The Developmental History of Athletes Questionnaire: Towards a comprehensive understanding of the development of sport expertise

Melissa Jayne Hopwood
Bachelor of Applied Science (Exercise and Sport Science) (Hons)

School of Sport and Exercise Science
Faculty of Arts, Education and Human Development
Victoria University

Submitted in fulfilment of the requirements of the degree of Doctor of Philosophy

August, 2013

Abstract

Knowledge of the training requirements and environmental factors associated with expert sports performance is essential for the design of effective sport programs and the creation of nurturing practice environments. However, several limitations in the sport expertise development literature restrict the application of current knowledge to practical settings. Results and recommendations from existing studies are inconsistent, with discrepancies likely related to a combination of small, homogenous sample sizes and differences in measurement tools. Furthermore, the questionnaires and interview guides utilized are generally poorly validated. To begin to address these limitations, the Developmental History of Athletes Questionnaire (DHAQ) was constructed and rigorously validated. Results highlighted a number of issues related to reliability of retrospective recall and sub-optimal questionnaire design. In an attempt to resolve these issues, modifications were made to the DHAQ, and it was converted to an online format to allow large scale distribution. Following modification and conversion, the DHAQ was subjected to further reliability analyses. While several reliability issues were still apparent and several further revisions were required, the amendments markedly strengthened the measurement tool and the DHAQ can now start to be utilised in investigations of sport expertise development. Continued assessments of validity of the updated, online DHAQ are recommended; however, it is proposed that the DHAQ be considered the emergent standard tool for the collection of athlete developmental history information, providing a major contribution towards a comprehensive understanding of the development of sport expertise.

Student Declaration

I, Melissa Jayne Hopwood, declare that the PhD thesis entitled –The Developmental History of Athletes Questionnaire: Towards a comprehensive understanding of the development of sport expertise" is no more than 100,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references, and footnotes. This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work.

Signature:  Date: August 23, 2013

Acknowledgements

There are many people whom I need to acknowledge for their contribution to this thesis and my journey throughout these past four years. First and foremost, I must thank with utmost respect and gratitude, my supervisors, Joe, Clare, and Damian - three wise monkeys who taught me so much; not only about the research process and becoming an academic, but also about myself. You each brought so much to the table for me to learn from and model myself off, yet each of your contributions was so unique. Joe, four years ago you welcomed me into your lab and into your home. Your knowledge of just about anything and everything constantly astounds me, and I will never understand how one person can write so many papers, read so many books, watch so many tv shows and movies, and even renovate their house, yet still find time to make his students feel like they are a top priority. To this thesis you were the theorist, always focussing on the processes and contributions from a theoretical perspective. To me, you were my academic mentor, teaching me how to navigate and succeed in the world of research. This PhD would not have been possible without your knowledge, guidance, and generosity. I can honestly say that packing my bags and coming to Canada was the one of the best decisions I have ever made, and that is largely thanks to you.

Clare, four years ago you took a chance on a girl who you barely knew, and agreed to take me on as one of your first PhD students. Not only that, but you agreed to let me abandon ship and move to Canada, leaving you at VU to take care of all my paperwork! To this thesis you were the quality controller, always making sure that things had been thought through thoroughly, that all the t‘s had been crossed and the i‘s dotted, and that everything I submitted was the best it could possibly be. To me, you were my personal mentor, always checking up on how I was doing, and often receiving a barrage of tears and complaints in return. Your friendship during this journey was invaluable and this PhD would not have been as complete and polished without your close attention to detail and concern for my wellbeing.

Damo, four years ago you set me up with the dream team of supervisors and passed me off into the hands of Clare and Joe to build on the skills and experiences I gained with you at the AIS. To this thesis you were the practical advisor, ensuring that our approach was realistic and would provide interesting, meaningful information for coaches and athletes. To me, you were my mentor for applied research, servicing, and knowledge translation, and I would not have the same passion for applied sport science if it weren‘t for you. This PhD certainly wouldn‘t have been possible without your connections and persuasive abilities to recruit participants, and your faith in me to keep the coaches, administrators, and athletes happy.

I would also like to thank my friends who accompanied me in this adventure. You widened my research horizons, showed me all of the fun things about doing a PhD that don‘t involve books and computers, and taught me numerous new moves on the dance floor. To my lab mates at York: Brad, Jane, Shilpa, Harmonie, Jared, Jason, Rachael, and Nick; Jessica Fraser-Thomas and the girls from the JFT lab: Natasha, Sarah, Mata, and Theresa; the Germans: Rebecca, Jörg, Bernd, and Maike; and the Brock girls: Kaitlyn, Tammy, Kate, and Lori – thank you for the memories and good times. Thanks also to Elmira, the FITS crew, and the Dragons girls for keeping me company and making me happy when I was avoiding my PhD.

I must also thank the people who made this PhD possible. Firstly, thank you to the organisations that provided funding to complete this research: Victoria University, York University, the International Council for Canadian Studies, the Ontario Ministry for Health Promotion, the Canadian Sport Centre Ontario, and the Centre for High Performance Sport at the University of Toronto. Secondly, thank you to my small army of research assistants for your help throughout various stages of my research: Allan, Cassim, Jeff, Justin, Luke, Melissa, Natasha, Samantha, Stephanie, and Young-Bin. Your persistence and commitment, particularly when the work was repetitive and un-exciting, was greatly appreciated. If it weren‘t for you, I would probably still be entering data myself! Finally, to the athletes, coaches, sport administrators, and parents who participated in this research – thank you for so graciously giving your time and sharing your stories.

Last, but most definitely not least, I would like to thank my rocks: Mum, dad, and Luke. Mum and dad, thanks for supporting my decisions to go back to university and move overseas, for acting as my secretary at home base, and for bank rolling my adventures while in Canada. Throughout my life you have instilled values of hard work and dedication, and have always encouraged me to strive for my dreams – all essential elements to begin and complete this PhD. You are my role models in life, and I hope I will grow to be as happy and successful as you both.

Luke, you are everything I could ever hope for in a partner. Throughout this journey you were always there to celebrate the triumphs, work through the challenges, wipe away the tears, and even help recruit participants and interpret results! I will be forever grateful for the sacrifices you made and the lengths you went to, to support me on this crazy ride. I could not have done it without you, and I can‘t wait to begin the next adventure with you.

To everyone who made this possible, once again, thank you.

List of Publications and Awards

Publications in Refereed Journals

Hopwood, M. (2012). The PhD process: An Australian perspective. *Ze-phir,* 19(1), 26-27.

Rienhoff, R., & **Hopwood, M.** (2011). SCAPPS Conference 2010: An international student's perspective. *Sportwissenschaft,* 1, 60-61.

Published Abstracts for Refereed Conference Presentations

Hopwood, M.J., Baker, J., Farrow, D., & MacMahon, C. (2012). Sport expertise development: Skill level differences in practice profiles during childhood and adolescence. *Journal of Sport & Exercise Psychology*, 34, S236.

Hopwood, M.J., Baker, J., MacMahon, C., & Farrow, D (2012). Faster, higher, stronger, …, and younger? Birth order, sibling sport participation and sport expertise. *Journal of Sport & Exercise Psychology*, 34, S235.

Hopwood, M.J., Farrow, D., Baker, J., & MacMahon, C. (2012). Questioning the importance of diversification for sport expertise development. *Journal of Sport & Exercise Psychology,* 34, S236.

Hopwood, M.J., MacMahon, C., Baker, J., & Farrow, D. (2012). Following in their footsteps? Sport expertise and parental participation in sport and physical activity. *Journal of Sport & Exercise Psychology*, 34, S48.

Hopwood, M.J., MacMahon, C., Baker, J., & Farrow, D. (2011). Organized sports, informal play, and the development of sport expertise: Reliability of retrospective recall. *Journal of Sport & Exercise Psychology,* 33, S158.

Hopwood, M., MacMahon, C., Baker, J., & Farrow, D. (2010). The Developmental History of Athletes Questionnaire (DHAQ): Considerations when collecting athlete training histories. *Journal of Sport & Exercise Psychology,* 32, S176.

Refereed Conference Presentations

Hopwood, M.J., MacMahon, C., Farrow, D., & Baker, J. (2012, November). *Sibling dynamics and sport expertise: Findings from the Pathways to the Podium Research Project.* Canadian Society for Psychomotor Learning and Sport Psychology Conference, Halifax, Canada.

Hopwood, M.J, Baker, J., MacMahon, C., & Farrow, D. (2012, March). *Sport expertise: A family affair?* Eastern Canadian Sport & Exercise Psychology Symposium, London, Canada.

Hopwood, M.J., Farrow, D., MacMahon, C., & Baker, J. (2011, October). *Sporting milestones and career progression of male Australian junior international level team sport athletes.* Canadian Society for Psychomotor Learning and Sport Psychology Conference, Winnipeg, Canada.

Hopwood, M., Baker, J., MacMahon, C., & Farrow, D. (2011, March). *Social media as a tool for participant recruitment and knowledge translation.* Eastern Canadian Sport & Exercise Psychology Symposium, Waterloo, Canada.

Hopwood, M., Baker, J., MacMahon, C., & Farrow, D. (2010, October). *Reliability and Validity of the Developmental History of Athletes Questionnaire (DHAQ).* Canadian Society for Psychomotor Learning and Sport Psychology Conference, Ottawa, Canada.

Hopwood, M., Baker, J., MacMahon, C., & Farrow, D. (2010, March). *Designing the DHAQ: Construction and validation of the Developmental History of Athletes Questionnaire.* Eastern Canadian Sport & Exercise Psychology Symposium, Montreal, Canada.

Hopwood, M., MacMahon, C., Baker, J. & Farrow, D. (2009, July). *A large scale examination of long-term athlete development: Towards a comprehensive understanding of long term athlete development.* 3rd Meeting of the Australasian Skill Acquisition Research Group, Brisbane, Australia.

Hopwood, M., MacMahon, C., Baker, J. & Farrow, D. (2009, March). *Long-term athlete development: A framework for sport expertise?* Eastern Canadian Sport & Exercise Psychology Symposium, Toronto, Canada.

Awards

Hopwood, M.J. 2012

Victoria University Secomb Conference and Travel Fund $350

Following in their footsteps? Sport expertise and parental participation in sport and physical activity

Baker, J. & **Hopwood, M.J.** 2011

York University Faculty of Health Minor Research Grant $2,500

Pathways to the Podium: A large scale examination of long-term athlete development

Baker, J., **Hopwood, M.J**., MacMahon, C., & Farrow, D. 2010

Research Program in Applied Sport Sciences Grant $10,000

Pathways to the Podium: A large scale examination of long-term athlete development

Hopwood, M.J. 2010

International Council for Canadian Studies Graduate Student Scholarship $3,500

A large scale examination of long-term athlete development: Towards a comprehensive understanding of the development of sport expertise

Hopwood, M.J. 2009 - 2012

Australian Postgraduate Award $22,500 per annum

A large scale examination of long-term athlete development: Towards a comprehensive understanding of the development of sport expertise

Table of Contents

List of Tables

List of Abbreviations

DHAQ	Developmental History of Athletes Questionnaire
DMSP	Developmental Model of Sport Participation
ICC	Intraclass correlation coefficient
PA	Percent agreement

Introduction

This chapter provides a brief review of the literature related to the development of sport expertise. Gaps in, and limitations to, our understanding of long-term athlete development are highlighted, providing a rationale for the direction of this thesis. An overview of the structure and content of the thesis is provided.

Expert sports performance is characterised by a variety of skills and abilities. For example, high performance athletes display superior physical abilities such as strength, speed, and balance, and execute advanced technical skills like shooting, striking, and acrobatics with ease. Skilled performers are also often described as having ‗eyes in the back of their head' or ‗all the time in the world' as they demonstrate outstanding perceptual-cognitive skills such as anticipation, decision making, and creativity. However, a question asked by many athletes, coaches, parents, sports fans, and sport scientists, is ‗how did these athletes become so great'?

Although there are studies of athlete development dating back to the mid-1800's (e.g. Galton, 1869), many trace the systematic study of the development of expertise to 1985 and the publication of ―Developing Talent in Young People" (Bloom, 1985). This landmark series of studies investigating the developmental experiences of exceptional musicians, artists, mathematicians, scientists, tennis players, and swimmers was the first to discuss World class athletes' pathways toward expertise. The investigation described participants' progression through three stages of development and provided a detailed account of characteristics within the home and sporting environments believed to be influential on their success.

Since Bloom's exploration of the development of sport expertise, research has typically fallen into four areas: 1) biological accounts of sport expertise development; 2) practice as the primary contributor to the development of sport expertise; 3) environmental and contextual factors influencing athlete development; and 4) multi-factorial approaches to the development of talent. In brief, biological accounts of sport expertise development emphasise the contribution of physical characteristics and capacities to expert performance. This line of research has typically examined anthropometric (Hoare, 2000; Keogh, 1999; Pienaar, Spamer, & Steyn, 1998), physiological (Gabbett, 2002; Keogh, Weber, & Dalton, 2003; Vaeyens et al., 2006), and/or genetic (Bray et al. 2010; Puthucheary et al., 2011; Tucker & Collins, 2012) correlates of superior performance, with the aim of developing biological indicators for talent identification.

Practice accounts of sport expertise development suggest that expert performance is the result of engagement in considerable volumes of high quality training. Research in this area has typically sought to identify the types, volumes, and conditions of practice in which expert athletes have engaged throughout the history of their involvement in sport, in order to understand the practice requirements to attain expertise (Baker, Côté, & Abernethy, 2003a; Baker, Côté, & Abernethy, 2003b; Deakin & Cobley, 2003; Helsen, Starkes, & Hodges, 1998; Hodge & Deakin, 1998; Hodges, & Starkes, 1996; Hodges, Kerr, Starkes, Weir, & Nananidou, 2004; Starkes, Deakin, Allard, Hodges, & Hayes, 1996; Ward, Hodges, Starkes, & Williams, 2007; Young & Salmela, 2002). This line of research has also explored the contribution of participation in multiple sports (Baker et al., 2003b; Baker, Côté, & Deakin 2005; Berry,

Abernethy, & Côté, 2008; Ford & Williams, 2008; Ward et al. 2007; Memmert, Baker, & Bertsch, 2010; Oldenziel, Gagné, & Gulbin, 2003; Soberlak & Côté, 2003) and playful sporting games (Berry et al., 2008; Ford, Ward, Hodges, & Williams, 2009; Memmert et al., 2010; Soberlak & Côté, 2003) to the development of sport expertise, generally highlighting a positive association between expertise and participation in these activities.

Environmental and contextual accounts of the development of sport expertise focus on external influences on athlete development. Examples of environmental and contextual factors found to be influential on the development of sport expertise include birth date (Barnsley, Thompson, & Barnsley, 1985; Cobley, Schorer, & Baker, 2008; Helsen, van Winckle, & Williams, 2005; Wattie, Baker, Cobley, & Montelpare, 2007), birthplace (Baker & Logan, 2007; Carlson, 1988; Curtis & Birch, 1987; Côté, MacDonald, Baker, & Abernethy, 2006; MacDonald, King, Côté, & Abernethy, 2009), familial support (Côté, 1999; Kay, 2000; Lauer, Gould, Roman, & Pierce, 2010; Sloane, 1985), coaching (Bloom, 1985; Durand-Bush & Salmela 2002; Gibbons, Hill, McConnell, Forster, & Moore, 2002; Oldenziel et al., 2003), deliberate programming and sport medicine / sport science support (Bullock, Gulbin, Martin, Ross, Holland, & Marino, 2009; Durand-Bush, & Salmela, 2002; Smith, 2003), cultural importance of sport (Baker & Horton, 2004; Salmela & Moraes, 2003), sport maturity (Baker & Horton, 2004), and depth of competition (Baker & Horton, 2004). An understanding of contextual influences such as these allows for the provision of recommendations regarding the environmental conditions required to promote optimal long-term athlete development.

Finally, multi-factorial approaches to the development of expertise recognise the complimentary and synergistic contributions of biological, practice, and/or contextual factors to sport expertise, and advocate holistic models of expertise development. Examples of multi-factorial models proposed to enhance our understanding of athlete development include Gagne‘s (2004) Differentiated Model of Giftedness and Talent, emphasising the influence of interpersonal, environmental, and chance factors on the transformation of gifts into talents; Bailey et al.‘s (2010) biopsychosocial model of development, emphasising the importance of biological, psychological, and social influences on athlete development; and Phillips, Davids, Renshaw, and Portus‘ (2010a) dynamic systems approach to talent development, emphasising the individual nature of developmental pathways. This line of research encourages researchers and practitioners to consider the wide variety of influences on the development of sport expertise when designing studies and practice environments.

Upon the recommendation of multi-factorial approaches to the development of sport expertise, this thesis addresses a variety of practice and contextual influences on athlete

development, ultimately aiming to extend our understanding of the associations and interactions between the attainment of sport expertise and these factors.

A brief review of the literature adopting a practice-based approach to the investigation of expertise development highlights some discrepancies in recommendations of optimal requirements and pathways for the development of sport expertise. Compare, for example, Ericsson and colleagues' deliberate practice framework (Ericsson, Krampe, & Tesch-Römer, 1993), Côté and colleagues' Developmental Model of Sport Participation (DMSP; Côté, Baker, & Abernethy, 2007; Côté & Fraser-Thomas, 2007; Côté & Hay, 2002), and Ford and colleagues' early engagement hypothesis (Ford et al., 2009).

The deliberate practice framework suggests that expert performance in a given domain is the result of large amounts of effortful practice, and individuals who engage in high levels of practice from an early age will outperform those who have completed similar practice, but did not start until a later age (Ericsson et al., 1993). The DMSP recognises that large amounts of deliberate practice from an early age may be necessary for expert performance in sports where athletes typically reach their peak prior to puberty (for example, gymnastics and figure skating), however, the model proposes an alternative three-stage approach for sports in which peak performance occurs later (Côté et al., 2007; Côté & Fraser-Thomas, 2007).

The first stage of the DMSP, known as the ‗sampling years', is characterised by participation in a variety of sports, large amounts of involvement in playful sporting games (termed deliberate play; see Côté & Hay, 2002), and small amounts of deliberate practice. During the ‗specializing years', a gradual reduction in activities and a balance of deliberate play and deliberate practice are typical, and it is not until the ‗investment years', usually encompassing ages 16 and beyond, that a focus on one sport and large amounts of deliberate practice are common. In contrast, the early engagement hypothesis proposes that the pathway to expert performance involves a combination of deliberate practice and sport specific deliberate play from a young age, but not necessarily participation in other sports (Ford et al., 2009).

Similar inconsistencies are observable in a brief review of the literature examining the development of sport expertise from a contextual approach, with equivocal results concerning the importance of birth date and birthplace on sport expertise development (Cobley, Baker, Wattie, & McKenna, 2009; Baker, Schorer, Cobley, Schimmer, & Wattie, 2009). Furthermore, the evidence pertaining to the influence of other environmental and contextual factors such as family, sport medicine / sport science support, and culture on the attainment of sport expertise is limited, warranting further investigation in these areas.

The lack of uniformity in current findings and the absence of clear conclusions relating to optimal conditions for the development of sport expertise may reflect the complexity of sport

expertise development and the diversity of potential pathways toward expert performance (Phillips et al. 2010a). However, there are a number of methodological differences and limitations within the studies on which the findings described above are based, and indeed within this field of research in general, which must be considered.

The majority of investigations relating to sport expertise development involve retrospective recall techniques where athletes are required to think back to as early as age five, and identify details of their participation in various sporting activities throughout the years. Many studies have adopted a qualitative retrospective recall approach, conducting a series of open-ended interviews with athletes, parents, and coaches (Baker et al., 2003a; Baker et al., 2003b; Baker et al., 2005; Berry et al., 2008; Bloom, 1985; Carlson, 1988; Côté, 1999; Durand-Bush & Salmela, 2002; Law, Côté, & Ericsson, 2007; Phillips, Davids, Renshaw, & Portus, 2010b; Soberlak & Côté, 2003; Weissensteiner, Abernethy, & Farrow, 2009). Although qualitative interviews generate large amounts of detailed information, this approach is particularly labour intensive and time consuming. As a result, sample sizes are typically small, including athletes from just one or a small number of sports, who are often members of the same sporting team. Therefore, the generalisability of findings to athletes from a more diverse range of sports and scenarios is limited.

Quantitative questionnaires provide an alternative approach to qualitative interviews to overcome the limitations associated with small sample sizes. Initially, the use of questionnaires in studies of sport expertise development were confined to simple investigations of hours spent participating in various activities (Helsen et al., 1998; Hodges & Starkes, 1996; Starkes et al., 1996) and in fact, the participant samples involved in these studies were still relatively small and homogenous. More recently, however, several groups have administered in-depth questionnaires to large numbers of athletes, occasionally from a wide variety of sports, in order to obtain a more detailed understanding of the factors associated with sport expertise development (Gibbons et al., 2002; Memmert et al., 2010; Moesch, Elbe, Hauge, & Wikman, 2011; Oldenziel et al., 2003; Ward et al., 2007). Despite these recent multi-dimensional, large scale investigations of sport expertise development, a noteworthy issue remains, namely that each study usually incorporates a unique measurement tool, making comparison of results and recommendations between studies difficult because athletes tend to be asked different questions, in different ways.

Another limitation of previous research regarding the development of sport expertise concerns the validity and reliability of the measurement tools utilised. It is important to note that the majority of studies have included a test of validity and/or reliability for their particular construct. Common methods for verification of the trustworthiness of the data have included

triangulation with parent and/or coach interviews (Baker et al., 2003a; Baker et al., 2003b; Berry et al., 2008; Durand-Bush & Salmela, 2002; Law et al., 2007; Oldenziel et al., 2003; Soberlak & Côté, 2003), comparison with training diary information (Baker et al., 2005; Helsen et al., 1998; Hodges, & Starkes, 1996; Ward et al., 2007), consultation of independent records such as news items and published results (Berry et al., 2008; Durand-Bush & Salmela, 2002; Law et al., 2007; Phillips et al., 2010a), and test-retest measures (Ford & Williams, 2008; Helsen et al., 1998; Memmert et al., 2010; Moesch et al., 2011; Ward et al., 2007). However, in almost all cases, only a small number of items within the full interview guide or questionnaire were selected for the validity and reliability analyses. In each of these studies, the establishment of validity and reliability of the measurement tools utilised has been secondary to the testing of hypotheses relating to the development of sport expertise.

Measures commonly assessed for validity and reliability include estimates of weekly practice hours (Baker et al., 2003b; Helsen et al., 1998; Hodges & Starkes, 1996; Moesch et al., 2011; Oldenziel et al., 2003; Ward et al., 2007), yearly practice hours (Baker et al., 2003a; Baker et al., 2003b; Baker et al., 2005; Helsen et al., 1998; Law et al., 2007; Ward et al., 2007), or cumulative practice hours across athletes‘ careers (Baker et al., 2005; Memmert et al., 2010; Soberlak & Côté, 2003); and occasionally, age at first participation in main sport (Memmert et al., 2010; Oldenziel et al., 2003; Ward et al., 2007), or total years of involvement in main sport (Law et al., 2007; Memmert et al., 2010). In studies examining participation in other organised sports and deliberate play, the number of activities reported (Baker et al., 2003b; Berry et al., 2008; Law et al., 2007; Memmert et al., 2010; Oldenziel et al., 2003; Soberlak & Côté, 2003), and hours of involvement in these activities (Memmert et al., 2010; Soberlak & Côté, 2003) are also occasionally assessed for validity and/or reliability.

In general, validity and reliability of recall for hours of involvement in practice activities is typically quite good, with Pearson correlation coefficients between the primary measure of interest and the relevant comparative measure ranging from .59-.97. Similarly, age at first participation in main sport and total years of involvement in main sport have been reported to demonstrate Pearson correlation coefficients of .79-.99, and on occasion, even 100% agreement between data sources. Agreement between primary and secondary data sources is also typically strong for number of sporting activities, with percent agreement values reported between 88-100%; although, Oldenziel and colleagues (2003) reported a low Pearson correlation coefficient for number of sporting activities of .26. While validity and reliability of recall for time involved in deliberate play has only been assessed in a small number of studies, it does not appear to be as strong as validity and reliability of recall for time involved in deliberate practice. For example, Memmert et al. (2010) reported a test-retest reliability Pearson

correlation coefficient of .79 for total cumulative hours of involvement in deliberate practice, but the Pearson correlation coefficient for cumulative hours of involvement in deliberate play was only .70. Similarly, Soberlak and Côté (2003) reported a 30% difference between athletes and parents accounts of time involved in deliberate play compared to just a 1% difference for time involved in deliberate practice. Despite the apparent lower levels of validity and reliability of recall for time involved in playful activities compared to practice activities, the values reported still indicate a reasonable level of agreement. Details pertaining specifically to athletes' main sport, on the other hand, appear to be typically recalled with good validity and reliability.

Although previous investigations of validity and reliability of retrospectively recalled athlete developmental history information have demonstrated reasonable consistency and agreement of measures, it remains problematic that validity and reliability assessments have typically only been conducted on a small number of items within each investigation, and establishment of validity and reliability of the measurement tools has been secondary to the testing of hypothesis relating to sport expertise development. Given the retrospective nature of research in this area, validity and reliability of recall is a key concern and deserves greater attention than it typically receives.

As such, there is great value in constructing and validating a measurement tool to collect quantitative, objective information related to athlete developmental histories from large, diverse samples of participants. A questionnaire is recommended over an interview-based tool, as this style is more suitable for large scale administration. As well, inclusion of items similar to those previously investigated is required to facilitate examination of the generalisability of existing knowledge; however, the addition of new items is also necessary to extend our understanding of influences on athlete development for which our knowledge is currently limited. Most importantly, it is essential that the instrument rates highly in scores of validity and reliability on all items in order to ensure maximal accuracy of responses.

When examining validity and reliability, it is important to consider concurrent validity, convergent validity, test-retest reliability, and internal consistency. Concurrent validity is demonstrated when a test score or measurement instrument correlates well with a criterion measure such as another tool that has been previously validated; while convergent validity indicates that measures designed to supposedly address the same underlying construct are, in fact, in agreement (Bryant, 2000). Test-retest reliability refers to the ability of the instrument to elicit the same results when completed by the same respondent, under the same conditions, on two separate occasions; and internal consistency signifies that items within a single tool that propose to measure the same general construct, converge on the same conclusions (Strube, 2000). It would also be valuable to assess the predictive validity of the measurement tool (i.e.

the ability of the instrument to predict a criterion measure assessed at a later time point; Bryant, 2000); however, in the context of sport expertise development, a considerable period of time following administration of the questionnaire would be required to ascertain whether athletes ultimately attained expertise. Therefore, assessment of predictive validity is outside the scope of this thesis, but should be considered in future research.

Thus, this thesis aimed to construct and validate a new measurement tool for collecting quantitative data pertaining to athlete developmental histories. Once validated, this instrument can then be used for subsequent large scale examinations of long-term athlete development to work towards a comprehensive understanding of the development of sport expertise. To achieve this aim, four methodological and experimental phases were completed. This thesis outlines each of the four phases in detail.

Study phase one involved construction of the Developmental History of Athletes Questionnaire (DHAQ). The related chapter of this thesis describes the original structure and content of the DHAQ, including a review of key literature influencing the design of the questionnaire, and evidence-based rationale to support the inclusion of each item. Study phase two aimed to establish the validity and reliability of the DHAQ, with the associated chapter describing the methodology and findings of the investigation. Based upon the results and recommendations of study phase two, phase three involved modifying the DHAQ to improve the quality of the instrument. In addition, to aid large scale distribution of the DHAQ, following modification, the questionnaire was converted from a paper-based format to an online format. Details of all revisions, along with an updated outline of the structure and content of the DHAQ are provided in the chapter titled ―Study Phase Three: Modification and Online Conversion of the Developmental History of Athletes Questionnaire". This chapter also includes information regarding the online conversion of the DHAQ, along with an outline of benefits and limitations of web-based measurement tools. Given the large number of modifications and online conversion of the questionnaire, the DHAQ was then subjected to further reliability analyses. As such, study phase four aimed to reassess test-retest reliability of the updated, online DHAQ. Procedures and results are described in the associated chapter.

Following presentation of the four phases of this research, a general discussion of findings is provided. Further, to contextualise this thesis within the wider body of related knowledge, the general discussion chapter also outlines major contributions of this research to the field of sport expertise development, as well as potential extensions and applications in additional domains. Afterward, future directions emerging from this investigation are suggested, and a brief overview and re-cap of the thesis is offered in conclusion.

It is important to note that assessment of validity and reliability of the DHAQ involved extensive statistical analyses of great amounts of data. As such, this thesis includes a large number of detailed tables outlining the results of these analyses. Many of these tables extend over multiple pages, with some spanning a particularly high number of pages. To assist interpretation of the data alongside discussion of results, all tables have been provided in a separate volume, Supplement One. Furthermore, to aid navigation, the structure and format of all results tables are identical, including consistent statistical notation. All results presented in tabular form within Supplement One are discussed in general terms within the text; however, the tables provide more comprehensive information regarding the classification of items as having very good, good, moderate, or poor reliability and/or validity.

Similarly, Appendices A-I are also provided in a separate volume, Supplement Two. Supplement Two includes details of ethical approval, participant information sheets and consent forms, and copies of the DHAQ. Presentation of appendices in this manner allows for simultaneous consultation of related documents, statistical data, and text, to maximise understanding and interpretation of methodological procedures and results.

First and foremost, however, an introduction to the structure and content of DHAQ must first be provided. This is the focus of the next chapter.

Study Phase One: Construction of the Developmental History of Athletes Questionnaire

This chapter describes the DHAQ and outlines previous research that was highly influential during its construction. Key literature influencing the general design of the DHAQ is outline, followed by a detailed account of the structure and content of the questionnaire. An evidence-based rationale is provided to support the inclusion of each item in the questionnaire.

Design of the Developmental History of Athletes Questionnaire

Interview guides, questionnaires, results, and recommendations of previous research were used as a basis for the construction of the DHAQ. In particular, the interview guide utilised by Côté and colleagues (2005) provided the major foundation for the DHAQ, as this interview procedure is one of the few measurement tools for the collection of athlete developmental histories that has undergone rigorous validity and reliability testing. Additional studies that were particularly influential on the design of the DHAQ included Baker et al. (2003a; 2003b), Helsen et al. (1998), Hodges and Starkes (1996), Starkes et al. (1996), Soberlak and Côté (2003), and Ward et al. (2007).

These studies were selected as a basis for the DHAQ as all involved collection of retrospective athlete developmental history information via partially validated measurement instruments; and all except one (Ward et al., 2007) were identified among the most prominent articles in the athlete development literature during Bruner and colleagues‘ (Bruner, Erikson, McFadden, & Côté, 2009; Bruner, Erikson, Wilson, & Côté, 2010) citation network analyses of the field. It is possible that the later publication date of the influential study by Ward et al. (2007) may have contributed to its exclusion from the list of prominent athlete development articles highlighted in the citation network analyses (Bruner et al., 2009; Bruner et al., 2010).

Noticeably absent from the identified studies influencing the design of the DHAQ are two large scale, multi-sport investigations of athlete development (Gibbons et al., 2002; Oldenziel et al., 2003). At the time of construction of the DHAQ, results from these organisational reports had not been published in peer-reviewed journals, so the measurement tools had not been subjected to the same level of scientific scrutiny as those utilised within other investigations [see Gulbin, Oldenziel, Weissensteiner, & Gagné (2010) for a recent publication of data from the National Athlete Development Survey]. Furthermore, many aspects of these questionnaires focused on athletes‘ perceptions of their talent, experiences, and the factors that influenced their development. It has previously been shown that subjective information is recalled less reliably than objective information, so where possible, objective measures should be obtained (Côté et al., 2005). Consequently, although reviewed and considered, the questionnaires utilised by Gibbons et al. (2002) and Oldenziel et al., (2003) were not highly influential on the design of the DHAQ.

Structure and Content of the Developmental History of Athletes Questionnaire

Amalgamation of the instruments, results, and recommendations of previous research resulted in the synthesis of ten distinct sections for the DHAQ, each relating to a different aspect of athlete development. Below is a summary of the information gathered in each section along with a brief description of its relevance for studies of sport expertise development. The full version of this initial questionnaire is provided in Appendix A.

1. Demographic Information

Section one focused on basic demographic information including sex, date of birth, nationality, and country of residence. This information facilitates comparative analyses of developmental histories between sexes, cultures, and countries. Date of birth is a particularly important variable as it can be utilised in examinations of the relative age effect (Baker et al., 2009; Barnsley et al., 1985; Musch & Grondin, 2001), and in investigations of generational trends in the development of sport expertise. Items in this section required short open responses or selection of the most appropriate option from a list.

