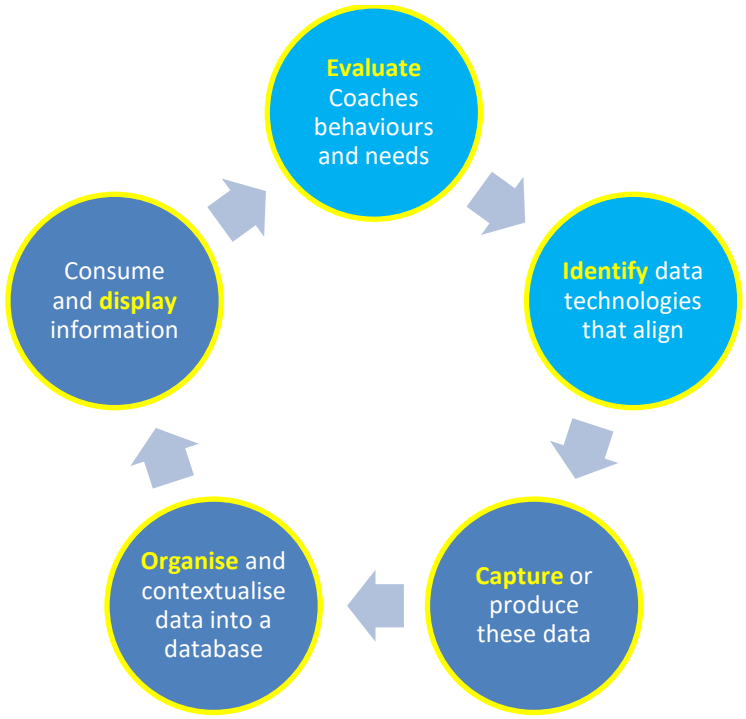
performance of a single team, this number should also be simple and provide context, both short term and across a season. A season represents the performance of the current squad of players, while a three-game average will show recent trends which the player should still remember details of.

Net promoters score and surveys

Net promoters score (NPS) is a metric that is commonly used in business for assessing user satisfaction and has been linked to future company growth (Reichheld & Markey, 2011). It has also been adopted in high-risk health-related research where client satisfaction is important (Koladycz, Fernandez, Gray, & Marriott, 2018). The NPS method meets the objective of this study to evaluate the ecological validity of a tool developed using the framework in Figure 24. By surveying the participants and asking them to score the likelihood of using the tool again, a NPS can be calculated.

Figure 24.

A complete framework for the development of tools to navigate large sports datasets.



7.2 Methods

7.2.1 Participants

Six elite netball coaching and sport science staff were recruited to give feedback on the tool developed using the framework in Figure 24. All participants were employed in professional netball teams and

used performance statistics in their roles. Two of the participants were female professional coaches with more than 10 years' experience. Two participants were male performance analysts who had worked in netball and other elite sports for at least 5 years. The final two male participants had strength and conditioning backgrounds and had worked in netball and other high-performance sports for at least 10 years. Each participant gave informed consent (Ethics application number 21/33) and was then given access to the tool.

7.2.2 Equipment and procedures

Participants were given access to the tool for two weeks, in which time the results of two rounds of elite netball were displayed. The participants were then provided with a survey which asked questions about the tools usefulness in making decisions. There were individual questions for 'movement', 'actions', 'positioning', 'possession outcomes', and strategy. It also asked them if they would use the tool again, 'future use', and gave an option to provide comment. The survey utilised a sliding scale from 'very helpful when decision-making' to 'not useful when decision making'. They were instructed to make a mark along the scale to indicate the usefulness of the tool. These scales were measured in units of 10 and a net promoter score (NPS) was applied, to the 'future use' scale to rate the satisfaction of the participants. The NPS was calculated by taking the number of ratings scored 9 or 10 and then subtracting the number of ratings 6 or below. The resulting value was calculated as a percentage between -100 and 100.

The stages below indicate how the tool was developed, utilising the framework (Figure 24) and the results of previous research.

7.2.3 Stage of development

Stage 1: 'Evaluate coaches behaviours and needs'.

Croft et al. (2020) identified five tactical dimensions of themes discussed by coaches during elite netball matches. These were identified using a grounded theory approach with a content analysis of their spoken language. These tactical dimensions were 'movement' (38.4%), 'actions' (23.9%), 'positioning' (17.3%), 'possession outcomes' (10.6%) and 'strategy' (9.8%). These dimensions were made up of 18 higher order themes the seven most common being: 'movement intensity', 'movement spatial', 'team positioning', 'player positioning', 'attacking actions', 'gains/losses', and 'defensive strategy'. These findings gave a strong indication of what information coaches were interested in during netball matches.

When revisited in another study (Croft & Spencer, 2021b) it was found that coaches engaged with statistical information similar to the themes they had discussed (58.3%). This was despite the

two studies being conducted three years apart. The Croft & Spencer (2021b) study also provided insight into the movements between categories of statistical information. It showed that there were layers of complexity in the information that needed to be understood and navigated by coaches. This research showed the need for specific categories of information to be integrated into tools which were layered in complexity, allowing coaches to 'drill down' through the layers as they required.

Stage 2: 'Identify data technologies that align'.

Potential data sources for the five tactical dimensions and seven higher order themes were investigated. Croft & Spencer (2021a), evaluated data and information producing systems. The systems that best aligned with the coaches themes included the Champion Data ™ group of websites and event coding software such as Hudl-Sportscode™. These sources aligned with operational criteria, were accessible, and had strong validity. They were available as XML and CSV file formats which could be integrated into a data pipeline. Hudl-Sportscode™ could export a live updated XML, however an Application Programming Interface (API) was not available for the Champion Data™ website. The website itself did display statistics in real-time during matches with minimal delay. However, this was without normalisation, contextualisation and the structure identified as important by Croft et al. (2021b). For this study Champion Data XML files were therefore combined with CSV data, which was exported from Hudl-Sportscode™.

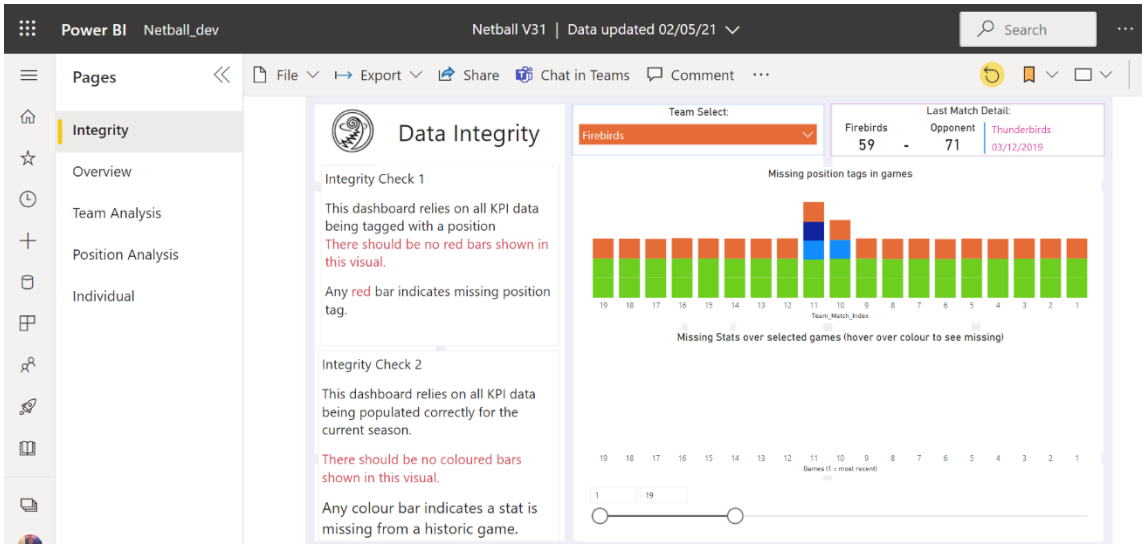
Stage 3: 'Capture or produce this data'.

Croft et al. (2021b), demonstrated a methodology for automated data processing. The first part of this process was to capture the data sources, XMLs and CSVs, into a Amazon™ cloud located 'bucket'. From this 'bucket' a data pipeline ran algorithms over the data and created variables into an integrated, aggregated, de normalised table structure. These variables could be considered 'information' in Kaushik (2016)'s data to wisdom continuum. Data accuracy checks were then implemented to ascertain if data was either missing or inaccurate by cross referencing the two data sources.

The tool developed had a 'data integrity' tab, so the coach or analyst could check for missing or inaccurate data (Figure 25).

Figure 25.

Data integrity screen on data consumption and display tool in PowerBI™.



Stage 4: 'Organise and contextualise data into a database'.

Croft et al. (2020), used an inductive content analysis to identify four layers of information in coach discussions (Table 16). Croft & Spencer (2021) also identified layers in the way coaches navigated categories of data. These insights encouraged the use of layers when displaying information for coaches to navigate. These layers required the creation of variables for 'team', 'position', and 'individual' player performance.

Table 16.

Content analysis levels aligned to data, information and coach interface.

Layers	(Croft, Spencer, & Robertson, 2020) Coach spoken themes	Variables created in database for coach tool.
1	General dimensions	Team information
2	Second order themes (higher)	Position information
3	First order themes (lower)	Individual player information
4	Raw data themes	Raw data captured from sources

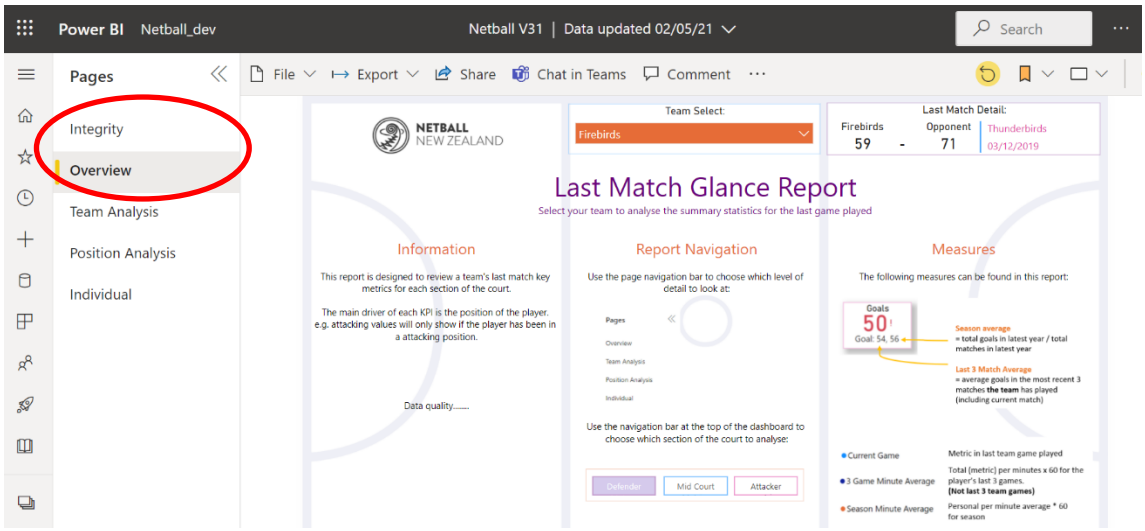
Croft et al. (2021b) identified that providing contextualisation, to information, helped coaches understand it. This was achieved by taking historical data and producing normative percentiles and presenting them as single number benchmarks. The tool developed in this manuscript presents each statistic with a three game and season average. These statistics were also standardised and presented as 'per 60 minutes played'. This allowed comparisons to be made between players and teams, regardless of the number of minutes played.

Stage 5: 'Consume and display information'.

PowerBI™ data visualisation software was chosen for this tool. It not only allowed an automated connection to the cloud data storage, but the licencing costs and accessibility were better suited to the participants. The visualisations were arranged in a series of layers, allowing the user to 'drill down' from broad team information to more complex position and individual player information. The top two screen tabs were the 'integrity' page and an 'overview' page, highlighted in Figure 26 and were additional to the information layers, 1 to 3 below. Layer four was raw data, however this was not presented to the coaches.

Figure 26.

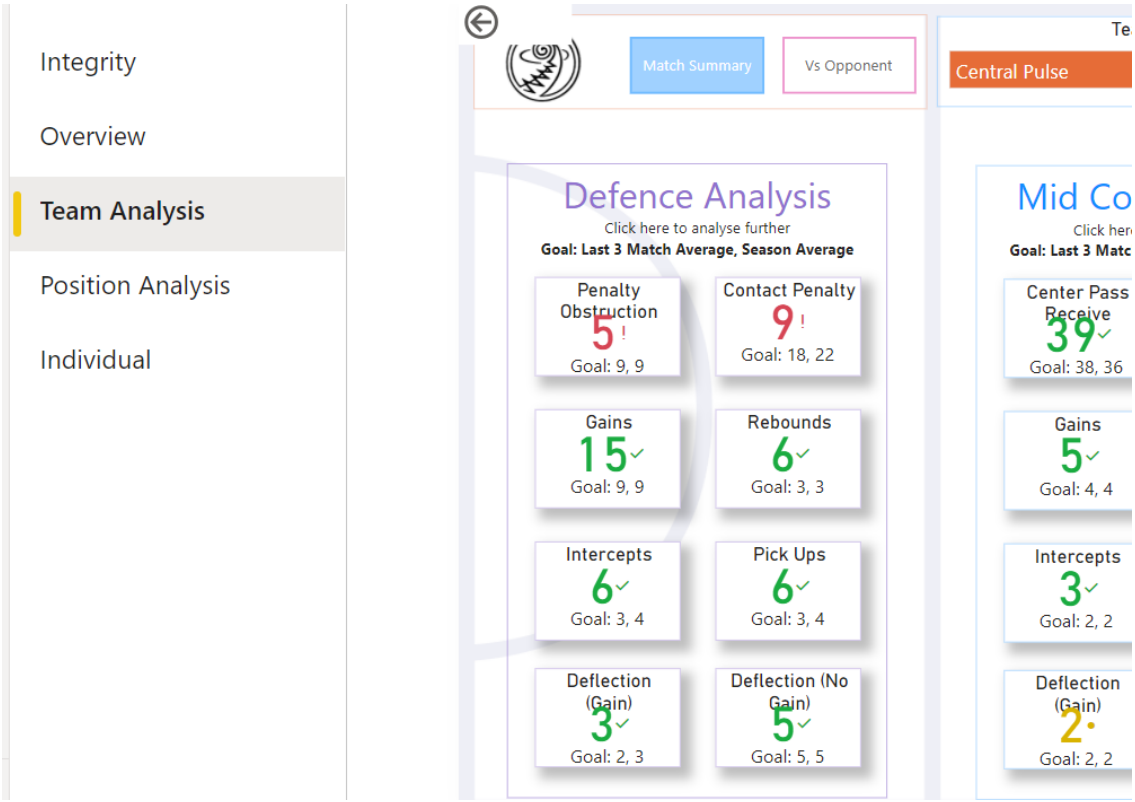
Overview screen on data consumption and display tool



The 'Team Analysis' screen (Figure 27) displayed information unique to 'defensive', 'midcourt' and 'attack' components of the team. These were contextualised with a three game and season average below each statistic, and a number colouring. The numbers were coloured green for better than both these averages, orange for equal to or between these numbers, and red for below. The information could also be viewed in comparison to an opponent's performance by selecting the 'Vs Opponent' tile.