2. Sporting Career and Milestones

Section two relates to general information about the athlete‘s sporting career. Participants indicated their main sport, the highest level of competition they had participated at for this sport, and their greatest sporting achievement. Athletes then provided the ages at which they first moved house for reasons related to sport, formed the idea to become an elite athlete, set themselves the goal of becoming an elite athlete, and reached the peak of their career (if applicable). Next, a brief timeline of involvement in various practice activities was collected via identification of the ages at which the athletes first participated in regular supervised practice, unsupervised practice, non-sport specific practice, and informal play related to their main sport, as well as the age at which they first participated in off-season or year-round practice.

This was followed by a brief timeline of progression through the various levels of competition, with athletes indicating the ages at which they first participated at each of the club, regional, state, national, and international levels of competition, as well as the ages at which they were first recognised among the top 5 performers for their age group at each of those levels, and ultimately the best performer for their age group (if applicable). Last, this section required athletes to identify particular times they noticed a change in their participation or

success in their main sport, marking a transition from one stage of their career to another. The items in this section were similar to those included in previous studies of sport expertise development (Abernethy, Côté, & Baker, 2002; Ford & Williams, 2008; Law et al., 2007).

Identification of athletes‘ main sport is essential for comparative analyses of developmental histories across sports, and information related to the highest level of competition and greatest sporting achievement can be used to differentiate athlete skill level. While it is common to examine expert-novice differences in sport performance, studies of sport expertise strive to identify the factors that differentiate the most successful performers from those who may still be classified as highly skilled, but are not recognised as the best in their field. Although highest level of competition may provide a crude indication of skill level, identification of each participant‘s greatest sporting achievement can assist in constructing a more detailed skill level classification system, allowing for a closer examination of skill level differences in developmental histories.

Milestones such as relocating for reasons related to sport, forming the idea to become an elite athlete, setting the goal to become an elite athlete, reaching the peak of a career, and transitional phases, provide markers to help differentiate the stages through which elite athletes progress on their path to expertise; and information regarding athletes‘ introduction to various practice activities and their progression through the representative levels of competition can assist in the identification of recommendations for developmentally appropriate coaching and sport programming. For example, Fraser-Thomas, Côté, & Deakin (2008) identified that former athletes who dropped out of sport during adolescence reached a variety of sporting milestones at a younger age than athletes still engaged in their sport, suggesting that early involvement in some activities may be detrimental to long-term athlete development. As the study by Fraser-Thomas et al. (2008) was limited to the sport of swimming, further research in this area is required. Finally, information obtained within this section could also be utilised to explore associations between involvement in practice activities (obtained within later sections of the questionnaire) and competitive success, contributing to our understanding of the relationship between practice and performance. All items in this section required short open responses or selection of the most appropriate option from a list.

3. Family Characteristics and Participation in Sport and Physical Activity

The next section of the DHAQ collected information concerning the athlete‘s immediate family and their involvement in physical activity and competitive sport. Previous research has identified the influential role of the family and home environment on the

development of sport expertise, proposing that without the appropriate provision of resources and support from the family, it is unlikely that expertise will be attained (Bloom, 1985; Côté, 1999). Additionally, Bloom's (1985) findings indicated that sport was particularly valued in the homes of developing athletes, opportunities for participation in sporting activities were plentiful, and sports were a popular family activity. To continue to explore familial influences on sport expertise development, in section three athletes indicated the regularity of their immediate family's participation in physical activity, listed the competitive sports that each member of their immediate family participated in, and recorded the highest level of competition they reached for each of the sports identified.

Demographic information including date of birth, nationality, and the ages during which the athlete lived with each family member was obtained, along with the sex of each sibling, and the occupations of each parent. This information is important in order to identify the contextual factors associated with the athlete's home environment during the early stages of their sporting career. Furthermore, knowledge of siblings' birth dates can be used to investigate the influence of birth order on the development of sport expertise. Limited evidence exists to suggest that order of birth among siblings has the potential to influence participation in sport (Landers, 1979) and attainment of success in sport (Clark, 1982), with previous research suggesting that later born siblings more likely to participate in and achieve success in sport than first born siblings. Therefore, further investigation in this area is warranted. Items in section three required short open responses or selection of the most appropriate option from a list.

4. Places of Residence

Section four was designed to provide further information regarding the environmental context of athletes' development. Specifically, this section addressed the influence of place of residence on the development of sport expertise. Previous investigations have indicated that a disproportionately large number of elite athletes grew up in small to medium sized towns with populations of approximately 1,000 to 500,000 people (Côté et al., 2006; MacDonald et al., 2009). It has been suggested that the physical and psychosocial environments of small to medium sized towns may be more conducive for plentiful and positive sporting experiences than very small towns or large cities (Côté et al., 2006).

Support for the birthplace effect, however, has been equivocal (Baker et al., 2009), and as such requires further investigation. Studies of the birthplace effect to date have typically extracted data pertaining to birthplace or home town from professional athletes' open source biographical records (Baker & Logan, 2007; Baker et al., 2009; Côté et al., 2006; Curtis &

Birch, 1987; MacDonald et al., 2009). As a result, investigations have failed to take into consideration the possibility that athletes may have relocated to a different residence prior to or during their early sporting experiences; in which case, another town, potentially of a different population, may have been more influential on their development than their birthplace. This issue was highlighted by Baker et al. (2009), who observed inconsistencies between the location of German athletes‘ birthplace and the location of their first club, hence questioning the validity of birthplace as a proxy for athletes‘ early developmental environments.

Accordingly, it is important to consider not only athletes‘ birthplace, but also place of residence during various stages of development. It would also be interesting to identify relocations for reasons relating to sport, including athletes‘ age and stage of development at the time of relocation, as well as the nature of the relocation (i.e. from a smaller city to a larger city, to a state or national training centre etc.), in order to explore the potentially reciprocal relationship between place of residence and the development of sport expertise in more detail. Therefore, items in this section required participants to recall the location of each home in which they have lived, the ages during which they lived there, and who they lived with. For each new place of residence, athletes were also required to specify the particular reason for relocation. This information allows for examination of typical city sizes during various stages of development, as well as identification of when, and how many times athletes relocated for reasons related to sport. All responses for section four were provided in chart format.

5. Organised Sport Practice History

This section of the DHAQ addresses athletes‘ participation in sports other than their main sport. One of the most debated areas in the literature relating to the development of sport expertise surrounds early specialisation or diversification as the optimal pathway for success. The deliberate practice framework (Ericsson et al., 1993) suggests that in order to attain expertise in any given domain, it is important to engage in large amounts of highly structured, highly effortful practice activities that are designed with the specific goal of performance improvement. The framework proposes that the effects of practice are cumulative, so a participant who initiates deliberate practice at a later age will be unable to catch up and surpass another who began at an earlier age. However, large investments in a single sport at a young age are also associated with a number of risks and detrimental effects that could outweigh the performance benefits of engaging in such high amounts of practice (Wiersma, 2000).

Early diversification has been proposed as an alternative pathway towards sport expertise (Baker, 2003). This approach suggests that it is more beneficial to engage in a number

of different sporting activities throughout the early stages of learning, as this encourages the development of a wider range of fundamental movement skills, physiological capacities, and cognitive abilities that can be transferred from one activity to another (Baker, 2003). In addition, a diversified approach to development has been associated with a reduction in the number of practice hours required to attain expertise in the eventual chosen domain (Baker et al., 2003b).

In section five, athletes recalled all of the organised sporting activities they participated in, other than their main sport. For each year of their involvement in each other sport, athletes were required to indicate the number of months per year they participated in this sport, along with the average number of hours per week they engaged in all practice and competition activities combined. Athletes were also asked to identify the highest level of competition they participated at during each year of their involvement in each of the organised sports listed. The information in this section provides a detailed view of not only the number and types of sports participated in, but also the depth and competitive level of involvement in each particular sport. The procedures in section five are similar to those utilised by Baker et al. (2003b; 2005), Berry et al. (2008), and Côté et al. (2005), requiring the completion of a single chart encompassing all of the aforementioned items.

6. Participation in Informal, Playful Sporting Games

Section six of the DHAQ was designed to investigate athletes‘ involvement in deliberate play. Deliberate play refers to those activities that are loosely structured and involve minimal adult supervision; participation is voluntary and the activities are enjoyable, intrinsically motivating, and provide immediate gratification (Côté & Hay, 2002). It has been suggested that these kinds of activities not only develop fundamental movement skills, but also provide the enjoyment and motivation necessary for continued involvement in organised sporting activities (Côté & Hay, 2002). Deliberate play has been identified as influential in the development of sport expertise within small samples of athletes (Baker et al., 2003b; Berry et al., 2008; Ford et al., 2009; Soberlak & Côté, 2003). As such, further investigation into the generality of this finding across a larger, more diverse sample of athletes is required.

Therefore, in this section, athletes were asked to recall details of their participation in any informal, playful sporting games that resembled competitive sports but may have involved modified rules and/or equipment, and did not involve formal instruction or supervision. Similar to the approach adopted by Baker et al. (2003b), Berry et al. (2008), and Côté et al. (2005), athletes identified all activities they engaged in that fit this description, along with the ages during which they regularly participated in each activity. For each year of their participation in

each informal, playful sporting game, the number of months per year and the average number of hours per week they typically engaged in this activity were also recorded.

Additionally, information regarding who athletes‘ tended to play these games with, the ages of these people (three or more years younger/older, one to two years younger/older, or around the same age), and where these games were generally played (back yard, local streets, local park, local school yard etc.) was collected in order to gain a more detailed understanding of the context of engagement in these informal, playful sporting games, This section aimed to identify the influence and relative importance of participation in deliberate play during the early stages of development, and the context in which deliberate play occurs. Once again, all responses for section six were provided in chart format.

7. Main Sport Practice and Competition History

Section seven of the DHAQ was the most extensive, as it covered various aspects of athletes‘ involvement in their main sport. This section included a number of sub-sections addressing different components of athletes‘ practice and competition history. A description of each of the sub-sections is provided below.

7.1 General involvement in main sport. Athletes were first asked to provide general information regarding their involvement in their main sport. For each year of their participation, athletes recalled the total number of months involved in their main sport, the average number of hours per week engaged in all practice and competition activities combined, the average number of practice sessions per week, and the average frequency of competitions. Additionally, for each year of involvement, athletes indicated the highest level of competition reached, any significant achievements obtained, and the main reason for their participation.

Details of practice and competition involvement are important to assess the nature of the relationship between hours of practice and performance (Ericsson et al., 1993; Simon & Chase, 1973). Coupled with indicators of highest level of competition and significant achievements, details of the number of hours per week and months per year involved in all activities relating to the athlete‘s main sport can be utilised to quantify the time commitment associated with different stages of development and increasing levels of expertise. This information can then be compared between sports, countries, sexes, and skill levels.

Identification of the main reason for the athlete‘s participation in their main sport during each year of their involvement was incorporated into the DHAQ to explore how motivations for participation change over time, and how these motivations may be related to time spent in practice and/or performance. This item was inspired by Bloom‘s investigations of the

development of talent (1985) which highlighted that exceptional performers transition through three stages of development across their career. Among others, one of the suggested markers indicating progression from one stage to the next included a change in motivation for participation, so we sought to investigate this concept in more detail. Information obtained within this sub-section can be used for both within- and between-subject analyses of involvement in main sport. Responses were collected in chart format, with the chart adapted from similar investigations conducted by Baker and colleagues (2003b; 2005).

7.2 Detailed history of involvement in main sport. This sub-section aimed to gain a more detailed understanding of the types of practice associated with expert performance. The 10-year rule (Simon & Chase, 1973) and the theory of deliberate practice (Ericsson et al., 1993) have been supported in a number of different sports; however, the specific types of activities that constitute deliberate practice have been debated (Helsen et al., 1998; Hodge & Deakin, 1998; Hodges & Starkes, 1996; Starkes et al., 1996; Ward et al., 2007). Identification of the patterns of participation in different types of practice activities expert athletes typically engage in may provide more precise recommendations for coaches and athletes regarding how best to structure practice in order to optimise development.

Items within this sub-section were loosely based upon questions commonly asked within studies of deliberate practice in sports (Côté et al., 2005; Ericsson et al., 1993; Helsen et al., 1998; Hodge & Deakin, 1998; Hodges & Starkes, 1996; Starkes et al., 1996; Ward et al., 2007), but were expanded to incorporate a wide range of practice types and practice conditions. For each year of their involvement in their main sport, athletes were required to identify the average number of sessions per week, hours per week, and months per year they participated in supervised group practice, supervised individual practice, unsupervised individual practice, unsupervised play related to their main sport, passive activities related to their main sport (for example watching matches live or on television, reading books about their sport, or having conversations about their sport with team mates or coaches), and/or organised competition.

Following this, athletes recalled the average number of hours per week and months per year they participated in a range of specific practice activities including technique/skills training, tactical/games-based training, physical conditioning/weights, mental/psychological skills training, recovery techniques, video analysis/review, study related to their main sport, and/or watching their main sport live or on television. Details regarding the nature and type of practice athletes participated in can be compared across sports, countries, sexes, skill levels, and stages of development, in order to identify trends relating to the characteristics of practice that appear to be associated with the development of sport expertise. Responses in this section were collected via a series of charts.

7.3 Detailed competition history for main sport. This sub-section aimed to examine athletes' patterns of participation in competition to explore associations between competition involvement and expert sport performance. It has been suggested that an overemphasis on competition during the early stages of learning is detrimental for athlete development and that the system of competition can "make or break an athlete" (Balyi, Cardinal, Higgs, Norris, & Way, 2005, p. 32). In addition, the Canadian Sport for Life Long-Term Athlete Development framework provides clear recommendations regarding the optimal training to competition ratios that should be adopted throughout various stages of development (Balyi et al., 2005). However, the validity of these recommendations has not been investigated, and competition involvement has only been considered within a small number of studies investigating sport expertise development (Baker et al., 2003a; Ford et al., 2007). This is surprising given that both coaches and athletes frequently report competition as a critical factor influencing development (Baker et al., 2003a; Ford et al., 2009; Singer & Janelle, 1999; Ward et al., 2007).

In section 7.3, for each year of involvement in their main sport, athletes were asked to identify all levels of competition they participated at (for example, school, club, state, national, or international), and for each level of competition, the age group classifications they competed in (for example, under 14 years, 14/15 years, or open age group). Further, for each age group classification and level of competition, athletes recalled the average frequency of competitions, and described their average success (for example, mostly among the top three competitors, mostly among the middle of the pack, or won matches more often than lost matches).

Time spent in competition was included within sub-section 7.2; thus this sub-section provided insight into the frequency of participation in competition, progression through the various levels of competition, and the occurrence of 'competing up' an age group against older athletes. Playing with and/or against older competitors is often expressed anecdotally by athletes as a critical factor for the development of sport expertise (Phillips et al., 2010a), and has been proposed as one of the several interacting mechanisms contributing to the birthplace effect (Côté et al., 2006). As for the other sub-sections relating to participation in main sport, responses were collected in chart format.

8. Coaching History

Section eight collected information regarding all the coaches athletes' had ever trained under. In retrospective accounts of their development, athletes commonly identify their coaches as particularly influential in their career (Bloom, 1985, Durand-Bush & Salmela 2002; Oldenziel et al., 2003). In many cases, a new coach is associated with a transition from one

stage of development to the next (Bloom, 1985). For each coach they had ever trained under, participants were asked to identify the role of the coach in the practice environment (e.g. head coach, assistant coach, position coach etc.), the ages during which they trained with that coach, and the various levels of competition participated at under the supervision of that coach.

This information can be utilised to investigate areas such as the timing of coach transitions in relation to the developmental process, the effects of a new coach on an athlete‘s practice schedule and practice composition, and how the number and roles of coaches change throughout development. Examination of these factors will provide a more detailed understanding of the influence coaches may have on the development of sport expertise. Responses within this section were completed in chart format.

9. Support Services

This section was designed to examine athletes‘ access to and utilisation of support services. In contemporary sport it is not uncommon for a coach to be accompanied by a team of support staff specialising in areas such as medicine, physiotherapy, psychology, physiology, and biomechanics. Several studies have identified the importance of these support staff in ensuring optimal performance (Durand-Bush, & Salmela, 2002; Smith, 2003); however, the availability and utilisation of sport science and sport medicine specialists throughout developmental has been relatively under-examined. Section nine of the DHAQ aimed to address this issue.

Specifically, athletes were asked to identify the support staff they had access to during each year of their involvement in their main sport, and to indicate services they utilised on a regular basis. The information obtained can be used to investigate the roles that sport science and sport medicine specialists play in the developmental process, and how these roles differ across sports, countries, sexes, skill levels, and stages of development. This section involved a chart in which a range of sport science and sport medicine services were listed, and athletes simply put a tick or a cross in the box to indicate their access to and utilisation of each service, across each year of their involvement in their main sport.

10. Injury, Illness, and Time Off

The final section of the DHAQ required athletes to identify times during which they were prevented from participating in their main sport as a direct result of injury or illness, as well as any other times during which they spent a significant period of time away from their main sport for some other reason. In earlier sections of the DHAQ, athletes indicate the

‗average' number of sessions per week, hours per week, and/or months per year they participated in various activities related to their main sport. These estimates are provided as an average over an entire year. To ensure accurate calculation of hours spent in practice and competition over a career span, it is important to factor in any periods during which participation did not conform to the ‗average'. As such, this section of the DHAQ was designed to identify any significant periods of time away from the athlete's main sport due to injury, illness, or some other reason, and to investigate how these periods may have affected their involvement and achievement in sport. Periods of non-involvement and reduced participation can then be factored into calculations of total time spent in practice and competition activities.

In this final section, athletes were required to indicate any injuries or illnesses endured, the age at which the injury or illness occurred, the total time completely unable to participate in sport, the total time restricted to a reduced training load, and the total number of competitions missed. For periods of significant time away from main sport for reasons unrelated to injury or illness, athletes were required to specify the age at which this time off occurred, the main reason for the time away, the total time away from training, and the total time away from competition.

Information obtained in this section is important not only for the accurate calculation of hours spent in practice and competition activities, but also to investigate the relationships between injuries, illnesses, and significant periods away from sport, and a variety of factors including performance, stage of development, practice scheduling, practice characteristics, and the utilisation of support staff. These relationships can then be compared between sports, countries, sexes, and skill levels. As for many other sections of the DHAQ, the information in section ten was collected via completion of a series of charts.

The DHAQ is an extensive measurement tool covering many different aspects of the development of sport expertise. There are, however, additional factors that have been shown to influence development that have not been included (for example motivation, personality, and other psychological factors; and the provision of financial, social, and tangible resources). This questionnaire focuses primarily on the patterns of sport participation and the characteristics of practice that are associated with expert sports performance, along with several additional environmental and contextual factors. Collection of this information has the potential to provide considerable insight into these specific influences on the development of sport expertise, and would provide a valuable contribution to this field of research. However, it is essential that the questionnaire elicits valid, reliable responses in order to make accurate conclusions regarding athlete development. Therefore, the DHAQ must be subjected to rigorous validity and reliability assessments before utilisation within large scale examinations of long-term athlete development.

Study Phase Two: Establishing Validity and Reliability of the Developmental History of Athletes Questionnaire

This chapter provides details of the investigation designed to assess the validity and reliability of the DHAQ. Procedures, including the development of a rigorous methodology and taxonomy for classifying validity and reliability of retrospective recall are outlined, results are discussed, and recommendations to improve the DHAQ are suggested.

Methodology

Participants

Participants included 15 Australian athletes, all of whom had competed in their sport at the national level or above. Four males and four females were recruited from the Victorian Institute of Sport men‘s and women‘s field hockey teams, and three males and three females were recruited from the Australian Institute of Sport swimming team. One additional male swimmer was recruited as a former member of the New South Wales Institute of Sport swimming team. At commencement of testing, the athletes had a mean age of 21.4 years (*SD* = 2.6 years), with an average of 13.8 years of experience in supervised activity for their main sport (*SD* = 3.3 years). Individual participant information is detailed in Table 1.

Male and female athletes from two sports were sought to ensure the DHAQ was validated for both team and individual sports, as well as for both sexes; however, sample size was restricted to 15 athletes from a single country due to the labour intensive nature of the procedures, and the necessity for direct researcher-participant contact throughout the course of the study. Although small numbers of homogenous participants was highlighted earlier as a major limitation of previous research, the current study aimed to validate the DHAQ rather than test hypotheses concerning the development of sport expertise. As such, a large sample size is not as pertinent in this study as it will be in future investigations that utilise the DHAQ to explore factors associated with sport expertise development. While a large sample size is always preferable, the detailed methodology of the current study placed a practical restriction on the number of participants that could be effectively managed to achieve the stated objectives. Notably, previous investigations related to sport expertise development have conducted validity and/or reliability analyses with as few as four participants (Soberlak & Côté, 2003) so although small, a sample size of 15 was greater than many similar assessments conducted in the past.

Ethical approval was obtained from Institutional Ethics Committees prior to the commencement of research (see Appendix B), and informed consent was received from all participants (see Appendix C).

Research Design

This phase of the investigation involved a repeated measures design to assess validity and reliability of the DHAQ. Both retest and alternative-form methods of repeated measures were adopted (Carmines & Zeller, 1979). Since the DHAQ is a retrospective questionnaire

requiring participants to think back and recall specific details related to the history of their involvement in sport and physical activity, establishing reliability of recall for events that occurred in the past is imperative before administering the DHAQ for investigational purposes. Additionally, establishing validity of the DHAQ is critical to ensure the instrument elicits high quality, trustworthy data. A repeated measures approach allows for examination of response consistency across multiple test occasions and criterion, hence is appropriate for establishing both validity and reliability of the measurement tool (Carmines & Zeller, 1979).

Measures

All items within all sections of the DHAQ, as described in study phase 2 were subjected to validity and reliability analyses. Concurrent validity was assessed via comparison of responses obtained from the DHAQ to those obtained from a similar, previously validated measurement tool; and convergent validity was assessed via comparison of responses obtained from the DHAQ to those obtained from a similar measurement tool administered to athletes' parents and coaches. Test-retest reliability was assessed via administration of the DHAQ to athletes on a second occasion; and internal consistency was assessed wherever possible via comparison of responses obtained from similar items within the DHAQ.

Procedures

Assessment of validity and reliability of the DHAQ involved quantitative and qualitative techniques. In addition to completing the DHAQ (as described in study phase one), qualitative semi-structured interviews were conducted with all participants. The interview guide was based on the procedure outlined by Côté et al. (2005), which has been previously validated, and utilised in a number of investigations of sport expertise development (Baker et al. 2003a; Baker et al., 2003b; Baker et al., 2005; Berry et al., 2008; Law et al., 2007; Soberlak & Côté, 2003). Therefore, comparison of responses from the DHAQ to those elicited by this established interview procedure allows for examination of concurrent validity of the questionnaire.

A number of minor changes were required to adapt the Côté and colleagues (2005) interview guide to the needs of the current investigation. First, items that failed to produce strong scores for validity or reliability during the 2005 investigation were removed. Second, the original interview guide did not address all areas encompassed within the DHAQ, so several additional items were incorporated to ensure that all quantitative responses from the DHAQ

could be compared to a qualitative response from the interview. The full athlete interview guide is provided in Appendix D.

Participants completed the DHAQ and the semi-structured interview in a quasi-randomised order such that two males and two females from each sport completed the interview followed by the questionnaire, while the remaining athletes completed the questionnaire before the interview. The questionnaire and interview procedures were completed on two separate occasions, with a minimum of one day between tasks (M = 5.20 days, SD = 6.44 days). Fourteen athlete questionnaires and twelve athlete interviews were completed in a private room at the athlete‘s training venue. Due to participant availability and logistical issues, one athlete questionnaire was completed unsupervised at the athlete‘s home, and three athlete interviews were conducted via telephone.

The DHAQ was presented as a booklet, and participants provided written responses in the form of short answers, check boxes, and charts. Participants who completed the DHAQ in a private room at their training venue did so under the supervision of the lead researcher, however, no additional instructions other than those documented within the written questionnaire were provided. Once validated, it is intended that the DHAQ will be suitable for administration to large numbers of athletes, in a range of environments. Consequently, on occasion the questionnaire will be required to be completed without the direct supervision of a researcher. It is therefore important to assess the validity and reliability of the DHAQ under similar conditions to which it will eventually be administered. Although in this case a researcher was present in the room during the completion of the DHAQ, they did not provide assistance to participants. The role of the researcher was to monitor the time taken to complete the questionnaire, and to note any questions that may arise as an indication of where instructional sets or response charts may need to be altered in order to improve clarity. Several participants did seek assistance while completing the questionnaire; however, on these occasions the researcher simply informed the athlete to interpret or provide information as they felt appropriate. Questions were recorded in the researcher‘s field notes for consideration during data analysis and revision of the DHAQ. The participant who completed the DHAQ unsupervised was encouraged to contact the lead researcher if they had any questions relating to the questionnaire; however, additional assistance was not requested. On average, the total time required to complete the DHAQ was approximately 1 hour and 10 minutes (SD = 18 mins).

Athlete interviews were conducted one-on-one with the lead investigator, using the semi-structured interview guide described above. The interview consisted of a series of main questions that encouraged detailed, open-ended responses. As in previous studies (e.g., Côté, 1999; Côté et al., 2005), these main questions were supported by probe and follow-up questions.

Probe questions were used to encourage participants to provide more detailed information on a particular topic or theme, and to clarify information that had been given. Follow up questions were used to pursue topics and themes that emerged from participants‘ responses throughout the interview process. Leading questions directing the participants‘ thoughts towards a particular response were avoided at all times, and in some cases, charts were used to assist in the data collection procedure. On average, the total time required to complete the athlete interview was 1 hour and 15 minutes (*SD* = 20mins). With written informed consent from all participants, interviews were recorded using a digital voice recorder, and were subsequently transcribed verbatim. Following transcription, the information was used to complete an additional copy of the DHAQ so responses from both the questionnaire and interview could be directly compared. This comparison allowed for assessment of concurrent validity of the DHAQ with a previously validated data collection procedure (i.e. Côté et al., 2005).

A second procedure commonly used to assess validity of measurement tools designed for the collection of athlete developmental histories is to interview not only athletes, but also others who are significantly involved in the athlete‘s sporting career; namely, parents and coaches (Baker et al., 2003a; Baker et al., 2003b; Berry et al., 2008; Côté, 1999; Côté et al., 2005; Durand-Bush & Salmela, 2002; Law et al., 2007; Oldenziel et al., 2003; Soberlak & Côté, 2003). Collecting information from parents and coaches allows for triangulation of the data to assess convergent validity of responses. As such, all athletes were asked to provide contact details for the parent most significantly involved in their sporting commitments, their current coach, and one former coach of their choice. Parents and coaches for whom details were provided were contacted and invited to become involved in the research, resulting in 13 parents and 22 coaches agreeing to participate in a telephone interview (refer to Table 1 for details of parent and coach involvement for each participant).

The guide for the parent and coach interviews was the same as that used for the athlete interviews; however, items that were not applicable to the interviewee were removed. For example, sections of the interview relating to the athlete‘s family and residential history, as well as their involvement in other organised sports and informal, playful sporting games were removed from the coach interview guide as it was unlikely that the coach would be familiar with this information (see Appendices E and F for full parent and coach interview guides). Parent interviews lasted approximately 1 hour and 5 minutes (*SD* = 26 mins), while coach interviews lasted approximately 40 minutes (*SD* = 12 mins). Coach interviews were considerably shorter than parent interviews because only the time period during which the athlete trained under each respective coach was addressed, compared to the athletes‘ entire career span for parent interviews.

As with athletes, following provision of written informed consent, parent and coach interviews were recorded using a digital voice recorder and were subsequently transcribed verbatim. Following transcription, all participating parents and coaches received a copy of the script, and were instructed to read the document carefully to verify the accuracy of information they provided. Upon approval of interview transcripts from parents and coaches, responses were again used to complete additional versions of the DHAQ in order to allow accurate comparison of data between measures. Interview transcript verification was not incorporated into the athlete interview procedures as a test-retest protocol was adopted for the questionnaire, and the verification process would have interrupted this protocol.

The test-retest protocol was adopted to assess test-retest reliability of the DHAQ. Approximately three to four months following the initial test occasion, all athletes were invited to complete the DHAQ a second time. This procedure allowed for assessment of consistency of responses when participants were asked the same questions, in the same format, on separate occasions following a significant delay. Participants received the second questionnaire by post, and were instructed to complete and return it in their own time. As mentioned previously, it is a realistic possibility that future studies utilising the DHAQ will require participants to complete the questionnaire unsupervised. As such, the second completion of the DHAQ was intentionally unsupervised to examine the quality and trustworthiness of data collected without the assistance of a researcher. Of the 15 questionnaires posted for the test-retest protocol, 11 were completed and returned. Participant details relating to the return of retest questionnaires are provided in Table 1. A minimum three month period between test-retest occasions was adopted to ensure minimal remembering of responses from the initial questionnaire and interview procedures.

One final procedure was adopted to assess internal consistency of the DHAQ. The questionnaire included several redundant items whereby the same information was collected using two or more slightly different questions. For example, an initial question required athletes to estimate the total number of hours per week they were involved in all types of practice activities for each year of participation in their main sport, and a subsequent question required athletes to identify the total number of hours per week involved in various types of practice activities during each year of their participation. The practice hours in each type of activity can be summed and compared to the estimate provided for total practice hours. Similarity of responses between redundant items indicates high internal consistency, representing high quality, trustworthy data. Assessment of response similarity between redundant items is a technique that has been utilised for establishing reliability of recall in a number of studies related to sport expertise development (Baker, et al., 2003a; Baker et al., 2003b; Baker et al. 2005; Hodges et al., 2004; Moesch et al., 2011; Oldenziel et al., 2003).

The procedures for this phase of the study allow for rigorous examination of the robustness of the DHAQ via assessment of concurrent validity, convergent validity, test-retest reliability, and internal consistency. Such intensive evaluation of a measurement tool for the collection of athlete developmental histories is quite unique within this field of research, and will provide valuable insight into the trustworthiness of retrospectively recalled athlete history information.

Statistical Analysis

Validity and reliability of each item within the DHAQ were established via a series of comparative analyses between responses from the initial completion of the DHAQ to those provided during the athlete, parent, and coach interviews, and the second completion of the questionnaire. Concurrent validity was assessed through comparison of responses from the first completion of the DHAQ with those from the athlete interview. Convergent validity was assessed through comparison of responses from the first completion of the DHAQ with those from each of the parent and coach interviews. Test-rest reliability was considered through comparison of responses from the first completion of the DHAQ with the second. In addition, internal consistency was assessed for the initial completion of the DHAQ through the comparison of responses provided to redundant items within the questionnaire. All statistical analyses were conducted using the SPSS statistical computer software package, Version 17.

The statistical analyses performed were consistent across all assessments of validity and reliability. This approach was adopted in order to allow for uniform interpretation and subsequent classification of questionnaire items. Percent agreement values (PA) and intraclass correlation coefficients (ICC) were calculated for all items within each of the comparative pairs outlined above. An exception occurs for items involving categorical responses, in which case, only PA was considered. These two methods of analysis were adopted in order to obtain indications of both absolute agreement and relative consistency between test occasions (Atkinson & Nevill, 1998).

Previous investigations of the development of sport expertise have typically assessed consistency of recall using Pearson product moment correlation analyses (Baker et al., 2003a; Baker et al., 2003b; Baker et al., 2005; Berry et al., 2008; Côté et al., 2005; Helsen et al., 1998; Hodges & Starkes, 1996; Memmert et al., 2010; Oldenziel et al., 2003; Ward et al., 2007). This approach was not utilised in the current study as the Pearson correlation was considered to be inappropriate because they are a bivariate statistic, intended for assessing the strength of the association between two different variables, for example, height and weight (Haggard, 1958).

Intraclass correlations however, are classified as a univariate statistic, and are more appropriate for assessing the strength of the relationship between multiple measurements of the same variable (Haggard, 1958). The ICC is considered a more suitable indicator of the relationship between multiple measurements of the same variable because both within-subject variance and total variance across all measures for all subjects are considered in its calculation (Haggard, 1958; Bartko, 1976). Intraclass correlations are also preferred over Pearson correlations when sample sizes are small (Garson, 2012), as in this investigation.

ICCs were calculated according to the two-way random effects model, with absolute agreement (Shrout & Fleiss, 1979). The two-way random effects ICC is most suitable for data sets in which the responses are provided by participants who are considered to be randomly selected from a larger sample, and for which the results are to be generalised to respondents other than those involved in the current investigation (Bartko, 1976; Bartko, 1966; Garson, 2012; Shrout & Fleiss, 1979; Krebs, 1986; Müller & Büttner, 1994). Additionally, the absolute agreement ICC is most suitable for data sets in which the similarity of responses across multiple measurements is of interest, and not just the relationship between the responses (Garson, 2012). In all cases the single measures ICC is reported, indicating that individual pairs of responses for each participant were incorporated in the calculation as opposed to group means (Garson, 2012; Ebel, 1951).

ICCs will in most cases range from 0-1, with larger values indicating lower levels of variance and a stronger relationship between responses provided across multiple test occasions (Garson, 2012; Costa-Santos, Bernardes, Ayres-de-Campos, Costa, & Amorim-Costa, 2011). In select cases, ICCs can be negative however in the majority of instances the interpretation of a negative ICC is unclear (Costa-Santos et al., 2011; Haggard, 1958; Müller & Büttner, 1994). Negative ICCs are often associated with errors in data entry or small sample sizes (Garson, 2012; Haggard, 1958). It is likely that in this study, negative correlation coefficients are attributable to an unfavourable combination of a small sample size with a high degree of total variance within the data set. As such, any negative ICCs were examined for data entry errors before being interpreted as having zero correlation (Haggard, 1958).

Despite their common application, correlation analyses have a major limitation associated with their use in studies of validity and reliability. Correlation coefficients provide valuable information regarding the strength of the relationship between two variables; however, they do not address the degree of similarity between them (Altman & Bland, 1983; Bland & Altman, 1986; Costa-Santos et al., 2011; Kottner & Dassen, 2008; Müller & Büttner, 1994). It is possible for two sets of values to be highly correlated, yet not similar in magnitude (Bland & Altman, 1986; Burdock, Fleiss, & Hardesty, 1963). To address this limitation within the current

study, PAs were calculated for all identified pairs of responses in addition to the ICC. The use of a combination of statistical methods for assessing validity and reliability has previously been recommended within the fields of nursing, medicine, and sports medicine (Atkinson and Neville, 1998; Kottner & Dassen, 2008; Kottner et al. 2011), and was adopted by Ford and Williams (2008) in a similar investigation of the development of sport expertise in Irish professional soccer players.

PAs are most typically calculated when data are categorical, indicating the proportion of responses that are identical across multiple test occasions (Bahrick, Hall, & Berger, 1996; Ropponen, Levalahti, Simonen, Videman, & Battie, 2001). Delta scores, percent difference, or percent error calculations are more common when assessing agreement between continuous variables (MacDonald et al., 2009; Soberlak & Côté, 2003). Considering percent difference values are the compliment of percent agreement, it was decided to report agreement for continuous variables rather than difference, so that interpretation and classification of all questionnaire items was consistent, regardless of the nature of the response. PA calculations have also been used in several previous investigations of sport expertise development to assess validity and reliability of retrospective recall of training history information (Baker et al., 2005; Law et al., 2007; Memmert et al., 2010). As such, PAs for categorical variables were calculated as the percentage of responses that were the same for both test occasions under investigation, and PAs for continuous variables were calculated by dividing the smaller value of the pair by the larger, and multiplying the result by 100 to indicate a percentage. PAs range from 0-100%, with larger values indicating greater similarity between responses.

At this point it is important to highlight that data from the current year were excluded from all analyses in order to avoid issues relating to incomplete data sets for this year. Further, in cases where data was available for fewer than five participants (e.g., if a particular item was not applicable to all athletes, fewer than five athletes had reached a particular age, or fewer than five coaches were able to provide comment on a particular item, etc.), statistical analyses were not performed. In these cases it was deemed that insufficient data was available to conduct meaningful analyses and interpretation. Items for which insufficient data was available for analyses are indicated throughout the text and in all relevant tables in Supplement One.

Classification of Validity and Reliability of the DHAQ

The validity and reliability of each item within the DHAQ was assessed and classified according to criteria based upon both PA and ICC values. As the statistical methods utilised to assess concurrent validity, convergent validity, test-retest reliability, and internal consistency

reliability were the same, a uniform classification criteria was applied across all validity and reliability components, for all items within the DHAQ. Classification involved several steps.

Step one involved the classification of PA. Items were classified as displaying very good, good, moderate, or poor agreement according to the criteria outlined in Table 2. In a study of sport expertise development in triathletes, Baker and colleagues (2005) reported PA values of 70% as ‗reasonable agreement' and values above 80% as ‗high agreement' for retrospective recall of hours involved in sport specific training activities. Similarly, a 20% change score for retrospective recall of hours involved in physical activity in a sample of master athletes was considered acceptable in an investigation of the reliability of a semi-structured interview procedure relating to predictors of physical activity in older adults, while percent agreements of 63% and 72% for recall of several categorical variables relating to physical activity where deemed to be ‗less reliable' (MacDonald et al., 2009). In the absence of established criteria for acceptable levels of PA, and given the relatively limited use of PA statistics in studies of sport expertise development, these examples were used as a basis for the classification criteria adopted in the current study.

Step two involved the classification of the ICC. Two factors were considered in this process: a) the strength of the correlation (i.e. the value of the correlation coefficient); and b) the significance of the correlation (i.e. the *p* value associated with the correlation coefficient). ICCs were considered very strong if the value of the coefficient was equal to or exceeded .80, strong if the coefficient ranged from .65 to .79, reasonable if the coefficient ranged from .50 to .64, and weak if the value of the coefficient was .49 or below. Interpretation of the magnitude of ICCs is inconsistent, with no apparent guidelines for the coefficients as they relate to analytical research goals Atkinson and Neville (1998), and Kottner and Dassen (2008) suggested that the comparison of ICCs between studies is limited because the value of the coefficient is influenced by differences in the characteristics of the study participants, and because there are multiple models available for the calculation of the coefficient. Additionally, Costa-Santos et al., (2011) observed that a sample of medical clinicians and biostatisticians were inconsistent in their interpretations of ICCs despite being presented with the same set of results.

A wide range of ICCs have been reported as ‗high' or ‗good' within reliability investigations across a variety of fields. In the medical investigation described above, correlation coefficients as low as .65 were rated as ‗good' (Costa-Santos et al., 2011), while values of .90 have been recommended as the cut-off criteria for ‗high' correlations in tests of physiological capacities (Lemmink, Elferink-Gemser, & Visscher, 2004). In investigations of reliability of retrospective recall for recreational and occupational physical activity, values above .65 tend to be considered ‗high' (Ropponen et al., 2001), while values of .35 to .55 have

been considered ‗moderate' (Ainsworth, Richardson, Jacobs Jr., Leon, & Sternfeld, 1999; Reis, Dubose, Ainsworth, Macera, & Yore, 2005). The classification criteria established for the strength of the ICCs obtained in this study were based upon the studies of physical activity recall described above, and a suggestion that ICCs can be interpreted according to the well established criteria for the interpretation of the kappa statistic (Garson, 2012).

In a similar fashion, ICCs were considered highly significant if the *p* value was less than or equal to .01, significant if the *p* value ranged between .02 and .05, approaching significance if the *p* value was in the range of .06 to .10, and non-significant if the *p* value exceeded .10. The *p* value indicates the chance that the null hypothesis is rejected when it is actually true (Tabachnick & Fidell, 2007), or in other words, the chance that a significant finding is reported when in fact it is not significant. In scientific research, a *p* value of .05 is generally accepted as a suitable criterion for determining statistical significance, as this indicates there is a 5% chance of incorrectly reporting a significant result (Tabachnick & Fidell, 2007). Values less than .01 are also commonly acknowledged, as these findings indicate the chance the null hypothesis has been incorrectly rejected is less than 1% (Hopkins, 2000a). The criteria adopted in this study for classification of the ICC as significant or highly significant, were based upon these conventions.

In many cases *p* values above .05 are considered non-significant, however, the significance of a correlation coefficient is highly dependent upon the number of participants involved in the investigation (Haggard, 1958; Hopkins, 2000b; Morrow Jr. & Jackson, 1993). Due to the small sample size involved in this study, *p* values ranging between .06 and .10 were classified as approaching significance in order to allow a small buffer before removing an item from the DHAQ on the basis of displaying a non significant ICC. ICCs for which the *p* value exceeded .10 were classified as non-significant, as questionnaire items with probabilities of incorrectly rejecting the null hypothesis greater than 10% were deemed to be too uncertain to be considered for further investigation.

Following classification of the strength and the significance of the ICC for each item within the DHAQ, correlations were given an overall rating of very good, good, moderate, or poor, according to the criteria outlined in Table 2. Once again, as the significance of a correlation coefficient is highly dependent upon sample size (Haggard, 1958; Hopkins, 2000b; Morrow Jr. & Jackson, 1993), the strength of the correlation was weighted more heavily than its significance in the classification of the ICC overall. As such, any items with a very strong, strong, or reasonable ICC were respectively classified as having a very good, good, or moderate correlation overall, providing the coefficient was either highly significant, significant, or approaching significance. All items displaying a non-significant ICC were classified as poor overall regardless of the strength of the correlation. Similarly, all items rated as having a weak

correlation coefficient were also classified as poor overall regardless of the associated significance, because the relationship between the values was considered too inconsistent to warrant further investigation.

The final step for assessment of the concurrent validity, convergent validity, test-retest reliability, and internal consistency reliability for each item within the DHAQ involved combining the ratings obtained for agreement and correlation to provide an overall classification. All items were classified as having very good, good, moderate, or poor concurrent validity/convergent validity/test-retest reliability/internal consistency reliability, according to the criteria outlined in Table 2. PAs were weighted more heavily than ICCs in the overall classification of each item because absolute agreement is considered to be of greater importance than relative consistency (Atkinson & Neville 1998; Bland & Altman, 1986; Müller & Büttner, 1994). It is argued that if responses provided across multiple test occasions are similar in magnitude, the systematic relationship between the test occasions matters less. For this reason, items noted as displaying very good or good agreement were classified as very good or good overall, regardless of the rating of the correlation coefficient for that item.

The ICC becomes of greater interest when PA values are lower. An item displaying moderate agreement but very good or good correlation is indicative that responses provided during one test occasion tend to be systematically lower or systematically higher than those reported during the other test occasion. These items are important to note because despite displaying lower absolute agreement, the systematic relationship for the recall of information relating to these items allows for continued analyses and interpretation of the results, providing this relationship is acknowledged and applied to the data (Bland & Altman, 1986). For example, an investigation involving a sample of wrestlers observed consistent over-estimation of training hours during completion of a retrospective practice history questionnaire compared to actual reports of practice time recorded in training diaries (Hodges & Starkes, 1996). If, for instance, a similar relationship is found between quantitative questionnaires and qualitative interviews, this can be taken into consideration when interpreting results and comparing data collected via questionnaires to data collected via interviews. Therefore, any items that displayed a moderate agreement and either a very good or good correlation were classified as moderate overall. However, items that displayed a moderate agreement but only a moderate or weak correlation, were not believed to have a strong enough relationship to be considered for inclusion in the questionnaire, and thus were rated as weak overall. Similarly, any items classified as having a poor PA were also rated as poor overall regardless of the rating of the associated ICC. It was decided in this case, that responses differing by more than 50% were too inconsistent to be considered for any further analyses.

As mentioned previously, for items in which responses were categorical, ICCs could not be calculated. As such, these items were classified according to PA only, using the same criteria outlined in the discussion of agreement values above and in Table 2. Classification of concurrent validity, convergent validity, test-retest reliability, and internal consistency reliability for each item within the DHAQ allows for recommendations regarding the suitability of each item for retention in the questionnaire and utilisation in future investigations.

Results and Discussion

Sporting Career and Milestones

PA values, ICCs, and validity and reliability classification information for items within the sporting career and milestones section of the DHAQ are presented in Table 3. Note that not all items were applicable to each participant (for example, age when first recognised as the best athlete at the international level) so in some cases insufficient data were available to conduct meaningful validity and reliability analyses. Milestones for which validity and reliability statistics could not be calculated due to insufficient responses are highlighted within the table. Additionally, convergent validity with coaches could not be established for several items because milestones relating to the early stages of athletes‘ careers and events occurring outside the practice and competition environment were beyond the extent of coaches‘ knowledge. These milestones are also highlighted in Table 3.

In general, validity and reliability for this section of the DHAQ was strong, with most items receiving very good or good classifications for concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability. Three items received moderate or poor classifications for one or more validity/reliability conditions including age at first regular participation in sport-specific informal play, age at first participation in regional level competition, and the number of career transitions experienced to date. All items receiving very good or good ratings of validity and reliability across all applicable comparisons are recommended for retention in the DHAQ. The three items receiving moderate or poor ratings for one or more validity and/or reliability conditions warrant further discussion regarding their inclusion or removal from the questionnaire.

Age at first regular participation in sport-specific informal play received very good classifications for both concurrent validity and test-retest reliability, but convergent validity with parents was moderate (convergent validity with coaches was unavailable). It is possible that convergent validity with parents was lower for this item because athletes will often engage

in this type of activity during their leisure time. Although parents may be aware their child is participating in some form of physical activity during this time, it is unlikely they will attend to the specific activities the child is engaging in. In some cases, the parent may not even be present during the time the athlete is participating in play, for example during school hours, or when at friends‘ houses. As such, recall of athlete participation in unsupervised, informal play activities may be more difficult for parents. Given the high level of consistency in athlete responses for this item, it is recommended that age at first participation in sport-specific informal play be retained in the DHAQ despite the moderate rating for convergent validity with parents.

Age at first participation in regional level competition received classifications of poor for concurrent validity, moderate for convergent validity with parents, very good for convergent validity with coaches, and very good for test-retest reliability. All other items relating to participation in regional level competition received very good classifications across all validity and reliability conditions. Despite a number of very good ratings, researcher observations during athlete interviews indicated that regional level competition was not a regular fixture for either of the sports examined, nor was it considered a compulsory stepping-stone required for qualification to participate at the state level. Therefore, participation in regional level competition did not appear to be a significant milestone for these athletes, explaining the lower ratings of concurrent validity and convergent validity with parents for the item concerning age at first participation. Given the irregular nature and low importance of participation in regional level competition, it is recommended that all items relating to this level competition be removed from the DHAQ.

Also assessed for consistency of recall within this section of the DHAQ was number of career transitions experienced to date. This value was deduced from an open-ended question requiring athletes to identify particular times they noticed a change in their participation or success in their main sport, marking a transition from one stage of their career to another. Although the number of career transitions reported was classified as having good concurrent validity and test-retest reliability, qualitative examination of responses indicated that specific details pertaining to the transitions, for example age at transition and catalyst for transition, were inconsistent between test occasions. It appears that athletes were not able to consistently identify and recall times at which they noticed a prominent change in their participation or success in their main sport. This is interesting given the emphasis placed on stages of development in the literature (Balyi et al., 2005; Bloom, 1985; Côté & Fraser-Thomas, 2007; Durand-Bush & Salmela, 2002).

The results of this study suggest that although assessment of athlete developmental history profiles may lead to the identification of clear stages in sport participation, athletes are

largely unaware of their progression through these stages. It is also very interesting that convergent validity with parents for both number of career transitions experienced to date, as well as the qualitative details of each transition, was poor, suggesting that parents view their children‘s participation and progression in sport quite differently to the athletes themselves. Although outside the scope of this research, this finding may have important implications for athlete psychosocial development, and the provision of resources (financial, tangible, time, and emotional) to the athlete from the family.

Due to the inconsistency of recall associated with items relating to career transition stages, it is recommended that these items be removed from the DHAQ. While theoretical examination of stages of sport participation may be of considerable interest to sport researchers, coaches, and administrators, since athletes do not seem to be overtly aware of their transition through these stages, there is little value in explicitly questioning them on this topic.

Family Characteristics and Participation in Sport and Physical Activity

Table 4 displays validity and reliability information for items within the family characteristics and participation in sport and physical activity section of the questionnaire. All participants reported living with both parents in intact families, so results pertain to biological parents and full siblings in all cases. The majority of participants (n = 10) reported having two siblings, so validity and reliability of recall for sibling characteristics and participation in sport and physical activity was assessed for siblings one and two only. For athletes reporting three or more siblings (n = 3), the two oldest siblings were selected for inclusion within reliability and validity analyses. Items pertaining to the family were not included within the coach interview guide, so convergent validity with coaches was not assessed for this section.

As expected, date of birth and nationality were recalled consistently for all family members with concurrent validity, convergent validity with parents, and test-retest reliability all classified as very good. The ages during which athletes lived with each family member were also rated very good for all validity and reliability conditions. Parental occupations were recalled less consistently, with father‘s occupation receiving a moderate classification for concurrent validity, and mother‘s occupation receiving moderate ratings for both concurrent validity and convergent validity with parents. Researcher observations during athlete and parent interviews highlighted that in many cases, participants‘ mothers held several jobs since the athlete‘s birth. Therefore, it may have been difficult for athletes and parents to be consistent in their recall of mother‘s occupation. Although father‘s occupations were more stable over time, observations during athlete interviews revealed that participants could not always classify their

father‘s occupation with certainty. In some cases, athletes indicated they were not exactly sure what their father‘s occupation involved. The basis for inclusion of this item within the DHAQ was to provide an indication of family socio-economic status. Given the inconsistent recall of parent‘s occupations, it is recommended that this item be replaced with an indication of the highest level of education each parent has completed. It is expected that recall of highest level of education completed will be more consistent given the reduced ambiguity of the available response options; however, validity and reliability of the revised item should be monitored.

Familial involvement in physical activity was examined first via a yes/no question: ―During the time you lived with this family member, did they participate in regular physical activity‖; followed by an open-response question: ―If yes, on average how often did they participate in physical activity‖. The majority of siblings were reported to have regularly participated in physical activity, resulting in very good or good classifications for this item across all validity and reliability conditions, for both siblings examined. Validity and reliability ratings for fathers were also very good or good, but convergent validity with parents for mother‘s participation in regular physical activity was classified as moderate. Discrepancies between responses in the DHAQ and parent interviews occurred when athletes indicated their parents did not participate in regular physical activity but parents indicated they did. It is possible that athletes and parents perceptions of what constitutes physical activity differ, or it may be that children are unaware of the physical activity pursuits of their parents, which is particularly likely if parents engage in physical activity during times when their children are at school, training, or playing with their friends.

Difficulties classifying constituents of physical activity and/or lack of awareness of the physical activity patterns of family members is reflected in relatively poor ratings of validity and reliability for the item related to frequency of physical activity participation. Classifications for this item were mostly moderate or poor. Given the lack of research investigating the influence of familial participation in physical activity on the development of sport expertise, it is recommended that these items be revised and monitored for validity and reliability in future investigations rather than be removed.

In an attempt to improve identification of the constituents of physical activity, it is recommended that athletes be prompted to identify the frequency of their family members‘ participation in each of three different forms of physical activity: 1) general fitness activities; 2) recreational sport/informal sporting games; and 3) competitive sport. It is also recommended that specific response options be provided (e.g. never, occasionally, 1-2 times per week, 3-5 times per week, more than 5 times per week) to remove response ambiguity and improve

consistency. It is believed that lower ratings of validity and reliability for these questionnaire items could be the result of sub-optimal item structure as opposed to poor recall per se.

The remaining items within this section of the DHAQ concerned familial involvement in competitive sport. Athletes first identified whether each parent and sibling had participated in competitive sport during any time in their lives. If yes, athletes then identified each sport in which their family member(s) participated, along with the highest level of competition reached for each sport. Consistency of recall for these items was mixed. For mothers, concurrent validity, convergent validity with parents, and test-retest reliability were very good or good for both identification of whether she participated in any competitive sports and the total number of competitive sports reported. For fathers, all conditions of validity and reliability were classified as very good or good for whether he participated in competitive sport, but concurrent validity with parents was poor for the total number of competitive sports identified. This was because parents tended to report a greater number of competitive sports for fathers than athletes. Researcher notes recorded during interviews indicated that parents referred to sports athletes‘ fathers participated in during high school and/or university at the inter-school level. It is possible that athletes‘ fathers had not discussed their minor involvement in these sports with their children, explaining the lower convergent validity with parents for these items.

For siblings, identification of whether they participated in competitive sport received ratings of very good or good for all validity and reliability conditions, but the number of competitive sports reported tended to differ between test occasions, with convergent validity with parents and test-retest reliability rated as poor for sibling two. In the majority of cases, siblings participated in most of their sports at the school or local club level. Unfortunately, the duration of their involvement in each of these sports was not collected. It may be that athletes‘ siblings were sampling a variety of sports at a relatively low competitive level. These sports may not have required a significant commitment from the siblings, making their involvement difficult to recall.

Validity and reliability of sport type and highest level of competition was assessed for up to three sports for each family member. Recall of this information was highly inconsistent with reliability and validity classifications ranging from very good to poor. Interestingly, identification of sport type and associated highest level of competition were more consistent for parents than siblings. This finding supports the suggestion that poor recall of sibling involvement in competitive sport may be related to a high level of sport sampling. Athletes are less likely to be aware of the sports their parents may have sampled for a short period of time, listing only those sports in which their parents were involved for a significant duration, thereby reducing the variability associated with responses relating to parental involvement in

competitive sport. Once again, given the lack of research examining the role of familial involvement in competitive sport in the development of sport expertise, it is recommended that these questionnaire items be revised, requiring athletes to identify only those competitive sports in which their family members have participated for a significant duration, for example three years or more. This revision would alleviate the possibility of responses being affected by high levels of sport sampling among family members.

Places of Residence

As seen in Table 5, validity and reliability of items pertaining to athletes‘ places of residence were reasonably good. As for family characteristics, these items were not included within the coach interview guide, so convergent validity with coaches was not examined for this section of the DHAQ. On average, participants reported three to four places of residence (DHAQ: M = 3.40, SD = 1.77; athlete interview: M = 3.80, SD = 1.86; parent interview: M = 3.77, SD = 1.83; DHAQ-retest: M = 3.91, SD = 1.70); and concurrent validity, convergent validity with parents, and test-retest reliability for total number of residences to date were all classified as very good. Although some participants reported up to eight places of residence, validity and reliability were only assessed for residences one to four as the number of participants reporting five or more residences (n = 3) was considered too small to conduct accurate analyses.

All information relating to the first home in which participants lived was recalled with excellent consistency. Location of residence and details of whom participants‘ lived with at this residence were both rated as having very good concurrent validity, convergent validity with parents, and test-retest reliability. Similarly, details of participants‘ second and third residences were also recalled consistently, with very good or good validity and reliability ratings for all items. Note that convergent validity with parents for items relating to residence three was slightly lower than concurrent validity and test-retest reliability; however, convergent validity with parents was still rated as good for all items.

Although most items for residence four were classified as having good validity and reliability, there were a few exceptions. Test-retest reliability for reason for relocation to residence four was classified as moderate, and identification of who participants lived with at this residence was classified as moderate across all validity and reliability conditions. Interestingly, athletes and parents were not consistent in their identification of the location of this residence, indicated by a poor rating for convergent validity with parents for this item.

For most athletes, residences one and two tended to be of longer duration than residences three and four. It is likely that details corresponding to residences of longer duration will be more memorable, hence recalled with greater consistency. This could explain the slightly lower validity and reliability scores for residences three and four compared to residences one and two. In addition, the majority of athletes reported moving out of the family home prior to their participation in this study. Understandably, when athletes moved to a residence away from their parents, convergent validity with parents declined. Even though convergent validity with parents was moderate or poor for several items relating to residence four, it is expected that athletes' responses will be more accurate than parents' responses when residing apart. Therefore, concurrent validity and test-retest reliability are considered more important indicators of trustworthiness than convergent validity with parents for these residences.

Reductions in validity and reliability scores with each successive residence were greatest for reason for relocation and whom participants lived with at each residence. Reason for relocation to each new place of residence was asked primarily to identify relocations related to sport training and competition. While test-retest reliability for reason for relocation to residence four was only moderate, it is important to highlight that concurrent validity, convergent validity with parents, and test-retest reliability were all very good for both age at first relocation for reasons related to sport and total number of relocations for reasons related to sport. Furthermore, internal consistency between age at first relocation for reasons related to sport as identified within this section of the DHAQ, and age at first relocation for reasons related to sport as identified within the sporting career and milestones section was perfect, with all athletes providing the same response on both occasions. As relocations for reasons related to sport are of primary interest, the lower test-retest reliability for reason for relocation to residence four is not considered to be an issue.

Regarding the moderate validity and reliability scores for whom participants lived with at residence four, qualitative examination of responses revealed that terms such as ‗friends', ‗team-mates', ‗university students', ‗housemates', and ‗other athletes' were used interchangeably, creating ambiguity in responses. At previous residences, responses typically included ‗family', ‗athlete residence', or ‗student residence', resulting in greater consistency between these more clearly defined groups of people. Given the potential for ambiguous, overlapping, and/or poorly defined responses to the question ―who did you live with at this residence", it is recommended that this item be removed from the DHAQ.

Considering the strong validity and reliability scores for the other items within this section, it is recommended that all other items be retained. However, a minor modification is

suggested. It was highlighted previously that residences of short duration were recalled with less consistency than residences of long duration. For the majority of instances during which the period of residence was brief, the relocation was to a new home within the same local area. Additionally, no participants cited sport as a reason for relocation to a new residence within the same local area as the previous residence. Therefore, to reduce issues concerning validity and reliability of information pertaining to residences of short duration, it is recommended the DHAQ be modified to collect information regarding only the different towns/cities in which participants have resided rather than each of the individual homes in which they have lived.

Organised Sport Practice History

On average, athletes reported participating in three to four organised sports in addition to their main sport (DHAQ: *M* = 3.67, *SD* = 1.76; athlete interview: *M* = 4.00, *SD* = 2.56; DHAQ-retest: *M* = 2.64, *SD* = 1.63); parents also reported that athletes participated in approximately four additional sports (*M* = 4.00, *SD* = 1.58). Note that coaches were not required to comment on athletes‘ involvement in other organised sports so convergent validity with coaches was not assessed for this section. PA values, ICC statistics, and validity and reliability classification information for all items with the organised sport practice history section of the DHAQ are provided in Table 6.

When completing the DHAQ, athletes identified all organised sports other than their main sport in which they had participated, as well as specific details of their involvement in each sport, for each year of their participation. These details included the total number of months per year they participated in each sport, the average number of hours per week they engaged in all practice and competition activities, and the highest level of competition they reached. During athlete and parent interviews, only a list of all organised sports and accompanying ages of participation were collected; specific details of athletes‘ involvement in each sport for each year of their participation were not obtained. Therefore, concurrent validity and convergent validity with parents could not be assessed for items relating to hours of participation in organised sports, or highest level of competition reached. As such, analysis of this information was restricted to examination of test-retest reliability.

Validity and reliability for total number of other sports reported, total hours of participation in all other organised sports combined, and highest level of competition for organised sports other than main sport were all classified as very good or good. However, these results were not as positive as they may first appear. Inspection of the sports identified during each test occasion revealed that even though participants consistently reported participating in

the same number of sports, the sports listed differed between test occasions. Only nine of the fifteen athletes provided the same list of sports during the interview as the initial completion of the DHAQ, and only three of the eleven athletes who completed the DHAQ for a second time provided the same list of sports as they did the first time. Most significantly, of the thirteen parents interviewed, none of the lists of sports provided by parents matched that provided by the athlete during the initial completion of the DHAQ. This is a significant finding as number of other organised sports, and total cumulative hours of participation in all other organised sports combined are frequently used as indicators of validity and reliability in studies of sport expertise development (Baker et al., 2003b; Berry et al., 2008; Memmert et al., 2010; Oldenziel et al., 2003; Soberlak & Côté, 2003). Our results suggest that these measures may be slightly misleading if congruence between sports reported is poor or not established.

Despite the inconsistencies in sports reported by participants, sufficient data were available to assess validity and reliability of information provided on a sport-by-sport basis for three organised sports. Recall of participation in ‗other organised sport one' (the sport in which the athlete had invested the most considerable amount of time) was reasonably strong, with responses for age started, age stopped, and highest level of competition receiving very good or good ratings for concurrent validity, convergent validity with parents, and/or test-retest reliability. Test-retest reliability for total hours of participation in ‗other organised sport one', however, was only classified as moderate.

With each successive sport, ranked in order of time invested into the activity, validity and reliability appeared to weaken. For ‗other organised sport two', concurrent validity and test-retest reliability were good for age started and age stopped, but convergent validity with parents for these items was poor. The same trend was true of ‗other organised sport three'. As with ‗other organised sport one', test-retest reliability for total hours of participation was moderate for both ‗other organised sport two' and ‗other organised sport three'; and although test-retest reliability for highest level of competition reached was good for ‗other organised sport two', it was poor for ‗other organised sport three'.

While not provided as a direct response from participants, age at specialisation was also calculated for each test occasion and assessed for consistency. Age at specialisation was considered to be the age at which the athlete no longer participated in any organised sports other than their main sport. Across all three measures of validity and reliability, age at specialisation was classified as very good.

These results highlight two major issues with recall of information relating to participation in organised sports other than athletes' main sport. First, when athletes participate in a variety of sports, it may be difficult to recall specific details of their involvement in each

sport. These difficulties are most apparent for recall of time spent in practice and competition activities, and more generally, for sports in which periods of involvement were brief.

Secondly, convergent validity with parents was poor. As mentioned previously, parents frequently identified different sports to the athletes, and even when the same sport was identified by both athletes and parents, convergent validity for the ages at which the athlete participated in these sports was poor. In this situation it is difficult to discern whether the athlete‘s account or their parent‘s account is more accurate. In defence of the athlete, it is likely easier to recall details of your own extra-curricular involvement than it is to recall the participation of others, particularly in multiple-child families where parents may confuse one child‘s involvement in recreational activities with those of another child. On the other hand, in most cases parents will be responsible for enrolling their child in sports programs, providing the necessary financial assistance, and will most likely provide transportation to and from practice and competitions. The investment required by parents in their children‘s sport participation therefore provides a strong argument for the accuracy of the parental account of children‘s involvement in sports other than their main sport. It is recommended that wherever possible, secondary sources such as official results statements, medals, certificates, news items, and/or photographs be used to assist accurate recall and validation of information relating to participation in organised sports other than athletes‘ main sport.

Considering the emphasis placed on participation in other organised sports in the sport expertise development literature, it is recommended that this section be retained in the DHAQ despite the issues discussed. It is imperative, however, that validity and reliability continue to be monitored closely in future investigations. Given the potential for inaccurate recall of these data, it is also suggested that results concerning participation in other organised sports, obtained both directly from investigations utilising the DHAQ and from other studies of sport expertise development, be interpreted cautiously.

Participation in Informal, Playful Sporting Games

Similar to organised sport practice history, the section of the DHAQ addressing participation in informal, playful sporting games was slightly more detailed than the same section of the athlete and parent interviews. Within the DHAQ, athletes were required to list all the informal, playful sporting games in which they had ever participated on a regular basis, and for each year of participation in each game listed, they provided the approximate number of months per year and average number of hours per week they played each game. Additionally, participants also indicated who they typically played these informal, playful sporting games

with (e.g. friends, siblings, team-mates), the average age of the people with whom they played (three or more years younger/older, 1-2 years younger/older, or around the same age), and where they typically played (e.g. backyard, local streets, local park, local school yard). While the interviews also required participants to list all informal, playful sporting games in which they/the athlete had ever participated on a regular basis, the corresponding ages of participation in each game, and information regarding whom they/the athlete played with and where, details of time involved in informal, playful sporting games were not collected. Moreover, this section was not included within the coach interview guide as it was not expected that coaches would be familiar with this information. PA, ICC, and reliability classification information for this section is provided in Table 7.

In general, validity and reliability of recall for participation in informal, playful sporting games was very weak, with the majority of items receiving moderate or poor classification ratings across almost all validity and reliability conditions. Although test-retest reliability for number of informal, playful sporting games identified was good, concurrent validity and convergent validity with parents were moderate and poor respectively. Further, as was the case with participation in organised sports, the lists of informal, playful sporting games that were provided varied considerably between test occasions. In fact, no athletes were found to have identical lists of informal, playful sporting games across all methods of data collection. Most likely as a result of providing different lists of sporting games on each test occasion, test-retest reliability of total hours accumulated in informal, playful sporting games to date was also classified as poor. In addition, identification of whom athletes played with, the ages of the people with whom athletes played, and where they played were also highly inconsistent.

Recall of participation in informal, playful sporting games was not only poor on a general level, but on a more specific game-by-game level as well. Sufficient data were available to assess validity and reliability of details of participation in three informal, playful sporting games on a game-by-game basis. As shown in Table 7, even when the same playful game was reported on multiple occasions, the ages at which athletes participated in these games and the time invested in these games were reported with relatively low consistency. It is interesting to note that eight athletes reported participating in informal, playful sporting games related to their main sport within this section during the initial completion of the DHAQ, however, of these eight athletes, only one listed main-sport related play on all test occasions. The remaining athletes reported participation in main-sport related play on only one or two test occasions, resulting in insufficient data to accurately assess validity and reliability of information pertaining to participation in informal, playful sporting games related to the athletes‘ main sport in isolation.

These findings are particularly noteworthy, but at the same time, perhaps not surprising. Details of participation in informal, playful sporting games may be difficult to recall because, by nature, participation in these activities is inconsistent. Children typically engage in play spontaneously, and the time involved in play can vary considerably between sessions. The sporting games played from one session to the next can also vary considerably, and even within a single session, a variety of activities may be played. The nature of participation in informal, playful sporting games differs from that of organised sports, as organised sports typically have clearly defined and regularly occurring practice times, competitions, and seasons. The structured nature of organised sports makes it much easier to recall past involvement, whereas the irregular, informal nature of play makes it more difficult to accurately quantify participation.