Figure 27.

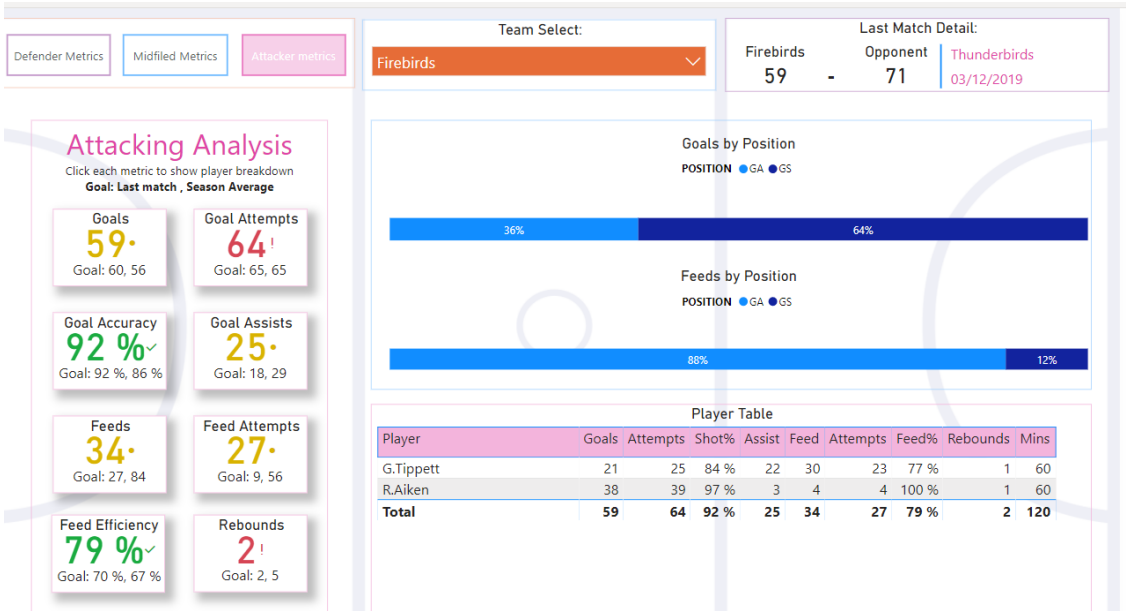
Team analysis screen on data consumption and display tool, showing contextualisation of information as three game and season average below each statistic.



The 'Positional Analysis' screen (Figure 28) displayed information for each position group: 'Attacking', 'Midcourt', and 'Defensive'. These could be displayed by selecting the tile at the top. By selecting a metric to the left, individual player contributions would then be displayed on the right side. Information was displayed as a column chart with each player represented by a group of three columns: current match, three match average and season average. This information was standardised as 'per 60 minutes played'.

Figure 28.

Position Analysis screen on data consumption and display tool.

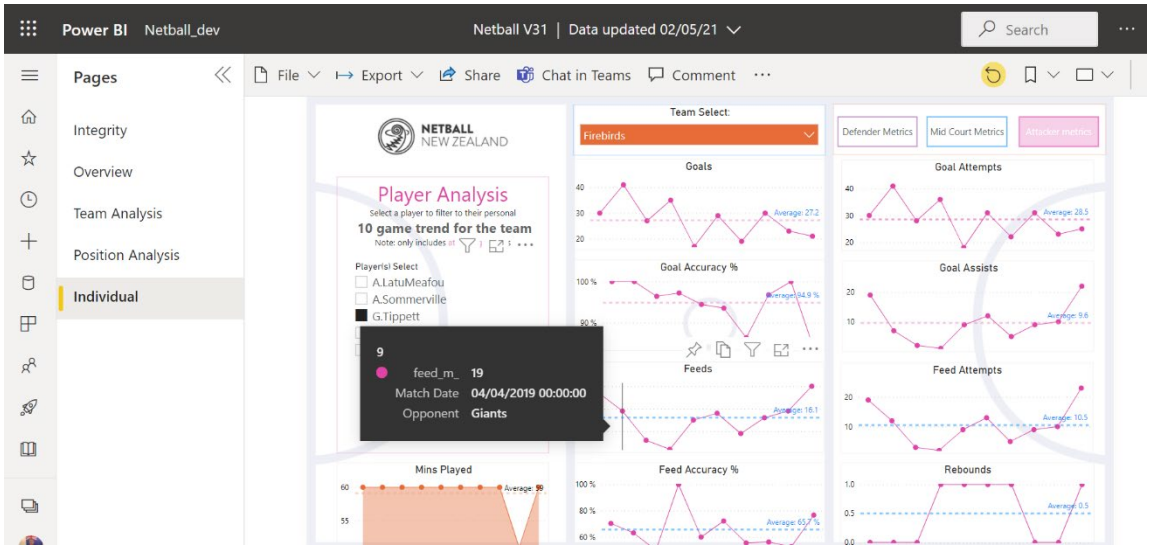


The final 'layer' of information displayed in the tool was the 'Individual' player analysis screen. This allowed the user to select a position group, on the top left (Figure 29) and then select an individual, or multiple players. The line plots then demonstrated the group metrics for the individual over the last 10 performances, also showing an average for these. As these were absolute values a 'Mins Played' plot showed the individuals playing time to help with standardisation.

On all these screens the user could choose one, or multiple teams, from the 'Team Select' drop down menu at the top of the page. This allowed for comparison between players, across different teams, or competitions.

Figure 29.

Individual Player Analysis screen on data consumption and display tool.

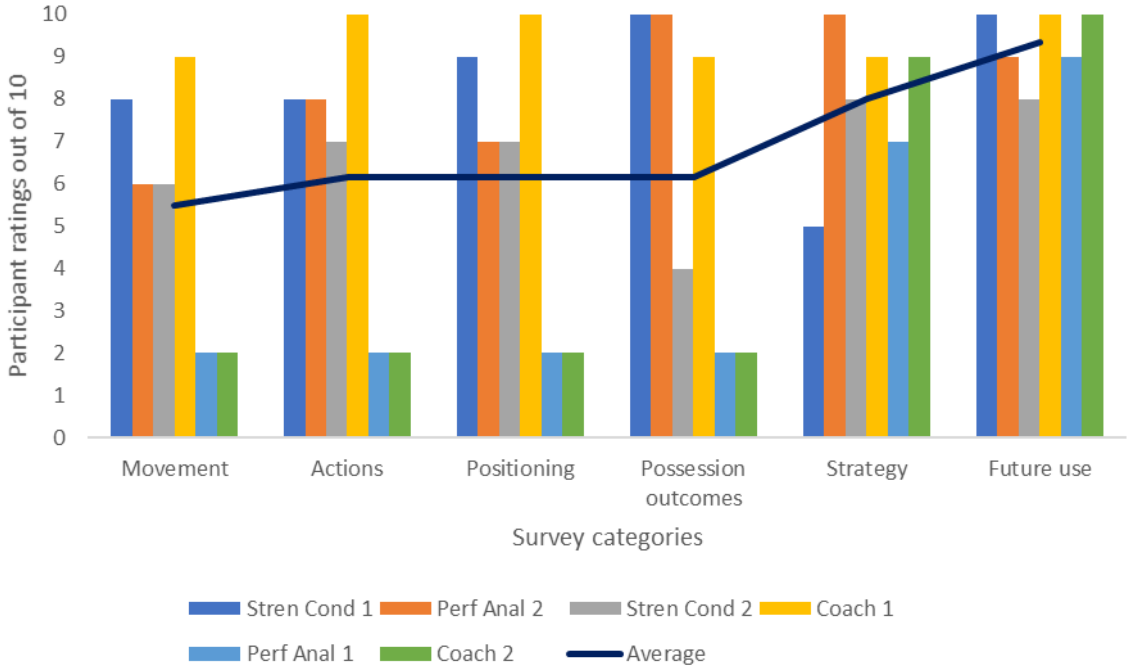


7.3 Results

The participants used the tool and then completed the survey. By measuring the marks made on the sliding scale for each survey question, values were generated for each category. Figure 30 shows the responses from each of the six participants and an average for the group. The first five category results represent the participants perception the usefulness of the tool to enhance decisions. In the final category, they indicated if they were likely to use the tool again in the future. Figure 30 displays the individual responses and mean response of each participant.

Figure 30.

Ratings by participants for each survey question with a line to demonstrate the category average.



The results in Figure 30 shows that for the 'movement', 'actions', 'positioning', and 'possession outcomes' categories, two participants scored the tool very low. When their survey comments were reviewed it appeared they saw value in the tool, however, did not use statistics directly for decision making:

> I use statistics to measure performance (not decision making) and as a tool for training. Stats also provide evidence to change behaviours and actions should be used in conjunction (Coach 2).

> The dashboard and statistics therein did provide me with some extra justification to add to discussions when they do happen (Coach 1).

The results for the other participants mostly scored at or above 7. The net promoters score for the question 'future use' was 100, all participants responding with a 9 or 10. Comments that supported this result include:

> Ability to have this data on hand, clear easy presentation using averages and having historical data supports the athlete's awareness of performance and the coach's capacity to make robust decisions and be supported with accurate feedback (Coach 2).

> I thought the system was user friendly and easy to navigate (Stren Cond 2).

> Hard to decipher what the strategy was but good at finding information. Great setup (Stren Cond 1).

Overall, the survey results provided evidence that some ecological validity was achieved with this tool. In contrast, the 'movement' category was rated the lowest which indicates better information needs to be included.

7.4 Discussion

It must be acknowledged that the results of this study are based on a small sample of 6 participants and therefore may not be a fair representation for all coaches in netball or other sports. With this limited sample of participants the study showed that the framework (Figure 24) assisted the development of a data visualisation tool that the participants liked and would promote to others or use again in the future. It showed that 'future use' was likely and that the tool was best for making decisions about 'strategy'. In a netball context the NPS is a useful measure for evaluating the ecological validity of a tool. This is because it measures the satisfaction of the user and their likelihood of using it again (Reichheld & Markey, 2011). Bendle & Bagga (2016) claimed that the use of net promoters score shifted organisation culture from bureaucratic to customer centric. This aligns with the objective of this study to create an ecologically valid tool. However, the NPS metric is only effective if it is used as a system, where further information is unpacked and acted upon, not purely a score (Reichheld & Markey, 2011). This infers that the other information in the survey must be analysed and acted upon.

Results from the other survey categories, 'movement', 'actions', 'positioning' and 'possession outcomes' rated lower than the 'future use'. This may be due to the survey question asking the participants if the tool would help them make decisions. The tool may add value to not only decision making, but also as evidence to a coach that changes or strategies are working. This relates to Hughes and Franks (2008) findings that some coaches might look for information that reinforces their opinions. This has often been viewed as contributing to conformation bias, however if the information is organised and represented in a valid manner, this may challenge their bias instead. As (Nash & Collins, 2006) explain, some coaches look to reduce cognitive load when making decisions, especially when time constrained, and this tool might assist them with this.

Other features of the tool were the layers that represented different levels of performance information e.g., 'Team Analysis', 'Positional Analysis' and 'Individual'. This followed the natural arrangement of discussed themes identified in Croft et al. (2020) from 'Team Analysis' to 'Individual' hierarchy. In Croft & Spencer (2021b), it was also observed that coaches spent time looking at 'Team Stats' and then would click down into position and individual information. The tool therefore also incorporated three game and season averages to contextualise the statistic being viewed, as seen in Croft et al. (2021).

There were limitations to the tool displayed in Figures 26-29. Due to the unavailability of an application programming interface (API) from the sources investigated in Croft and Spencer (2021), this dashboard did not update live as a netball match was being played. The Champion Data, (2020) statistics website investigated by Croft and Spencer (2021), is an alternative that displays some important information. It did not however allow for the same "drill down" data hierarchy proposed in this tool, nor contextualisation of the statistics demonstrated in Croft and Spencer (2021).
The tool could be modified in the future to allow a 'coder' to use Hudl-Sportscode™, or similar notational analysis software. This would add live data to the database during a game. Updates to the visualisation software PowerBI™ would also be required.

Another limitation of this tool was that data about the spoken player 'movement' themes identified in (Croft et al., 2020), was not available in a compatible format. This is reflected in the participant feedback where it rated the lowest of the survey categories. Future research and development need to be conducted to bring this information into statistical tools. A recent review of movement sensor systems (Homayounfar & Andrew, 2019) found that these technologies still require development as they were to intrusive and not capable of long-term, continuous monitoring.

7.5 Conclusion

This chapter has demonstrated the embodiment of a framework, for performance analysts, in a tool for coaches to navigate large complex datasets. The tool received positive feedback from users. The

chapter has also acknowledged that there are areas of research and development that need to be addressed to make this tool a live and fully aligned solution. The five-stage framework that it follows, has shown some benefit as a guide for practitioners in the field of performance analysis in sport.

8.0 Discussion, Conclusions, Limitations/Delimitations and Future Research

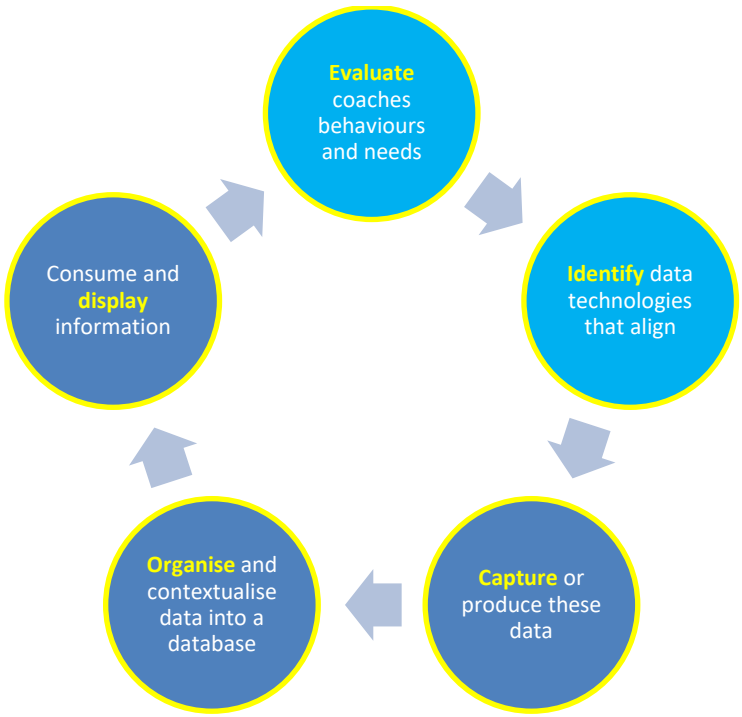
8.1 Discussion and Practical Implications

This PhD set out to answer the question, 'What process, or framework, will present large datasets to coaches in a meaningful and valid manner to assist their decision making?'. To answer this, a review of literature was followed by the four studies in chapters 3 to 6, and an example tool was built and evaluated. The four studies sought to understand: 1) What tactical and technical comments do coaches make during sports matches? 2) What data sources are available in netball that align to these? 3) How to organise, normalise and contextualise these data automatically, and 4) What coach behaviours were exhibited when engaging with this information during a live game? The final study (manuscript 6) constructed the embodiment of the framework, in a specially designed tool, which received positive feedback from users.

Chapters 3, 4, 5, and 6 attempted to demonstrate the key stages in developing a framework that guides performance analysts when designing statistical tools. The components of this framework attempt to align with the stages of a system design methodology; 1) Research and planning, 2) Specific analysis, 3) System design and 4) System testing and redevelopment. It also adopted many elements of an Action Research Design methodology. Both the ADR and systems design approaches to developing tools were generalised for a range of settings and applications. However, for managing large datasets a more specialised approach was required. This PhD implemented a new framework, developed specifically for presenting large datasets to coaches, with an attempt to maintain ecological validity. The five stages, seen in Figure 31, included:

1. Evaluate the coaches behaviours and needs though observation and questioning or surveys.
2. Identify data technologies that can provide information relevant to the needs of the coach.
3. Capture or produce data and information about these needs and ensure they are valid and reliable.
4. Organise, normalise, and contextualise these data into a database for easy access.
5. Consume and display this data in a data visualisation tool which the coach can navigate with ease.
6. Repeat task 1 to evaluate the coaches behaviours with the tool and develop if needed.

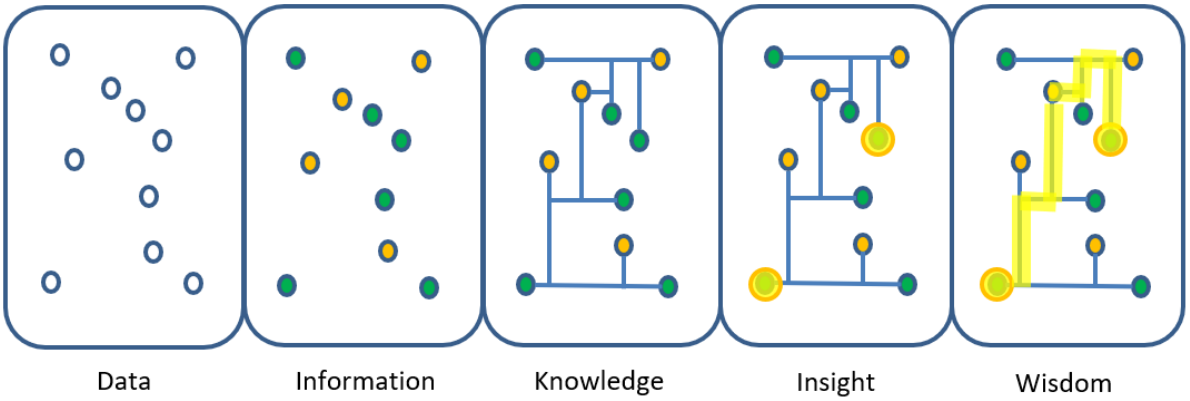
Figure 31.
A complete framework for the development of tools to navigate large datasets.



The title of this PhD 'Addressing the problem of 'big data' in sports: A framework for performance analysts' acknowledges the performance analyst as an information user who create tools from complex datasets and presents them to coaches who use these 'tools' when making decisions. The framework developed in this PhD was designed to assist performance analysts when developing these tools. The tool developed in chapter 7 was an example of the embodiment of this framework and was designed for this purpose.

As seen in Figure 32, coaches use tools at the 'information' stage (Kaushik, 2016), and then use their experiences and learnings to make connections and create 'knowledge'. As a coach's career progresses, this knowledge can become 'insights' where they could see the beginning and the endpoint but are not completely sure of the path. As more insights develop, they may start to understand the path to the endpoint and can steer their players along it. This is what is defined as 'wisdom'; however, the process is initiated by converting 'data' into good 'information'. Frameworks and tools are needed to help coaches and performance analysts transform 'data' into good 'information'.

Figure 32. *A visualisation of the continuum from data to wisdom, adapted from (Kaushik, 2016).*



Tacit knowledge has been described as the 'art of coaching', however some dispute this (Nash & Collins, 2006) as it is built upon experiences and is often implicit and unarticulated (Kabir & Carayannis, 2013). These experiences allow coaches to identify the importance of cues and gain greater insights from these cues. As tacit knowledge was acquired from the information presented to them via their experiences, in a non-verbal manner, the quality and range of information they observe and learn from was important (Jakubowska, 2017). Kabir and Carayannis, (2013) explained that tacit knowledge was embedded in 'big data' and there were opportunities for its discovery. Therefore, information constructed from data sources in sport could be important to coach development and decision making. It could help form a coach's tacit knowledge as well as their future insights and wisdom. Essentially the framework in this PhD, if applied to tool construction underpinned by 'big' or 'large data', could support the development of a coach's tacit knowledge.

The framework, and its embodiment in a tool, could assist performance analysts in supporting the coach by streamlining observation and feedback aspects of the coaching process (Williams & Kendall, 2007; Drust, 2010; Nicholls et al., 2018). It has been explained that poor communication between sports scientists and coaches, contributed to a gap, and was reciprocated through coaches' understanding of the sport science fields (Waters et al., 2019). By observing and understanding the coach's requirements, this gap could be reduced or even closed.

Chapter 3 elucidated the tactical themes discussed by coaches during matches and used these observations to build an inductive model explaining each of the themes and the layers within each theme. The tactical themes observed, movement; positioning; actions; possession outcomes; strategy, built upon previous research (Horton et al., 2005; Zetou et al., 2011) which found coaches discussed tactical instruction the most followed by general instructions; and technical instruction with encouragement, motivation, other comments also present. More recent research has emphasised this with Batista et al., (2018) finding, in practice situations in football, strategic instruction was the most impactful and constrained the technical, tactical, and physical demands of the task.

Chapter 4, explored information technologies and sources that aligned with the tactical conversations identified in chapter 3, making recommendations on the use of operator controlled notational analysis software (Hudl-SportsCode™, Nacsport™, DartFish™) as they were deemed best suited to providing live information and data to coaches and players during netball matches. Many studies have utilised these systems to produce information about team sports like Australian football and hurling (Lee et al., 2017; Aarons et al., 2021; O'Brien et al., 2021). Recommendations were also made around the use of commercial data providers, link Champion Data™, who publish their statistics online publicly. These data providers statistics have been utilised in research studies into team sports like Australian football (Sullivan et al., 2014) and netball (Ofoghi et al., 2021).

Chapter 5 used data collected from the tools and sources evaluated in manuscript 3 and constructed a technical workflow to demonstrate how a performance analyst might use this information to construct context for performance statistics. This information was presented as normative data tables and visualisations that were guided with feedback by coaches. This normative information adds to the body of literature in netball that already exists in physical (Thomas et al., 2016), physiological (Venter et al., 2005) and biomechanical/anthropometrical domains (Ferreira & Spamer, 2010)

In chapter 6 the coach's behaviours were revisited. This time through their engagement with statistics that were related to the themes determined in chapter 3. Coaches' patterns of viewing of the statistics were observed during live games, with movements between categories and the levels of complexity in the statistics or information observed. Although other studies directly related to these behaviours do not exist, research does exist that reports the access and usage of quantitative match data (Wright et al., 2012; Painczyk et al., 2017; Kraak et al., 2018). Chapter 6's findings many add additional insights to these studies.

In chapter 7 a tool was developed to embody the framework, with layers of complexity arranged for coaches to navigate. By aligning the tool with the coach's behaviours, both spoken and statistical, user satisfaction was observed. It is surmised that, due to these factors, the tool complimented coach's tacit knowledge and exhibited ecological validity. This tool and framework compliments research (Ofoghi et al., 2013) which has also looked to develop a framework in performance analysis and data mining as it brings the coaches requirements into the process of data collection. Other frameworks offer a more theoretical approach to performance analysis (Travassos, et al., 2013; Glazier, 2017) incorporating ecological dynamics or constraints based (Newell, 1986) approaches. The framework proposed in this PhD my share elements of these other frameworks, however its intension is to be highly practical and applicable to the performance analyst practicing in netball and other team sports.

8.2 Conclusions

In a fast-changing world, 'big data' or 'large data' sets are both an opportunity and a challenge. Poor use of data can create confirmation biases, yet good use of data may inform decision making. The solution to this opportunity or challenge was to create a process and framework for performance analysts, and other sport science staff to use as a guide. The framework that was established in this PhD incorporated the principles of the ADR and the Systems Design approach. Each stage of the framework was developed with the insights detailed in each manuscript. Some of the most interesting insights were:

- There was a relationship between what coaches discussed tactically and the statistics they viewed.
- Further work needs to be conducted to identify suitable movement data sources available to fulfil netball coach's needs.
- Benchmarks or targets from normative data may need to be presented as simply as possible.
- Coaches may navigate complex information by moving between layers of complexity.
- Information was not always used for decision making, and in some cases, it could be used to reinforce observations.

This framework was then practically demonstrated within a working 'tool' for netball coaches, with positive feedback received regarding its usefulness and ecological validity. It was acknowledged that the designed 'tool' was limited to netball, however the framework could have been applied across many other sports. The PhD has therefore created a framework for performance analysts when designing tools for coaches to use.

8.3 Limitations and delimitations.

As with all research there are limitations and delimitations with this research and its application. The limitations of this research are:

- Due to COVID19 restrictions in 2020 and the associated changes to professional sport in 2021 it was difficult to attend all data collections and ensure correct methods were adopted. Data collection was also delayed completely from 2020 until 2021. This limited access to participants and some observations were not suitable for inclusion. In manuscript 6 a mixture of coaches and sport science staff were used to overcome the issue of access to coaches as participants.
- There was no available Application Programming Interface (API) for bringing data in real-time into the tool constructed in chapter 7. This is a common issue with data sources, or providers,

and would enable automation during live matches. Other sports, such as rugby and cricket, do have access to API's through companies like OPTA Sport Data™, however these services do not currently exist for netball.

- There was a lack of usable 'movement' data to turn into information aligned with the coach discussions in chapter 3. It is acknowledged that there are various movement measuring systems available, however these are not suitable for use long-term, continuous, and unobtrusive use (Homayounfar & Andrew, 2019).

The delimitations of this research are:

- The sample sizes in chapter 3, 6, and 7 were only six participants. This was due to the limited population size for elite netball coaches in New Zealand. It was decided to keep the sample size the same throughout each of these studies to allow an appropriate comparison.
- Although the framework developed, can be applied in any sport, the example given in this thesis is netball. Netball is a team sport, invasive, and has very specific tactical requirements of players and teams. The tool developed is specific to the sport of netball and would not be suitable for other sports.
- The tool was developed with netball data, collated from school to international level. It was however, only trialled (in chapter 7) with elite professional coaches and sport science staff. It cannot be assumed therefore that this tool would be ecologically valid if used by school level coaches.
- Not all coaches or performance analysts are the same and therefore the validity of the framework and tools may not be as effective for different individuals.

8.4 Future Research

Throughout this PhD there were several observations that were not reflected in the manuscripts, as they were not part of the original question. One such observation was that in New Zealand high performance netball the strength and conditioning coaches tend to work independently of the performance analysts. This is an important gap to acknowledge, as the strength and conditioning coach captures, organises and presents the 'movement' information. The performance analyst does not usually access this data or information. This may not be the case in other countries or sports. As found in chapter 3 ("What tactical and technical comments do coaches make during sports matches? A content analysis in netball."), the 'movement' category is the most frequently discussed by coaches to players. This would suggest that better integration of this movement information is an area for improvement and future research in tool development.

The observational methods implemented in chapter 3 and 6, for understanding coach behaviours, were only applied in a netball context. These methods are the first and last stages of tool development yet have not been conducted in other sports. Future research could look at coach behaviours between sports. An understanding of the influence of the physical environment, courtside vs. coach's box, would also be insightful, as this may limit the information a coach can utilise during a sports match.

In chapter 4 and 5, 'data literacy' was required to identify, organise, contextualise and present information (Bersin & Zao-Sanders , 2020). Researchers could look to investigate the 'data literacy' of performance analysts in sport. Many performance analyst are proficient in using products like Hudl-Sportscode™ or Dartfish™, however, these tools do not currently integrate large multidimensional data sources. When performance analysts begin working with data in software like Microsoft Excel™, they may be making mistakes when they manipulate the data, for example averaging percentages, averages, non-standardisation etc.

9.0 References

Aarons, M. F., Young, C. M., Bruce, L., & Dwyer, D. B. (2021). The effect of team formation on defensive performance in Australian football. *Journal of Science and Medicine in Sport.* https://doi.org/10.1016/j.jsams.2021.09.002

Abdulla, U. A., Taylor, K., Barlow, M., & Naqshbandi, K. (2013). *Measuring walking and running cadence using magnetometers*. Retrieved from: https://ieeexplore.ieee.org/document/6681000

Adderley, R. (2004). The use of data mining techniques in operational crime fighting. In Intelligence and Security Informatics (pp. 418-425). Springer Berlin Heidelberg.

Alhoniemi, E., Himberg, J., Parviainen, J., & Vesanto, J. (2012, December). *SOM Toolbox 2.1 for MATLAB.* Retrieved from: https://github.com/ilarinieminen/SOM-Toolbox

Arastey, G. M. (2020, April 21). *Setting up performance analysis equipment on matchday*. Retrieved from Sport Performance Analysis: www.sportperformanceanalysis.com/article/setting-up-performance-analysis-equipment-matchday

Attias, R. (2014). Big Data: The next revolution of sport. *Huffington Post Sport*. Retrieved from: http://www.huffingtonpost.com/richard-attias/big-data-the-next-revolut_b_5800342.html

Barrick, M., & Mount, M. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology, 44*(1), 1-26. Retrieved from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1744-6570.1991.tb00688.x

Barris, S., & Button, C. (2008). A review of vision-based motion analysis in sport. *Sports Medicine, 38*(12), 1025-43.