The results are important however, because a number of investigations of sport expertise development have found participation in informal, playful sporting games to differentiate highly skilled from lesser skilled athletes in a variety of sports (Berry et al., 2008; Ford et al., 2009; Memmert et al., 2010). In the past, validity and reliability of recall for participation in informal, playful sporting games has not been assessed to the same degree as this study, however Soberlak and Côté (2003) also reported considerably larger discrepancies between athlete and parent responses for hours of participation in deliberate play compared to deliberate practice and other organised sports.

Given the previous research in this area, it is recommended that this section of the DHAQ be modified in an attempt to improve recall of participation in informal, playful sporting games. While participants would still be required to provide a list of all informal, playful sporting games in which they have participated on a regular basis, it is recommended that instead of having to provide details of the ages and duration of their involvement in each game reported, they simply provide an overall estimate of the ages and duration of their involvement in all informal, playful sporting games combined. It is hypothesised that this more general approach may improve recall, as participants will only have to remember broad details of their involvement in sporting play, rather than their involvement in specific sporting games. It is, however, recommended that items relating to whom participants engaged in informal, playful sporting games with, and where they played, be removed from the questionnaire. Modifications to this section of the DHAQ should be monitored closely.

General Involvement in Main Sport

Validity and reliability for items within the general involvement in main sport section of the DHAQ were assessed overall, and where applicable, on an age-by-age basis for ages 4-23.

For age-by-age analyses, convergent validity with coaches was only assessed for ages 17-23, as few coaches worked with the athletes before this time. Results are provided in Table 8.

Practice history. Concurrent validity, convergent validity with parents, and test-retest reliability were all very good for age at first participation in main sport. Internal consistency between age at first participation in main sport as reported within this section of the questionnaire and the sporting career and milestones section was also very good (PA = 97.78, ICC = .98, $p < .01$). Time involved in main sport was examined via multiple measures. Participants recalled the total number of months involved in their main sport each year from age four to the present, along with the average number of hours per week for all types of practice and competition activities combined. The average number of hours per week and months per year were multiplied to calculate total hours of involvement in main sport per year of participation, which were subsequently combined to provide the total cumulative hours of participation in main sport across athletes‘ careers to date. The average number of practice sessions per week and the average frequency of competitions were also collected to gain a more detailed understanding of how time involved in main sport was distributed. Validity and reliability of recall for each of these measures will be discussed separately.

Concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability were all very good across the majority of ages for recall of total number of months involved in main sport per year. For average number of hours per week involved in main sport, concurrent validity, convergent validity with coaches, and test-retest reliability were all very good or good for nearly all ages; however, convergent validity with parents was poor for several years. With regards to recall of average number of practice sessions per week, concurrent validity, and test-retest reliability were very good for most ages, convergent validity with coaches was good, and while convergent validity with parents was also good for the majority of ages, once again, several years were classified as poor. Responses provided for hours per week and months per year of involvement in main sport translated to very good or good concurrent validity classifications for hours per year across all ages, as well as very good or good test-retest reliability classifications for all ages except six and seven.

Convergent validity with coaches for hours of participation in main sport per year of involvement was very good or good for the majority of applicable ages, and although convergent validity with parents was mostly good from ages 14 onward, prior to this point it was generally quite poor. In addition, concurrent validity and test-retest reliability were very good for total cumulative hours of involvement in main sport to date, and convergent validity with parents was good. Convergent validity with coaches could not be determined for total

cumulative hours of participation in main sport to date as no coaches were familiar with the athlete throughout their entire career.

The results obtained within this section thus far highlight a number of interesting issues regarding validity and reliability of recall for time involved in main sport. The DHAQ appears to be a very reliable measurement tool for the collection of general practice history information, as test-retest reliability was very good or good for all items, across almost all ages. The only exceptions were moderate to poor ratings for average number of hours per week and, subsequently, total number of hours per year at ages six and seven. Interestingly, ages 6-7 correspond to the average age of first involvement in main sport for participants in this study. It is therefore possible that recall for hours of participation during the first year or two of involvement in a new sport may be challenging until a regular schedule is in place. To support the strong reliability ratings, concurrent validity was also very good or good for all items across all ages. The higher concurrent validity for hours of participation at ages six and seven compared to test-retest reliability might be reflective of a recency effect given the shorter time delay between test occasions. In the event of uncertainty, participants may have remembered their response from the previous test occasion, which could explain the stronger consistency for these years. Consequently, athlete reports of hours of participation in main sport during the first two years of involvement should be interpreted with caution.

As already discussed, parents are frequently used as a source of validation for athlete developmental history information. It is expected that since parents are typically highly involved in children‘s extracurricular activities during the childhood and adolescent years, they could provide an accurate account their child‘s experiences. In this study however, convergent validity with parents was relatively poor. Interestingly, convergent validity with parents for total cumulative hours of participation in main sport to date was classified as good, but convergent validity with parents for average number of sessions per week, average number of hours per week, and total number of hours per year, were poor across a range of ages.

As responses for average number of hours per week are utilised in the calculation of total number of hours per year and, ultimately, total cumulative hours of participation in main sport to date, the good convergent validity with parents for total cumulative hours of participation in main sport to date is slightly misleading. This is important as a number of investigations of sport expertise development have addressed total cumulative hours of participation in validity and reliability analyses without reporting validity and reliability of the constituents of this measure, that is, hours per week, months per year, and hours per year (Berry et al., 2008; Memmert et al., 2010; Soberlak & Côté, 2003). Furthermore, the poor convergent validity with parents for several measures raises the question of which response is more accurate

– parents‘ or athletes‘? Before answering this question, we must first discuss results pertaining to convergent validity with coaches.

Convergent validity with coaches was very good or good for recall of months per year during each age they were involved with the athlete, average number of hours per week, and average number of sessions per week. Convergent validity with coaches for total hours of involvement in main sport per year was also very good or good with just one exception at age 19. Given the high degree of involvement of the coach in an athlete‘s training schedule, it was expected that convergent validity with coaches would be strong for time involved in main sport across all ages. During instances of lower consistency, it may have been the case that coaches were recalling information based upon their own schedule or the schedule of their ‗average‘ athlete, rather than the personalised schedule of the athlete in question. Nonetheless, convergent validity with coaches was generally good, and undoubtedly stronger than convergent validity with parents. Therefore, considering the strong concurrent validity, test-retest reliability, and convergent validity with coaches for items relating to time involved in main sport, it can be assumed that athlete recall of this information is more accurate than parent recall. As such, the use of parents as sources of validation for training history information may be problematic.

Competition involvement. In addition to practice history information, basic details of competition involvement were also collected within this section. Age at first participation in competition was recalled with excellent consistency, with concurrent validity, convergent validity with parents, and test-retest reliability all classified as very good. Similarly, highest level of competition also received very good classifications for concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability, as did the age at which participants reported reaching their highest level of competition. Additionally, age at first participation in competition, highest level of competition, and age when first reached highest level of competition all displayed very good internal consistency with the sporting career and milestones section of the DHAQ (age at first participation in competition: PA = 95.83, ICC = .81, $p < .01$; highest level of competition: PA = 93.33; age when first reached highest level of competition: PA = 99.21, ICC = .99, $p < .01$).

On an age-by-age basis, validity and reliability of recall for highest level of competition varied. Concurrent validity was very good or good for almost all ages, but test-retest reliability received moderate classifications for several years. Convergent validity with coaches was very good or good, but convergent validity with parents ranged from very good to poor. In general, validity and reliability were particularly weak during the teenage years. Interviews revealed that during these years, athletes frequently participated in a variety of school, club, regional, state, and national level competitions. During this busy period of development, it appears that it may

be difficult to recall details of competition involvement with accuracy. This finding is also reflected in results concerning average frequency of competitions. When asked to identify the average frequency of competitions for their main sport (e.g. one per week, one per month, four per year), participants' responses were highly inconsistent. Prior to age 11, recall was reasonable, however after this point, most validity and reliability ratings were moderate or poor.

Accompanying results regarding highest level of competition indicate that ages 11-12 mark the typical progression from club competition to participation in regional, state, and/or national level competition. This progression results in a more complex competition schedule compared to that associated with participation at the club level alone. Not only does the volume of competition increase with progression to higher levels of representation, but each level of competition tends to follow a different schedule, rendering it difficult to provide an ‗average' frequency of competitions throughout the year. Therefore, it is likely that validity and reliability ratings were poor for this item because the structure of the question did not adequately cater to the complexity of the issue. Discussion regarding recall of information pertaining to competition involvement will continue during presentation of results for the detailed history of involvement in main sport and detailed competition history sections of the questionnaire.

Significant achievements. In an attempt to explore skill level and success on a more refined scale than simply highest level of competition, participants were asked to indicate their most significant achievement relating to their main sport, for each year of their involvement. It was intended that this information could be used to investigate associations between success and patterns of participation in practice and competition activities, as well as the relative importance of early success for the development of sport expertise. Unfortunately, concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability were poor for almost all years beyond age 11. Qualitative observation of responses and participant feedback suggested that at times it may have been difficult to recall relatively minor successes; while at other times, it may have been difficult to single out the most significant achievement among a number of notable accomplishments.

Within the sporting career and milestones section of the DHAQ, participants were required to indicate their most significant achievement across their career to date. As discussed previously, validity and reliability of recall for the athletes' single most significant achievement were good. Therefore, major achievements such as making a national team or winning an Olympic gold medal appear to be recalled accurately, but comparatively minor achievements such as winning a medal at state championships or being voted most valuable player on a team appear to be less memorable and hence recalled with less consistency. Côté and colleagues (2005) suggested that subjective information is recalled with less reliability than objective

information, and our results support this. Although an indication of success throughout the various stages of an athlete‘s career would be very informative, retrospective recall of this information is not consistent enough to warrant inclusion of this item in future investigations. It is therefore recommended that this item be removed from the DHAQ.

Reasons for participation. In a similar manner, validity and reliability of recall for the main reason athletes‘ participated in their main sport was also very poor. A change in motivation for participation has been suggested as a marker of the transition from one stage of development to another (Bloom, 1985), so this item was included to allow for the investigation of the relationships between motivations for participation, time involved in practice and competition, relative success, and stage of development. However, recall of reasons for participation were very inconsistent. Concurrent validity, convergent validity with parents, and test-retest reliability were all very poor for the majority of ages, particularly past 11 years old. This item was not included within the coach interview guide, so convergent validity with coaches could not be assessed. Once again, 11 years old is identified as an age at which consistency of recall exhibits a noticeable change, suggesting that this age marks a significant milestone within an athlete‘s career. In this case it is likely that the progression to higher levels of competition at this age leads to a shift in reasons for participation from fun and enjoyment to factors such as improvement, to win, and/or to be selected for a team.

Although this information is particularly interesting, the specific responses provided for this item are not recalled with sufficient consistency to justify retention in the questionnaire. As for recall of most significant achievement, it is possible that the subjective nature of this item may be contributing to the inconsistent recall of main reason for participation. It may also be that reasons for participation are multifaceted and difficult to express in a single word within a questionnaire. It is therefore recommended that this item also be removed from the DHAQ, and this line of questioning be restricted to investigations specifically targeting motivations for participation in sport and physical activity.

Detailed History of Involvement in Main Sport

The detailed history of involvement in main sport section of the DHAQ aimed to gain a greater understanding of the different types of practice in which athletes participate, and the conditions under which they engage in practice.

Conditions of practice. First, information regarding the various conditions of practice in which participants engaged was collected. For each year of involvement in their main sport, athletes were required to identify the number of months per year, sessions per week, and hours

per week they engaged in each of five practice conditions: 1) supervised group practice; 2) supervised individual practice; 3) unsupervised individual practice; 4) unsupervised sport-specific play; and 5) passive sport-specific activities (e.g. watching matches live or on television, reading sport-specific news and information, or having sport-specific conversations with others). Details of time involved in competition were also collected. During each method of data collection, the opportunity to include additional practice conditions was available; however, no further conditions were identified by any participants during any test occasion.

Responses provided for months per year and hours per week were multiplied to indicate the approximate number of hours per year of involvement in each practice condition, and hours per year were subsequently aggregated to provide a cumulative total for hours of participation in each practice condition across athletes‘ careers to date. Hours per year and total cumulative hours of participation in each practice condition were also summed to provide values for participation in main sport overall. Due to the large number of practice conditions addressed within this section of the questionnaire, validity and reliability results are only presented for hours per year and total cumulative hours. Validity and reliability results for age at first participation in each practice condition are also presented. Results related to participation in main sport according to practice condition are provided in Table 9.

Supervised group practice. Age at first participation in supervised group practice activities was recalled very consistently, with concurrent validity, convergent validity with parents, and test-retest reliability all classified as very good. Concurrent validity for hours of participation per year of involvement in this type of practice was very good or good across all ages, and both convergent validity with coaches and test-retest reliability were very or good across most ages, with just a few exceptions (note that, as for the general involvement in main sport section, convergent validity with coaches was only assessed for ages 17-23). Conversely, convergent validity with parents for hours of participation in supervised group practice activities was poor or moderate for the majority of ages, particularly prior to age 18.

Despite poor ratings for convergent validity with parents for hours of participation in supervised group practice activities per year of involvement, convergent validity with parents for total cumulative hours of participation in this practice type was good, thus further highlighting that validity and reliability assessments of total cumulative hours of participation in sport may be misleading. Concurrent validity and test-retest reliability for total cumulative hours of participation in supervised group practice activities were both classified as very good.

Supervised individual practice. Age at first participation in supervised individual practice activities was also recalled very consistently, reflected by very good classifications for concurrent validity, convergent validity with parents, and test-retest reliability for this item.

Interestingly, although concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability were all classified as very good or good across all ages, concurrent validity and convergent validity with parents for total cumulative hours of participation in supervised individual practice activities were both poor. Test-retest reliability for total cumulative hours of participation in supervised individual practice activities was classified as good, but yet again discrepancies between validity and reliability classifications for hours per year and total cumulative hours are observed.

Unsupervised individual practice. Recall of details pertaining to participation in unsupervised individual practice activities was not as strong as recall for supervised practice activities. Age at first participation in unsupervised practice activities was rated as good for test-retest reliability, but poor for concurrent validity and convergent validity with parents. Similarly, test-retest reliability for hours per year for this practice type was very good or good for almost all ages, while concurrent validity and convergent validity with parents was generally poor from approximately age 13 onward. Convergent validity with coaches was mostly good between ages 17-23. Overall, classifications for total cumulative hours of participation in unsupervised individual practice activities were poor for concurrent validity, convergent validity with parents, and test-retest reliability.

Unsupervised sport-specific play. Likewise, details of participation in unsupervised sport-specific play were also recalled with low consistency. While concurrent validity and test-retest reliability were very good and convergent validity with parents was good for age at first participation in this practice type, concurrent validity, convergent validity with parents, convergent validity with coaches and test-retest reliability for hours of participation per year of involvement were generally poor between ages 10-18. Although validity and reliability ratings were very good outside of these ages, participants did not typically engage in sport-specific play before age 10 or after age 18. As such, recall of time engaged in sport-specific play during the years in which athletes were most highly involved in this activity was poor. In support of this finding, concurrent validity, convergent validity with parents, and test-retest reliability for total cumulative hours of participation in unsupervised sport-specific play were all classified as poor.

Passive sport-specific activities. Mixed results were observed for participation in passive sport-specific activities such as watching main sport live or on television and reading main-sport related news and information. For age at first participation in this practice condition, test-retest reliability was very good, concurrent validity was good, but convergent validity with parents was poor. Concurrent validity and test-retest reliability for hours of involvement in passive sport-specific activities per year were very good or good across all ages, and convergent validity with parents was very good or good prior to age 19; however, convergent validity with

parents was poor beyond this age. Convergent validity with coaches for hours of involvement in passive sport-specific activities per year was very good or good across the majority of applicable ages. Overall, concurrent validity for total cumulative hours of participation in passive sport-specific practice activities was good, but convergent validity with parents and test-retest reliability were poor.

Competition. With regards to competition activities, athlete, parent, and coach participants all expressed difficulties quantifying competition involvement. As discussed previously, athletes commonly participate in club, regional, state, national, and/or international level competitions simultaneously, and each competitive level often involves a different competition schedule. For example, club level competitions frequently occur once per week, while state and national level competitions may only occur once or twice per year. Additionally, some competitions involve a single match or event, lasting just a few hours, whereas other competitions involve multi-day tournaments. Furthermore, competitions typically include time spent actively involved in competition, active warm-up and cool-down activities, and down-time between matches/events. From conversations with participants during interview procedures, it became apparent that the DHAQ did not adequately cater to the complexities of competition scheduling. Participants expressed uncertainty in how to provide an average frequency for their varying competition schedule, and were unsure whether they were expected to include time in which they were at the competition venue but not necessarily actively engaged in competition, within their estimates of time involved. As a result, the manner in which participants responded to items regarding time involved in competition activities was highly inconsistent.

Consequently, validity and reliability of recall for hours per year and total cumulative hours of participation in competition were not assessed. Concurrent validity, convergent validity with parents, and test-retest reliability were, however, all very good for age at first participation in competition. Internal consistency for age at first participation in competition was also very good between this section of the DHAQ and the sporting career and milestones (PA = 91.91, ICC = .90, $p < .01$) and general involvement in main sport (PA = 89.23, ICC = .72, $p < .01$) sections of the questionnaire.

All practice conditions. Note that as a result of the complications encountered regarding quantification of competition involvement, details of participation in competition were not included within the following assessments of validity and reliability for all practice conditions combined. However, combining details of participation in supervised group practice, supervised individual practice, unsupervised individual practice, unsupervised sport-specific play, and passive sport-specific activities together, concurrent validity, convergent validity with parents,

and test-retest reliability for age at first participation in main sport overall were all very good. Internal consistency for age at first participation in main sport with the sporting career and milestones (PA = 98.67, ICC = 1.00, $p < .01$) and general involvement in main sport sections were also very good (PA = 96.44, ICC = .98, $p < .01$).

Validity and reliability of recall for hours of participation in all practice conditions combined per year of involvement was rated as very good or good across the majority of ages for concurrent validity, convergent validity with coaches, and test-retest reliability, but convergent validity with parents was poor or moderate for most years. Additionally, concurrent validity, convergent validity with parents, and test-retest reliability for total cumulative hours of participation in main sport under all practice conditions combined were all classified as good.

Interestingly, internal consistency between responses in this section and those within the general involvement in main sport section of the DHAQ was very good or good for total hours of participation in main sport per year across all ages (see Table 10), as well as for total cumulative hours of involvement in main sport overall (PA = 88.89, ICC = .97, $p < .01$). It is surprising that internal consistency for these items was so strong because within the general involvement in main sport section of the questionnaire, participants were instructed to consider all practice and competition activities combined when providing estimates of time involved in main sport. As previously discussed, within the current practice condition analysis, time involved in competition activities was not included in calculations of overall time involved in main sport, suggesting that participants may not have accounted for time engaged in competition during their earlier responses within the general involvement in main sport section.

It is important to highlight that a number of items within this section of the DHAQ were classified as having very good or good validity and/or reliability on the basis of a strong percent agreement score alone. As indicated in Table 9, many items displayed very good or good agreement, yet the corresponding ICC was poor. As PA is weighted more heavily than ICC in the overall classification system, these items were considered to have very good or good validity and/or reliability, despite having low, zero, negative, or non-significant ICCs. The combination of a high PA and a low ICC suggests that responses provided on both test occasions of interest were reasonably similar in magnitude; however, some athletes provided responses during the initial completion of the DHAQ that were slightly lower than those provided during the comparative data collection method (i.e. athlete interview, parent interview, coach interview, or retest occasion), whereas other athletes provided responses during the initial completion of the DHAQ that were slightly higher. The poor ICC indicates that there was no systematic occurrence of under- or over-reporting during one method of data collection compared to the other, but this does not matter given the similarity in magnitude of responses. This scenario is

an example of why PA is weighted more heavily than ICC in the overall classification system, because the correlation matters less when responses are in close agreement.

Results obtained regarding conditions under which athletes practiced highlight some important considerations for the collection of practice history information. First, details of participation in supervised group practice activities were recalled most consistently, followed by supervised individual practice activities; recall of participation in unsupervised practice activities was noticeably less consistent, particularly for unsupervised individual practice activities. Supervised practice activities tend to be scheduled at regular times, and are typically of fixed duration. In comparison, unsupervised practice activities often occur spontaneously and for varying amounts of time. It is therefore not surprising that recall of details relating to unsupervised practice activities was somewhat inconsistent, as participation in these activities can itself be inconsistent, thus more difficult to accurately recall.

Despite the issue of poor recall, previous research has identified that individual unsupervised practice activities are highly relevant for performance improvement (Baker et al., 2003a; Helsen et al., 1998; Hodge & Deakin, 1998; Hodges & Starkes, 1996). Therefore, rather than remove these items from the DHAQ, it is recommended they be retained but monitored closely for validity and reliability. If validity and reliability of recall continue to be poor, examinations of participation in individual unsupervised practice activities may need to be reserved for longitudinal, prospective studies in which athletes complete training diaries to record their involvement as it occurs. Moreover, it is advised that results relating to hours of participation in unsupervised practice activities obtained both directly from investigations utilising the DHAQ and other external investigations involving retrospective recall techniques, be interpreted with care.

Second, concerning the issues surrounding quantification of time involved in competition activities, it is recommended that details of participation in competition be removed from this section and incorporated into the detailed competition history section. This would allow for a more in-depth investigation of time involved in competition than the current broad measures of hours per week and months per year. It is suggested that the detailed competition history section be modified to address involvement in various formats of competition (for example regular fixture format competition, tournament format competition, and occasional competitions), and to differentiate between time spent actively involved in competition, and time spent at the competition venue but not physically engaged in a match/event.

Regarding assessment of validity and reliability of recall for practice history information in future investigations, results obtained within this section further support our earlier proposals that a) parents may not be the most appropriate source for validation of these

items so wherever possible coaches and/or training diaries should be consulted to corroborate athlete responses regarding involvement in various conditions of practice and b) the use of total cumulative hours of participation in various practice activities for validity and reliability assessments without concurrent investigation of at least hours of participation per year of involvement may be misleading. Furthermore, despite being prompted to include time involved in competition within overall estimates of time involved in main sport, it appears that athletes may not do so. Therefore, when collecting athlete developmental history information it is important to explicitly address practice and competition involvement separately.

Similar conclusions can also be drawn from results obtained within the second half of the detailed history of involvement in main sport section of the questionnaire, where validity and reliability of recall for participation in a range of practice types was assessed.

Types of practice. Analogous to the first half of this section concerning the conditions under which athletes practiced, the number of hours per week and months per year participants engaged in each of the following types of practice were collected: 1) technique/skills training; 2) tactical/games-based training; 3) physical conditioning/weights; 4) mental/psychological skills training; 5) recovery techniques; 6) video analysis/review; 7) study related to main sport; and 8) watching main sport live or on television. These values were multiplied to give hours of participation in each practice type and in all practice types combined per year, which were subsequently added to provide total cumulative hours of involvement in each practice type and all practice types combined, across athletes‘ careers to date. As several of these activities can be combined within a single practice session, information regarding the number of sessions per week participants engaged in each practice activity was not collected.

During all data collection procedures, space was provided for participants to include any other types of practice in which they engaged on a regular basis. This resulted in the identification of one additional practice activity: flexibility training, yoga and/or Pilates; however, as flexibility training, yoga and/or Pilates was not included within the initial version of the DHAQ, validity and reliability for recall of participation in this type of practice could not be assessed. Since flexibility training, yoga and/or Pilates was highlighted as a regular component of practice for several participants, it is recommended that this activity be included within future investigations involving the collection of detailed practice history information.

Once again, due to the large number of practice types included within this section, validity and reliability results are only presented for total hours of participation in each practice type per year, and as a total cumulative value (see Table 11). Results pertaining to age at first participation in each practice type are also provided.

Sport-specific physical practice. The two most common types of practice for team and individual sports alike were technique/skills training and tactical/games based training. Although it was intended to consider these practice types independently, feedback from participants indicated that it was rare to work on these components in isolation. Participants expressed difficulty separating their involvement in the practice types and did not feel they could provide accurate representations of time involved in each of the categories independently. Therefore, in response to the needs of participants, technique/skills training and tactical/games based training were combined into a single category titled sport-specific physical practice.

The age at which participants first engaged in sport-specific physical practice was recalled consistently, with concurrent validity and test-retest reliability classified as very good and convergent validity with parents classified as good. Somewhat surprisingly given that sport-specific physical practice is the main form of training for the majority of athletes, validity and reliability of recall for hours of participation in this practice type was generally poor. On an age-by-age basis, test-retest reliability was very good or good prior to age 14 when hours of involvement were relatively low, but beyond this age, test-retest reliability was poor. Concurrent validity, convergent validity with parents, and convergent validity with coaches were all classified as poor for the majority of years, with just several exceptions at young ages, again when participation was limited. Consequently, total cumulative hours of involvement in sport-specific physical practice was also classified as poor for concurrent validity, convergent validity with parents, and test-retest reliability.

Examination of values for hours of participation in sport-specific physical practice revealed that, in general, responses provided during the athlete interview, parent interview, coach interview, and DHAQ-retest were relatively similar, but were much lower during the initial completion of the DHAQ. As a number of participants completed the questionnaire before the decision was made to collapse technical/skills training and tactical/games-based training into a single category, values for participation in sport-specific physical practice for these athletes were calculated via summation of responses for the two original practice types. It is possible that the difficulties associated with separating involvement into technical/skills training and tactical/games-based training for these athletes led to inaccurate accounts of practice history. Since all validity and reliability comparisons were conducted against responses provided during the initial completion of the DHAQ, this could explain the seemingly poor recall for hours of involvement in sport-specific physical practice. It is highly likely that improved item structure, combining technical/skills training and tactical/games-based training into a single sport-specific physical practice category, would lead to more consistent recall of details of participation in this practice type in future investigations.

Physical conditioning/weights. Recall of involvement in physical conditioning/weights was slightly better than sport-specific physical practice. Athletes did not typically report participation in this type of practice before age 12, so validity and reliability of recall for hours of participation before this age was very good. Relatedly, identification of age at first participation in physical conditioning/weights was also classified as very good for concurrent validity, convergent validity with parents, and test-retest reliability.

Beyond age 12, test-retest reliability for hours of participation in physical conditioning/weights was mostly good or very good, but concurrent validity was poor between the ages of 13-18. Descriptions of practice schedules provided during athlete interviews highlighted that during this time, physical conditioning activities were usually incorporated within sport-specific physical practice sessions, whereas at older ages, physical conditioning activities were conducted as separate stand-alone sessions. When incorporated into sport-specific physical practice sessions, scheduling of physical conditioning/weights was not always consistent or for a fixed duration, explaining the poor concurrent validity for these ages.

Convergent validity with parents for hours of involvement in physical conditioning/weights was classified as poor from age 15 onward. This corresponds to the time at which athletes‘ practice schedules tend to become more complex, and also when parents are less likely to attend their child‘s practice sessions. Resultantly, it is around this time that parents will become less aware of the activities in which athletes are engaging during training. Furthermore, as already mentioned, the majority of participants had moved out of the family home prior to their involvement in this study, so when athletes are no longer living with their parents, it becomes even less reasonable to assume that parents will be accurate sources for validation of information relating to practice scheduling and composition. When evaluating hours of participation in various practice activities, particularly during the late adolescent and adult years, coaches are considered to provide a more accurate record of involvement, and hence convergent validity with coaches becomes an important indicator for accuracy of athlete recall. In regards to recall for details of participation in physical conditioning/weights, convergent validity with coaches was very good or good across almost all applicable ages.

Overall, recall for total cumulative hours of participation in physical conditioning/weights across participants‘ careers to date was good for concurrent validity and test-retest reliability, but poor for convergent validity with parents.

Mental/psychological skills training. As national and international level athletes, it was expected that participants would engage in regular sessions with a sport psychologist to build mental skills and psychological strategies for performance. While a number of participants reported engaging in mental/psychological skills training during the initial and/or retest

completions of the DHAQ, no athletes reported any involvement in this form of practice during their interviews. It is unclear why this was the case considering all athletes were prompted to discuss their involvement in each practice type during all test occasions. Moreover, no parents or coaches reported their athletes to have participated in mental/psychological skills training either. This finding was very surprising given the high skill level of participants involved in this phase of the study. The fact that no parents or coaches reported athlete participation in this type of practice may indicate that mental/psychological skills training was not a scheduled, compulsory activity for the athletes involved in this study, and that athletes may have chosen to engage in this type of practice voluntarily, in their own time, without the knowledge of their coach or parents. This does not explain, however, why eight participants reported engaging in mental/psychological skills training during the initial and/or retest completions of the DHAQ but not the athlete interview.

Although validity and reliability statistics and classification information for items relating to involvement in mental/psychological skills training are provided in Table 11, these results are slightly misleading because all athletes including both those who did and did not report any hours of involvement in this type of practice are incorporated into the analyses. When athletes did not report participation in a particular practice type at a given age, responses for hours per week, months per year, and consequently hours per year, were entered as 0. Therefore, when responses were 0 for both test occasions involved in validity and reliability analyses, agreement was calculated to be 100%. For ages during which the majority of participants displayed 100% agreement between test occasions, the mean PA was high, so validity/reliability was classified as very good or good, even when several athletes displayed 0% agreement because participation was reported during completion of the DHAQ but not the interview. For this reason, concurrent validity and convergent validity with parents were rated as very good or good from ages 15-20 despite the fact that no athletes or parents reported involvement in mental/psychological skills training during their interviews. Had only participants who reported involvement in mental/psychological skills training during the initial completion of the DHAQ been included in the analyses, concurrent validity, convergent validity with parents, and convergent validity with coaches would have been poor for all ages beyond 15.

It is noteworthy that test-retest reliability was classified as very good for age at first participation in mental/psychological skills training, and was good for hours of involvement in mental/psychological skills training per year beyond age 15, once participation generally commenced. However, test-retest reliability for total cumulative hours of participation in mental/psychological skills training throughout athletes‘ careers to date was poor suggesting that the over-representation of athletes who did not participate in this type of practice may have

inflated validity and reliability results for hours per year during the test-retest comparison also. It could therefore be concluded that when athletes engaged in mental/psychological skills training, recall of details of their involvement was generally quite poor.

Recovery techniques. A similar situation to the validity and reliability of reported participation in psychological training/mental skills was apparent for participation in recovery techniques. In this case, almost all athletes reported engaging in recovery techniques during the initial and/or retest completions of the DHAQ, but no athletes, parents, or coaches reported involvement during the interviews. Since the majority of athletes reported their participation in this type of practice during the initial completion of the DHAQ, the validity and reliability results presented in Table 11 are not affected by the same issues described above relating to over-representation of athletes reporting no involvement in this activity. Therefore, validity and reliability results discussed herein can be considered true reflections of the data obtained.

Test-retest reliability for age at first participation in recovery techniques was very good. Athletes reported participating in this activity from as early as age 13, although most did not begin their involvement until approximately 17-18 years of age. As such, prior to this point, validity and reliability of recall for hours of participation in this type of practice per year of involvement in main sport were very good or good. Beyond age 17, concurrent validity, convergent validity with parents, and convergent validity with coaches were poor as no participation in recovery techniques was reported during the interviews. Test-retest reliability was moderate to poor for ages 18-20, but was very good or good at all other ages.

Despite strong consistency between responses provided for hours of participation in recovery techniques at most ages, the discrepancies at ages 18-20 were sufficient to result in poor test-retest reliability for total cumulative hours of participation in this activity across athletes‘ careers to date. As was the case for participation in mental/psychological skills training, it appears that when athletes engaged in recovery techniques, recall of specific details regarding their involvement was relatively poor.

Video analysis/review. Participation in video analysis/review activities was reported for most athletes during each data collection occasion. Age at first participation in video analysis/review activities was recalled consistently, with concurrent validity and test-retest reliability for this measure classified as very good and concurrent validity with parents classified as good. The typical age at first participation in this type of practice was approximately 15, so concurrent validity, convergent validity with parents, and test-retest reliability for recall of hours of participation in video analysis/review activities per year of involvement in main sport up to this age was very good. Test-retest reliability remained very good or good until age 18, at which point classifications lowered to moderate or poor; and

concurrent validity was poor from age 17 onward. Convergent validity with parents was poor from age 15, and convergent validity with coaches was poor for all relevant ages (i.e. 17-23). Although involvement in video analysis/review activities was reported across all data collection occasions, hours of involvement were reported to be considerably higher during both completions of the DHAQ than each of the interview procedures. This led to poor ratings of concurrent validity and convergent validity with parents for total cumulative hours of participation in video analysis/review activities across athletes' careers to date, and moderate test-retest reliability.

Again, it is possible that athlete responses for hours of participation in this type of practice were higher than parent and coach responses because they could be engaging in these activities unsupervised, in their own time, without the knowledge of their parents or coach. Yet, the discrepancies between athlete responses within the DHAQ and athlete responses during the interview are still unexplained. It may be that during interview procedures athletes experienced a form of fatigue whereby they felt they had already shared enough information about their practice schedule and were anxious to move on to the next section. The interview procedures employed in this study were exhaustive, requiring participants to provide significant amounts of detailed information about their involvement in their main sport. Although the interview guide followed the structure of the DHAQ closely, and the time to complete both procedures was quite similar, the discussional nature of the interview may have been more taxing on the individuals, requiring greater sustained attentional focus. It could be the case that in an attempt to progress through the interview faster, participants were not as thorough in their descriptions of their practice activities as they were when completing the DHAQ.

Alternatively, it may simply be that involvement in practice activities such as video analysis/review, recovery techniques, and mental/psychological skills training are difficult to recall with accuracy because they are engaged in to a lesser degree than activities such as sport-specific physical practice and physical conditioning/weights. The lower consistency of recall may merely reflect the fact that these activities are less prominent features of athletes' involvement in their main sport.

Study related to main sport. Along the same line of argument, very few participants reported engaging in study related to their main sport. Activities provided as examples for this type of practice included reading news or information about their sport or talking about their sport with others. Just four athletes reported participating in this type of practice, but it is noteworthy that these athletes each indicated participation on multiple test occasions. However, similar to other types of practice that appear to be less integral components of athletes'

involvement in their main sport, no parents or coaches reported their athletes to have engaged in sport related study.

For athletes who reported involvement in study related to their main sport, concurrent validity and test-retest reliability for age at first participation were both good. Given the high proportion of athletes who did not report involvement in this activity, concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability for hours of participation in sport related study per year of involvement in main sport were classified as very good or good for most ages. Concurrent validity and test-retest reliability were also rated as good for total cumulative hours of participation across athletes‘ careers to date, although convergent validity with parents was poor as a result of the fact that no parents reported their child participating in this activity. Once again, closer inspection of responses provided by those athletes who did engage in this activity suggests that specific details of participation were not recalled with great consistency, similar to other minor supplementary practice activities.

Watching main sport live or on television. The final type of practice addressed within this section of the questionnaire was watching main sport live or on television. Most participants reported spending time watching their main sport during the initial completion of the DHAQ, approximately half of participants reported watching their main sport during the retest completion of the DHAQ, but only two participants reported spending time watching their main sport within the athlete interviews. Additionally, no parents or coaches reported athlete involvement in this activity during their interviews. Due to the varying number of athletes reporting participation in this type of practice, validity and reliability results were mixed.

Concurrent validity for age when first started watching main sport live or on television was very good, but test-retest reliability was poor. Furthermore, concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability for hours spent watching main sport per year all ranged from very good to poor, and concurrent validity, convergent validity with parents, and test-retest reliability for total cumulative hours spent watching main sport were all poor. Watching main sport live or on television appears to be yet another practice activity that was either ignored during participation in interview procedures, and/or was too difficult to recall with accuracy because it did not form a major component of athletes‘ practice schedules.

All practice types. Given the relatively poor recall for hours of involvement in each of the various types of practice across most ages, it is not surprising that total hours of participation in all practice types combined per year of involvement in main sport received mostly poor ratings for concurrent validity, convergent validity with parents, and convergent validity with

coaches. Test-retest reliability was slightly better with a number of years receiving classifications of good, although ages 16-21 were still rated as having poor consistency.

Interestingly, despite large inconsistencies between values reported for total cumulative hours of participation in each practice type during the initial completion of the DHAQ, the athlete interview, and the retest completion of the DHAQ, the overall total for cumulative hours of participation in all practice types combined across athletes‘ careers to date was classified as having good concurrent validity and test-retest reliability. Convergent validity with parents for total cumulative hours of participation in all practice types combined across athletes‘ careers to date was classified as poor, but this is not surprising given parents reported nil participation for many of the activities. The finding regarding seemingly consistent recall between athlete test occasions for total cumulative hours of participation in all practice types combined supports earlier propositions that assessments of validity and reliability of recall involving only more global measures of participation such as total cumulative hours, without accompanying consideration of validity and reliability of recall for more specific details, may be misleading, and more rigorous analyses should be incorporated into future investigations to be confident of the trustworthiness of the data.

Also of interest are results relating to internal consistency for total hours of participation in main sport. Values calculated for total hours of participation in all types of practice combined per year of involvement in main sport were compared to similar values obtained within the general involvement in main sport section of the questionnaire, as well as the first half of the detailed history of involvement in main sport section concerning the conditions under which athletes practiced. Even though hours of participation in main sport were collected using different approaches within each of these sections, it was expected that the total number of hours per year, and consequently the total cumulative hours of involvement across athletes‘ careers to date, would be similar. So far, comparison of total hours per year and total cumulative hours of involvement between the general involvement in main sport and conditions of practice sections have revealed very good internal consistency ratings. However, internal consistency for hours of participation in main sport between these earlier sections of the DHAQ and the type of practice section was not nearly as good (see Table 10).

On an age-by-age basis, internal consistency for hours of participation in main sport between the type of practice section and the general involvement in main sport section was primarily very good or good prior to age 11 and after age 19. Between these years however, internal consistency was moderate to poor. Furthermore, internal consistency for hours of participation in main sport between the type of practice section and the conditions of practice section was noticeably worse, with almost all ages classified as poor. Important to note is that

internal consistency for total cumulative hours of participation in main sport across athletes' careers to date between the type of practice section and both earlier sections was classified as good (general involvement in main sport: PA = 68.87, ICC = .77, $p < .01$; conditions of practice: PA = 68.24, ICC = .70, $p < .01$), again supporting the argument that total cumulative hours of participation should not be utilised within validity and reliability investigations to assess accuracy of recall for athlete developmental histories, because validity and reliability statistics for this value are misleading.

Closer inspection of the data for total hours of participation in main sport revealed that values reported within the type of practice section were considerably lower than those reported in earlier sections of the questionnaire across all ages. As already discussed, this is likely to be the result of under-reporting of hours involved in sport-specific physical practice due to difficulties associated with differentiating involvement in technical/skills training and tactical/games-based training. Therefore, the low internal consistency for hours of participation in main sport per year of involvement is probably related more to sub-optimal item structure than poor athlete recall per se.

On a more positive note however, age at first participation in any type of practice as reported within this section of the questionnaire received classifications of very good for concurrent validity and test-retest reliability, and convergent validity with parents was good. Age at first participation in any type of practice for main sport was also rated as having very good internal consistency with similar items within the sporting career and milestones (PA = 82.90, ICC = .32, $p = .08$), general involvement in main sport (PA = 85.12, ICC = .30, $p = .10$), and conditions of practice (PA = 81.56, ICC = .33, $p = .07$) sections.

As was the case for the conditions of practice section of the DHAQ, a number of items relating to practice type displayed very good or good PA values, thus were classified as having very good or good validity and/or reliability overall, despite the corresponding ICC being rated as poor. As previously explained, this pattern of results reflects that there was no systematic over- or under-reporting of responses provided during the initial completion of the DHAQ compared to the comparative data collection method, but the responses provided on both test occasions were relatively similar in magnitude (i.e. some athletes reported slightly higher values during the initial completion of the DHAQ compared to the alternative data collection method, while other athletes reported slightly lower values during the initial completion of the DHAQ). When responses are similar in magnitude, the systematic relationship between test occasions becomes less important; hence, it is fair to classify validity and reliability as very good or good on the basis of a very good or good PA value alone, even when the corresponding ICC is low, zero, negative, or non-significant.

Overall, the comparatively poor results for validity and reliability of recall for information pertaining to the types of practice in which athletes have participated across their careers to date were unexpected. While poor convergent validity with parents for this section is understandable given parents are unlikely to attend all practice sessions and know exactly what activities athletes are engaging in during training, particularly during the late adolescent and adult years, the poor consistency between athlete test occasions, and poor convergent validity with coaches were surprising. There are two key factors that appear to be driving these results.

First, many of the practice types addressed within this section seem to form very minor components of athletes‘ training schedules. Sport-specific physical practice, physical conditioning/weights, and video analysis/review activities were the only practice types for which athletes were reported to have participated during all test occasions. Participation in mental/psychological skills training, recovery techniques, study related to main sport, and watching main sport live or on television were only reported by athletes, not coaches or parents, and hours of participation in these activities were typically very low. Since parents and coaches were unaware of athletes‘ involvement in these activities, it is assumed that participation was both voluntary and unsupervised. Details of involvement in activities of this nature, particularly those undertaken for relatively short durations are likely more difficult to recall with accuracy compared to scheduled, supervised, regularly occurring practice activities, which could explain the low validity and reliability of recall for these practice types.

The second factor that may be responsible for low consistency of recall between athlete test occasions is poor experimental design. Two features of the experimental design that may have contributed to low validity and reliability of recall for hours of participation in various practice types have already been discussed: 1) difficulties associated with differentiating technical/skills training and tactical/games-based training, leading to under-reporting of time involved in sport-specific physical practice; and 2) respondent fatigue during interviews, leading to incomplete accounts of practice history during this test occasion, and subsequently, poor concurrent validity. Furthermore, the main sport practice and competition history section of the DHAQ required athletes to provide details of their involvement in their main sport multiple times. First, athletes provided general details of their involvement in practice and competition activities, next they provided details of the conditions under which they practiced, and then they identified the various types of practice in which they participated. The questionnaire was designed in this manner to obtain a more detailed understanding of athletes‘ practice profiles, and also to allow for assessments of internal consistency so reliability of recall for this information could be explored in great depth. While these considerations were theoretically sound, in a practical sense the questions appeared very repetitive to participants, and they

expressed frustration with having to provide so much detail about their involvement in their main sport over and over again. This frustration could have led to inaccurate reporting of details within this third sub-section concerning details of practice involvement, as participants may have rushed through their responses to move on to a new line of questioning. Contrary to intentions, rather than facilitating examinations of internal consistency, the repetitive structure of the questionnaire may actually have negatively influenced validity and reliability results.

Together, issues relating to accuracy of recall for minor practice activities and poor experimental design may explain the substandard level of validity and reliability for recall of details pertaining to participation in various types of practice throughout athletes‘ careers. An alternative explanation for the lower consistency of recall for participation in specific practice types compared to general involvement in main sport overall, could be related to the complexity of high performance athletes‘ training schedules. Athletes participating at the national and international level are likely to follow a carefully periodised training plan taking into consideration the appropriate balance of practice activities to ensure attainment of peak performance at critical stages of the competition season. As such, while the overall number of training sessions and hours of practice may remain relatively constant throughout the year, the specific practice activities in which athletes engage during this time may vary considerably between general preparation, specific preparation, competition, and transition phases. Therefore, it may be more difficult to provide an indication of the ‗average‘ hours of participation in various activities during each year of involvement, because the composition of practice is so highly variable. Consequently, a more general indication of overall hours of practice per week may be more consistent throughout the year, and hence a more accurate indication of involvement in main sport.

Despite the issues concerning accuracy of recall for participation in specific types of practice, it is important the study of sport expertise development progresses beyond global measures of sport involvement to look more closely at how expert athletes spend their time during practice sessions, and how this may differ from lesser skilled athletes. Therefore, it is recommended the DHAQ continues to address participation in specific practice types and practice conditions, but a major restructure for the main sport practice and competition history section is required. As already discussed, it is recommended that items within the general involvement in main sport section relating to significant achievements and main emphasis of participation be removed, and that all items within the same sub-section that concern competition be relocated to the detailed competition history sub-section. Additionally, it is recommended that items relating to practice history within the general involvement in main

sport, conditions of practice, and types of practice sub-sections be merged into a single line of questioning to reduce repetition and avoid respondent fatigue.

To combine the general involvement in main sport, conditions of practice, and types of practice sub-sections, it is suggested that participation in the various conditions and types of practice be collapsed and addressed according to the following structure: a) sport-specific physical practice completed under group supervised, individual supervised, group unsupervised, and individual unsupervised conditions; b) physical preparation activities (including physical conditioning/weights, recovery techniques, and flexibility/yoga/Pilates) completed under group supervised, individual supervised, group unsupervised, and individual unsupervised conditions; c) mental preparation activities (including mental/psychological skills training, video analysis/review activities, sport-related study, and watching main sport live or on television) completed under group supervised, individual supervised, group unsupervised, and individual unsupervised conditions; and d) sport-specific informal play completed under group unsupervised and individual unsupervised conditions. It is expected that the proposed re-structure will reduce repetition, improve clarity of definitions for each practice type, and allow participants to combine their involvement in minor supplementary activities together, which will hopefully lead to decreased respondent fatigue and increased response accuracy. It is imperative that validity and reliability of recall following re-structure of this section be re-evaluated before utilising the DHAQ in the future. Specific recommendations regarding the structure of the competition history for main sport sub-section of the questionnaire will be provided following presentation of validity and reliability results from the existing detailed competition history for main sport items.

Detailed Competition History for Main Sport

Prior discussions in this chapter surrounding recall of competition involvement highlighted the complex nature of competition scheduling and structure, and the difficulties this creates for accurate reporting of competition information. While these complexities may have been overlooked in the design of earlier sections of the DHAQ, the detailed competition history section attempted to address the issue by requiring athletes to provide specific details of their involvement in competition at various representative levels (i.e. school, club, regional, state, national, and international).

For each year of their involvement at each representative level, participants indicated the age group classification in which they competed (e.g. under 14 years, 14/15 years, open age group, etc.), and for each age group classification at each representative level, they provided the

average frequency of competitions (e.g. one per week, one per month, two per year, etc.), and their typical level of success (e.g. won most matches/events, finalist, middle of the pack etc.). This elicited a very comprehensive record of participants' competition involvement, resulting in a large number of factors to assess for validity and reliability. Detailed results to accompany the discussion that follows are provided in Table 12. Note that coaches were only able to comment on their athlete's involvement in competition at the specific ages and level of competition for which they worked together, so convergent validity with coaches could not be assessed for all measures, all ages, or all representative levels.

General involvement in competition. To begin with, validity and reliability were determined for recall of age at first involvement in competition at any representative level, highest level of competition in which athletes participated both overall and at each age of involvement in main sport, and age when first participated in highest level of competition overall. Each of these measures was assessed for concurrent validity, convergent validity with parents, and test-retest reliability, as well as internal consistency with similar items from earlier sections of the questionnaire.

Age at first participation in competition at any representative level was classified as very good for concurrent validity, convergent validity with parents, test-retest reliability, and internal consistency with items from the sporting career and milestones (PA = 86.71, ICC = .77, $p < .01$), general involvement in main sport (PA = 85.43, ICC = .69, $p < .01$), and detailed history of involvement in main sport (PA = 90.85, ICC = .93, $p < .01$) sections of the DHAQ. Similarly, highest level of competition reached throughout participants' careers to date received very good classifications for concurrent validity, convergent validity with parents, test-retest reliability, and internal consistency with the sporting career and milestones (PA = 100.00) and general involvement in main sport (PA = 100.00) sections. Concurrent validity, convergent validity with parents, and test-retest reliability were also very good for age when first participated in highest level of competition, as was internal consistency with items from the sporting career and milestones (PA = 95.37, ICC = .87, $p < .01$), and general involvement in main sport (PA = 94.57, ICC = .85, $p < .01$) sections.

When examined on an age-by-age basis, recall for the highest level of competition reached during each year of involvement in main sport was not as strong. Test-retest reliability was very good or good across most ages, but concurrent validity, convergent validity with parents, and convergent validity with coaches were all rated as moderate or poor for several years. In addition, internal consistency with similar items from the general involvement in main sport section of the questionnaire was mostly moderate prior to age 16, but was very good from age 17 onward (see Table 13).

The results for this section so far support earlier suggestions that, particularly during the late childhood and adolescent years, athletes appear to have a very complex competition schedule involving participation in events at a variety of different representative levels. In a general sense, it appears as though athletes were able to consistently recall significant competitive milestones such as age at first involvement in competition, highest level of competition reached overall, and age at which they first participated at this representative level, but more specific details of competition involvement on a yearly basis seem to be difficult to recall with accuracy. The depth of information provided within this section of the DHAQ allowed us to explore this notion more closely.

Specifically, recall of age at first participation in competition at each representative level and whether athletes engaged in competition at each representative level across each year of their involvement in their main sport were assessed for validity and reliability; and for years during which athletes participated in competition at the various representative levels, recall of the age group classification in which they participated, and the average frequency of competitions were also examined. When asked to report their typical level of competitive success during both the initial and retest completions of the DHAQ, many participants did not provide a response. During interviews, participants expressed difficulty answering this question as their success often varied between competitions. As such, validity and reliability assessments were not conducted on responses relating to typical level of success, and it is recommended that this item be removed from the questionnaire. Remaining results relating to participation in competition are discussed according to representative level and the broad age group categories of junior (i.e. restricted age group) and open (unrestricted age group) competition.

Participation in competition at the high school level. A number of athletes reported participating in competition for their main sport at the high school level. Concurrent validity, convergent validity with parents, and test-retest reliability for age at first participation in school level competition were each classified as very good or good; and concurrent validity, convergent validity with parents and test-retest reliability for whether athletes participated in competition at the school level during each year of involvement in their main sport were also very good or good across all ages.

During the years athletes participated in competition at the high school level, concurrent validity, convergent validity with parents, and test-retest reliability for recall of the average frequency of school level competitions were very good, but identification of the age group classifications in which athletes competed was not as strong. Concurrent validity for age group classification for competition was very good, but test-retest reliability was mostly moderate, and convergent validity with parents was mostly poor. It appears that athletes were able to

consistently recall whether they participated in competition at the school level, and how often they participated in competition at this level, but recollection of the age group classifications in which they participated was less accurate.

Participation in competition at the local club level. Almost all athletes reported extensive competition experience at the local club level within both junior and open age group categories. Concurrent validity for age at first participation in junior club level competition was good, convergent validity with parents was very good, but test-retest reliability was moderate. Concurrent validity, convergent validity with parents, and test-retest reliability for whether athletes participated in competition at the junior club level during each year of their involvement in their main sport were very good or good across most ages, however, concurrent validity and convergent validity with parents were moderate to poor for ages 13-15.

Similarly, during the years athletes participated in junior club level competition, concurrent validity, convergent validity with parents, and test-retest reliability for frequency of competitions at this level were very good or good across most ages, with the exception of ages 13, 15, and 16 for which concurrent validity was moderate. With regards to identification of the age group classifications in which athletes participated during each year of their involvement this level of competition, concurrent validity and test-retest reliability for were both very good or good for most ages, with exceptions again at ages 15 and 16, and convergent validity with parents was also moderate or poor across several ages.

Age at first participation in open club level competition received very good or good classifications for concurrent validity, convergent validity with parents, and test-retest reliability, and identification of whether athletes participated in open club level competition during each year of their involvement in their main sport was also reasonably good; convergent validity with parents and test-retest reliability were classified as very good or good across all ages, but concurrent validity was rated as moderate for ages 13-15 and 22-23. For the years during which athletes participated in open club level competition, recall of the average frequency of competitions was excellent, with concurrent validity, convergent validity with parents, and test-retest reliability all classified as very good or good for all ages.

Results pertaining to participation in club level competition suggest that, in general, recall of involvement in this level of competition is relatively good; however, identification of the junior age group categories in which athletes compete at the club level were not recalled with great consistency during the teenage years. This period corresponds to instances of ‗competing up' in higher age groups as well as the transition from junior to open club level competition, so participation in multiple age group classifications during these years may be responsible for the moderate ratings observed.

Participation in competition at the regional level. As mentioned during discussion of the sporting career and milestones results, regional level competition does not appear to be a significant component of athletes‘ competition history. Only half of participants involved in this study reported competing at the regional level, with only one participant indicating involvement in regional competition within an open age group category. Therefore, assessment of validity and reliability of recall for details relating to participation in regional level competition was restricted to involvement in junior age group classifications.

Concurrent validity and test-retest reliability for age at first participation in junior regional level competition were good, but convergent validity with parents was moderate. Similarly, concurrent validity and test-retest reliability for whether athletes participated in junior regional level competition during each year of their involvement in their main sport were very good or good across all ages, but convergent validity with parents was moderate for ages 12-15. During the years in which athletes participated in junior regional level competition, both frequency of competitions and the age group classification in which athletes competed were rated as having very good or good concurrent validity, convergent validity with parents, and test-retest reliability, with just a few exceptions at several ages.

The results indicate that parent recall of involvement at this level of competition was not consistent with athlete recall, but as discussed previously, it is likely that athletes‘ responses regarding involvement in their main sport are more accurate than parents‘. Although athlete recall within this section of the DHAQ was reasonably good, relatively few participants reported engaging in competition at this level. Therefore, the recommendation stands that all items relating to participation in competition at the regional level be removed from the questionnaire.

Participation in competition at the state level. Recall of age at first participation in junior state level competition received classifications of very good for concurrent validity and test-retest reliability but moderate for convergent validity with parents. Identification of whether athletes participated in junior state level competition during each year of their involvement in their main sport was rated as very good or good across nearly all ages for concurrent validity, convergent validity with parents, and test-retest reliability, as was identification of the age group classifications in which athletes competed. Frequency of competitions for this level, however, were recalled with very low consistency, with moderate or poor classifications for concurrent validity, convergent validity with parents, and test-retest reliability across many ages.

Recall of age at first participation in open state level competition was classified as very good across all three validity and reliability conditions, and although test-retest reliability for whether athletes participated in open state level competition during each year of their involvement in their main sport was very good across all ages, concurrent validity and

convergent validity with parents were moderate for ages 21-23. Additionally, convergent validity with coaches for whether athletes participated in open state level competition during the time they worked together was mixed, ranging from good to poor for ages 17-23. Similar to junior state level competition, concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability for the average frequency of open state level competitions were each rated moderate or poor across a range of ages.

These results suggest that athletes are relatively consistent in their recall of when they participated in state level competition, and in which age group classifications they competed, but are less sure of the frequency with which competitions at this level occur. It is possible that scheduling of state level competitions may not be as regular as the school and/or club level competitions already explored, rendering it difficult to provide an ‗average' frequency.

Participation in competition at the national level. For national level competition, age at first participation in both the junior and open age group categories were recalled very consistently, with concurrent validity, convergent validity with parents, and test-retest reliability all rated as very good for both age groups. Similarly, identification of whether athletes participated in national level competition during each year of their involvement in their main sport was rated as very good or good across nearly all ages for concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability, again for both the junior and open age group categories.

During the years athletes participated in junior national level competition, the age group classifications in which they competed were recalled consistently, with concurrent validity, convergent validity with coaches, and test-retest reliability all rated as very good across all relevant ages, but convergent validity with parents received ratings of moderate or poor for several years. As for state level competition, recall of frequency of national level competitions for both the junior and open age group categories was mixed, with classifications for all validity and reliability conditions ranging from very good to poor.

Once again it is apparent that convergent validity with parents was slightly lower across all measures relating to competition involvement at the national level compared to concurrent validity, convergent validity with coaches, and test-retest reliability. It is also evident that, similar to state level competition, athletes were able to consistently recall the ages at which they were involved in national level competition, but providing an average frequency of competitions at this level was more troublesome.

Participation in competition at the international level. Finally, although several athletes reported participating in international level competition within junior age group categories, insufficient data were available to conduct meaningful analyses on validity and

reliability of recall for participation in competition at this level. As such, results relating to competition involvement at the international level are restricted to the open age group category.

Concurrent validity, convergent validity with parents, and test-retest reliability were all very good for age at first participation in open international level competition. Identification of whether athletes engaged in competition at the open international level during each year of their involvement in their main sport was very good for most ages; however, convergent validity with parents, convergent validity with coaches and test-retest reliability were moderate or poor for ages 22-23. As was the trend for state and national level competition, identification of the frequency with which international competitions occurred was highly inconsistent, with concurrent validity, convergent validity with parents, convergent validity with coaches, and test-retest reliability all receiving moderate to poor classifications across most ages. It is unusual that validity and reliability of recall for participation in international level competition was slightly lower at ages 22-23 as these represent the most recent years of involvement in main sport for most participants. Since several athletes were under the age of 23 at the time of data collection, the lower validity and reliability scores for these ages may reflect greater variability within a smaller sample size than earlier years.

Combined, these results indicate that athletes were able to recall the ages at which they participated in various representative levels of competition with acceptable validity and reliability; however, more specific details of competition involvement were difficult to recall consistently. When discussing details of their involvement in school and club level competition, regular scheduling of events (typically one competition per week) allowed for consistent recall of competition frequency, but the frequent occurrence of ‗competing up' and competing in multiple age group classifications simultaneously led to difficulties identifying the age group categories in which athletes participated. Conversely, state, national, and international level competitions did not appear to follow a regular schedule, making it more difficult to respond to items requesting the ‗average' frequency with which these competitions occurred. However, at these higher levels of competition, athletes typically reported participating in a single age group category, resulting in higher validity and reliability scores for identification of the age group classifications in which they competed.

It is interesting that, once again, scores for convergent validity with parents were consistently lower than scores for concurrent validity and test-retest reliability. Although, where available, convergent validity with coaches was also slightly lower than concurrent validity and test-retest reliability. While concurrent validity and test-retest reliability were relatively strong for a number of measures, the inconsistencies between athlete, parent, and coach responses call to question the accuracy of athlete responses.

Given the results highlighted above, it is recommended that items concerning average frequency of competitions be removed from the DHAQ, but identification of the ages at which athletes participated in each of the various levels of competition be retained. Based on suggestions arising from the general involvement in main sport and detailed history of involvement in main sport sections of the questionnaire, it is also recommended that this section be modified to address time involved in competition, which would require consideration of the various formats in which competitions are structured. Due to concerns surrounding the accuracy of athlete responses pertaining to competition involvement, it is recommended that validity and reliability of this section of the questionnaire be re-evaluated and monitored closely.

Coaching History

Table 14 outlines validity and reliability results for items within the coaching history section of the DHAQ. Although coaches interviewed for this study were required to provide details of their involvement with participating athletes, this section of the questionnaire encompassed information regarding all coaches with whom participants had trained; as such, convergent validity with coaches was not assessed for these items.

Convergent validity with parents and test-retest reliability for total number of coaches overall and total number of head coaches were classified as good, but concurrent validity was poor. Validity and reliability for total number of assistant/specialist coaches was also moderate to poor. Observations and researcher notes from athlete and parent interviews, and qualitative examination of responses to items within this section indicate that when completing the DHAQ athletes typically only identified coaches who they considered to be their ‗main' coaches (i.e. coaches of long duration and/or coaches of great significance to the athlete). Parents also tended to only identify several key coaches throughout athletes' careers. However, during athlete interviews, following prompting from the researcher, participants frequently recalled additional coaches with whom they trained for short periods of time. This could explain the lower concurrent validity for total number of coaches overall and total number of head coaches compared to convergent validity with parents and test-retest reliability.

Very few athletes reported training under the supervision of an assistant or specialist coach. In the rare cases that assistant or specialist coaches were identified, the duration of involvement with the coach was usually brief. Consequently, these coaches were not considered by athletes to be among their ‗main' coaches, so their details were recalled with less consistency. The low number of assistant/specialist coaches reported coupled with the relative

insignificance of these coaches to participating athletes' careers could explain the lower ratings of validity and reliability for total number of assistant and specialist coaches.

The duration of athletes' involvement with each coach was recalled with relatively good consistency. Validity and reliability ratings for average duration with a coach, longest duration with a coach, and shortest duration with a coach were mostly very good or good; the only exception being convergent validity with parents for shortest duration with a coach, which was classified as poor. As mentioned previously, coaches of short duration were often not identified by parents, contributing to this lower classification. It appears that when asked to identify the coaches under which they have trained, athletes were able to consistently recall details of their involvement with coaches who have been integral to their career, but do not necessarily mention coaches who have made more minor contributions to their development.

Considering the inability of the DHAQ to elicit complete coaching history information, it is recommended that this section be removed. The role coaches play in the development of sport expertise is complex and the current items within the DHAQ do not sufficiently explore the contribution of coaching to expert performance. Rather than modify and extend the DHAQ to gather more detailed information about coaching history, it is recommended that questions regarding associations between coaching and sport expertise be reserved for more specific investigations in the future.

Support Services

Validity and reliability of recall for information concerning participants' access to, and utilisation of, a variety of support services is detailed in Table 15. Insufficient data were available to assess validity and reliability of recall for strength and conditioning, skill acquisition, or career and education services, as few participants reporting having access to these specialists; however, details regarding recall of accessibility and utilisation of medicine, physiotherapy, massage therapy, psychology, nutrition, physiology, biomechanics, and performance analysis services are provided. Athletes were also asked to identify any additional services to which they had access throughout their sporting careers, but no further support services were discussed.

During completion of the questionnaires and interviews, participants identified ages at which they had access to each of the above mentioned support services, along with ages at which they utilised each service. The total number of years for which support services were accessible and utilised was calculated post-hoc to assist with validity and reliability investigations. Although coaches were required to identify the support services available to

athletes during the time training together, they were unable to comment on the age at which participants‘ first had access to each service, first utilised each service, or how many years they had access to and utilised these services in total. As such, convergent validity with coaches was not assessed for this section.

Overall, the age at which support services became accessible to participants was recalled with very good consistency by both athletes and parents, as was the age at which participants first began to utilise support services. Athletes were also consistent in their identification of how many support services were accessible and utilised throughout their career to date; however, convergent validity with parents for total number of support services accessible to athletes throughout their career was poor.

Looking more closely at recall for accessibility and utilisation of specific support services, validity and reliability classifications ranged from very good to poor. Test-retest reliability was very good or good for almost all support services, but concurrent validity and convergent validity with parents were highly variable. It is understandable that parents may not be aware of all support services available to athletes, particularly if athletes do not utilise the available services and/or if they have moved away from the family home. Therefore, the relatively low convergent validity with parents for this section is not surprising. The moderate to poor concurrent validity is, however, more problematic. Reasons for the strong test-retest reliability but moderate concurrent validity for items within this section are unclear, especially considering the interview procedures involved completion of a very similar table to that incorporated in the DHAQ.

In view of the inconsistent reporting of information relating to accessibility and utilisation of support services, it is recommended that this section be removed from the DHAQ. Future investigations wishing to explore associations between the provision of support services and sport expertise development should seek alternative methods for gathering these data.

Injury, Illness, and Time Off

Table 16 provides validity and reliability statistics for the total number of chronic injuries, acute injuries, and illnesses experienced by athletes throughout their careers, and the total number of days affected by injury and illness. The total number of days affected by injury and illness is also broken down according to the total number of days completely unable to participate in sport, total number of days on reduced training load, and total number of competitions missed. Additionally, validity and reliability statistics are provided for the total number of periods in which athletes took a break away from their main sport due to reasons

other than injury or illness, along with the total number of days away from both practice and competition during these periods. As was the case for support services, coaches were required to provide details of any injuries, illnesses, and periods away from the sport experienced by athletes during the time they trained together; however, as the presented validity and reliability statistics are an aggregate of all injuries, illnesses, and time away from sport across athletes' complete careers, convergent validity with coaches for these measures could not be assessed.

Combining details of all injuries and illnesses together, concurrent validity was good for all measures, test-retest reliability was good for all measures except number of competitions missed, and convergent validity with parents was moderate to poor. Looking at acute injuries, chronic injuries, and illnesses separately, it appears that athletes and parents were able to recall details of illnesses very consistently, but details of acute and chronic injuries were recalled with less certainty. Chronic injuries seem to be more memorable than acute injuries, as concurrent validity and test-retest reliability were very good or good for all measures relating to chronic injuries, whereas several measures relating to acute injuries received poor classifications for concurrent validity and test-retest reliability. Convergent validity with parents was slightly stronger for chronic injuries than acute injuries, although most measures received poor ratings for this comparison. Regarding periods of time away from sport due to reasons other than injury or illness, concurrent validity was good, convergent validity with parents was poor, and test-retest reliability produced mixed results.

Although athletes were able to recall details of chronic injuries and illnesses with relatively good consistency, their recall of acute injuries and periods away from sport was more variable. Additionally, convergent validity with parents was poor for the majority of measures. It is probable that athletes' recall of information concerning injuries, illnesses, and time off is more accurate than parents, as these events and their consequences are likely to be more significant to the athlete than their parents. Therefore, the low convergent validity with parents is not of great concern, but the poor concurrent validity and test-retest reliability for several items relating to acute injuries and periods away from sport is troublesome. Several athletes also expressed uncertainty and vagueness regarding items within this section during participation in the interview procedures.

The primary reason for including this section in the DHAQ was to adjust reports of time in practice and competition activities according to periods during which athletes' ‗average' schedule may have been disrupted. During the interviews however, it was apparent that when reporting details of their practice and competition history, most participants automatically accounted for altered training loads due to significant injuries, illnesses, and time off, even without prompting from the researcher. This section may therefore be redundant; although, the

secondary reason for its inclusion was to explore relationships between injuries, illnesses, and periods away from sport with factors including performance, stage of development, and utilisation of support services. However, considering the slightly inconsistent recall of details relating to injuries, illness, and time off, it is suggested that investigations of these associations may not produce trustworthy results.