Bartlett, R. (2001). Performance analysis: can bringing together biomechanics and notational analysis benefit coaches? *International Journal of Performance Analysis in Sport, 1,* 122-126.

Bartlett, R. (2006) Artificial intelligence in sports biomechanics: New dawn or false hope? *Journal of Sports Science and* Medicine, 5, 474-479.

Batista, J., Goncalves, B., Sampaio, J., Castro, J., Abade, E., & Travassos, B. (2019). The influence of coaches' instruction on technical actions, tactical behaviour, and external workload in football small-sided games. *Montenegrin Journal of Sports Science and Medicine, 8(1)*, 29.

Baxter, K., Courage, C., & Caine, K. (2015). Evaluation Methods. In K. Baxter, C. Courage, & K. Caine, *In interactive technologies, understanding your users (Second Edition)* (pp. 430-446). Morgan Kaufmann. doi: https://doi.org/10.1016/B978-0-12-800232-2.00014-6

Bellman, R. E. (2003). *Dynamic Programming.* Courier Dover Publications.

Bersin , J., & Zao-Sanders , M. (2020, February 12). *Boost your team's data literacy*. Retrieved from Harvard Business Review: https://hbr.org/2020/02/boost-your-teams-data-literacy

Bloom, G., Crumpton, R., & Anderson, J. (1999). A systematic observation study of the teaching behaviours of an expert basketball coach. *The Sport Psychologist, 13,* 157-170.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101.

Brown, L. (2020, June 3rd). *Observational Field Research*. Retrieved from faculty Washington EDU: https://faculty.washington.edu/wpratt/MEBI598/Methods/Collecting%20Data%20Through%20Observation.htm

Bruce, L., Brooks, E., & Woods, C. (2018). Team and seasonal performance indicator evolution in the ANZ Championship netball league. *Journal of Sports Sciences*. doi:10.1080/02640414.2018.1473099

Carter, A. (1996). Time and motion analysis and heart rate monitoring of a back-row forward in first class rugby union football. In Hughes, M. (Eds.), *Notational Analysis of Sport*, I & II (pp. 145-160). Cardiff: UWIC.

Cassidy, T., Jones, R., & Potrac, P. (2009). *Understanding sports coaching: The social, cultural and pedagogical foundations of coaching practice (2nd ed.).* London: Routledge.

Casual Inference. (2019, April 13). *DIY Metrics: Why do we normalize*. Retrieved from Casual Inference: https://casual-inference.com/post/diy-metrics-why-normalize/

Champion Data. (2020, December 10). *ANZ Premiership*. Retrieved from Champion Data: https://mc.championdata.com/anz_premiership/

Chandler, P., Pinder , S., & Curran, J. (2014). Physical demands of training and competition in collegiate netball players. *Journal of Strength and Conditioning Research., 28*(10), 2732–7.

Cheolkon. J, & Joongkyu, K. (2008). GNAVI: Golf navigation system based on player information. In *Proceedings of the 2008 IEEE International Conference on Multimedia and Expo* (pp. 1413 – 1416), Hannover: Germany.

Chow, J., Davids, K., Button, C., & Koh, M. (2005). Organization of motor system degrees of freedom during the soccer chip: An analysis of skilled performance. *International Journal of Sport Psychology, 37*, 207–229.

Chi, E. (2008) Sensors in sport. In Dabnichki, P., & Baca, A. (Eds.), *Computers in Sport* (pp. 249-267), Southampton: WIT Press.

Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM, 13*(6), 377-387.

Connell, G., & Spencer, K. (2020). A player effectiveness analysis system in elite football using an action design research framework. *South African Journal for Research in Sport, Physical*

Education and Recreation, 42(3), 1-10. Retrieved from: https://www.ajol.info/index.php/sajrs/article/view/202658

Coutts, A., & Duffield, R. (2010). Validity and reliability of GPS devices for measuring movement demands of team sports. *Journal of Science and Medicine in Sport, 13*(1), 133-135.

Cox, T., & Dunn, R. (2002). An analysis of decathlon data. *Journal of the Royal Statistical Society, 51*, 179–187. doi: 10.1111/1467-9884.00310

Croft, H., & Spencer, K. (2021b). Coach engagement with live netball statistical information. *Journal of Sport Sciences (submitted)*.

Croft, H., Spencer, K., & Robertson, S. (2020). What tactical and technical comments do coaches make during netball matches? A content analysis in netball. *International Journal of Performance Analysis in Sport*. doi:DOI: 10.1080/24748668.2020.1846112

Croft, H., & Spencer, K. (2021a). An evaluation of data sources in netball performance analysis. *(Under review) The Journal of Sport and Exercise Science*.

Croft, H., Lamb, P., & Middlemas, S. (2015). The application of self-organising maps to performance analysis data in rugby union. *International Journal of Performance Analysis in Sport, 15*, 1037–1046.

Croft, H., Spencer, K., Robinson, S., & Cronin, J. (2021). The problem with big data and coaching: closing the gap – a review. *Submitted to Journal of Human Sport and Exercise* , In review.

Croft, H., Spencer, K., Taurua, N., & Wilton, E. (2021). Creating a live and flexible normative dataset for netball. *Frontiers in Sport and Active Living*. https://doi.org/10.3389/fspor.2021.743612

Croft, H., Willcox, B., & Lamb, P. (2018). Using performance data to identify styles of play in netball: An alternative to performance indicators. *International Journal of Performance Analysis in Sport, 17*(6), 1034-1043.

Cushion, C. (2007). Modelling the complexity of the coaching process. *International Journal of Sports Science & Coaching, 2*(4), 395-401. doi:10.1260/174795407783359650

Cushion, C. J., & Townsend, R. C. (2018). Technology-enhanced learning in coaching: a review of literature. *Educational Review*, 1-21. doi:10.1080/00131911.2018.1457010

Cushion, C., Harvey, S., Muir, B., & Nelson, L. (2012). Developing the coach analysis and intervention system (CAIS): Establishing validity and reliability of a computerised systematic observation instrument. *Journal of Sport Science, 30*(2), 201-16. doi:10.1080/02640414.2011.635310

Data, C. (2019, 12 10). *ANZ Premiership*. Retrieved from Champion Data: https://mc.championdata.com/anz_premiership/

Davids, K. (1988). Ecological validity in understanding sport performance: Some problems of definition. *Quest, 40*(2), 126-136. Retrieved from: https://tandfonline.com/doi/abs/10.1080/00336297.1988.10483894

Davidson, A., & Trewartha, G. (2008). Understanding the physical demands of netball: A time-motion investigation. *International Journal of Performance Analysis in Sport, 8*, 1–17.

Davies, D., & Bouldin, D. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 1*(2), 224–227.

Delextrat, A., & Goss-Sampson, M. (2010). Kinematic analysis of netball goal shooting: A comparison of junior and senior players. *Journal of Sports Sciences, 28*(12), 1299–1307.

Drust, B. (2010). Performance analysis research: Meeting the challenge. *Journal of Sports Sciences, 28(9),* 921–922. https://doi.org/10.1080/02640411003740769

Edwards, S. (2017). Reflecting differently. New dimensions: reflection-before-action and reflection-beyond-action. *International Practice Development Journal, 7*(1), 1-14. doi:10.19043/ipdj.71.002

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing, 62*(1), 107-115. Retrieved from: https://ncbi.nlm.nih.gov/pubmed/18352969

English Institute of Sport. (2021). *Sports Performance Analysis*. Retrieved May 10, 2021, from https://www.eis2win.co.uk/service/performance-analysis/

Erickson, K. (2009). *State space grids: First application of a novel methodology to examine coach athlete interactions in competitive youth sport.* (Master Thesis, Queen's University, Kingston, Ontario, Canada) Retrieved from: https://qspace.library.queensu.ca/bitstream/handle/1974/5357/Erickson_Karl_T_200912_MA.pdf;jsessionid=47AA9103FA44C2FC5AB8DC2DD3946159?sequence=1

Escot, R. (2003). '1987–2003. How the world cup has changed', *L'Equipe Magazine, (115),* 46–54.

Etzioni, O., Tuchinda, R., Knoblock, C., Yates, A. (2003). To buy or not to buy: mining airfare data to minimize ticket purchase price. In *KDD '03 Proceedings of the ninth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 119-128), Washington, DC: USA.

Ferreira, M., & Spamer, E. (2010). Biomechanical, anthropometrical and physical profile of elite university netball players and the relationship to musculoskeletal injuries. *South African Journal for Research in Sport, Physical Education and Recreation, 32*(1), 57-67.

Fox, D., O'Malley, E., & Blake, C. (2014). Normative data for the Functional Movement Screen™ in male Gaelic field sports. *Physical Therapy in Sport, 15*(3), 194-199.

Fox, S., Spittle, M., Otago, L., & Saunders, N. (2013). Descriptive analysis of landings during international netball competition : Enhancing ecological validity of laboratory testing environments. *International Journal of Performance Analysis in Sport, 13*(3), 690-702. Retrieved from: http://dro.deakin.edu.au/view/du:30058926

Franks, I. (2004). The need for feedback. In M. Hughes & I. Franks (Eds.), *Notational Analysis of Sport* (pp. 8-16). London and New York: Routledge.

Franks, I., & Miller, G. (1991). Training coaches to observe and remember. *Journal of Sports Sciences, 9*(3), 285-297.

Friendly, M., & Denis, D. (2001). *Milestones in the history of thematic cartography, statistical graphics, and data visualization.* Web document, http://www.datavis.ca/milestones/. Accessed: February 9, 2017.

Forzoni, R. (2006). The psychological effects of video. *Sport & Exercise Scientist, 7*, 10-11.

García, J., Ibáñez, S., De Santos, R., Leite, N., & Sampaio, J. (2013). Identifying basketball performance indicators in regular season and playoff games. *Journal of human kinetics*, *36*, 161.

Gibbs, G., Friese, S., & Mangabeira, W. (2002). The use of new technology in qualitative research. *Qualitative Social Research, 3*(2). Retrieved from: http://www.qualitative-research.net/index.php/fqs/article/view/847/1840

Gil-Arias, A., Moreno, M., García-Mas, A., Moreno, A., García-González, L., & del Villar, F. (2016). Reasoning and Action: Implementation of a Decision-Making Program in Sport. *The Spanish Journal of Psychology, 19*, E60. https://doi.org/10.1017/sjp.2016.58

Gil-Arias, A., Garcia-Gonzalez, L., Del Villar Alvarez, F., & Iglesias Gallego, D. (2019). Developing sport expertise in youth sport: A decision training program in basketball. *PeerJ, 7,* e7392. https://doi.org/10.7717/peerj.7392

Glazier, P. (2010). Game, set and match? Substantive issues and future directions in performance analysis. *Sports Medicine, 40* (8), 625-634.

Glazier, P. S. (2017). Towards a grand unified theory of sports performance. *Human Movement Science, 56*, 139-156. doi:10.1016/j.humov.2015.08.001

Glory League. (2021, June 29). *What are the benefits of glory league?* Retrieved from: https://gloryleague.basketball/

Greenwald, R., Gwin, J., Chu, J., and Crisco, J. (2008). Head impact severity measures for evaluating mild traumatic brain injury risk exposure. *Neurosurgery, 62*, 789–798.

Groom, R., & Cushion, C. (2005). Using of video based coaching with players: A case study. *International Journal of Performance Analysis in Sport(e), 5*(3), 40-46.

Groom, R., Cushion, C., & Nelson, L. (2011). The delivery of video-based performance analysis by England youth soccer coaches: Towards a grounded theory. *Journal of Applied Sport Psychology, 23*, 16-32. doi:10.1080/10413200.2010.511422

Groom, R., Cushion, C., & Nelson, L. (2012). Analysing coach–athlete 'talk in interaction' within the delivery of video-based performance feedback in elite youth soccer. *Qualitative Research in Sport, Exercise and Health, 4(3)*, 439–458. https://doi.org/10.1080/2159676X.2012.693525

Guadagnoli, M., Holcomb, W., & Davis, M. (2002). The efficacy of video feedback of learning the golf swing. *Journal of Sports Sciences, 20*, 615–622.

Hale, S., & O'Donoghue, P. (2007). Addressing turning and direction changes when using the Bloomfield Movement Classification. *International Journal of Performance Analysis in Sport, 7*, 84–89.

Hartson, R., & Pyla, P. S. (2012). Chapter 14 - Rigorous Empirical Evaluation: Preparation. In R. Hartson, & P. S. Pyla, *The UX Book* (pp. 503-536). Burlington, Massachusetts, United States: Morgan Kaufmann. doi:10.1016/C2010-0-66326-7

Henderson, K. (2011). Post-positivism and the pragmatics of leisure research. *Leisure Sciences, 33*(4), 341-346. doi:10.1080/01490400.2011.583166

Hewit, J., Cronin, J., Button, C., & Hume, P. (2011). Assessing agility in netball players : A thesis submitted to Auckland University of Technology in fulfilment of the degree of Doctor of Philosophy (PhD), April 2011.

Hetherington, S., King, S., Visentin, D., & Bird, M. (2009). A kinematic and kinetic case study of a netball shoulder pass. *International Journal of Exercise Science, 2*, 243-253.

Higgins, T., Naughton, G., & Burgess, D. (2009). Effects of wearing compression garments on physiological and performance measures in a simulated game-specific circuit for netball. *Journal of Science and Medicine in Sport, 12(*1), 223–6.

Hiscock, D., Dawson, B., Heasman, J., & Peeling, P. (2012) Game movements and player performance in the Australian Football League, International Journal of Performance Analysis in Sport, 12:3, 531-545, DOI: 10.1080/24748668.2012.11868617

Hodges, N., & Franks, I. (2004). The nature of feedback. In M. Hughes & I. Franks (Eds.), *Notational Analysis of Sport,* 17-39. London and New York: Routledge.

Holt, N. L., & Sparkes, A. C. (2001). An ethnographic study of cohesiveness in a college soccer team over a season. The Sport Psychologist, 15(3), 237–259.

Horton, S., Baker, J., Deakin, J. (2005). Experts in action: A systematic observation of 5 national team coaches*. International Journal of Sport Psychology, 36(4*), 299-319.

Holleman, G., Hooge, I., Kemner, C. & Hessels, R. (2020). The 'real-world approach' and its problems: A critique of the term ecological validity. *Frontiers in Psychology, 11*. doi: https://doi.org/10.3389/fpsyg.2020.00721

Homayounfar, Z., & Andrew, T. (2019). Wearable sensors for monitoring human motion: A review on mechanisms, materials, and challenges. *SLAS Technology, 25*(1), 9-24. doi:10.1177/2472630319891128

Hopkins, W. (1997). *Sportscience: New view of stats*. Retrieved from Percentile ranges: https://sportscience.sportsci.org/resource/stats/index.html

Horton, S., Baker, J., Deakin, J. (2005). Experts in action: A systematic observation of 5 national team coaches. *International Journal of Sport Psychology, 36(4*), 299-319.