Given the inconsistent recall for several items, noted redundancy, and negative participant feedback, it is recommended that this section be removed from the DHAQ. While the occurrence of injuries and illnesses may significantly impact athletic careers, the DHAQ does not appear to be an appropriate tool to explore the influence of injuries, illnesses, and time off on sport expertise development.

General Discussion

This study aimed to establish validity and reliability of the DHAQ, a new tool for the collection of athlete developmental history information. In contrast to the measurement tools utilised for collecting athlete developmental history data in the majority of previous investigations, validity and reliability was scrutinised for every item within the DHAQ, rather than just a small selection. Overall, validity and reliability results were mixed. While many items were classified as having acceptable validity and reliability for retention in the questionnaire, other items were recalled with such poor consistency that they are recommended for removal. In addition, several items that did not meet validity and reliability standards but were considered of critical importance to the study of sport expertise development were suggested to undergo major revisions and be reassessed for validity and reliability before using the DHAQ in future investigations.

In addition to the specific findings already discussed within each section of the questionnaire, a number of general observations regarding validity and reliability of retrospective recall are particularly noteworthy. First, our rigorous approach to establishing trustworthiness of information collected by the DHAQ highlighted that validity and reliability checks incorporated within previous investigations of sport expertise development may not have been detailed enough to be certain of recall accuracy.

For example, validity and reliability assessments involving measures of total cumulative hours of involvement in practice activities over a given period of time may be misleading, as on several occasions, validity and reliability classifications for total cumulative hours were very good or good, despite moderate or poor validity/reliability for hours per year across the majority of ages. This is likely due to differing magnitudes of the values involved in each assessment.

Since PA was adopted as the primary measure of validity and reliability, the 20% error margin associated with a very good classification provides a more generous range of variability acceptability during assessments involving large values compared to smaller values. Consistency of recall for hours of involvement in practice activities per year of involvement is imperative when analysing skill level differences in practice history profiles at various ages and stages of athletes‘ careers, and also when assessing changes in practice involvement over time. Therefore, validity and reliability of recall for hours of participation per year, and wherever possible, hours per week should be established instead of, or as well as, validity and reliability of recall for total cumulative hours of involvement.

Similarly, when assessing validity and reliability of recall for number of organised sports, playful sporting games, and/or practice activities in which athletes have engaged, it is essential to compare the lists of activities provided during each test occasion and examine congruence first. Consistency of recall for the number of other sporting or practice activities in which athletes have participated is meaningless if the specific types of activities identified vary from one data collection occasion to the next; as was the case in the present study within both the organised sport practice history and informal, playful sporting game sections.

Furthermore, our results highlighted considerable variability in consistency of recall between different sections of the questionnaire, and also between different items within the same section. Previous studies of sport expertise development have typically assessed validity and reliability of recall for several items within their measurement tool, and assumed acceptable accuracy across all remaining items. This finding emphasises the importance of assessing validity and reliability for all items within developmental history questionnaires and interview guides, rather than a select few. Therefore, when constructing novel instruments in the future, researchers should attempt to be more thorough in their assessments of validity and reliability. Alternatively, existing measurement instruments with established validity and reliability, such as the DHAQ, could be utilised to ensure confidence in accuracy of recall, and also to allow more direct comparisons of findings between investigations.

Another general trend observed throughout this study was that the teenage years appeared to be a particularly complex phase of athlete development, resulting in comparatively weaker validity and reliability ratings for many items than during the initial introductory and later adult years of involvement. Prior to age 11, athletes‘ participation in their main sport typically involved small amounts of sport specific physical practice completed in a group under the direct supervision of a coach, and competition at the local club level. The relative simplicity of this phase resulted in reasonably good consistency of recall. Between ages 11 and 17, athletes reported participating in several additional practice activities, and started to engage in

competition at the regional, state, and national levels of competition, often within both junior and open age group categories. Additionally, during these years, athletes were often participating in several other organised sports, and were likely balancing school and non-sport related extra-curricular and social activities as well. Consequently, details of involvement in main sport, other organised sports, and informal sporting play during this busy time were not remembered clearly, leading to less consistent recall and lower ratings of validity and reliability.

Following the teenage years, the volume and variety of practice activities in which athletes participated tended to increase, but competitions appeared to occur less frequently due to a reduction in the number of competitive levels and age group categories in which they were involved, plus, participation in other organised sports typically ceased. By this age athletes had often committed to the goal of becoming an elite athlete, and sport had become the focus of their attention. As a result of the less complex competition schedule and increased importance of their main sport during this phase, it is likely athletes‘ involvement became more stable and a more significant component of their everyday lives. These factors, along with the closer proximity of these years to the present, could have facilitated more accurate recall, leading to improved validity and reliability of responses.

Interestingly, these phases align with those outlined in the DMSP (Côté & Fraser-Thomas, 2007). The initial introductory phase described above corresponds to the sampling years of the DMSP, while the demanding teenage period corresponds to the specialising years. The final phase during adulthood, characterised by increased commitment to main sport, is reflective of the DMSP‘s investment years. Thus, the sampling years seem to represent a highly variable period of sport involvement, resulting in less consistent recall of details of participation during this time. Therefore, information pertaining to athlete developmental histories during the teenage years should be interpreted with caution.

One final broad observation from this phase of the investigation pertains to the specificity of information able to be recalled consistently. General information such as age at first participation in various sporting activities, most significant achievement to date, and overall hours of practice per week tend to be recalled quite consistently. However, the more specific the requirements of the question, for example, levels of competition and age group classifications participated in at each age, and involvement in particular types of practice, the less certain responses become. While this finding may be somewhat intuitive, it also limits our capacity to fully understand the sport participation profiles associated with expertise development. While general trends and patterns of participation are important to understand, knowledge of more specific details such as practice composition and competition scheduling would allow for more precise recommendations regarding optimal conditions for the development of sport expertise.

In order to continue to advance this field of study, researchers should endeavour to explore methods for more accurate collection of specific details of sport involvement, for example completion of detailed training diaries or longitudinal tracking.

Two main findings of this study were particularly surprising given the previous research in this area. The relatively low validity and reliability of recall for participation in other organised sports and informal, playful sporting games was not predicted since these factors are thought to be strongly associated with sport expertise development (see Baker, 2003 and Côté, Lidor, and Hackfort, 2009 for reviews). Although possible reasons for these results have already been discussed, it is necessary to reinforce the potential for low accuracy of recall for this information considering the attention diversification, specialisation, and deliberate play have received in the sport expertise literature. This is not to say that participation in a variety of sports and informal sporting games is not important or related to the development of sport expertise, but rather to highlight the precision with which this information is recalled may be questionable; and greater attention ought to be paid to the assessment of validity and reliability for recall of involvement in these activities.

Additionally, the relatively low convergent validity with parents was unexpected since many studies have incorporated parent interviews as a key indicator of data trustworthiness (Baker et al. 2003a; Baker et al., 2003b; Berry et al., 2008; Côté, 1999, Côté et al., 2005; Durand-Bush & Salmela, 2002; Law et al., 2007; Oldenziel et al., 2003; Soberlak & Côté, 2003). Our results suggest that while parents may be suitable sources for validating demographic, residential, and familial information, they may not be appropriate for validating practice history details, including involvement in main sport, other organised sports, and informal, playful sporting games. Oldenziel et al. (2003), also observed relatively low convergent validity with parents for a number of items within their National Athlete Development Survey, citing a variety of potential reasons behind the inconsistencies including insufficient understanding of questions by athletes and/or parents, different interpretations of questions, different definitions, for example, of what constitutes training, and insufficient knowledge to accurately answer some questions. Each of these factors is also possible within the current study. Although coaches may be an apt substitute for validating information pertaining to participation in main sport, at least for the years during which they were involved with the athlete, more objective records such as training diaries, official competition programs and results guides, news items, photographs, and mementos ought to be consulted wherever possible to consolidate participants‘ memories, and confirm as many details as possible.

A number of methodological considerations are also pertinent to this discussion. As noted earlier in greater depth, several items within the DHAQ were not sufficiently detailed

and/or optimally structured to facilitate accurate recall of developmental history information. As such, a number of recommendations have been made to improve the robustness of the DHAQ. During modification of the DHAQ it is important to ensure that all directions, examples, and response options are very specific, as this will lead to greater understanding of requirements, improved ease of completion, and more valid, reliable results. Related to improving response quality, while redundant items built into the questionnaire allowed for the assessment of internal consistency, the repetitive nature of these items appeared to be frustrating for participants, and may have been detrimental to the accurate completion of the questionnaire. It is recommended that the number of redundant items be reduced to ensure participant satisfaction and diligent reporting of information. Given the extent of the suggested modifications to the DHAQ, it is essential that the updated version undergo additional validity and reliability checks to ensure all modified and new items do, in fact, elicit consistent recall.

As further validity and reliability investigations are required to ensure the trustworthiness of the DHAQ, it is recommended that a more diverse participant sample be employed so the findings regarding the psychometrics of the DHAQ can be generalised to a wider population than was involved in the current study. As only 15Australian national and international level swimmers and field hockey players were recruited for participation in this investigation, it could be argued that the results obtained are applicable only to Australian athletes of national and international calibre, in these specific sports. Therefore, future investigations of the validity and reliability of the DHAQ should seek to recruit athletes from a variety of sports, and a range of skill levels so it can be determined whether the measurement tool is suitable for use with a broader range of participants.

Despite the unexpected findings and methodological limitations apparent in this investigation, perhaps the greatest contribution of this study to the sport expertise literature is the proposed methodology and taxonomy for establishing trustworthiness of data collected via retrospective recall. Not only does the calculation of PAs and ICCs result in more appropriate assessments of response consistency than the more common Pearson correlation coefficients, the integration of PA and ICC statistics into an explicit classification system ensures uniform interpretation of validity and reliability results, and provides a standard to which future investigations can adhere. Given the lack of firm criteria for assessing validity and reliability of retrospective recall data, and indeed the interpretation of PA and ICC statistics in isolation, the proposed methodology and taxonomy is a great strength of this study. Furthermore, the classification system could easily be adapted for use in investigations involving retrospective recall techniques in domains outside of expertise and skilled performance, including, but not limited to, examination of health behaviours and participation in physical activity.

Summary and Conclusions

This study highlighted a number of issues relating to the validity and reliability of retrospective recall for athlete developmental history information. As a result, a number of modifications are required before the DHAQ is suitable for use in hypothesis testing investigations pertaining to the development of sport expertise. A summary of the suggested modifications is outlined below.

Items within the sporting career and milestones section were generally recalled with very good consistency, and only minor modifications are suggested. Items relating to participation in regional level competition and career stages and transitions are recommended for removal, however all other items may be retained.

Family characteristics were also recalled consistently, with the exception of parental occupations. This item should therefore be replaced with an indication of the highest level of education each parent completed. Recall for frequency of familial participation in sport and physical activity was relatively poor, however items within this section were found to be insufficiently detailed to elicit accurate responses. It is therefore recommended that this section be updated with more specific lines of questioning and response options.

Ages and locations of places of residence in which participants lived for extended periods of time were recalled consistently, but validity and reliability ratings for residences of shorter duration were slightly lower. Since short duration residences were usually within close geographical proximity to extended duration residences, it is recommended that this line of questioning be modified to collect information pertaining to each city/town of residence rather than each individual home. Furthermore, identification of whom participants lived with at each place of residence proved to be troublesome during the adult years, as it was sometimes difficult to classify co-habitants into specific categories. As such, this item is recommended for removal.

Recall of details of participation in other organised sports and informal, playful sporting games was generally quite poor; however, given the emphasis on these factors as considerable contributors to the development of sport expertise, it is recommended these sections be retained in the DHAQ, with the caveat that related results be interpreted cautiously. While no modifications are suggested for the organised sport practice history section, it is recommended that the participation in informal, playful sporting games section be updated to collect a list of all informal, playful sporting games in which participants engaged on a regular basis, but require only general indications of total hours involved in sporting play overall, rather than estimates of time involved in each individual activity. Furthermore, items relating to whom

participants typically engaged in sporting play with, and where these games were usually played, are recommended for complete removal.

Regarding recall of information relating to participation in main sport, athletes and coaches were very consistent in their recall of general involvement in main sport, but convergent validity with parents was slightly weaker. Age at first participation in practice and competition overall, age at first participation in various conditions/types of practice and various levels of competition, and overall hours of involvement in main sport per year of involvement were recalled reasonably consistently, but most significant achievement and main reason for participation during each year of involvement were recalled quite poorly, prompting recommended removal of these items from the questionnaire. Athletes were able to consistently recall their involvement in supervised practice activities, however validity and reliability of recall for unsupervised activities was not as good. Similarly, athletes could consistently recall their involvement in sport-specific physical practice, physical conditioning/weights, and video analysis/review activities, but participation in additional supplementary practice types such as mental/psychological skills training, sport-related study, and recovery techniques appeared to be more sporadic and more difficult to recall with consistency.

Specific details of involvement in competition, for example, competitive levels and age group categories participated in during each year of involvement, frequency of competitions, and time engaged in competition activities, were recalled quite poorly; however, it became clear throughout data collection that the DHAQ did not adequately cater to the complexity of competition involvement so it is expected that recall will improve following major modifications to this section. Recommended modifications to the detailed competition history section include: a) removal of all items concerning frequency of competitions; b) division of competition involvement into regular season, occasional, and tournament format competitions; and c) separation of time involved in competition into time spent actively engaged in the event, and time at the competition venue but not physically competing.

Furthermore, it is recommended that the general involvement in main sport, conditions of practice, and types of practice sections be combined into a single detailed practice history for main sport section, encompassing involvement in sport-specific physical practice, physical preparation activities, mental preparation activities, and sport-specific play, each under supervised and unsupervised conditions, in both group and individual settings.

Validity and reliability of recall for items within the coaching history, support services, and injury, illness, and time off sections was mixed, with classifications ranging from very good to poor. While recall for many items within these sections was sufficient to warrant inclusion in the DHAQ for future investigations, it was deemed that each of these potential influences on

sport expertise development deserved greater attention to detail than was currently received within the questionnaire. Rather than extend the DHAQ to address these areas more closely, it is recommended that the coaching history, support services, and injury, illness, and time off sections be removed from the questionnaire and reserved for more specific investigations in the future.

This phase of the investigation sought to establish the validity and reliability of the DHAQ. Through the development of a stringent methodology and taxonomy for assessing validity and reliability of retrospective recall, each item was clearly classified as being suitable for retention, recommended for removal, or in need of modification. Given the large number of revisions required for the DHAQ to be suitable for use in large scale examinations of long-term athlete development, reassessment of reliability of the questionnaire is recommended following implementation of the suggested modifications.

Study Phase Three: Modification and Online Conversion of the Developmental History of Athletes Questionnaire

This chapter outlines amendments to the DHAQ based on the results and recommendations of the initial validity and reliability assessments. Details of all revisions, along with an updated outline of the structure and content of the DHAQ are provided. Moreover, this chapter describes the conversion of the DHAQ from a paper-based questionnaire to an online format. Benefits and limitations of web-based measurement tools are discussed.

Results of the initial validity and reliability assessment of the DHAQ highlighted some important modifications required for the questionnaire to be suitable for use in hypothesis testing investigations related to the development of sport expertise. In his book on questionnaire design, Oppenheim (1992) indicated ―questionnaires have to be composed and tried out, improved and tried out again, often several times over, until we are certain that they can do the job for which they are needed". Oppenheim (1992) further suggested that this piloting process should address item relevance, wording, order, and layout, as well as administrative and analysis procedures. Question wording, type, and order were also highlighted as fundamental issues in questionnaire design by Murray (1999); and assessment of validity and reliability, item amendment and reduction, and measure reassessment have been outlined as key steps within the questionnaire design process by Hinkin (1998) and Rattray and Jones (2007). Each of these critical components of questionnaire design was addressed during the modification process for the DHAQ, as outlined in detail below.

The first step of the modification of the DHAQ was eliminating all items that had insufficient validity/reliability ratings, and/or were considered unsuitable for retention in the measurement tool. This led to the deletion of the coaching history, support services, and injury, illness, and time off sections of the questionnaire, as well as several items within the sporting career and milestones, places of residence, and participation in informal, playful sporting games sections. Second, alterations were made to sections deemed to have sub-optimal item structure, including reconfiguration of the participation in informal, playful sporting games, and main sport practice and competition history sections, and adaptation of several items within the family characteristics and participation in sport and physical activity section.

Finally, the sections of the DHAQ were reordered to improve the flow of questioning and ensure that items requiring the most detailed responses were completed during early phases of participation when attentional focus is likely to be greatest. This was an important consideration since aversion to continue, distractibility, and mental fatigue have been shown to increase, and performance shown to decrease, with increasing time on task (Boksem, Meijman, & Lorist, 2005). It was expected that these negative effects would be minimised if the final stages of the questionnaire could be completely quickly and easily, with minimal cognitive effort. A summary of the updated DHAQ is below, outlining changes from the original version, and a complete version of the revised questionnaire is provided in Appendix G.

1. Demographic Information

This section remained relatively unchanged from the original DHAQ, collecting basic demographic details including sex, date of birth, main country of residence, and highest level of education completed.

2. Sporting Career

The sporting career and milestones section of the original DHAQ was divided into two separate sections within the updated DHAQ. Similar to the original DHAQ, the restructured sporting career section required respondents to indicate their main sport, highest level of competition reached to date, and whether they had reached the peak of their career.

Newly added to the modified version, participants who had reached the peak of their career estimated the age when they reached their peak, and indicated whether they were still participating in practice and competition activities. If participation in practice and/or competition activities had ceased, ages of cessation of participation in these activities were provided. Participants who had not reached the peak of their career predicted the age at which they thought they might reach their peak, and the highest level of competition at which they realistically expected to participate for their main sport. More detailed information regarding career peak was included within the updated DHAQ to gain a greater understanding of each athlete‘s current situation in relation to their full or predicted career span. Following review of the initial questionnaire, it was apparent that such information was absent, yet would assist in the identification of athletes‘ current stage of development.

3. Sporting Milestones

Similar to the original version of the questionnaire, this section required athletes to indicate the ages at which they reached various sporting milestones including first participation in main sport in any format, regular supervised practice for main sport, regular unsupervised practice for main sport, non sport-specific practice activities, and off-season or year-round training. Athletes also identified the age when they stopped involvement in all other sports to concentrate on their main sport, first developed the idea of becoming an elite athlete, made a conscious decision to become an elite athlete, noticed that all leisure time was spent on activities relating to main sport, first moved house for reasons relating to main sport, and first established a close and extended relationship with a coach for their main sport.

A timeline of progression through the various levels of competition was then provided through indication of the age at which athletes first participated, started to make competition finals, started to place in the top three, and started to place first at the local, state/provincial, national, and international levels for individual sports, or when they first participated, became a regular starting player, became one of the top five players on a team, became the best player for their position on a team, and became the best player overall on a team at each of the representative levels for team sports. These competitive milestones were collected for participation in both the junior and open age group categories for all representative levels.

4. Representative History

Sections four, five, and six of the updated DHAQ were the result of major modifications to the main sport practice and competition history section of the original DHAQ. In this restructured representative history section, participants provided details of the representative levels and age categories in which they participated for each year of involvement in their main sport. To facilitate clear recall, the section was structured according to representative level. That is, athletes indicated the ages at which they participated in competition against others within the local area, and the age group categories in which they competed. This process was repeated for competition at each of the state/provincial, national, and international levels of competition, giving a clear indication of the ages at which participants were involved at each of the representative levels, and highlighting any times at which athletes may have ‗competed up' in a higher age group than expected for their age. This information provides valuable insight into athletes' involvement in competition, and the timeline and pathway of progression through representative levels.

5. Practice History

The general involvement in main sport, conditions of practice, and types of practice sections from the original DHAQ were combined into a single practice history section for the updated DHAQ. Within this section, participants indicated the ages at which they participated in the following practice activities for their main sport: a) sport-specific physical practice completed with others under the direct supervision of a coach, one-on-one with a coach, with others but without a coach, and on their own; b) physical preparation activities completed with others under the direct supervision of a coach/specialised instructor, one-on-one with a coach/specialised instructor, with others but without a coach/specialised instructor, and on their

own; c) mental preparation activities completed with others under the direct supervision of a coach/specialised instructor, one-on-one with a coach/specialised instructor, with others but without a coach/specialised instructor, and on their own; and d) sport-specific play completed with others and on their own. For each age of participation in each practice activity, athletes provided the average number of hours per week and total number of months per year they were engaged in the activity. Participants also indicated the ages at which they participated in training camps for their main sport, as well as the total number of camps attended each year, and the average duration of each camp.

6. Competition History

Following recommendations from initial validity and reliability assessments of the DHAQ, the competition history section of the updated DHAQ was reorganised to address involvement in regular season competitions, occasional competitions, and tournament format competitions separately. For each year of involvement in regular season competitions (i.e., regularly occurring events/matches), participants identified the number of months in the season, average number of events/matches per month, average time spent at the competition venue each event/match, and average time actively participating in competition during each event/match. For each year of involvement in occasional competitions (i.e., single day competitions occurring less often than once per month), participants indicated the total number of occasional competitions in which they participated, the average number of events/matches per occasional competition, average time spent at the competition venue on each day of occasional competition, and average time spent actively competing during each day of occasional competition. Finally, for each year of involvement in tournament format competitions (i.e., competitions held over two or more consecutive days), athletes reported the total number of tournaments in which they participated, average number of events/matches per tournament, average time spent at the competition venue each event/match, and average time actively participating in competition during each event/match. These measures allowed calculation of the total number of events in which athletes competed and the total number of hours involved in competition activities per year of participation in their main sport, while at the same time allowing athletes to be more precise in their descriptions of competition involvement by addressing each type of competition separately.

7. Participation in other organised sports

This section remained unchanged from the original DHAQ. Participants listed all organised sports they participated in (other than their main sport), and for each year of involvement in each sport, they provided the total number of months per year and average number of hours per week engaged in all practice and competition activities combined, along with the highest level of competition reached.

8. Participation in informal playful sporting games

Validity and reliability of information obtained from this section of the questionnaire during initial assessments was generally poor. In an attempt to improve recall of involvement in these activities, this section was modified for the updated DHAQ. Participants first provided a list of all informal playful sporting games they typically engaged in. Then, instead of providing estimates of time involved in each sporting game separately, for each age they regularly participated in informal sporting play, athletes estimated the total number of hours per week and months per year they engaged in sporting play overall, combining involvement in all informal sporting games together.

To gain an understanding of the activities in which participants engaged most often, athletes were required to identify the two sporting games they participated in most frequently, and then estimated the percentage of total play time devoted to each of these two games. Items concerning whom participants played these informal sporting games with and where were removed from this version of the DHAQ.

9. Family

The section of the updated DHAQ relating to familial characteristics and participation in sport and physical activity incorporated several minor changes from the original version. For each member of their immediate family, participants provided date of birth, main country of residence, highest level of education, and the ages they lived together. The sex of each sibling was also collected. Athletes then indicated the frequency with which each family member participated in general fitness activities, recreational sport, and competitive sport during the time living together, as well as a list of all competitive sports each family member participated in during any time in their lives, and the highest level of competition reached for each sport.

Information was collected for biological parents as well as step-parents and/or other legal guardians, along with full-, half-, step-, and adoptive- siblings.

10. Places of Residence

The original DHAQ required participants to provide information relating to each home in which they had lived. The updated DHAQ required participants to instead identify each town/city in which they had lived. Athletes first provided the location of their birthplace, then for each subsequent town/city in which they lived, participants indicated the age at which they moved, and the reason for their relocation. The item pertaining to whom participants lived with at each location was not retained in the updated DHAQ.

General modifications to the updated DHAQ intended to increase the quality of information elicited included: a) enhanced detail and specificity within item descriptions, including examples, to improve clarity and maximise participant understanding; b) replacement of open-ended responses with clear response options to improve uniformity of responses and aid comparison of results between participants; and c) addition of comment boxes for participants to provide additional details where necessary. It was expected these modifications would result in improved validity and reliability of the revised DHAQ relative to the original version.

Following modification, the DHAQ was converted to an online format. With the exception of a recent investigation by Moesch et al. (2011), studies of sport expertise development have typically involved interviews or paper-based questionnaires; however, web-based surveys have a number of advantages over traditional data collection methods, including increased reach, rapid distribution and response times, and automated data entry and analysis procedures (Duffy, 2002; van Selm & Jankowski, 2006; Wyatt, 2000). Online questionnaires also allow participants to complete their responses at a time and location of their choosing, and interactive feed-forward programming can direct respondents to the next applicable item, allowing convenient, customised completion (van Selm & Jankowski, 2006; Wyatt, 2000). Furthermore, online questionnaires can be updated quickly and easily, and reminder, follow-up, and/or thank you notices can be distributed automatically (van Gelder, Bretveld & Roeleveld, 2010; Wyatt, 2000).

Despite these benefits, web-based questionnaires have several limitations; most notably, participants must be computer literate and have access to the Internet. Additionally, completion conditions are non-standardised, and researchers are unable to observe participants during data collection to evaluate their understanding of questionnaire items, focus, and fatigue (Duffy, 2002; Manfreda, Batagelj, & Vehovar, 2006). Web-based questionnaires are also known to have

lower response rates than email- and mail-based unsupervised data collection methods (Duffy, 2002; Manfreda et al., 2006; van Gelder et al., 2010; van Selm & Jankowski, 2006).

Since the DHAQ was intended for administration to large populations in multiple locations, the benefits of online data collection were considered to outweigh the limitations, so the questionnaire was converted to a web-based format. To overcome some of the limitations outlined, data collection sessions involving supervised completion of the online tool could be conducted; and a paper-based version of the updated DHAQ could be produced in cases where Internet access is not available or online completion is not convenient.

The online DHAQ was created using SurveyMonkey, a web-based questionnaire development and management service (http://www.surveymonkey.com). Items were constructed using the survey design tools, incorporating a selection of multiple choice, matrix of choices, matrix of drop-down menus, and text box response types. Key words and phrases within questions were emphasised using a combination of bold, italic, and underlined text, and page breaks were inserted at logical points to minimise the degree of scrolling required on each page. Where appropriate, question logic was applied to direct participants to the next applicable item based upon their response. For example, if a participant indicated they did not have any siblings, all items regarding sibling characteristics and participation in sport and physical activity were automatically skipped, and the respondent was directed to the next section. A progress bar was inserted on each page to provide an indication of completion status, and an automatic save feature was applied to ensure all responses are saved upon progression to a new page. Participants were able, however, to browse back through completed pages and edit previous responses if necessary.

The online DHAQ has potential to be access-controlled and password protected. Upon activation of these security measures, access to the questionnaire can be gained via a personalised email containing a hyperlink to the DHAQ and a participant identification code, which doubles as the access password. Each hyperlink can be uniquely associated with the respondent‘s email address, so in the event a registered participant forwards their email containing the hyperlink to an unregistered participant, the unregistered participant would receive an error message upon attempt to access the questionnaire. These security features can assist participant screening procedures, minimise the risk of participants completing the questionnaire multiple times, and also allow respondents to pause their participation at any stage and return to the questionnaire at a later time point, as clicking on the personalised hyperlink in the email would direct them to the page upon which they exited during the previous visit. The access-control features also allow for automatic monitoring of completion status, as periodic

reminder emails can be distributed to non-responders and partial-responders, and thank you emails can be distributed to participants who complete the questionnaire in its entirety.

A negative aspect of the access-control feature is that participants‘ responses can be traced to their email address. To maximise anonymity and confidentiality in this situation, only the lead researcher should have access to the server and database containing email addresses, and all responses should be immediately transferred to an external database for storage and analyses. The external database should contain only alphanumeric identification codes to distinguish between participants. When access-control and password security features are not activated, anonymous access to the questionnaire can be obtained via a public URL.

Given the large number of modifications and online conversion of the DHAQ, reassessment of the reliability of the instrument is required to ensure the trustworthiness of the updated measurement tool. Moreover, it is recommended that a diverse sample of participants including athletes from a variety of sports and a range of skill levels be recruited during the next phase of reliability testing to allow for greater generalisability of findings and recommendations relating to continued use of the DHAQ.

Study Phase Four: Reassessment of Test-Retest Reliability of the Developmental History of Athletes Questionnaire

This chapter describes the final study phase, which aimed to reassess test-retest reliability of the updated, online DHAQ. Procedures and results are discussed in detail, and further modifications are recommended for the DHAQ to be suitable for use in large scale examinations of athlete development.

Methodology

Participants

Participants involved in this phase of investigation included 20 athletes from Australia (n = 13) and Canada (n = 7), none of whom were involved in the initial assessment of validity and reliability of the DHAQ. Athletes were male (n = 6) and female (n = 14), aged 15-38 years (M = 24.33; SD = 6.80), and were current (n = 17) or former (n = 3) participants in Australian rules football (n = 1), baseball (n = 1), basketball (n = 2), disc sports (n = 1), diving (n = 2), soccer (n = 2), gymnastics (n = 1), netball (n = 1), swimming (n = 3), track and field (n = 3), triathlon (n = 2) and volleyball (n = 1). The highest levels of competition at which athletes participated for their main sport were relatively balanced across the local to international levels. Participants were recruited via advertisements placed on sporting organisation websites, social media pages, and newsletters, inviting athletes of all skill levels to participate in an investigation of factors associated with the development of sport expertise. Ethical approval was obtained prior to participant recruitment (see Appendix H), and informed consent was received from all athletes (see Appendix I). Parental consent was obtained for participants under 18 years of age.

Research Design

This study involved a repeated measures test-retest design, whereby participants completed the same procedures on two occasions. Since the DHAQ requires participants to retrospectively recall details of their past involvement in sport and physical activity, responses should remain constant over time. Therefore, a test-retest design allows for assessment of how consistently athletes answer the same questions on separate occasions, reflecting the ability of participants to accurately recall various characteristics of their sporting history, as well as the capacity of the measurement tool to elicit repeatable responses.

Measures

All items within all sections of the DHAQ, as described in study phase three, were subjected to test-retest reliability analyses. Consistency of responses between test occasions was the primary measure of interest, rather than examination of responses in the context of sport expertise development (i.e., reliability rather than construct validity). Following confirmation of test-retest reliability of the updated online DHAQ, responses may be utilised for hypothesis

testing investigations related to the development of sport expertise; however, the focus of the current study is limited to establishing test-retest reliability of the measurement tool.

Procedures

All participants completed the online DHAQ on two separate occasions. On each occasion, the questionnaire was completed unsupervised at a time and location of the participants‘ choosing. Access-control and password protection security features were activated on the DHAQ, so to complete the questionnaire, athletes were first emailed an information package and consent form, and were alerted to check their email for a subsequent message containing a hyperlink to the DHAQ and their personalised access password. Completion status was monitored on a regular basis, and reminder emails were distributed periodically to participants who had not finished their responses. A thank you email was sent to each athlete upon full completion of the questionnaire.

The average duration between test occasions was approximately 6 months (*SD* = 2 months; *range* = 2-10 months). Unfortunately, it is difficult to obtain an accurate indication of completion duration for the updated, online DHAQ, as questionnaires were completed unsupervised. Although the online host for the DHAQ, SurveyMonkey, does record time stamps for initial entry into the questionnaire and final submission of responses, many participants chose to complete the DHAQ over multiple sessions, rather than all in one sitting. As such, the time stamp for submission of responses was often on a different day to the time stamp for initial entry into the questionnaire. For those participants who did appear to complete the questionnaire in a single sitting, it appears that despite the removal of a number of sections from the original DHAQ, completion duration for the updated, online DHAQ is still approximately 60 minutes.

Statistical Analysis

Responses from the first completion of the DHAQ were compared to those provided during the second completion, and examined for similarity. Consistency of responses involving categorical data was assessed via calculation of PA, with PA reflecting the percentage of responses that were the same on both test occasions. Test-retest reliability of categorical items was classified as very good, good, moderate, or poor according to the criteria for agreement outlined in Table 2. For responses involving continuous variables, both PA and single measures absolute agreement ICCs were calculated, however, in this case PAs were determined by dividing the smaller value of the pair by the larger, and multiplying the result by 100%. Test-

retest reliability for these items was classified based upon a combination of the PA and ICC results (see Table 2). Wherever possible, internal consistency was also assessed for similar items within separate sections of the questionnaire. Statistical analysis for determination of internal consistency was the same as for test-retest reliability.

As in Phase 2 of the investigation, statistical analyses were not performed on data from the current year to avoid any issues related to incomplete data sets for this year. Analyses were also not performed on items for which data was available for fewer than five participants, because in these cases it was considered that there was insufficient data to draw meaningful conclusions from the results. All instances for which insufficient data was available for analyses are highlighted throughout the text and in all relevant tables in Supplement One. All statistical analyses were conducted using the SPSS statistical computer software package, Version 17.

Results and Discussion

Sporting Career

Test-retest reliability was very good for identification of highest level of competition reached to date and whether athletes felt they had passed the peak of their career. Among athletes who indicated they were past the peak of their career, identification of age at which they reached their peak was recalled with very good consistency, as was identification of whether they were still involved in practice and/or competition activities for their main sport. Unfortunately, insufficient participants reported having ceased involvement in practice and/or competition activities for their main sport to conduct accurate assessments of test-retest reliability for age at cessation of participation. Among athletes who had not reached the peak of their career, test-retest reliability was very good for both predicted age at which they expected to reach their peak, and highest level of competition at which they expected to participate.

Therefore, athletes were very consistent in their recall of general information regarding their sporting career to date, as well as their predictions for the extent of their career in the future. As such, no further modifications to this section are required. More detailed information regarding test-retest reliability results for the sporting career section is provided in Table 17.

Sporting Milestones

Table 17 also outlines PA and ICC statistics for test-retest reliability of all items within the sporting milestones section of the DHAQ. Consistency of recall for age when respondents

reached each of the general sporting milestones was strong, with all items classified as having very good or good test-retest reliability. Additionally, athletes involved in individual sports recalled their progression through the various levels of competition quite consistently, with almost all items displaying very good or good test-retest reliability. Only two exceptions were evident, with age when started to place first at the junior local and junior state/provincial levels of competition classified as having moderate test-retest reliability. Similarly, test-retest reliability was very good or good for all team sport milestones except age when first recognised as the best player on their team at the junior local level, age when first became a starting player on their team at the junior state/provincial level, age when recognised as one of the top five players on their team at the junior state/provincial level, and age when recognised as the best player for their position on their team at the junior state/provincial level, all of which were classified as having moderate test-retest reliability.

Combined, these results suggest that athletes could consistently recall the ages at which they first participated in each of the various levels of competition for their main sport, and could consistently judge and recall their level of success within higher, more recent stages of competition, but recollection of success during the earlier years of involvement in main sport was more difficult. Given the generally strong test-retest reliability of the sporting milestones section, it is recommended all items be retained in the DHAQ; however, caution should be exercised when interpreting results pertaining to success at the junior local and junior state/provincial levels of competition, as this information appears to be recalled less consistently than other milestones.

Representative History

Test-retest reliability results for the representative history section of the DHAQ are provided in Table 18. Athletes were highly consistent in identifying the ages at which they first participated in competition at the local, state/provincial, national, and international levels of competition, with test-retest reliability for age at first participation in both junior and open age group categories at each of these competitive levels classified as very good or good. Additionally, internal consistency with responses from the sporting milestones section of the DHAQ was very good for age at first participation in all competitive levels except the open local level, for which internal consistency was good (junior local: PA = 95.77, ICC = .98, $p <$ .01; open local: PA = 65.99, ICC = .20, p = .20; junior state: PA = 92.89, ICC = .95, $p < .01$; open state: PA = 87.49, ICC = .87, $p < .01$; junior national: PA = 98.75, ICC = 1.00, $p < .01$;

open national: PA = 92.89, ICC = .92, $p < .01$; junior international: PA = 95.00, ICC = .90, $p < .01$; open international: PA = 97.63, ICC = .95, $p < .01$).

Moreover, on an age-by-age basis, test-retest reliability for whether or not athletes participated in the various levels of competition was very good or good across all ages, for both junior and open age group categories, at all representative levels. Note that test-retest reliability for whether athletes participated in the various levels of competition was assessed across all ages 5-27 for both junior and open age group categories. Since the upper age limit for junior competition varies from sport to sport, it was decided not to impose an upper age limit for analyses of involvement in junior competition; and similarly, since the age at which athletes first participate in open competition varies, it was decided not to impose a lower age limit for analysis of involvement in open competition. Since test-retest reliability for whether athletes participated in junior or open competition at a particular age was assessed on the basis of yes, the athlete indicated participation, or no, they did not, it is not problematic to analyse and present results for this information across all ages. Rather, the high and often perfect agreement at ages 24-27 for whether athletes participated in junior competition across the various representative levels reflects that in the majority of instances, athletes were indicating that no, they did not participate in junior competition at those ages; and vice versa, the high and often perfect agreement at ages 5-11 for whether athletes participated in senior competition across the various representative levels reflects that in the majority of instances, athletes were indicating that no, they did not participate in open competition at those ages.

Although athletes were able to consistently identify whether they participated in each level of competition at all ages, recall of the age group categories in which they competed was generally poor. The total number of age group categories in which athletes reported competing was recalled with very good or good reliability across almost all ages, at all representative levels; however, test-retest reliability for identification of the specific age group categories in which athletes competed was moderate or poor for the majority of ages at each of the four representative levels. These results indicate that during each year of involvement in their main sport, athletes were able to remember whether they competed at a particular representative level and how many different age group categories they participated in, but could not consistently identify whether, for example, they had competed in the 12/13 years age group category, or the under 14 years age group category. Information regarding the age group categories in which athletes competed was collected to identify when they were competing up in a higher age group category than expected for their age. Engaging in practice and competition activities with older participants has been suggested to be influential on sport expertise development (Côté et al., 2006; Phillips et al. 2010a), so it was thought that the DHAQ could be a tool to investigate this

association in greater detail. Unfortunately, since recall of age group categories was not reliable, it is recommended that these items be removed from the questionnaire, and alternative methods be sought to explore the role that older teammates and/or competitors may play in the development of sport expertise.

While details of age group categories were not recalled consistently, identification of the representative levels at which athletes competed proved to be highly reliable. Therefore, it is recommended that the representative history section of the DHAQ be restructured and simplified. A basic grid requiring athletes to indicate all representative levels at which they competed during each year of their involvement in their main sport would facilitate examination of progression through the various levels of competition, providing important information regarding athletes‘ pathways toward expertise.

Practice History

Test-retest reliability was reasonably good for recall of details of involvement in practice activities. Results are discussed according to practice type, with accompanying PA and ICC statistics provided in Table 19. Note that reliability for measures evaluated on an age-by-age basis was only assessed for ages 5-27 as there were insufficient participants aged 28 or above to allow accurate interpretation of results.

Sport specific physical practice. In addition to examination of reliability of athletes‘ responses for average number of hours per week and total number of months per year engaged in sport specific physical practice during each year of involvement in the four sport specific physical practice conditions (in a group under the direct supervision of a coach, one-on-one with a coach, with others but without a coach, and on their own), several additional measures were calculated and assessed for reliability. These measures included age at first participation, total number of hours per year, and total cumulative hours of involvement across participants‘ careers to date. Moreover, details of participation in each of the four conditions of sport specific physical practice were combined to provide an indication of age at first involvement, total number of hours per year, and total cumulative hours of participation in sport specific physical practice overall.

Note that due to the large number of practice conditions addressed within this section of the questionnaire, Table 19 displays validity and reliability results for age at first participation, total hours per year, and total cumulative hours for each practice type/condition; however, PA values and ICC statistics are not presented for average number of hours per week or total number of months per year.

Sport specific physical practice completed in a group under the direct supervision of a coach. Test-retest reliability was very good for recall of age at first participation in sport specific physical practice completed in a group under the direct supervision of a coach. Recall of average number of hours per week and total number of months per year for this type of practice was very good or good across all ages, and reliability of total number of hours per year was also very good or good for all ages except 12, for which test-retest reliability was moderate. Reliability of total cumulative hours of participation was good.

Sport specific physical practice completed one-on-one with a coach. Test-retest reliability for age at first participation in sport specific physical practice completed one-on-one with a coach was very good, and reliability of recall was very good or good across all ages for hours per week, months per year, and hours per year. Test-retest reliability for total cumulative hours of participation in this type of practice was good.

Sport specific physical practice completed with others but without a coach. Age at first participation in sport specific physical practice completed with others but without a coach, as well as average number of hours per week, total number of months per year, and total number of hours per year across all ages, were all rated as having very good or good test-retest reliability. However, test-retest reliability for total cumulative hours of participation was poor. The main factor contributing to these mixed results is that hours of participation in sport specific physical practice completed with others but without a coach were relatively low compared to other sport specific physical practice conditions. As a result, the magnitude of the discrepancies between test occasions for hours of participation at a particular age tended to quite small, leading to a high overall PA for that age. However, when hours were accumulated to produce total hours of participation in this activity overall, the magnitude of both the values and discrepancies increased, producing a poor reliability rating for this measure. This same trend was observed on several occasions throughout various sections of the questionnaire, particularly when hours of participation were relatively low.

Individual sport specific physical practice. All measures pertaining to participation in individual sport specific physical practice displayed very good or good test-retest reliability.

Sport specific physical practice overall. Test-retest reliability for age at first participation in all conditions of sport specific physical practice combined was very good. Reliability of recall for total hours per year was very good or good for all ages except 12, which was moderate. Total cumulative hours of participation in sport specific physical practice overall was rated as having good reliability.

Physical preparation activities. Reliability assessments for involvement in physical preparation activities were the same as those conducted for participation in sport specific physical practice.

Physical preparation activities completed in a group under the direct supervision of a coach/specialised instructor. Test-retest reliability for age at first participation in physical preparation activities completed in a group under the direct supervision of a coach/specialised instructor was good. Details pertaining to hours per week, months per year, and hours per year of involvement in this practice type were recalled with very good or good reliability across all ages except 26 for which all measures were poor. Reliability for total cumulative hours of participation was also poor.

Physical preparation activities completed one-on-one with a coach/specialised instructor. All measures of reliability for participation in physical preparation activities completed one-on-one with a coach/specialised instructor were classified as very good. It is important to note, however, that only three athletes reported participating in this type of practice.

Physical preparation activities completed with others but without a coach/specialised instructor. Age at first participation in physical preparation activities completed with others but without a coach/specialised instructor was recalled with very good reliability. Reliability of recall for hours per week, months per year, and hours per week was very good or good across all ages except 24 and 26 for which several measures were rated as moderate or poor. Despite this, total cumulative hours of participation was still considered to have good test-retest reliability.

Individual physical preparation activities. Reliability of recall for age at first participation in individual physical preparation activities was moderate. Hours per week, months per year, and hours per year of involvement in this type of practice were recalled with very good or good reliability for most ages, although recall for hours per year at age 16, and hours per week and hours per year at age 26 was poor. Total cumulative hours of involvement in individual physical preparation activities was classified as having only moderate test-retest reliability.

Physical preparation activities overall. Combining all conditions of physical preparation activities together, reliability of recall for age at first participation was very good. Internal consistency between this item and a similar item within the sporting milestones section of the DHAQ was also very good (PA = 83.16, ICC = .36, p = .07). Reliability of recall for total hours per year was very good for most ages, with the exception of ages 16 and 26, for which reliability was poor. Total cumulative hours of participation in physical preparation activities overall was also rated as having poor test-retest reliability.

Mental preparation activities. As for sport specific physical practice and physical preparation activities, reliability was assessed for age at first participation, average number of hours per week, total number of months per year, total number of hours per year, and total cumulative hours of involvement in mental preparation activities under four conditions (in a group under the direct supervision of a coach/specialised instructor, one-on-one with a coach/specialised instructor, with others but without a coach, and on their own). Age at first participation, total hours per year, and total cumulative hours overall were also assessed for reliability for all conditions of mental preparation combined.

Mental preparation activities completed in a group under the direct supervision of a coach/specialised instructor. Age at first participation in mental preparation activities completed in a group under the direct supervision of a coach/specialised instructor was recalled with good reliability, and hours per week, months per year, and hours per year were recalled with very good or good reliability across all ages. Despite these positive results, test-retest reliability for total cumulative hours of participation in this type of practice was poor.

Mental preparation activities completed one-one-one with a coach/specialised instructor. Test-retest reliability was very good for age at first participation in mental preparation activities completed one-on-one with a coach/specialised instructor, as well as for hours per week, months per year, and hours per year across all ages. Reliability for total cumulative hours of involvement was good. It is noteworthy, however, that only 6 participants reported engaging in this type of practice.

Mental preparation activities completed with others but without a coach/specialised instructor. Similarly, only six athletes reported participating in mental preparation activities completed with others but without a coach/specialised instructor. Nevertheless, age at first participation was recalled with very good reliability, and reliability for hours per week, months per year, and hours per year was very good or good across all ages. Test-retest reliability for total cumulative hours of involvement in this type of practice was good.

Individual mental preparation activities. Although the majority of participants reported participating in individual mental preparation activities, reliability of recall for age at first participation was poor. Recall of hours per week, months per year, and hours per year was very good or good for most ages; however, a number of measures at ages 21-26 had poor reliability. Test-retest reliability for total cumulative hours of participation was also poor.

Mental preparation activities overall. Test-retest reliability for age at first participation and total cumulative hours of involvement in mental preparation activities overall was poor. Reliability of recall for hours per year was also poor for ages 20-22, and 24-26, but was very good or good across all other ages.

Sport specific play. The same measures of test-retest reliability were assessed for involvement in sport specific play as for the previous types of practice; however, only two conditions of sport specific play were addressed – with others, and on their own.

Sport specific play with others. Age at first participation in sport specific play with others was recalled with good reliability. Hours per week, months per year, and hours per year were recalled with very good or good reliability across all ages except 24, for which reliability of all measures was poor. Test-retest reliability for total cumulative hours of participation in this type of practice was also poor.

Individual sport specific play. All measures of reliability for individual sport specific play were very good; however, only four athletes reported participating in this type of practice.

Sport specific play overall. Test-retest reliability for age at first participation in sport specific play overall was good. Reliability for total hours of participation in both conditions of sport specific play combined per year was very good or good for all ages except 24, which was poor. Reliability for total cumulative hours of participation in sport specific play overall was moderate.

Total hours of practice overall. When details of involvement in all types and conditions of practice are combined together, age at first participation in practice activities for main sport was recalled with very good reliability. Additionally, internal consistency was good between this item and a similar item within the sporting milestones section (PA = 79.24, ICC = .91; $p < .01$). Test-retest reliability for total number of hours engaged in practice activities per year of involvement in main sport was very good or good across all ages except 12, for which reliability was moderate. Reliability for total cumulative hours of participation in all types and conditions of practice combined was good.

Training camps. In addition to the assessment of test-retest reliability for total number of training camps attended throughout each year of involvement in main sport and the average duration of training camps attended each year, age at first participation in training camps, total number of weeks in training camps per year, and total cumulative weeks in training camps throughout participants‘ careers to date were also calculated and assessed for reliability. Age at first participation in training camps was recalled with good reliability, and number of training camps per year, average duration of each training camp per year, and total number of weeks in training camps per year were all recalled with very good or good reliability across all ages. Reliability for total cumulative number of weeks in training camps throughout participants‘ careers to date was, however, poor.

Similar to results obtained within sections relating to involvement in practice activities for main sport during phase two of this investigation, a number of items in this section of the

revised DHAQ were classified as having very good or good test-retest reliability because the PA value was rated as very good or good, despite the corresponding ICC being considered poor. As was the case previously, for these items the ICC was low, zero, negative, or non-significant because some athletes reported slightly higher values during the initial completion of the DHAQ compared to the re-test, while other athletes reported slightly lower values during the initial completion of the DHAQ compared to the re-test. As such, there was no correlation between responses, but this does not matter because the strong PA reflects that the magnitude of responses provided on both test occasions were reasonably similar.

In general, reliability of recall for details related to athletes‘ history of involvement in practice activities was good. Recall of age at first participation was good or very good for all practice types, with the exception of individual unsupervised physical and mental preparation activities, for which reliability was rated as moderate and poor respectively. These results support earlier findings from the validation of the original DHAQ indicating that recall of participation in unsupervised practice activities appears to be less consistent than recall of participation in supervised practice activities.

On an age-by-age basis, details of involvement in practice activities were recalled well for almost all ages, across all practice types. Several minor exceptions were evident where reliability of recall for average number of hours per week, total number of months per year, and/or total number of hours per year was rated as moderate or poor. Interestingly, the majority of these instances occurred after the age of 20. It might be expected that recall at more recent ages would be more consistent than at younger ages, however, this was not the case. It is likely that a steady reduction in the number of participants with each age above 15 influenced the reliability results at these older ages because the reduced sample size was typically associated with increased variability, leading to lower PA values and non-significant ICC statistics.

Although reliability of recall for details of participation in practice activities throughout each year of athletes‘ involvement in their main sport was generally good, test-retest reliability for total cumulative hours of participation was moderate or poor for a number of practice types. This finding was unexpected because during the initial validation of the DHAQ, validity and reliability of total cumulative hours of involvement in several practice activities was found to be good even though validity and reliability of its component measures (i.e. hours per week, months per year, and hours per year) were recalled poorly. Here the reverse trend was observed, as reliability of recall for hours per week, months per year, and hours per year was stronger than reliability for total cumulative hours of involvement for most practice types. While it is likely that the previously highlighted lower reliability of recall for hours per week, months per year, and hours per year above age 20 could be negatively influencing reliability of total cumulative

hours of involvement, these results highlight the complex nature of practice history information, and the necessity to validate all items within developmental history measurement instruments in order to be confident in the trustworthiness of data reported.

Despite several moderate and poor test-retest reliability classifications within the practice history section of the updated DHAQ, reliability improved noticeably compared to the equivalent sections of the original questionnaire. This is particularly true for details of involvement in sport specific physical practice activities, and sport specific play. Therefore, the modifications outlined appear to have markedly strengthened the measurement tool. It is acknowledged that recall of autobiographical information over an extended period of time is difficult, so forgetting and confusion of details will occur (Bradburn, Rips, & Shevell, 1987). When memories are uncertain or fragmented, individuals tend to rely on inferences and approximations to fill in the gaps (Bradburn et al., 1987). Such approximations are likely to result in slightly inconsistent responses when recall is repeated on multiple occasions, particularly when the information required is as precise as practice hours. Therefore, it is not expected that further modifications to this section of the DHAQ would improve reliability. As such, it is suggested that the practice history component of the questionnaire is now suitable for use within future investigations related to sport expertise development.

Competition History

Items related to competition history within the original DHAQ were not adequately structured to sufficiently cater to the complexities of involvement in competition. As such, major modifications were made to address involvement in three different competition formats (regular season, occasional, and tournament format competitions), and to differentiate between time actively involved in competition activities and time spent at the competition venue. It was expected that a more detailed approach to the collection of competition history information would facilitate more accurate memory recall and hence improve reliability of this section. As detailed below, the modifications did slightly improve test-retest reliability, however, a number of measures still had moderate to poor reliability classifications. Results are discussed according to competition format, with PA values, ICC statistics, and classification information provided in Table 20. Note that, as for practice history, test-retest reliability of age-by-age measures was only assessed for ages 5-27 due to insufficient participants aged 28 or above.

Regular season competitions. In addition to assessing reliability for total number of months per year involved in regular season competitions, average number of regularly occurring events/matches per month, average time spent at the competition venue per event/match, and

average time spent actively competing per event/match for each year of participants' involvement in regular season competition, several other measures were subjected to test-retest reliability analyses. These included total number of regular season events/matches, total time spent at the competition venue, and total time spent actively competing in regular season competitions both per year, and cumulative across participants' careers to date, as well as age at first participation in regular season competitions.

Given the large number of items assessed for reliability within the competition history section of the questionnaire, Table 20 only displays PA values and ICC statistics for age at first participation in regular season competitions, total number of regular season events/matches per year and overall, total time spent at the competition venue per year and overall, and total time actively engaged in competition per year and overall. However, classification ratings for all items are discussed below.

Age at first participation in regular season competitions was recalled with good reliability, and the total number of months per year during which athletes participated in regular season competitions was recalled with very good or good reliability across all ages. The average number of events/matches in which participants competed during each month of their involvement in regular season competitions was recalled with very good or good reliability for the majority of ages, however, reliability was poor for ages 13-16, and 26. As a result, test-retest reliability for total number of regularly occurring events/matches in which participants competed per year was also poor at these same ages, as well as age 17.

With regards to test-retest reliability for time involved in regular season competitions, recall for average time spent at the competition venue per event/match was very good or good for all ages except 15 and 19, for which reliability was poor, but reliability for total time spent at the competition venue per year was moderate or poor at ages 12-17, 19, and 26. Similarly, reliability of recall for average time spent actively competing per event/match was poor for ages 14-17, but was very good or good across all other ages; however, test-retest reliability for total time spent actively competing per year was poor for ages 12-17, 21, 23, 24, and 27.

Combining details of participation in regular season competitions across all ages of involvement, test-retest reliability was poor for all cumulative measures including total cumulative number of events/matches in which participants competed, total cumulative time spent at the competition venue, and total cumulative time spent actively competing in regular season events/matches.

Occasional competitions. Measures related to participation in occasional competitions were similar to those described above for regular season competitions. Test-retest reliability was assessed for all responses, including total number of occasional competitions, average number

of events/matches per occasional competition, average time spent at the competition venue per occasional competition, and average time spent actively competing per occasional competition, for each year of involvement in main sport, as well as for several calculated measures, including age at first participation in occasional competitions, and total number of events/matches, total time spent at the competition venue, and total time spent actively competing per year and cumulative across participants‘ careers to date. As for regular season competitions, Table 20 displays PA values and ICC statistics for age at first participation, total number of events/matches per year and overall, total time spent at the competition venue per year and overall, and total time actively engaged in competition per year and overall; although, classification ratings for all responses are addressed below.

Test-retest reliability for age at first participation in occasional competitions was poor, but reliability was very good or good for all age-by-age measures. Despite strong reliability ratings across all ages, test-retest reliability for total cumulative occasional competitions, total cumulative events/matches, total cumulative time spent at the competition venue, and total cumulative time spent actively competing in occasional competitions overall was poor.

Tournament format competitions. Test-retest reliability was assessed for total number of tournament format competitions, average number of events/matches per tournament, average time spent at the competition venue per event/match, and average time spent actively competing per event/match across each year of participants‘ involvement in their main sport, as well as for total number of events/matches, total time spent at the competition venue, and total time spent actively competing per year and cumulative across participants‘ careers to date. Test-retest reliability for age at first participation in tournament format competitions was also assessed. As above, classification information will be discussed for all items, but Table 20 presents only PA values and ICC statistics for age at first participation in tournament format competitions, total number of tournament format events/matches per year and overall, total time spent at the competition venue per year and overall, and total time actively engaged in competition per year and overall.

Age at first participation in tournament format competitions was recalled with good reliability. The number of tournament format competitions athletes participated in per year, the average number of events/matches per tournament, and subsequently, the total number of tournament format events/matches per year, were all rated as having very good or good reliability across all ages. Additionally, reliability of recall for average time spent at the competition venue per event/match was very good or good across all ages, leading to very good or good reliability ratings for total time spent at the competition venue per year for all ages. Similarly, reliability for average time spent actively competing per event/match, and total time

actively competing per year was very good or good across most ages, although, at age 16 these reliability measures were poor and moderate respectively. Overall, reliability for total cumulative tournaments in which athletes participated, and total cumulative tournament format events/matches was good, but reliability for total cumulative time spent at the competition venue and total cumulative time spent actively competing in tournament format competition was poor.

Competition overall. Combining participation in all competition formats together, test-retest reliability for age at first participation in competition overall was very good. Internal consistency for this measure was also very good when comparing responses between this section of the questionnaire and both the sporting milestones (PA = 90.73, ICC = .95, $p < .01$) and representative history sections (PA = 90.62, ICC = .94, $p < .01$). Test-retest reliability for total number of competitive events/matches in which athletes participated per year was very good or good across all ages except 17, for which reliability was moderate, but reliability for total cumulative number of competitive events/matches across athletes‘ careers to date was poor.

Reliability for total number of hours spent at the competition venue per year was mixed, with very good or good reliability ratings for ages 5-11, 19-24, and 27, but poor reliability ratings for ages 12-18, and 25-26. Similarly, test-retest reliability for total number of hours actively engaged in competition was very good or good for ages 5-11, 19, 22-23, and 17, but was poor for all other ages. Subsequently, reliability for total cumulative hours spent at the competition venue overall and total cumulative hours actively engaged in competition overall was poor. PA values and ICC statistics are provided for all measures pertaining to participation in all formats of competition combined, in Table 20.

Test-retest reliability assessments for the competition history section of the updated DHAQ produced mixed results. The modified structure of items relating to details of participation in competition allowed athletes to discuss their competition involvement more successfully than the previous version of the DHAQ; however, there were still several issues surrounding the reliability of this section.

The most notable observation was that recall of involvement in regular season competitions appeared to be particularly problematic during the teenage years. Athletes were unable to consistently recall the average number of regularly occurring events/matches per month of their involvement in these types of competitions, leading to inconsistent calculations of hours of involvement in regular season competitions per year during adolescence. Poor recall of time involved in regular season competitions per year of participation undoubtedly contributed to the poor reliability ratings for total hours of involvement in competition activities

overall. Responses within the representative history section of the DHAQ revealed that during the teenage years, athletes commonly competed simultaneously in several different age group categories, at a range of representative levels. This complex competition schedule is likely to have contributed to the poor recall during these years, as athletes may have experienced difficulty remembering how many events/matches they participated in during a typical month. It is also a possibility that the number of regularly occurring events/matches varied significantly from one month to the next, making it difficult to provide an accurate indication of the average number of events/matches per month.

It is interesting that reliability ratings for time involved in regular season competitions, and subsequently time involved in competition overall, were also poor for several ages above 20 years old. It is not clear why this is the case because responses from the representative history section of the questionnaire indicated that by this age athletes were typically competing in just the open age group category at each of the representative levels, rather than multiple age group categories like during the teenage years; plus, a recency effect (Bjork & Whitten, 1974; Whitten & Leonard, 1981) was expected, such that details of participation in practice and competition activities during recent years were expected to be recalled with greater consistency than at younger ages. As proposed within the discussion of practice history results, the poor reliability for total hours of involvement in regular season competitions per year at these ages may be an artefact of the lower number of participants included in the reliability analyses, since athletes ranged in age from 15-35.

Involvement in occasional and tournament format competitions was recalled more consistently than involvement in regular season competitions. It is possible that these one-off events are more memorable than regularly occurring events because they could be of greater significance to the athlete. For example, local club level competitions frequently occur on a weekly or monthly basis, whereas national level competitions often only occur once or twice a year. It is understandable that athletes could more accurately recall their involvement in major events such as a national championship compared to more minor events such as a local league.

Interestingly, even though reliability ratings were very good or good for most age-by-age measures of involvement in occasional and tournament format competitions, the majority of measures pertaining to total cumulative involvement in these competition activities displayed poor test-retest reliability. This echoes previous findings from the practice history section of the questionnaire, where the same trend was observed. These findings suggest that when collecting practice and competition history information via the DHAQ, greater attention should be devoted to the discussion and interpretation of age-by-age results than total cumulative values. Positively, the majority of previous investigations involving examination of past practice and/or

competition hours have reported a combination of age-by-age and cumulative values (Baker et al., 2003b; Helsen et al., 1998; Hodge & Deakin, 1998; Hodges et al., 2004; Hodges & Starkes, 1996; Starkes et al., 1996; Ward et al., 2007); thus, it is recommended that future investigations continue to report results in this manner, however, the lower reliability of cumulative values should be noted.

Although participation in occasional and tournament format competitions was recalled with acceptable reliability, regular season competitions form a major component of the competitive timetable for many sports, so the poor reliability of this section is a concern. Therefore, further modifications are required before the competition history section of the DHAQ can be considered suitable for retention. As competition formats and scheduling can vary considerably between sports, it is possible that reliability of recall for competition history information may improve if items are catered more specifically to the competitive structure of the participants' sport. Since the DHAQ was designed to be a generic measurement tool suitable for completion by participants of any sport, separate sport-specific questionnaires may need to be developed in order to accurately address involvement in competition activities. Given the identification of competition as an important contributor to the development of sport expertise (Baker et al., 2003a; Ford et al., 2009; Singer & Janelle, 1999; Ward et al., 2007), continued efforts to construct a validated measurement tool for the collection of competition history information are essential. In the meantime, while these efforts are ongoing, it is recommended that the current competition history section of the DHAQ be removed, at least temporarily.

Participation in Other Organised Sports

Although the participation in other organised sports section of the updated DHAQ remained unchanged from the original version, test-retest reliability during this study was less than ideal. Table 21 displays detailed results of the test-retest reliability analyses for all items within this section of the questionnaire.

Reliability of recall for total number of organised sports other than the main sport in which athletes participated was rated as good, but as was the case in the initial validity and reliability study, only seven of the 20 athletes reported the same list of sports on both test occasions. Other athletes tended to identify between one and three sports on both test occasions, plus between one and six additional sports on just one of the two test occasions. Two athletes reported completely different lists of sports each time. This result once again highlights a potentially significant limitation associated with investigations of diversification and sport expertise, and the importance of in-depth examinations of reliability of recall during studies of

sport expertise development. Studies reporting total number of sports and measures of participation in all organised sports combined that do not discuss reliability of sports identified may be misleading, and should be interpreted cautiously.

With the limitations associated with inconsistent reporting of organised sports in mind, test-retest reliability of recall for age at first participation in all other organised sports combined was good, as was age at specialisation. Age at specialisation was calculated as the age at which participation in all other organised sports ceased, and the athlete began to concentrate on their main sport alone. Internal consistency between age at specialisation as calculated within this section of the DHAQ, and that reported within the sporting milestones section was also good (PA = 78.45, ICC = .03, p = .05).

Still combining participation in all organised sports together, test-retest reliability for total number of hours of involvement in other organised sports per year was good or very good for ages 6 and 19 to 27, but was moderate or poor for all other ages. The mean age at first participation in other organised sports was approximately seven years old, and by age 19 most athletes had ceased participation in other organised sports to specialise in their main sport. Therefore, hours of involvement in other organised sports were typically zero outside of these ages, leading to good reliability ratings. On the other hand, between the ages of 7 and 18 athletes were commonly participating in a variety of sports. Given the inconsistent reporting of sport types already discussed, it is not surprising that test-retest reliability for total hours of involvement in other organised sports across these ages was moderate to poor. It is consequently not surprising that reliability for total cumulative hours of participation in all other organised sports combined throughout participants‘ careers to date was also only moderate. Note that as for the practice and competition history sections, test-retest reliability for age-by-age measures was only assessed up to age 27 as insufficient participants aged 28 or older were involved to draw accurate conclusions from the data at ages above 27. Finally, test-retest reliability for highest level of competition reached in other organised sports overall was moderate.

Despite inconsistent reporting of the total number of organised sports in which athletes participated, an average of two sports were identified consistently on both test occasions, therefore, test-retest reliability was able to be assessed on a sport-by-sport basis for two sports.

For other organised sport one, even though sport type was identified consistently, reliability of recall for age at first participation in this sport was poor. Age at cessation of participation in this sport was, however, good. Test-retest reliability for average number of hours per week, total number of months per year, and the calculated value for total hours per year of involvement in other organised sport one demonstrated a similar trend to that described above, whereby reliability was moderate or poor between the ages of 8 and 17 when

participation was most common, but was very good or good outside of these ages (N.B. In the interest of space, test-retest reliability results for average number of hours per week and total number of months per year for other organised sport one are not included in Table 21, however, PA values and ICC statistics are provided for all other items). Reliability for total cumulative hours of participation in other organised sport one was also only moderate. With respect to highest level of competition reached for organised sport one, test-retest reliability was very good or good at the majority of ages, but was moderate at ages 9-12 and 16. Reliability was also moderate for highest level of competition reached for other organised sport one overall.

For other organised sport two, test-retest reliability for age at first participation and age of cessation of participation in this sport were both poor, but surprisingly, reliability for hours per week, months per year, and hours per year were very good or good across all ages except age 12, for which measures were recalled with poor reliability. Overall, however, test-retest reliability for total cumulative hours of participation in other organised sport two was poor. Similarly, test-rest reliability for highest level of competition was very good or good across all ages, but was only moderate for highest level of competition reached overall. Note that as for other organised sport one, reliability statistics are not provided in Table 21 for average number of hours per week and total number of months per year for involvement in other organised sport two, but all other results pertaining to participation in other organised sport two are presented. Since only two sports were identified consistently, reliability analyses were not conducted on any additional sports.

The relatively poor results for this section were similar to those observed during the validation of the original DHAQ; however, they were still surprising given the plethora of research investigating the associations between participation in a variety of organised sports and the development of sport expertise (Baker et al., 2003b; Baker et al., 2005; Berry et al., 2008; Ford & Williams, 2008; Ward et al., 2007; Memmert et al., 2010; Oldenziel et al. 2003; Soberlak & Côté, 2003). The structure adopted for this section of the DHAQ was very similar to that utilised in earlier investigations of diversification and sport expertise development, so it is unlikely that the low reliability is a result of sub-optimal item structure. Rather, limitations of memory recall are most likely responsible, and hence unavoidable in retrospective investigations such as this. Therefore, researchers should monitor reliability of recall for details of participation in organised sports closely, and should take care when interpreting related results and drawing conclusions.

Furthermore, the results of the current study highlight the need for researchers to re-visit the role of diversification in sport expertise development and consider setting criteria for reporting involvement in other organised sports. For example, perhaps only sports in which

participants were involved for two or more seasons should be included; or possibly only sports in which participation exceeded four hours per week during any given year of involvement. These examples are arbitrary, though discussions among researchers in the field are essential to define what constitutes diversification, so that more appropriate measurement instruments can be created to facilitate accurate recall of participation in other organised sports, and hence provide meaningful recommendations to parents, coaches, and sport administrators. To determine such criteria, experimental approaches exploring the nature and extent of skill transfer may be required.