Hua-Yong Liu, H. (2005). A sports video browsing and retrieval system based on multimodal analysis: In *SportsBR. Proceedings of 2005 International Conference on Machine Learning and Cybernetics*, (pp. 5077-5081), Guangzhou: China.

Hughes, A., Barnes, A., Churchill, S., & Stone, J. (2017). The use of performance indicators that discriminate winning and losing in elite mens and womens rugby union. *International Journal of Performance Analysis in Sport*. doi:10.1018/24748668.2017.1366759

Hughes, M., & Bartlett, R. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences, 20*(10), 739-754. Retrieved 5 10, 2021, from https://shapeamerica.tandfonline.com/doi/abs/10.1080/026404102320675602

Hughes, M., & Bartlett, R. (2008). What is performance analysis? In: M, Hughes., & I, M, Franks. (Eds.). *The essentials of performance analysis: An introduction (pp.* 8-20). London: Routledge.

Hughes, M., & Bartlett, R. (2019). What is performance analysis? In M. Hughes, I. Franks, H. Dancs, & 3rd (Ed.), *Essentials of Performance Analysis in Sport* (p. 10). Routledge.

Hughes, M., & Bell, K. (2001). Performance profiling in cricket. In M. Hughes, F. Tavares and C. Moya (Eds.) *Notational Analysis of Sport IV* (pp. 201- 208). Porto: Portugal.

Hughes, M.& Franks, I. (2004). Notational Analysis of Sport: Second edition. London: Routledge.

Hughes, M., & Franks, I. (2015). Essentials of Performance Analysis in Sport: Second edition. United Kingdom: Taylor and Francis.

Hughes, M., & Franks, I. (2015). *Essentials of performance analysis in sport* (2nd ed.). London: Routledge.

Jain, A. (2016, September 17). *The 5 V's of big data*. Retrieved July 6, 2021, from Watson Health Perspectives: https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/

Jakubowska, H. (2017). Skill Transmission, Sport and Tacit Knowledge: A Sociological Perspective (Vol. 4). Taylor & Francis.

Jaric, S. (2002). Muscle strength testing: use of normalisation for body size. *Sports Medicine(NZ), 32*(10), 615-31. doi:10.2165/00007256-200232100-00002

Jones, N., Mellalieu, S., & James, N. (2004). Team performance indicators as a function of winning and losing in rugby union. *International Journal of Performance Analysis in Sport, 4*(1), 61-71.

Kabir, N., & Carayannis, E. (2013). Big data, tacit knowledge and organizational competitiveness. *Journal of Intelligence Studies in Business, 3*, 54-62. doi:10.13140/RG.2.1.4183.6960

Kaski, S., & Lagus, K. (1996). Comparing self-organizing maps. *International Conference on Artificial Neural Networks* (pp. 809–814). Berlin: Springer.

Kaski, S. (1997). World poverty map. Neural networks research centre. Helsinki University of Technology, Finland. Retrieved from: http://www.cis.hut.fi/research/som-research/worldmap.html

Kaushik, A. (2016, April 4). *A Great Analyst's Best Friends: Skepticism and Wisdom.* Retrieved July 1, 2021, from Occam's Razor: https://www.kaushik.net/avinash/great-analyst-skills-skepticism-wisdom/

Keim, D., Mansmann, F., Schneidewind, J., & Ziegler, H. (2006). Challenges in visual data analysis. *Information visualization,* 9-16. DOI: 10.1109/IV.2006.31

Kelleher, W. (2015, December 5). Daily Mail. *Glasgow Warriors' Mike Blair to be first rugby player to wear a microphone in Sky Sports coverage of PRO12 match*. Retrieved from: http://www.dailymail.co.uk/sport/rugbyunion/article-3346878/Glasgow-Warriors-Mike-Blair-Rugby-player-wear-microphone-Sky-Sports-coverage-PRO12-match.html

Kinnunen, H., & Heikkila, I. (1998). The timing accuracy of the Polar Vantage NV heart rate monitor. *Journal of Sports Sciences, 16*, 5107-5110.

Kitchin, R. (2015). Big data and official statistics: Opportunities, challenges and risks. *Statistical Journal of the IAOS, 31*, 471-481.

Krippendorff, K. (2004). Content analysis: An introduction to its methodology. California: Sage. pp. 87–89. ISBN 0-7619-1544-3.

Kohonen, T. (2013). Essentials of the self-organizing map. *Neural Networks, 37*, 52–65.

Koladycz, R., Fernandez, G., Gray, K., & Marriott, H. (2018). The Net Promoter Score (NPS) for insight into client experiences in sexual and reproductive health clinics. . *Glob Health Sci Pract, 6*(3), 413-424. doi:10.9745/GHSP-D-18-00068

Koltay, T. (2019). Research data management and data literacy as we see them today . *Library and Information Science XXVIII*, 7-17. Retrieved from: https://www.researchgate.net/publication/337155442_Research_Data_Management_and_Data_Literacy_as_We_See_Them_Today

Kraak, W., & Welman, K. (2014). Ruck-play as performance indicator during the 2010 six nations championship. *International Journal of Sports Science & Coaching, 9*(3), 525-537. doi:https://doi.org/10.1260/1747-9541.9.3.525

Kraak, W., Venter, R. and Coetzee, F. (2016). Scoring and general match profile of Super Rugby between 2008 and 2013. *International Journal of Performance Analysis in Sport 16(2). Retrieved from: http://www.tandfonline.com/doi/citedby/10.1080/24748668.2016.11868923?scroll=top&needAccess=true*

Kraak, W., Magwa, Z., & Terblanche, E. (2018). Analysis of South African semi-elite rugby head coaches' engagement with performance analysis. *International Journal of Performance Analysis in Sport, 18*(2), 350-366.

Laird, P., & Waters, L. (2008). Eyewitness Recollection of Sport Coaches. *International Journal of Performance Analysis in Sport*, *8*(1), 76–84. https://doi.org/10.1080/24748668.2008.11868424Lamb, P., Bartlett, R., & Robins, A. (2010). Self-organising maps: An objective method for clustering complex human movement. *International Journal of Computer Science in Sport, 9*, 20–29.

Lamb, P., & Croft, H. (2016). Visualizing rugby game styles using self-organising maps. *IEEE Computer Graphics & Applications, 36*(6), 11–15.

Lamb, P., Bartlett, R., & Robins, A. (2011). Artificial neural networks for analyzing inter-limb coordination: The golf chip shot. *Human Movement Science, 30*, 1129–1143.

Lapham, A., & Bartlett, R. (1995) The use of artificial intelligence in the analysis of sports performance: A review of applications in human gait analysis and future directions for sports biomechanics, *Journal of Sports Sciences, 13*(3), 229-237. DOI: 10.1080/02640419508732232

Lago-Peñas, C., Lago-Ballesteros, J., & Rey, E. (2011). Differences in performance indicators between winning and losing teams in the UEFA Champions League. *Journal of human kinetics*, *27*(1), 135-146.

Laukkanen, R., & Virtanen, P. (1998). Heart rate monitors: State of the art. *Journal of Sports Sciences, 16*, 53-57.

Lee, J., Nam, D., Lee, J., Moon, S., Kim, K., & Kim, H. (2017, February). A study on composition of context-based soccer analysis system. In 2017 19th International Conference on Advanced Communication Technology (ICACT) (pp. 886-889). IEEE.

Lees A. (2002). Technique analysis in sports: a critical review. *Journal of Sports Sciences,* 20, 813-28

Legg, P., Chung, D., Parry, M., Jones, M., Long, R., Griffiths, I., & Chen, M. (2012) Matchpad: Interactive glyph-based visualization for real-time sports performance analysis. *Computer Graphics Forum, 31*, (3), 1255–1264. DOI: 10.1111/j.1467-8659.2012.03118.x

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

Lowe J., & Matthee M. (2020). Requirements of data visualisation tools to analyse big data: A structured literature Review. In: Hattingh M., Matthee M., Smuts H., Pappas I., Dwivedi Y., Mäntymäki M. (eds) Responsible design, implementation and use of information and communication technology. I3E 2020. Lecture Notes in Computer Science, vol 12066. Springer, Cham. https://doi.org/10.1007/978-3-030-44999-5_39

Mackenzie, R., & Cushion, C. (2013). Performance analysis in football: A critical review and implications for future research. *Journal of Sports Sciences, 31*(6), 1-38. DOI:10.1080/02640414.2012.746720

Marr, M. (2015). Big data: The winning formula in sports. *Forbes/Tech.* Retrieved from: http://www.forbes.com/sites/bernardmarr/2015/03/25/big-data-the-winning-formula-in-sports/

Martínez, J., Gómez-López, P., Femia, P., Mayorga-Vega, D., & Viciana, J. (2016). Effect of augmented verbal and visual feedback on efficiency in skiing teaching. *Kinesiology, 48(1)*, 49–57. https://doi.org/10.26582/k.48.1.12

Mason, R., Farrow, D., & Hattie, J. (2021). An exploratory investigation into the reception of verbal and video feedback provided to players in an Australian Football League club. *International Journal of Sports Science & Coaching, 16(1)*, 181–191. https://doi.org/10.1177/1747954120951080

Mayer-Schönberger, V., & Cukier, K. (2014). Big data: a revolution that will transform how we live, work, and think. *American Journal of Epidemiology, 179*(9), Pages 1143–1144. doi:doi.org/10.1093/aje/kwu085

McLean, S., Salmon, P., Gorman, A., Read, G., & Solomon, C. (2017). What's in a game? A systems approach to enhancing performance analysis in football. *PLoS ONE, 12(2)*, 1–16. https://doi.org/10.1371/journal.pone.0172565

Mclean, S., Hulme, A., Mooney, M., Read, G., Bedford, A., & Salmon, P. (2019). A systems approach to performance analysis in women's netball: Using work domain analysis to model elite netball performance. *Frontiers in Psychology, 10(201)*, 1–13. https://doi.org/10.3389/fpsyg.2019.00201

McNerney, S. (2011). on September 22, 2011. Cognitive biases in sports: the irrationality of coaches, commentators and fans. Scientific American. Retrieved from: https://blogs.scientificamerican.com/guest-blog/cognitive-biases-in-sports-the-irrationality-of-coaches-commentators-and-fans/

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and brain sciences, 34*, 57–111.

Middleton, F. (2020, June 19th). *The four types of validity*. Retrieved from Scribbr: https://www.scribbr.com/methodology/types-of-validity/

Middlemas, S. (2014). *The impact of video-based practice on the development of elite youth footballers* (Doctoral Thesis. Loughborough University, England). Retrieved from: https://dspace.lboro.ac.uk/2134/16215

Middlemas, S., Croft, H., & Watson, F. (2017). Behind closed doors: The role of debriefing and feedback in a professional rugby team. *International Journal of Sport Science and Coaching*, 1-12. doi:https://doi.org/10.1177/1747954117739548

Middlemas, S., & Harwood, C. (2018). No Place to Hide: Football Players' and Coaches' Perceptions of the Psychological Factors Influencing Video Feedback. *Journal of Applied Sport Psychology, 30(1)*, 23–44. https://doi.org/10.1080/10413200.2017.1302020

Mishori, R., Singh, L., Levy, B., & Newport, C. (2014). Mapping physician twitter networks: describing how they work as a first step in understanding connectivity, information flow, and message diffusion. *Journal of Medical Internet Research. 16*(4). DOI: 10.2196/jmir.3006

Mooney, R., Corley, G., Godfrey, A., Osborough, C., Newell, J., Quinlan, L., & ÓLaighin, G. (2016). Analysis of swimming performance: perceptions and practices of US-based swimming coaches. *Journal of Sports Sciences*, *34*(11), 997–1005. http://search.ebscohost.com/login.aspx?direct=true&db=s3h&AN=113738933&site=ehost-live

More, K., & Franks, I. (1996). Analysis and modification of verbal coaching behaviour: The usefulness of a data-driven intervention strategy. *Journal of sport science, 14*, 523-543.

Nash, C., & Collins, D. (2006). Tacit knowledge in expert coaching: science or art? *Quest, 58*(4), 464-476. doi:10.1080/00336297.2006.10491894

Nelson, L., & Groom, R. (2012). The analysis of athletic performance: some practical and philosophical considerations. *Sport, Education and Society, 17*(5), 687-701. http://dx.doi.org/10.1080/13573322.2011.552574

Newell, K., Carlton, L., & Antoniou, A. (1990). The interaction of criterion and feedback information in learning a drawing task. *Journal of Motor Behaviour, 22*, 8-20.

Neuendorf, K. (2002). The content analysis guidebook. Thousand Oaks, CA: Sage. Retrieved from: http://academic.csuohio.edu/kneuendorf/SkalskiVitae/SkalskiNeuendorfCajigas17.pdf

Nicholls, S. & Worsfold, P. (2006). The observational analysis of elite coaches within youth soccer: the importance of performance analysis. *International Journal of Sports Science and Coaching,11,* 825–831.

Nicholls, S., & Worsfold, P. (2016). The observational analysis of elite coaches within youth soccer: The importance of performance analysis. *International Journal of Sports Science & Coaching, 11(6)*, 825–831. https://doi.org/10.1177/1747954116676109

Nicholls, S., James, N., Bryant, E., & Wells, J. (2018). The implementation of performance analysis and feedback within Olympic sport: The performance analyst's perspective. *International Journal of Sport Science and Coaching, 14*(1), 63-71. doi:https://doi.org/10.1177/1747954118808081

O Brien, P., Martin, D., & Bradley, J. (2021). Differences in performance indicators between winners and losers in senior inter county hurling championship. *International Journal of Performance Analysis in Sport,* 1-11.

O'Connor, P. (1990). Normative data: their definition, interpretation, and importance for primary care physicians. . 1990;22(4):307-311. *Fam Med, 22*(4), 307-311.

O'Donoghue, P. (2005). Normative profiles of sports performance. *International Journal of Performance Analysis in Sport, 5*(1), 104-119. doi:10.1080/24748668.2005.11868319

O'Donoghue, P. (2009). *Research methods for sports performance analysis*. Taylor & Francis: United Kingdom. Retrieved from: https://books.google.co.nz/books?id=9Kd5AgAAQBAJ

O'Donoghue, P., & Holmes, L. (2015). *Data Analysis in Sport* (Vol. Routledge Studies in Sports Performance Analysis). London: Routledge.

O'Donoghue, P., Mayes, A., Edwards, K., & Garland, J. (2008). Performance norms for british national super league netball. *International Journal of Sports Science & Coaching, 3*(4), 501-511.

O'Donoghue, P. (2006). The use of feedback videos in sport. *International Journal of Performance Analysis in Sport, 6*(2), 1-14. doi:DOI: 10.1080/24748668.2006.11868368

O'Donoghue, P. (2015). *An introduction to performance analysis of sport. Routledge Studies in Sports Performance Analysis.* London: Routledge.

O'Donoghue, P., Holmes, L., & Robinson, G. (2018). *Doing a research project in sport performance analysis.* Abingdon, Oxon, United Kingdom: Routledge.

O'Donoghue, P., Mayes, A., Edwards, K., & Garland, J. (2008). Performance norms for British National Super League Netball. *International Journal of Sports Science & Coaching, 3*(4), 501–511.

Ofoghi, B., Zeleznikow, J., MacMahon, C., & Raab, M. (2013). Data mining in elite sports: a review and a framework. Measurement in Physical Education and Exercise Science, 17(3), 171-186.

Ofoghi, B., Chenaghlou, M., Mooney, M., Dwyer, D. B., & Bruce, L. (2021). Team technical performance characteristics and their association with match outcome in elite netball. *International Journal of Performance Analysis in Sport,* 1-13.

Orad, A.: 2019 Gartner Magic Quadrant (2019). https://www.sisense.com/gartner-magic-quadrant-business-intelligence/. Accessed 10 July 2021

Paget, N., Spencer, K., & Kilding, A. (2015). Performance analysis of highperformance netball umpires for match-play and fitness demands: a thesis submitted to Auckland University of Technology.

Painczyk, H., Hendricks, S., & Kraak, W. (2017). Utilisation of performance analysis among Western Province Rugby Union club coaches. *International Journal of Performance Analysis in Sport*, *17*(6), 1057-1072.

Palmius, J. (2007). Criteria for measuring and comparing information systems. In *30th Information Systems Research Seminar in Scandinavia* .IRIS (30). Retrieved from: http://www.palmius.com/joel/text/IRIS-30-final.pdf

Parthiban, P. (2020, May 26). *My experience with a doctoral thesis in format two*. Retrieved from Thesislink: https://thesislink.aut.ac.nz/?p=8077

Passfield, L., & Hopker, J. G. (2017). A mine of information: can sports analytics provide wisdom from your data? . *International journal of sports physiology and performance*, 851-855. doi:10.1123/ijspp.2016-0644.

Petersson, A. M., & Lundberg, J. (2016). Applying action design research (ADR) to develop concept generation and selection methods. *Procedia CIRP 50* (pp. 222-227). Elsevier. Retrieved from https://reader.elsevier.com/reader/sd/pii/S2212827116304772?token=7F668E190424689F64402BE3DD718BC4A674E0F07B27C8422DD5C0E2DE69180C649C55901C5F9FDD685290FB20EEFA8A&originRegion=us-east-1&originCreation=20210630025814

Point 9 Analytics. (2020, December 10). *2019 ANZ preseason*. Retrieved from Point 9 Analytics: www.point9analytics.co.nz/2019-anz-preseason/

Pollard, R. (2002). Charles Reep (1904-2002): Pioneer of notational and performance analysis in football, *Journal of Sports Sciences, 20*(10), 853-855. DOI: 10.1080/026404102320675684

Provost, L. (2016). Web Query Excel 2016: Importing data from a website to your spreadsheet. https://www.youtube.com/watch?v=2bYXvtpSeEg&feature=emb_title (accessed: 11/8/2020). Wayne State University.

Ransdell, S. E. (1993). Educational software evaluation research: Balancing internal, external, and ecological validity. *Behavior Research Methods Instruments & Computers, 25*(2), 228-232. Retrieved 2 24, 2021, from https://link.springer.com/article/10.3758/bf03204502

Reed, D. & Hughes, H. (2006) An exploration of team sport as a dynamical system. *International Journal of Performance Analysis in Sport, 6(2)*, 114-125. DOI: 10.1080/24748668.2006.11868377

Reichheld , F. F., & Markey, R. (2011). How net promoter companies thrive in a customer-driven world. *Harvard Business Review Press*.

Rein, R., Davids, K., & Button, C. (2010). Adaptive and phase transition behaviour in performance of discrete multi-articular actions by degenerate neurobiological systems. *Experimental Brain Research, 201,* 307–322.

Robinson, G. and O'Donoghue, P. (2007). A weighted kappa statistic for reliability testing in performance analysis of sport. *International Journal of Performance Analysis in Sport 7(1),* pp. 12-19(8).

Ryan, A. B. (2006). *Post-Positivist Approaches to Research.* Retrieved May 13, 2021, from MURAL - Maynooth University Research Archive Library: http://mural.maynoothuniversity.ie/874/

Sandelowski, M. (2000). Whatever happened to qualitative description? *Focus on Research Methods, 23*(4), 334-340. doi:10.1002/1098-240X(200008)23:4<334::AID-NUR9>3.0.CO;2-G

Schoeman, J. H., Matthee, M. C., & van der Merwe, P. (2006). The viability of business data mining in the sports environment: Cricket match analysis as application. *South African Journal for Research in Sport, Physical Education & Recreation, 28*(1), 109-125.

Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly, 35*(1), 37-56. doi:10.2307/23043488

Shaw, B., & Jebara, T. (2009). Structure preserving embedding. *In Proceedings of the 26th Annual International Conference on Machine Learning.* Montreal. doi:10.1145/1553374.1553494

Silva, A. J., Costa, A. M., & Oliveira, P. M. (2007). The use of neural network technology to model swimming performance. *Journal of Sports Science and Medicine. 6*(1), 117–125.

Smith, B., & Sparkes, A. C. (2006). Narrative inquiry in psychology: Exploring the tensions within. Qualitative Research in Psychology, 3(3), 169–192.

Smith, P. D., & Bedford, A. (2020). Automatic Classification of Locomotion in Sport: A Case Study from Elite Netball. *International Journal of Computer Science in Sport, 19(2),* 1–20. https://doi.org/10.2478/ijcss-2020-0007

Spencer, K., Natasha, P., Farley, O. R., & Kilding, A. E. (2019). Activity profile of elite netball umpires during match play. *Journal of Strength and Conditoning Research*.

Sport and Exercise Science New Zealand. (2019). *Performance Analysis*. Retrieved May 10, 2021, from https://sesnz.org.nz/accreditation/accredited-exercise-and-sport-science-practitioner/performance-analysis/

Sternberg, R. J., & Horvath, J. A. (1999). *Tacit knowledge in professional practice: Researcher and practitioner perspectives.* Mahwah, NJ, USA: Lawrence Erlbaum Associates Publishers.

Stratton, G., Reilly, T., Williams, A., & Richardson, D. (2004). Science of youth soccer. London: Routledge.

Sullivan, C., Bilsborough, J. C., Cianciosi, M., Hocking, J., Cordy, J., & Coutts, A. J. (2014). Match score affects activity profile and skill performance in professional Australian Football players. *Journal of Science and Medicine in Sport, 17(3)*, 326-331.

Swinnen, S. P., Schmidt, R. A., Nicholson, D. E., & Shapiro, D. C. (1990). Information feedback for skill acquisition: Instantaneous knowledge of results degrades learning. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*(4), 706-716.

Target Internet. (2020, December 10). What is data scraping and how can you use it? BLOG POSTS, ANALYTICS & DATA. Retrieved from Target Internet: https://targetinternet.com/what-is-data-scraping-and-how-can-you-use-it

Taylor, M., & Hughes, M. (1998). Analysis of elite under-19 tennis players. In A. Lees, I. Maynard, M.D. Hughes and T. Reilly (Eds.), In *Science and Racket Sports II.* (pp. 211- 220). London: E & FN Spon.

Tharp, R. G., & Gallimore, R. (1976). What a coach can teach a teacher. *Psychology Today, 9*(8):75-78

The ibi Team. (2017, November 30). *Contextualization: What is it, and how does it help you manage big data?* Retrieved from Information Builders: https://www.ibi.com/blog/ibi-team/contextualization-what-it-and-how-does-it-help-you-manage-big/

Thomas, J., & Cook, K. (2005). Illuminating the path: The R&D agenda for visual analytics. *National Visualization and Analytics Center, 30*, 4.

Thomas, C., Ismail, K. T., Comfort, P., Jones, P. A., & Dos'Santos, T. (2016). Physical profiles of regional academy netball players. *Journal of Trainology, 5(2)*, 30-37.

Tomkinson, G. R., Carver, K. D., Atkinson, F., Daniell, N. D., Lewis, L. K., Fitzgerald, J. S., Ortega, F. B. (2018). European normative values for physical fitness in children and adolescents aged 9–17 years: results from 2 779 165 Eurofit performances representing 30 countries. *British Journal of Sports Medicine, 52*, 1445-1456.

Travassos, B., Davids, K., Araújo, D., & Esteves, T. P. (2013). Performance analysis in team sports: Advances from an Ecological Dynamics approach. *International Journal of Performance Analysis in Sport*, *13*(1), 83-95.

Trinity, J., & Annesi, J. J. (1996). Coaching with video. *Strategies, 9*, 23-25.

UmaMaheswari, P., & Rajaram, M. (2009). A novel approach for mining association rules on sports data using principal component analysis: For cricket match perspective. In *Proceedings of the IEEE International Advance Computing Conference* (pp. 1074-1080), Patiala, India.

Vanwanseele, B., Stuelcken, M., Greene, A., & Smith, R. (2014). The effect of external ankle support on knee and ankle joint movement and loading in netball players. *Journal of Science and Medicine in Sport, 17*, 511-515.

Vaz, L., Rooyen, M. V., & Sampaio, J. (2010). Rugby game-related statistics that discriminate between winning and losing teams in IRB and super twelve close games. *Journal of Sports Science and Medicine, 9,* 51–55.

Venter, R. E., Fourie, L., Ferreira, S., & Terblanche, E. (2005). Physical and physiological profiles of Boland netball players. *South African Journal of Sports Medicine, 17(2)*, 3-7.

Vermeulen, E., & Yadavalli, S. V. (2018). Big data in sport analytics: applications and risks. *Proceedings of the International Conference on Industrial Engineering and Operations Management* (pp. 1-10). Pretoria / Johannesburg, South Africa: IEOM Society. Retrieved from http://ieomsociety.org/southafrica2018/papers/14.pdf

Vesanto, J., & Alhoniemi, E. (2000). Clustering of the self-organizing map. *IEEE Transactions on Neural Networks, 11*(3), 586–600.

Vincent, J. S. (2009). The role of databases in sport science: Current practice and future potential. *International Journal of Computer Science in Sport, 8*(2).

Vincent, J., Stergiou, P., & Katz, L. (2009). The role of databases in sport science: Current practice and future potential. *International Journal of Computer Science in Sport, 8*(2), 50–66.

Waters, A., Phillips, E., Panchuk, D., & Dawson, A. (2019). The coach–scientist relationship in high-performance sport: Biomechanics and sprint coaches. *International Journal of Sports Science & Coaching, 14*(5), 617-628. doi:10.1177/1747954119859100

Weed, M. (2009). Research quality considerations for grounded theory research in sport & exercise. Psychology. Psychology of Sport and Exercise (10). pp.502–510

What is sports performance analysis. (2020, December 18). Retrieved from iSport Analysis: www.isportanalysis.com/sport-performance-analysis.php

Williams, J. S., & Kendall, L. (2007). Perceptions of elite coaches and sports scientists of the research needs for elite coaching practice. *Journal of Sports Sciences, 25*(14), 1577-1586. doi:10.1080/02640410701245550

Wilson, B. D., Pharmy , A., & Nadzrin, M. (2008). Video technology and coaching. In F. K. Fuss, A. Subic, & S. Ujihashi, *The Impact of Technology on Sport II* (pp. 897-904). London: Taylor & Francis Grou

Winkler, W. (1996). Computer/video analysis in German soccer. In Hughes, M. (Eds.). *Notational Analysis of Sport, I & II (pp. 19-31)*. Cardiff: UWIC.

Wright, C., Atkins, S., & Jones, B. (2012) An analysis of elite coaches' engagement with performance analysis services (match, notational analysis and technique analysis), International Journal of Performance Analysis in Sport, 12:2, 436-451, DOI: 10.1080/24748668.2012.11868609

Young, C. M., Luo, W., Gastin, P., Tran, J., & Dwyer, D. B. (2019). The relationship between match performance indicators and outcome in Australian Football. *Journal of science and medicine in sport*, *22*(4), 467-471.

Zetou, E., Amprasi, E., Michalopoulou, M., & Aggelousis, N. (2011). Volleyball coaches behavior assessment through systematic observation. *Journal of Human Sport Exercise*, *6*(4), 585–593. DOI: https://doi.org/10.4100/jhse.2011.64.02

Zins, C. (2007). Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology, 58*(4), 479–493. DOI:10.1002/asi

Appendix I. Ethics Approval

24 August 2016

Kirsten Spencer
Faculty of Health and Environmental Sciences

Dear Kirsten

Re Ethics Application: **16/267 Determining what topics or themes are important to performance for team sport coaches in real-time.**

Thank you for providing evidence as requested, which satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC).

Your ethics application has been approved for three years until 23 August 2019.

As part of the ethics approval process, you are required to submit the following to AUTEC:

- A brief annual progress report using form EA2, which is available online through http://www.aut.ac.nz/researchethics. When necessary this form may also be used to request an extension of the approval at least one month prior to its expiry on 23 August 2019;
- A brief report on the status of the project using form EA3, which is available online through http://www.aut.ac.nz/researchethics. This report is to be submitted either when the approval expires on 23 August 2019 or on completion of the project.

It is a condition of approval that AUTEC is notified of any adverse events or if the research does not commence. AUTEC approval needs to be sought for any alteration to the research, including any alteration of or addition to any documents that are provided to participants. You are responsible for ensuring that research undertaken under this approval occurs within the parameters outlined in the approved application.

AUTEC grants ethical approval only. If you require management approval from an institution or organisation for your research, then you will need to obtain this.

To enable us to provide you with efficient service, please use the application number and study title in all correspondence with us. If you have any enquiries about this application, or anything else, please do contact us at ethics@aut.ac.nz.

All the very best with your research,

Kate O'Connor
Executive Secretary
Auckland University of Technology Ethics Committee

Cc: Hayden Croft, hayden.croft@op.ac.nz; John Cronin

14 March 2018

Kirsten Spencer
Faculty of Health and Environmental Sciences

Dear Kirsten

Re Ethics Application: **18/68 Data sources and collection in sport: (Study 2)**

Organising, normalising and analysing different type of sports data: (Study 3)

Thank you for providing evidence as requested, which satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC).

Your ethics application has been approved for three years until 13 March 2021.

Standard Conditions of Approval

1. A progress report is due annually on the anniversary of the approval date, using form EA2, which is available online through http://www.aut.ac.nz/researchethics.
2. A final report is due at the expiration of the approval period, or, upon completion of project, using form EA3, which is available online through http://www.aut.ac.nz/researchethics.
3. Any amendments to the project must be approved by AUTEC prior to being implemented. Amendments can be requested using the EA2 form: http://www.aut.ac.nz/researchethics.
4. Any serious or unexpected adverse events must be reported to AUTEC Secretariat as a matter of priority.
5. Any unforeseen events that might affect continued ethical acceptability of the project should also be reported to the AUTEC Secretariat as a matter of priority.

Please quote the application number and title on all future correspondence related to this project.

AUTEC grants ethical approval only. If you require management approval for access for your research from another institution or organisation then you are responsible for obtaining it. If the research is undertaken outside New Zealand, you need to meet all locality legal and ethical obligations and requirements. You are reminded that it is your responsibility to ensure that the spelling and grammar of documents being provided to participants or external organisations is of a high standard.

For any enquiries, please contact ethics@aut.ac.nz

Yours sincerely,

Kate O'Connor
Executive Manager
Auckland University of Technology Ethics Committee

Cc: hayden.croft@op.ac.nz; John Cronin

22 March 2021

Kirsten Spencer
Faculty of Health and Environmental Sciences

Dear Kirsten

Re Ethics Application: **21/33 Determining what statistical information is important to team sport coaches in real-time**

Thank you for providing evidence as requested, which satisfies the points raised by the Auckland University of Technology Ethics Committee (AUTEC).

Your ethics application has been approved for three years until 22 March 2024.

Standard Conditions of Approval

1. The research is to be undertaken in accordance with the Auckland University of Technology Code of Conduct for Research and as approved by AUTEC in this application.
2. A progress report is due annually on the anniversary of the approval date, using the EA2 form.
3. A final report is due at the expiration of the approval period, or, upon completion of project, using the EA3 form.
4. Any amendments to the project must be approved by AUTEC prior to being implemented. Amendments can be requested using the EA2 form.
5. Any serious or unexpected adverse events must be reported to AUTEC Secretariat as a matter of priority.
6. Any unforeseen events that might affect continued ethical acceptability of the project should also be reported to the AUTEC Secretariat as a matter of priority.
7. It is your responsibility to ensure that the spelling and grammar of documents being provided to participants or external organisations is of a high standard and that all the dates on the documents are updated.

AUTEC grants ethical approval only. You are responsible for obtaining management approval for access for your research from any institution or organisation at which your research is being conducted and you need to meet all ethical, legal, public health, and locality obligations or requirements for the jurisdictions in which the research is being undertaken.

Please quote the application number and title on all future correspondence related to this project.

For any enquiries please contact ethics@aut.ac.nz. The forms mentioned above are available online through http://www.aut.ac.nz/research/researchethics

(This is a computer-generated letter for which no signature is required)

The AUTEC Secretariat
Auckland University of Technology Ethics Committee

Cc: hayden.croft@op.ac.nz

Appendix II. Participant Information Sheets, Survey and Consent Forms

Participant Information Sheet

Project Title

Determining what topics and themes are important to performance for team sport coaches in real-time.

An Invitation

My name is Hayden Croft and I am seeking participants for this research project. It will look to identify what types of information are important to coaches of field and court sports, when making decisions during a live match. This study is the first part of my PhD which will develop a tool to guide coaches, objectively, through large data sets allowing them to answer questions related to performance. Participation is voluntary and you may withdraw at any time however your data may only be withdrawn up until the December 31st 2017. As I already work with the Otago Rugby Team and Southern Steel Netball Team, upon request, a third party will manage your information should you wish to retain anonymity from myself. If you do not participate it will neither advantage nor disadvantage you.

What is the purpose of this research?

To gain an understanding of the tactical and technical information that coaches draw on when making decisions about team sport performance. This will allow fast identification of what information should be captured and presented to coaches during a game in an objective way. It will also identify themes that are similar between coaches allowing a more in-depth understanding of coach thought processes.

How was I identified and why am I being invited to participate in this research?

You have been approached as you meet the selection criterion being an elite level, field or court team sport coach. Your contact details have been obtained by either a high performance manager within your sport or because you know one of the researchers.

What will happen in this research?

You will be asked to wear an audio recording devices during games played by a team that you coach. At a later date, if required, you may be asked via a phone or Skype interview to explain the meaning of any language, terminology of jargon used. The captured audio will be transcribed and a content analysis will identify the different types of themes or aspects of the game that you discuss. These will be categorised and combined with other coaches data and an analysis will identify which themes are common across sports and coaches. These findings will indicate what types of information would be best provided to coaches during a match.

What are the benefits?

This research will allow us to identify what types of data or information are useful to coaches during a live sports match, assisting the creation of a tool that assists decision making. It will also identify what themes are universal for all court and field team sports. Finally this research will contribute to my PhD.

How will my privacy be protected?

As the objective of this study is to identify themes and topics that are similar across various coaches and sports, individual data will not be reported. All participant information and data will be store securely during and after data collection. All information and data that is stored will be de-identified with the used of participant numbers, to prevent outside persons identifying participants.

What are the costs of participating in this research?

There is no cost associated with participating in this research other than the time to conduct the interview, post-game, via phone or Skype if required. This would be no more than 20 minutes.

What opportunity do I have to consider this invitation?

You have two weeks, upon receipt, to consider this invitation. The cut off for participation in this study is December 31st 2017.

How do I agree to participate in this research?

To participate please complete the Consent Form, either attached or supplied by the research or you High Performance Analysis Manager. You will also need to indicate when a where an appropriate game gave will occur for the audio recording to take place.

Will I receive feedback on the results of this research?

Yes, upon request, at the completion of the data analysis, a report will be sent to you via email or you can talk directly with the Primary researcher Hayden Croft.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Kirsten Spencer, *kspencer@aut.ac.nz,* 09 921 9999 ext 7239 .

Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTEC, Kate O'Connor, *ethics@aut.ac.nz* , 09 921 9999 ext 6038.

Whom do I contact for further information about this research?

Please keep this Information Sheet and a copy of the Consent Form for your future reference. You are also able to contact the research team as follows:

Researcher Contact Details:

Hayden Croft, hayden.croft@op.ac.nz, 021760 395, Forth Street, Private Bag 1910, Dunedin 9054, New Zealand.

Project Supervisor Contact Details:

Kirsten Spencer, kspencer@aut.ac.nz, 09 921 9999 ext 7239, 17 Antares Pl, Rosedale, Auckland 0632, New Zealand

John Cronin, Jcronin@aut.ac.nz, 09 921 921 ext 7523, 17 Antares Pl, Rosedale, Auckland 0632, New Zealand

Approved by the Auckland University of Technology Ethics Committee on *24 August 2016*, AUTEC Reference number *16/267*.

Participant Information Sheet

Project Title

Study 2: DATA SOURCES AND COLLECTION IN SPORT and STUDY 3: ORGANISING, NORMALISING AND ANALYSING DIFFERENT TYPE OF SPORTS DATA

An Invitation

My name is Hayden Croft and I am seeking participants for this research project. The research will look to identify what types of information are important to coaches of field and court sports, when making decisions during a live match. This study is the first part of my PhD which will develop a tool to guide coaches, objectively, through large data sets allowing them to answer questions related to performance.

Participation is voluntary and you may withdraw at any time however your data may only be withdrawn up until the December 31st 2018.

As I have previously worked with the Otago Rugby Team and Southern Steel Netball Team, upon request, a third party will manage your information should you wish to retain anonymity from myself. However due to the size of the potential participant pool being limited, confidentiality may be limited. If you do not participate it will neither advantage nor disadvantage you.

What is the purpose of this research?

Study 2, will determine and evaluate what information systems and types of data are available, in real-time, during evasive team sport matches, that align with the topics and themes identified in Study 1. It will then evaluate, against set criteria, the suitability of these information systems for use in assisting coaches decision-making and make recommendations therewith.

Study 3 is to design a platform for capturing, processing and organising data in real-time so it can be used in future research. A database will be designed that best manages the data types allowing comparisons across data types and sources.

How was I identified and why am I being invited to participate in this research?

You have been approached as you meet the selection criterion being an elite level, field or court team sport coach, player or performance analyst. Your contact details have been obtained by either a high performance manager within your sport or because you know one of the researchers.

How do I agree to participate in this research?

Your participation in this research is voluntary (it is your choice) and whether or not you choose to participate will neither advantage nor disadvantage you. You are able to withdraw from the study at any time. If you choose to withdraw from the study, then you will be offered the choice between having any data that is identifiable as belonging to you removed or allowing it to continue to be used. However, once the findings have been produced, removal of your data may not be possible.

What will happen in this research?

The research, Hayden Croft, will observe the types of technology used in your team's environment, in preparation or during a match. A sample of each data type will be collected. The technologies suitability for live, in-game, feedback will be evaluated as well as how well they inform decision making about the tactical themes identified, from coach observations, in Study 1.

The observation will be for the duration of one training session or match and any additional interviews will be for no longer then 30mins. Interviews will be recorded and transcripts will be stored for confirmation.

What are the benefits?

This research will provide an understanding of the technologies that available in sport, which can inform the tactical and technical themes identified in Study 1, related to decision making about team sport performance. Additionally it will provide recommendations about how to capture, organise and normalised the data these technologies produce, providing framework for future research into presenting information to coaches.

There is also a benefit to Hayden Croft as this research will assist him in gaining his PhD qualification.

How will my privacy be protected?

As the objective of this study is to identify technologies and data used by teams in sport, individual data will not be reported. All participant information and data will be store securely during and after data collection. All information and data that is stored will be de-identified with the used of participant numbers, to prevent outside persons identifying participants.

Due to the researcher conducting the data collection, Hayden Croft, having connections to other sports teams in rugby and netball there could be a perceived conflict of interest. If requested a 3rd party can be organise to collect the data, however It is acknowledged that due to the small sample of potential participants, confidentiality will be very limited, and this may influence your involvement in the study.

What are the costs of participating in this research?

There is no cost associated with participating in this research other than the time to conduct communications, post-game, via phone or Skype if required. This would be no more than 30 minutes.

What opportunity do I have to consider this invitation?

You have two weeks, upon receipt, to consider this invitation. The cut off for participation in this study is June 31st 2018.

How do I agree to participate in this research?

To participate please complete the Consent Form, either attached or supplied by the research or you High Performance Analysis Manager. You will also need to indicate when a where an appropriate game gave will occur for the audio recording to take place.

Will I receive feedback on the results of this research?

Yes, upon request, at the completion of the data analysis, a report will be sent to you via email or you can talk directly with the Primary researcher Hayden Croft.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Kirsten Spencer, *kspencer@aut.ac.nz,* 09 921 9999 ext 7239 .

Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTEC, Kate O'Connor, *ethics@aut.ac.nz* , 09 921 9999 ext 6038.

Whom do I contact for further information about this research?

Please keep this Information Sheet and a copy of the Consent Form for your future reference. You are also able to contact the research team as follows:

Researcher Contact Details:

Hayden Croft, hayden.croft@op.ac.nz, 021760 395, Forth Street, Private Bag 1910, Dunedin 9054, New Zealand.

Project Supervisor Contact Details:

Kirsten Spencer, kspencer@aut.ac.nz, 09 921 9999 ext 7239*,* 17 Antares Pl, Rosedale, Auckland 0632, New Zealand

John Cronin, Jcronin@aut.ac.nz, 09 921 921 ext 7523, 17 Antares Pl, Rosedale, Auckland 0632, New Zealand

Approved by the Auckland University of Technology Ethics Committee on -- *August 2018*, AUTEC Reference number *18/68*.

Participant Information Sheet

Date Information Sheet Produced:

15 March 2021

Project Title

Determining what topics or themes are important to performance for team sport coaches in real-time.

An Invitation

We are seeking participants for this research project which will look to identify what types of performance information coaches of netball teams find beneficial when making decisions during a live match. This study is the last part of my PhD which will look to develop a tool to assist coaches, objectively, through large data sets, to answer questions related to performance. Participation is voluntary and participants may withdraw at any time prior to the completion of data collection (May 31st 2021). To avoid conflicts of interest, upon request, a third party will deidentify your information should you wish to retain anonymity. If you do not participate it will neither advantage nor disadvantage you.

What is the purpose of this research?

To gain an understanding of the tactical and technical statistical information that coaches draw on when making decisions about team sport performance. This will allow fast identification of what information should be captured and presented to coaches during a game in an objective way.

The findings of this research may be used for academic publications and presentations.

How was I identified and why am I being invited to participate in this research?

You have been approached as you meet the selection criterion being an elite level netball coach. Your contact details will be obtained upon you responding to this invitation, which will have been sent to you via through your Netball New Zealand high-performance manager. Your response, agreeing to be a participant, will not be disclosed to NNZ.

How do I agree to participate in this research?

To participate please complete the Consent Form, either attached or supplied by the research or you High Performance Analysis Manager. You will also need to indicate when a where an appropriate game gave will occur for the audio recording to take place.

Your participation in this research is voluntary (it is your choice) and whether or not you choose to participate will neither advantage nor disadvantage you. You are able to withdraw from the study at any time. If you choose to withdraw from the study, then you will be offered the choice between

having any data that is identifiable as belonging to you removed or allowing it to continue to be used. However, once the findings have been produced, removal of your data may not be possible.

What will happen in this research?

You will be given a tablet device, to use during an ANZ Premiership or Beko National League netball match that you are watching. After the match you will be asked to complete a survey questionnaire to explain how useful the live statistics were to you during the game. Your interactions with the tablet will be screen recorded, however you will not. We will use this recording to analyse the amount and depth of interaction with the statistics during the game. These will be categorised and combined with other coach's data and an analysis will identify how useful this information was during a match. These findings will provide insights into what types of information would be best provided to coaches during a match.

You will be sent a protocol sheet which will give you step by step instructions for using the tablet device.

What are the discomforts and risks?

There are no known risks or discomforts associated with participating in this research.

What are the benefits?

The benefits of this research include; participants, an understanding of their utilisation of statistics during a game and how this may be improved in the future; the researcher, to gain a PhD qualification and create a tool and framework (stats presentation in real-time) that can be used in future professional practice; wider community, to enable a process or framework for sport coaches and performance analysts to follow around the selection, organisation and presentation of performance information.

How will my privacy be protected?

As the objective of this study is to identify themes and topics that are similar across various coaches, individual data will not be reported. All participant information and data will be store securely during and after data collection. All information and data that is stored will be de-identified with the used of participant numbers, to prevent outside persons identifying participants.

What are the costs of participating in this research?

There is no cost associated with participating in this research other than the time to watch the netball game and complete the survey questionnaire.

What opportunity do I have to consider this invitation?

You have at least one month to consider this invitation.

Will I receive feedback on the results of this research?

At the completion of the data analysis one page report of results will be sent to you via email or you can talk directly with the primary researcher Hayden Croft or supervisor Kirsten Spencer.

What do I do if I have concerns about this research?

Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Kirsten Spencer, *kspencer@aut.ac.nz,* 09 921 9999 ext 7239.

Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTEC, Dr Carina Meares, *ethics@aut.ac.nz* , 09 921 9999 ext 6038.

Whom do I contact for further information about this research?

Please keep this Information Sheet and a copy of the Consent Form for your future reference. You are also able to contact the research team as follows:

Researcher Contact Details:

Hayden Croft, hayden.croft@op.ac.nz, 021760 395, Forth Street, Private Bag 1910, Dunedin 9054, New Zealand.

Project Supervisor Contact Details:

Kirsten Spencer, kspencer@aut.ac.nz, 09 921 9999 ext 7239, 17 Antares Pl, Rosedale, Auckland 0632, New Zealand

Approved by the Auckland University of Technology Ethics Committee on *type the date final ethics approval was granted*, AUTEC Reference number *type the reference number*.

Consent Form

Project title: ***Determining what topics or themes are important to performance for team sport coaches in real-time.***

Project Supervisor: ***Kirsten Spencer, Sam Robertson and John Cronin***

Researcher: ***Hayden Croft***

○ I have read and understood the information provided about this research project in the Information Sheet dated 19th May 2016.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that notes will be taken during the interviews and that they will also be audio-taped and transcribed.

○ I understand that audio recordings will be taken during games and transcribed.

○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ If I withdraw, I understand that all relevant information including tapes and transcripts, or parts thereof, will be destroyed.

○ I understand that I can request a copy of the transcribe recording for review and editing.

I understand that is may receive a summary of the research findings if desired.

○ I agree to take part in this research.

○ I wish to receive a summary of the research findings (please tick one): Yes○ No○

Participant's signature: .....

Participant's name: .....

Participant's Contact Details (if appropriate):

.....

.....

.....

Date:

Approved by the Auckland University of Technology Ethics Committee 24 August 2016. AUTEC Reference number 16/267

Note: The Participant should retain a copy of this form

Consent Form

Project title: ***Study 2: DATA SOURCES AND COLLECTION IN SPORT and STUDY 3: ORGANISING, NORMALISING AND ANALYSING DIFFERENT TYPE OF SPORTS DATA***

Project Supervisor: ***Kirsten Spencer, Sam Robertson and John Cronin***

Researcher: ***Hayden Croft***

○ I have read and understood the information provided about this research project in the Information Sheet dated 13th February 2018.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that notes will be taken during the interviews and that they will also be audio-taped and transcribed.

○ I understand that observations and data samples will be taken during games and or trainings.

○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ If I withdraw, I understand that all relevant information including transcripts and data files or parts thereof, will be destroyed.

○ I understand that I can request a copy of the transcribe recording for review and editing.

I understand that is may receive a summary of the research findings if desired.

○ I agree to take part in this research.

○ I wish to receive a summary of the research findings (please tick one): Yes○ No○

Participant's signature: .....

Participant's name: .....

Participant's Contact Details (if appropriate):

.....

.....

.....

Date:

Approved by the Auckland University of Technology Ethics Committee 24 August 2016. AUTEC Reference number 16/267

Note: The Participant should retain a copy of this form

Consent Form

Project title: ***Determining what topics or themes are important to performance for team sport coaches in real-time.***

Project Supervisor: ***Kirsten Spencer***

Researcher: ***Hayden Croft***

○ I have read and understood the information provided about this research project in the Information Sheet and Protocol sheet dated 15th March 2021.

○ I have had an opportunity to ask questions and to have them answered.

○ I understand that survey information will be analysed for patterns and responses.

○ I understand that screen recordings will be taken during games and analysed.

○ I understand that I may withdraw myself or any information that I have provided for this project at any time prior to completion of data collection, without being disadvantaged in any way.

○ If I withdraw, I understand that all relevant information including screen recordings and questionnaires will be destroyed.

○ I agree to take part in this research.

○ I wish to receive a summary of the research findings (please tick one): Yes○ No○

Participant's signature: .....

Participant's name: .....

Participant's Contact Details (if appropriate):

.....

.....

.....

.....

Date:

Approved by the Auckland University of Technology Ethics Committee on type the date on which the final approval was granted AUTEC Reference number 21/33

Note: The Participant should retain a copy of this form.

Make a dash on the line for how each level of statistics aided your decision-making

Movement

Very helpful when decision-making _____ Not used when decision-making

Actions

Very helpful when decision-making _____ Not used when decision-making

Positioning

Very helpful when decision-making _____ Not used when decision-making

Possession outcome

Very helpful when decision-making _____ Not used when decision-making

Strategy

Very easy to use and understand _____ Very confusing to use and understand

Make a dash on the line for how likely you would be to use this interface in the future, if it was available

Future use

Very likely to use again _____ Very unlikely to use again

Please add any further comments that will help clarify your ratings above…….

….

Appendix III. Abstracts of Chapters as Published, in Press, or in Review

Manuscript 1:

Croft, H., Spencer, K., Robertson, S., & Cronin, J. (2021). The problem with big data and coaching: closing the gap – a review. *Submitted to Journal of Human Sport and Exercise (4th June 2021)*

Author contributions: Croft, H (85%), Spencer, K (5%), Robertson, S (5%), & Cronin, J. (5%).

Abstract

Mayer-Schönberger and Cukier (2013) describe that there has been an "explosion" of data since the 2000's due to new technologies that enable data to be generated about almost anything. This creates a problem for coaches and players in that this saturation of information allows them, like all humans, to use it to support their beliefs (Mercier., H, & Sperber., D. 2011). This phenomenon is known as conformational bias and leads to decision making based on samples of the evidence, taken out of context, unfairly representing situations (McNerney, S., 2011). This paper will begin to address this issue, reviewing the knowledge that currently exists around the processes of performance analysis. It will identify gaps in the literature and make recommend future research to resolve these. Literature was reviewed under the following topics: 1) What information is used by coaches when making decisions during a match, 2) Data sources and collection in sport, 3) Organising and analysing different type of sports data; 4) Visualizing and feeding back information to coaches. The review concludes that the first step in closing the gaps between coaches and scientists is to examine what coaches discuss or communicate during sport matches and then to follow a process of providing information on these themes and topics. This process needs to include methods for capturing, organising and presenting this information back to coaches in an ecologically valid way.

Manuscript 2:

Croft, H. G., Spencer, K., & Robertson, S. (2020). What tactical and technical comments do coaches make during netball matches? A content analysis in netball. *International Journal of Performance Analysis in Sport*. doi:DOI: 10.1080/24748668.2020.1846112

Author contributions: Croft, H (85%), Spencer, K (10%), Robertson, S (5%)

Abstract

Purpose: In sport an 'explosion' of information (Mayer-Schönberger and Cukier, 2013) has meant that coaches may not have the time or cognitive resources to integrate it effectively into their decision-making. Research suggests there is a disconnect exists between the information provided from the 'sport scientist to coach', and subsequent 'coach to player' (William and Kendall, 2007). Further investigation is required to determine the detail within the tactical and technical instruction, with the aim of creating better alignment and implementation between the performance analyst data collected and actual coach requirements during a game.

Methods: Using inductive content analysis, with an independent researcher acting as a 'critical friend', themes and dimensions were extracted from audio recordings of six coaches' (ANZ Premiership netball) conversations during competition.

Results: Analysis of the transcribed recordings found that the coaches discussed the technical and tactical elements of the game 695 times, within five general dimensions; movement (38.4%), actions (23.9%), positioning (17.3%), possession outcomes (10.6%) and strategy (9.8%). Within these 5 dimensions there were 17 higher order themes including; and 56 lower order themes.

Conclusion: These results will be discussed in relation to the latest performance analysis literature and recommendations will be made for future research.

Keywords: Netball, Performance Analysis, Tactics, Strategy, Content Analysis

Manuscript 3:

Croft, H., & Spencer, K. (2021a). An Evaluation of Data Sources in Netball Performance Analysis. *(Under review) The Journal of Sport and Exercise Science.*

Author contributions: Croft, H (90%), Spencer, K (10%)

Abstract

Introduction: This paper evaluated what information systems and types of data are available, in real-time, during netball sport matches. Feedback is recognised as an important part of the coaching process and if the vast quantities of data and information that exist in sport can be align with what coaches are trying to achieve than better decision-making will result.

Method: A review of research and public domain data sources was conducted and each was evaluated against two validity and two compatibility criteria. Ratings were assigned for each source and recommendations made for their use.

Results: Data sources identified included ubiquitous computing devices like heart rate monitors, accelerometers, and GPS. Notational analysis systems were found to be the most customisable and best suited for live feedback, while public domain web sources had excellent alignment and potential for compatibility if an API could be connected to.

Discussion: There are tactical themes in live netball performances that current information technologies do not produce data on. There are larger publicly available sources of data which could be used to provide context about live performance. Further research needs to look at both of these outcomes. These results also provides a useful guide for performance analysts when designing information technologies for use during matches.

Manuscript 4:

Croft, H., Spencer, K., Taurua, N., & Wilton, E. (2021). Creating a live and flexible normative dataset for netball. *(In preparation for submission) Frontiers in Sport and Active Living.*

Author contributions: Croft, H (85%), Spencer, K (10%),Wilton, E (3%), Taurua, N (2%)

Abstract

Previous research has identified large data and information sources which exist about netball performance and align with coaches' discussion during games. Normative data provides context to measures across many disciplines including fitness testing, physical conditioning, and body composition. These data are normally presented in tables as representations of the population categorised for benchmarking. Normative data does not exist for benchmarking or contextualisation in netball yet coaches and players use performance statistics. A systems design methodology was adopted for this study where a process for automating the organisation, normalisation, and contextualisation of netball performance data was developed. To maintain good ecological validity a case study utilised expert coach feedback on the understandability and useability of the visual representations of netball performance population data. This paper provides coaches with benchmarks for assessing the performances of players, across competition levels against player positions for performance indicators. It also provides insights to a performance analyst around how to present these benchmarks in an automated "real-time" reporting tool.

Manuscript 5:

Croft, H. G., & Spencer, K. (2021b). Coach engagement with live netball statistical information. *(In preparation for submission) International Journal of Performance Analysis in Sport*

Author contributions: Croft, H (85%), Spencer, K (10%), Robertson, S (5%)

Abstract

Limited research has investigated how coaches use statistical information, as a means of feedback to players (Nicholls et al., 2018; Middlemas., 2017). Interestingly this has not been investigated in a live sports match context, when decision making is time constrained. Observation provides a depth of information about behaviours that is not achieved though interventions (Brown, 2020). This study investigated the behaviours of six elite coaches' when using a statistical information screen during live netball matches. The coaches tended to favour combined team statistics, however this varied between participants, as did their navigation between these and individual player statistics. The coaches showed tendencies to navigate layers of complexity in the statistics information. Strong alignment was found between the information coaches viewed and the themes discussed by the same population (Croft et al., 2020). Recommendations are made about the design of live statistical information tools as well as the need for integration of movement information.

Manuscript 6:

Croft, H., Spencer, K., & Bowden, M. (2021). Developing information tools for presenting coaches with large datasets. (In preparation for submission) Journal of Computer Science and Sport.

Author contributions: Croft, H (80%), Spencer, K (10%), Bowden, M (10%)

Abstract

Frameworks are important for guiding practitioners in many areas of sport science and coaching. The 'coaching process' and 'systems design' approach are examples of frameworks that can be used for coaches and performance analysts. A framework does not exist for performance analysts developing 'large' or 'big data' tools which assist coaches decision making. This manuscript implements a new framework specifically designed for this problem and demonstrates a tool for netball coaches and sport science staff to navigate datasets in an ecologically valid way. The tool incorporated findings from previous research and utilised PowerBI™ to present information about team and player performance. The tool was evaluated by six participants using a survey and a net promoter score was calculated. The tool was rated high, by participants, and other insights were discovered about its benefits in decision-making and feedback about actions. The inclusion of better 'movement' data was identified as important for future research.

Appendix IV. Additional Related Research Outputs Since Starting the Doctor of Philosophy

Croft, H., & Spencer, K. (2020). An Evaluation of Data Sources in Netball Performance Analysis. *Otago Polytechnic, Institute of Sport, Exercise and Health, Intensive Week Symposium*, 10th November 2020, Sargood Centre.

Croft, H. (2020). An evaluation of data sources in netball performance analysis. *Sport and Exercise Science New Zealand Conference*, Christchurch, New Zealand November 25 - 27th.

Al-Attar, J., **Croft, H.**, Bloxham, K., Blair, M. (2020). Cohort study - training load and its progression from amatuer to semi-professional rugby. *Journal of Australian Strength & Conditioning. 28(07):*6-15. ASCA.

Croft,H. (2019). Analysing the Numbers (Performance Analysis). *High Performance Sport Symposium, Otago Polytechnic*, Dunedin, 11 – 12 November 2019.

Croft, H., Warren, P., & Healy, J. (2019) *NZ Data Science + Analytics Forum event focusing on sports analytics.* Expert panel discussion. Invited Speaker. Conference Centre, Spark City, 167 Victoria Street West. Tuesday 29 October, 2:30pm - 5:30pm

Ramsey, C., Handcock, P., Humphrey, R. & **Croft, H.** (2019.) Introducing the Certified Footwear Analyst: A multisport applied research project. *Sport and Exercise Science New Zealand Annual Conference*, 27-29 November 2019 Massey University, Palmerston North.

Middlemas, S., **Croft,H.,** Spencer,K., & Robertson,S. (2018). The first annual Oceania performance analysis conference. *Scope Contemporary Research Topics (Health & Wellbeing)*, 3, Occupation.

Croft, H., Spencer, K., & Robinson, S. (2018). What tactical and technical comments do coaches make during sports matches? A qualitative analysis in netball. *World Congress of Performance Analysis of Sport 12*. Opatija, Croatia. September 19-23. 2018. https://www.researchgate.net/publication/327667680_ISPAS-2018-final

Croft, H. (2017). Integrated learning in Sports Analysis. *Scope Contemporary Research Topics, Health and Wellbeing, 1*, Activity. 18-23. www.thescopes.org

Croft, H., Wilcox, B., and Lamb, P. (2017). Using Performance Data to Identify Styles of Play in Netball: An Alternative to Performance Indicators*. International Journal of Performance Analysis in Sport.* Accepted and posted online, December 2017.

Blair, M., Body. S., & **Croft, H.** (2017). Relationship between physical metrics and game success with elite rugby sevens players. *International Journal of Performance Analysis in Sport*, DOI: 10.1080/24748668.2017.1348060 http://dx.doi.org/10.1080/24748668.2017.1348060