In the meantime, given the importance placed on the relationship between diversification and sport expertise in previous literature (Baker et al., 2003b; Côté & Fraser-Thomas, 2007; Berry et al., 2008), it is recommended that this section of the DHAQ be retained in the questionnaire. However, it is important to be aware of the lower reliability of this section in its current format, and to acknowledge these reliability issues when conducting investigations utilising the DHAQ.

Participation in Informal Playful Sporting Games

Recall of information pertaining to participation in informal playful sporting games was very poor during the initial validity and reliability assessments of the DHAQ. In an attempt to improve recall, major modifications were made to this section when updating the questionnaire. Unfortunately, these modifications did not elicit more consistent responses from participants as test-retest reliability was poor for all items within the informal playful sporting games section.

Measures for which test-retest reliability was poor included total number of playful sporting games in which athletes reported participating, age at first participation in informal playful sporting games, and age at cessation of participation in all playful sporting games. The average number of hours per week and total number of months per year of participation in informal playful sporting games were recalled with poor reliability across almost all ages, leading to poor test-retest reliability for total number of hours per year across all ages as well. Test-retest reliability for total cumulative hours of participation in informal playful sporting games was also poor.

Only six athletes reported the same sporting game on both test occasions for the game they played most often, and only three athletes reported the same sporting game on both test occasions for the game they played second most frequently. The typical percentages of play time devoted to the most frequent and second most frequent sporting games played were also recalled with poor reliability. In fact, no athletes reported the same list of informal playful

sporting games in which they participated on both test occasions, which could have contributed to the inconsistent recall of ages of involvement in informal sporting play, hours of involvement, and identification of the two most frequent sporting games played.

PA values, ICC statistics, and classification information for items within the informal playful sporting games section of the DHAQ are provided in Table 22; however, note that as for other sections of the questionnaire, reliability results for average number of hours per week and total number of months per year are not included.

It is very interesting that with the exception of Memmert et al. (2010), and Soberlak and Côté (2003), previous investigations of associations between participation in sporting play and the development of sport expertise have failed to report reliability of recall for this information. Memmert and colleagues (2010) reported a significant Pearson correlation coefficient of .70 for test-retest reliability of time spent in play activities, and Soberlak and Côté (2003) reported a 30% difference between athletes‘ and parents‘ accounts of participation in sporting play. Authors of both papers considered reliability of recall for hours of involvement in sporting play to be acceptable, however, in each case, reliability of recall for hours of involvement in organised sporting activities was greater than for involvement in play. As discussed earlier, the unstructured, irregular nature of participation in informal sporting games is likely to be the most significant contributor to the lower reliability of recall for this information. As participation in sporting play is typically spontaneous and inconsistent, details of involvement are more difficult to recall compared to more formal, regularly scheduled organised sporting activities.

These poor results suggest that recall of involvement in informal, playful sporting games may not be strong enough to allow accurate interpretation of data obtained from this section of the questionnaire. Since modifications to the structure of this section of the DHAQ did not improve reliability of recall for details pertaining to participation in informal, playful sporting games, it is recommended that it be removed from the questionnaire. It is therefore important that if associations between participation in informal, playful sporting games and sport expertise continue to be explored in the future, researchers must carefully construct their measurement tools and conduct detailed assessments of reliability of recall to ensure trustworthy reporting of results.

Family

Within this section of the DHAQ, all participants provided demographic information and details of participation in sport and physical activity for their biological mother and biological father, however, no participants reported having lived with a step-mother, step-father,

or any other male or female legal guardian. Therefore, parental reliability assessments were restricted to biological parents. Similarly, although the majority of participants reported having one or two siblings, only four participants reported having three or more siblings. As such, sibling reliability assessments were restricted to participants‘ two oldest siblings only. Test-retest reliability results are discussed according to family member, with accompanying PA, ICC, and classification information provided in Table 23.

Biological mother. Date of birth, main country of residence, highest level of education and ages at which participants lived with their biological mother were each recalled with very good or good reliability. Test-retest reliability was poor for mother‘s frequency of participation in general fitness activities, moderate for mother‘s frequency of participation in recreational sport, and good for mother‘s frequency of participation in competitive sport during the time living together.

Although the total number of competitive sports in which athletes‘ mothers participated throughout their lives was recalled with good test-retest reliability, the actual sports identified were not recalled consistently, with reliability for sport type only moderate to poor for each of the sports provided. Positively, identification of whether mothers had participated in the same sport as the athlete‘s main sport was rated as having good test-retest reliability, and reliability for highest level of competition reached for this sport was also good. Unfortunately, reliability of recall for highest level of competition reached in other sports was poor, leading to a poor reliability classification for highest level of competition reached overall.

Biological father. Reliability results for recall of details pertaining to participants‘ biological fathers were very similar to those for mothers. Test-retest reliability was very good or good for recall of father‘s date of birth, country of residence, and highest level of education, as well as for the ages at which athletes lived with their father. Reliability was moderate for fathers‘ frequency of participation in both general fitness activities and recreational sport during the time living with the athlete, but was very good for fathers‘ frequency of participation in competitive sport.

Regarding athletes‘ fathers participation in competitive sport throughout any time in their lives, reliability of recall was generally poor. Reliability of recall for the total number of sports in which fathers participated, the particular sports they played, and the highest level of competition they reached, both on a sport-by-sport basis and overall, was moderate or poor, but recall of whether fathers participated in the same sport as the athlete‘s main sport and the highest level of competition reached in this sport was very good and good respectively.

Sibling one. Reliability of recall for information relating to participants‘ oldest sibling was mixed. Athletes were very consistent in recalling this sibling‘s sex, relationship to the

athlete (i.e. full sibling, half-sibling, step-sibling, etc.), date of birth, main country of residence, highest level of education, and the ages at which they lived together, with very good or good reliability ratings for all these items. Reliability of recall for this sibling's participation in sport and physical activity during the time living with the athlete was not as strong, with reliability for frequency of participation in general fitness activities only moderate, and reliability for frequency of participation in both recreational and competitive sports poor.

Test-retest reliability for total number of sports in which this sibling participated was also poor. However, it appears as though this sibling may have participated in one main sport because reliability for sport type and highest level of competition was very good and good respectively for one sport, but poor for all other sports. Reliability for whether this sibling participated in the same sport as the athlete's main sport was very good, but reliability for the highest level of competition reached in this sport was only moderate. Despite this, reliability for highest level of competition reached overall was classified as good.

Sibling two. Reliability results for demographic information and details of participation in sport and physical activity for athletes' second oldest sibling were similar to those for sibling one. All demographic items displayed very good test-retest reliability including sex, relationship to the athlete, date of birth, main country of residence, highest level of education, and ages at which the athlete lived with this sibling. Reliability of recall for this sibling's frequency of participation in both general fitness activities and recreational sport was moderate, while reliability of recall for frequency of participation in competitive sport was good.

Similar to sibling one, test-retest reliability for the total number of sports in which this sibling participated was poor; however, once again, one sport was identified consistently on both test occasions. In this case though, reliability of recall for highest level of competition reached was poor for all sports identified. While test-retest reliability was very good for identification of whether this sibling participated in the same sport as the athlete's main sport, reliability for the highest level of competition reached in the athlete's main sport was poor. Overall, reliability of recall for the highest level of competition reached for all sports combined was also poor for sibling two.

Test-retest reliability results were relatively uniform across all family members. In general, demographic information was recalled with very good consistency, but recall of details relating to family members' involvement in sport and physical activity was not as strong. When indicating the frequency with which family members participated in sporting activities during the time the athlete lived with them, recall for frequency of participation in competitive sport was typically more consistent then recall for frequency of participation in general fitness activities and recreational sport. This finding could reflect issues already discussed relating to

recall of structured versus unstructured sporting activities. Family members‘ participation in competitive sport is likely to involve regularly occurring practice sessions and competitions, while participation in general fitness activities and recreational sport may be more informal and irregular. Therefore, it may be more difficult to accurately indicate the frequency with which family members participated in these unstructured sporting activities compared to competitive sport activities.

In responding to these items, athletes were required to select the frequency of each family member‘s participation in the three types of sport and physical activity according to a five point scale including the options: never, occasionally, 1-2 times per week, 3-5 times per week, or more than 5 times per week. Rather than remove these items from the questionnaire completely, it might be beneficial to simplify the question to a dichotomous response, asking whether family members participated in general fitness activities/recreational sport/competitive sport on a regular basis, that is, one or more times per week, most weeks of the year, throughout most of the years that you lived with them.

If the current results are collapsed into two categories indicating regular participation (response options: 1-2 times per week, 3-5 times per week, and 5 or more times per week), or irregular participation (response options: never and occasionally) in general fitness activities, recreational sport, and competitive sport, test-retest reliability improves such that reliability classifications become very good or good for all three activity types across all four family members (biological mother: PA general fitness activities = 75.00, PA recreational sport = 75.00, PA competitive sport = 90.00; biological father: PA general fitness activities = 75.00, PA recreational sport = 65.00, PA competitive sport = 80.00; sibling one: PA general fitness activities = 65.00, PA recreational sport = 70.00, PA competitive sport = 70.00; sibling two: PA general fitness activities = 73.33, PA recreational sport = 80.00, PA competitive sport = 66.67). While this approach elicits less detailed information, we can be more confident in the trustworthiness of the data, and hence draw more accurate conclusions.

Reliability of recall was also relatively poor for items relating to family members‘ participation in organised sport throughout any time in their lives. In general, athletes were not able to consistently identify the types and total number of sports in which family members had participated. Test-retest reliability for whether family members participated in the same sport as the athlete‘s main sport was good, but identification of any additional sports was highly inconsistent. Additionally, reliability for the highest level of competition family members reached in the various sports in which they participated ranged from poor to very good. Consequently, further modifications to this section of the DHAQ are required.

Within the current version of the DHAQ, athletes indicated whether family members had participated in any competitive sports any time in their lives for an extended duration (i.e., three years or more). If yes, athletes then indicated the name of each sport and the highest level of competition reached for each sport. It is recommended these items be replaced with a series of similar, yet simplified questions. Instead of listing all sports in which each family member participated, it is recommended that athletes respond yes or no to the questions –did this family member participate in the same sport as your main sport for an extended duration (i.e. three years or more) at any time in their lives?" and —id this family member participate in any other sports for an extended duration (i.e. three years or more) at any time in their lives?".

In an attempt to improve reliability of recall for highest level of competition reached by family members, it is recommended that the competitive level classification system be collapsed from 15 to three response options based upon the guidelines of the Gulbin et al. (2010) Athlete Development Triangle: elite (open international level competition); pre-elite (junior international or open national level competition); and non-elite (encompassing all other levels of competition). Further, it is recommended that only highest level of competition reached in the same sport as the athlete‘s main sport plus highest level of competition reached overall for all other sports combined be required for each family member (where applicable), rather than highest level of competition reached for each competitive sport in which the family member participated. It is expected that simplification of items relating to familial involvement in competitive sport will lead to improved reliability due to the reduced level of detail required.

When the recommended modifications are applied to the current data set and responses are reanalysed, test-retest reliability for whether family members participated in any organised sports other than the athlete‘s main sport was good or very good for all four family members (biological mother: PA = 75.00; biological father: PA = 80.00: sibling one: PA = 89.47; sibling two: PA = 73.33). Test-retest reliability classifications for highest level of competition reached in the same sport as the athlete‘s main sport did not change following the simplification of the competitive level response options (biological mother: PA = 66.67; biological father: PA = 66.67; sibling one: PA = 62.50; sibling two: PA = 42.86), but reliability for highest level of competition reached overall for all other sports combined did improve for fathers and both siblings (biological father: PA = 80.00; sibling one: PA = 86.67; sibling two: PA = 70.00). Reliability for highest level of competition reached overall for all other sports combined, however, remained poor for mothers (PA = 46.15).

Although reliability classifications for several items pertaining to familial involvement in sport and physical activity did not reach acceptable standards for all family members, following the recommended modifications to this section, the majority of proposed items were

rated as having very good or good test-retest reliability. Therefore, it is suggested that all demographic items within the family section of the questionnaire be retained within the DHAQ, and items pertaining to familial involvement in sport and physical activity be updated according to the proposed modifications. Reliability of updated items should continue to be monitored.

Places of Residence

Within this section of the DHAQ, participants identified all towns/cities in which they have lived, along with the ages at which they lived in each town/city, and the reason for relocation to each new town/city. The number of towns/cities in which athletes reported living displayed very good test-retest reliability, with sufficient information available to examine reliability of recall for specific details of the first four towns/cities of residence. PA, ICC, and reliability classification information for items within the places of residence section of the DHAQ are provided in Table 24.

As expected, all athletes reported the same location for their birthplace on both test occasions. For those athletes who moved away from the town/city in which they were born, reliability of recall for the location of the second town/city in which they lived was also very good. The age at which participants relocated to this second town/city was recalled with very good reliability, and identification of the main reason for relocation to this town/city also had good reliability.

Unfortunately, recall of the third town/city in which participants lived was not as strong, with test-retest reliability for the location of this residence rated as poor. Reliability of recall for age at relocation to this town/city was very good, as was reliability for the main reason for relocation to this town/city, indicating that athletes could remember when they moved from their second place of residence and why, but could not accurately recall where they moved to at this time. Closer inspection of the ages at which participants‘ lived in this third town/city indicated that the duration of residence was typically very short, in some cases less than a year. It is, therefore, possible that athletes were inconsistent in whether they reported residences of short duration on both test occasions.

Following inconsistent reporting of the third town/city of residence, details of the fourth town/city in which participants lived were also recalled poorly. Identification of the location of the fourth town/city of residence, age at relocation to this place of residence, and reason for relocation to this place of residence were all recalled with poor reliability. Of particular note, however, the number of times participants relocated to a new town/city for reasons relating to sport, and the age at which they first relocated to a new town/city for reasons relating to sport

were recalled with very good and good reliability respectively. Internal consistency for age at first relocation for reasons relating to sport was also good between this section of the questionnaire and the sporting milestones section (PA = 79.41, ICC = 0.33, p = .05).

Although reliability results for this section of the DHAQ are not entirely positive, the consistent recall of details pertaining to the first two places of residence in which athletes lived is encouraging, especially considering these places of residence tended to be of relatively long duration. This is because these first two places of residence typically represented the towns/cities in which athletes lived throughout their childhood and adolescent years, during which time they are likely to have obtained their early, and potentially formative, sporting experiences. It is also positive that athletes were able to consistently identify when they relocated for reasons relating to sport, as these relocations could represent key transition points within athletes‘ development, marking progression from one stage of their career to another. Furthermore, it would be interesting to examine the influence of these relocations on athletes‘ involvement in practice and competition activities, and subsequently, their performance. Given the strong reliability of recall for the most critical components of the places of residence section of the DHAQ, it is recommended that this section be retained in the questionnaire without the need for further modifications.

General Discussion

Following its initial construction and validation, a number of modifications to the DHAQ were recommended before the measurement tool was deemed suitable for use in investigations of sport expertise development. This study therefore aimed to implement the recommended changes and reassess test-retest reliability for the updated DHAQ. The questionnaire was also converted to an online format to allow its administration to large samples of athletes. The results indicate the modification and online conversion of the DHAQ were reasonably successful, and following several further adjustments, the questionnaire will be appropriate for continued use.

The majority of items within the updated online DHAQ received test-retest reliability ratings of very good or good, reaching acceptable standards for retainment in the questionnaire. In general, test-retest reliability for the sporting career, sporting milestones, practice history, and places of residence sections was strong, with no additional modifications necessary for any items. Minor modifications were suggested to address reliability issues within the representative history and family sections of the DHAQ, while the competition history and participation in informal, playful sporting games sections of the questionnaire were recommended to be

removed completely due to major reliability concerns. The participation in other organised sports section of the questionnaire failed to reach acceptable reliability standards, however, it was recommended for retention in the DHAQ without further modification due to its prominence in the literature as a highly influential factor on the development of sport expertise.

It is interesting that despite the removal of several sections from the original DHAQ, the approximate completion duration for the updated, online DHAQ appears to remain at roughly one hour. Given the improvements in test-retest reliability of the revised questionnaire, it does not seem as though participant fatigue is a limiting factor of the revised version of the questionnaire in the same manner that is was for the original DHAQ. This is likely related to the reduced repetition and improved clarity of questioning in the updated, online DHAQ, as well as the provision of the opportunity for participants to complete the questionnaire in their own time, over multiple sessions if desired. The time stamp information obtained from the online questionnaire host, SurveyMonkey, certainly indicates that the majority of participants chose to complete their responses over several days, presumably in short periods at a time. The ability for such a detailed questionnaire to be completed over multiple sessions rather than during a single sitting appears to be a highly beneficial feature of the updated, online DHAQ, and is a major consideration for future investigations of the development of sport expertise.

While test-retest reliability of the updated, online DHAQ improved considerably from its initial version, a number of recurring issues surrounding reliability of recall were apparent. First, as researchers, we often seek to advance our understanding of a topic area by exploring increasingly specific questions. The DHAQ, in particular, attempted to address a range of factors associated with the development of sport expertise in considerable detail. Results from the earlier validity and reliability investigation as well as the current study indicated that athletes could consistently recall general information regarding their involvement in sport and physical activity and other related contextual factors, but as items became more specific, reliability of recall decreased. For example, athletes were able to consistently identify the ages at which they participated in competition at each of the various representative levels, but they were unable to accurately recall the age group categories in which they competed. Similarly, while athletes were able to consistently identify whether their immediate family members regularly participated in general fitness activities, recreational sport, and/or competitive sport, they were unable to accurately estimate the frequency of their family‘s participation in these activities. Therefore, when collecting data via retrospective recall methods such as the DHAQ, at times a compromise must be made, sacrificing level of detail, and potentially interesting information, for trustworthy responses.

Degree of detail was also identified as a major factor influencing accuracy of recall in epidemiological studies, and was highlighted as an important consideration for study design and questionnaire development in retrospective investigations (Coughlin, 1990). It is likely that two main factors contribute to these limitations in specificity of memory recall. First, as explained by Williams et al. (2007) in a review of autobiographical memory specificity, several theories of memory retrieval suggest that memories are stored in a hierarchical fashion, with event representations organised from more general to more specific. Williams et al. (2007) go on to indicate that general event representations are assumed to be the "preferred or default level of access into the autobiographical memory knowledge base" (p. 131). Based on this hierarchical storage system, it appears that additional activation processes are required to access more specific memories. So, in the case of the DHAQ, it is possible that athletes were either experiencing difficulties accessing these more specific memories, and/or curtailed their search for more specific information in order to move on to the next item in the questionnaire more quickly. Alternatively, limitations in specificity of memory recall may be associated with the encoding and storage of information in the first place. Indeed, Tulving and Thomson (1973) emphasised that remembering involves an interaction of "information stored in the past and information present in the immediate cognitive environment of the rememberer" (p. 352). Moreover, event significance has been associated with accuracy of recall in previous research (Coughlin, 1990). Therefore, it is possible that at the time of occurrence, such specific details of sport involvement were not considered to be significant to the athlete, so the information may not have been encoded into memory in the first instance. As such, it is possible that athletes were simply guessing their responses to some of the more specific items within the questionnaire opposed to actually recalling the information from memory.

As a result of limitations relating to specificity of memory recall, several compromises were made within the DHAQ where items that could have provided novel contributions to our understanding of sport expertise development were simplified or removed from the questionnaire because athletes were unable to respond consistently. The specificity of information that can be reliably recalled is an important limitation to consider when conducting retrospective research, with the results from the present study highlighting the importance of confirming validity and reliability for all items within retrospective measurement instruments and not just several general representative measures.

A second recurring reliability concern was the discrepancy between reliability of recall for structured versus unstructured sporting activities. Although test-retest reliability for the practice history section of the questionnaire was generally very good and no further modifications were required, recall for details of participation in supervised practice activities

tended to be slightly stronger than recall for details of participation in practice activities that were not supervised by a coach. This was particularly true for identification of age at first participation in the various practice activities. Similarly, athletes were able to recall the frequency of family members‘ participation in organised sporting activities during the time living together with greater consistency than frequency of participation in general fitness activities and recreational sport, which are presumably less structured. Most notably, reliability of the participation in informal playful sporting games section of the questionnaire was exceptionally weak, with almost all items rated as having poor test-retest reliability.

Organised sports and supervised practice activities are typically scheduled on a regular basis, at a set time, for a fixed duration. As such, they become part of a routine, which appears to become consolidated in long-term memory. On the other hand, unsupervised, unstructured sporting activities are typically self-directed and spontaneous. The timing and duration of unstructured sporting activities tend to be irregular, so participation is less habitual, hence more difficult to recall with accuracy. Therefore, while participation in unsupervised, unstructured sporting activities may be a critical factor contributing to the development of sport expertise, the results of this study suggest that retrospective recall is not a suitable method for collecting this information. Alternative approaches such as activity diaries and longitudinal monitoring should be considered as more appropriate methods for exploring associations between sport expertise development and participation in unstructured, unsupervised sporting activities.

Not only does this finding have important implications for future research, but also for our current understanding of relationships between sport expertise and participation in unsupervised, unstructured sporting activities such as deliberate play (Baker et al., 2003b; Berry et al., 2008; Ford et al., 2009; Soberlak & Côté, 2003) and self-directed practice (Baker et al., 2003a; Helsen et al., 1998; Hodge & Deakin, 1998; Hodges & Starkes, 1996). Since past investigations of sport expertise development involving retrospective recall techniques have involved significantly less detailed validity and reliability assessments compared to the DHAQ, the reported psychometrics of earlier measurement instruments may be misleading. Consequently, our interpretation and application of previous results could potentially be misguided. Therefore, the need for suitable, robust approaches for collecting information relating to participation in unsupervised, unstructured sporting activities is especially pertinent, in order to clarify the associations between involvement in these activities and the development of sport expertise. That the findings of the current study have implications for both the conduct of future research and our understanding of past research is particularly noteworthy.

Although considered structured sporting activities, recurring reliability issues were also apparent within the competition history and other organised sports sections of the DHAQ. The

competition section of the updated, online questionnaire was considerably different to the initial version because the original DHAQ failed to cater to the complexity of competition involvement. While the modifications led to significant improvements in the reporting of competition history information, recall of participation in regular season format competitions was still highly inconsistent, particularly during the teenage years.

The format and scheduling of competitions varies considerably between sports and even within sports the format and scheduling of competitions often differs depending upon representative level. It is, therefore, difficult to attend to the intricacies of competition involvement within a generic questionnaire. It is acknowledged that, while beneficial in many respects, a non sport-specific questionnaire may actually restrict reliability of recall for certain information as sport-specific subtleties may not be sufficiently addressed. As a result, participants may experience difficulty applying the generic items to their sport, decreasing the accuracy with which they are able to respond. The strong test-retest reliability results for the practice history section of the DHAQ indicate this is not a problem when reporting details of involvement in practice activities, but it is possible that the generic nature of the questionnaire may be contributing to the poor recall of involvement in competition activities. It is also possible that, like unstructured sporting activities, the scheduling of competitions is too irregular to report with an acceptable degree of consistency. This is especially likely during the teenage years when athletes are competing for several different teams at a range of representative levels simultaneously, leading to a complex competition schedule.

Patterns of participation in competition and the balance of practice and competition activities associated with sport expertise would be a very interesting area to explore, but it does not appear that the DHAQ is an appropriate tool for collecting this information. Further efforts are therefore required to establish a more suitable approach for examining competition involvement and the development of sport expertise. It appears as though a sport-specific questionnaire for the collection of competition history information may be required in order to ensure the acquisition of meaningful, trustworthy data.

A particularly notable recurring reliability issue was apparent within the participation in other organised sports section of the questionnaire, which has important implications for our understanding of the role diversification plays in the development of sport expertise. A number of previous investigations have explored associations between participation in organised sports and the attainment of sport expertise (Baker et al., 2003b; Berry et al., 2008; Memmert et al., 2010; Moesch et al. 2011; Oldenziel et al., 2003; Soberlak & Côté, 2003). While reliability checks were incorporated into the majority of these investigations, these checks were typically restricted to the total number of sports athletes participated in and total cumulative hours of

involvement in all sports combined. In most cases, reliability for these measures met reliability standards, but the results from our study indicate that these measures may be misleading.

Both validity and reliability assessments of the DHAQ revealed that although reliability for the total number of organised sports reported was good, very few athletes provided the same list of sports on both test occasions. In the current study, only 35% of participants provided the same list of sports on both test occasions, while 10% of participants provided two completely different lists of sports on each test occasion. The remaining participants tended to list one or two sports on both test occasions, plus several sports on one occasion but not the other. As athletes appear to be inconsistent in their identification of the sports in which they have participated, the value for total number of sports has the potential to be highly misleading. Additionally, if different sports are reported on each test occasion, reliability for total cumulative hours of involvement in all organised sports combined will also be misleading, because hours of participation in different sports will contribute to the overall cumulative value.

Consequently, if reliability checks were not conducted on the particular sport types identified within previous investigations of diversification and sport expertise development, we cannot be completely confident in the accuracy of the data presented, and hence must interpret the results with caution. Since discussions relating to reliability of recall within the majority of previous investigations were typically quite brief, it is unclear whether reliability checks on sport types were conducted. Therefore, in the future, it is imperative that researchers are particularly vigilant in confirming reliability of recall for sport type, as well as reliability of measures pertaining to hours of involvement and highest level of competition reached in other organised sports, before further analyses are conducted on the data.

Although not evident during the validation of the original DHAQ, an additional reliability concern from the current study is noteworthy. Within each of the practice history, competition history, and participation in other organised sport sections of the questionnaire, it was observed that while test-retest reliability for hours of participation in a variety of activities per year was classified as very good or good, test-retest reliability for total cumulative hours of involvement was only moderate or poor. This was most apparent for activities in which hours of involvement per year were relatively low. As explained earlier, the discrepancy in reliability ratings occurs because when values for hours of involvement are low, the percent difference between responses from one test occasion to the next is small; however, as the magnitude of the values increases, differences become more noticeable. So, when all hours per year are added to provide a total cumulative value for hours of involvement in the activity overall, the initially small differences accumulate into a larger difference, at times resulting in a lower reliability classification. In light of this finding, it is important to be aware that information pertaining to

details of involvement in particular sporting activities is likely to be more reliable on an age-by-age basis than as a cumulative measure across an extended period of time. Therefore, when interpreting results concerning hours of participation in sporting activities, greater weight should be placed on analysis of hours of involvement per year rather than total cumulative hours of involvement overall.

At this point it is important to acknowledge a major limitation of this phase of the investigation. Following implementation of the recommended amendments, the updated, online DHAQ will have established test-retest reliability for all items; however, validity of this version of the questionnaire has not been determined. Due to the extensive procedures involved in assessing concurrent validity and convergent validity with parents and coaches, a complete validation of the updated, online DHAQ was outside the scope of this thesis. Given the general similarity in test-retest reliability and concurrent validity results during the initial validation of the original DHAQ, it could be expected that concurrent validity of the updated, online DHAQ would also be acceptable; however, it is important to confirm this assumption in the future. Similarly, despite limitations discussed during the initial validation of the original DHAQ regarding the suitability of parents and coaches as sources of validation for some developmental history information, it would also be important to confirm convergent validity with parents and coaches for appropriate sections of the questionnaire. Therefore, despite marked improvements in the test-retest reliability of the DHAQ, further validation testing is required before it can be said that the tool has established validity and reliability for all items.

More positively, as the participant sample recruited for this investigation of test-retest reliability of the updated, online DHAQ included athletes from two countries, a range of skill levels, and a variety of sports, it could be assumed that the results obtained are generalisable across a diverse range of participants. Hence, it is not unreasonable to suggest that following implementation of the suggested modifications, the DHAQ could be considered a reliable measurement tool for the collection of athlete developmental history information from athletes of any sport and any skill level. However, more deliberate attempts to affirm this statement are required. While individual differences in memory recall abilities are likely (Rubin, 1988; 1995), it is not expected that inter-individual variability in the ability to recall autobiographical information would differ greatly between athletes from different sports. Expert performance though, has been associated with more efficient storage and retrieval of domain specific knowledge in short- and long-term working memory (Ericsson & Delaney, 1998). Therefore, it is possible that more highly skilled athletes could recall details of their sporting history with greater ease and accuracy than lesser-skilled athletes. In fact, Starkes et al. (1996) reported slightly stronger correlations between estimated practice hours and hours of practice recorded in

training diaries for international level wrestlers compared to club level wrestlers. Additionally, some associations have been found between sex and autobiographical memory (Aizpurua & Koutstaal, 2010), with females appearing to recall autobiographical information slightly more accurately than males. Individual differences in retrospective recall were not considered within the current investigation, so it would be advisable to consider potential influences of skill level, sport type, and sex on validity and reliability of recall for athlete developmental history information in the future.

Summary and Conclusions

Regardless of the limitations discussed, test-retest reliability for the sporting career, sporting milestones, practice history, and places of residence sections of the DHAQ reached an acceptable standard for retention in the questionnaire without further modification. Following minor amendments, the representative history and family sections of the DHAQ will also be suitable for continued use. With removal of the competition history and participation in informal playful sporting games sections of the questionnaire, the only section of the DHAQ with questionable reliability that requires continued close monitoring is that related to participation in other organised sports. Notwithstanding the marked improvements in test-retest reliability however, further testing is required to confirm validity of the updated, online version of the DHAQ.

Once the suggested recommendations are implemented, the DHAQ can be recognised as an emergent quantitative measurement tool for the collection of detailed athlete developmental histories, with established reliability for all items. Following the recommended concurrent and convergent validity testing, it is proposed that the questionnaire could act as a standard for retrospective investigations of sport expertise development, as at present, no other athlete developmental history instrument has been subjected to such rigorous validity and reliability assessments.

General Discussion

This chapter provides a summary of findings, and discusses the major contributions of this thesis.

The aim of this thesis was to construct and validate a quantitative measurement tool for the collection of athlete developmental histories, suitable for large scale, unsupervised data collection. The vast literature that influenced the structure and content of the DHAQ during its initial construction phase was described within the chapter pertaining to study phase one. Interview guides, questionnaires, results, and recommendations of previous studies related to the development of sport expertise were consulted and integrated into a comprehensive paper-based questionnaire addressing participant demographics, attainment of sporting milestones, practice and competition history, coaching history, access to and utilisation of support services, details of involvement in other organised sports and informal playful sporting games, residential history, familial characteristics and participation in sport and physical activity, and a record of injuries, illness and time off.

Once constructed, the DHAQ was subjected to rigorous validation procedures to establish concurrent validity, convergent validity, test-retest reliability, and internal consistency. The results of this investigation, outlined in the chapter pertaining to study phase two, highlighted a number of issues concerning validity and reliability of recall. Several of these issues were associated with sub-optimal questionnaire design, while others were related to limitations of memory recall. As such, recommendations were made to modify and/or restructure particular items or sections within the questionnaire, while other items and sections were recommended for complete removal from the DHAQ. Given the large number of recommended modifications to the DHAQ, following implementation of the changes, reliability was reassessed again using a test-retest design.

To facilitate unsupervised, large scale distribution of the DHAQ, the questionnaire was also converted to an online format. Details of the modifications and online conversion of the updated DHAQ were provided in the chapter pertaining to study phase three, and results of the test-retest reliability investigation were discussed in the chapter related to study phase four. While several reliability issues were still apparent and a few additional minor changes were recommended, the majority of items in the updated, online DHAQ reached acceptable test-retest reliability standards. It was advised, however, that further testing take place to confirm concurrent validity and convergent validity with parents and coaches for the updated, online version of the DHAQ, as well as to reassess test-retest reliability for items still requiring modification. Therefore, this thesis did not completely achieve the stated aim of constructing and validating a quantitative measurement tool for the collection of athlete developmental histories, suitable for large scale, unsupervised data collection. The tool is, however, well on the way to having established, acceptable validity and reliability for all items, and could be used for

hypothesis testing investigations of sport expertise development, providing researchers conduct appropriate validity and reliability checks with the sample at hand.

Despite the need for ongoing validity and reliability testing, this thesis still provides three major contributions to the field of sport expertise development: 1) the creation of the DHAQ; 2) the development of a rigorous methodology and taxonomy for assessing validity and reliability of retrospective recall; and 3) a detailed understanding of consistency of recall for athlete developmental history information. This chapter will focus on these three major contributions, highlighting the novelty and significance of each. First, however, the findings of this thesis are reviewed and discussed.

Discussion of Findings

Validity and reliability for the attainment of significant sporting milestones was generally very good. Athletes, parents, and coaches could consistently recall the ages at which milestones such as first participation in practice activities for main sport, specialisation in main sport, and first relocation for reasons relating to sport were reached; and they were also able to consistently recall ages at which athletes first participated in competition at the various representative levels. The results from both validity and reliability investigations within this thesis indicate that we can be confident in athlete recall of significant sporting milestones; and, that parents and coaches are reasonable sources of validation for this information.

Athlete recall of practice history information was also quite good. During the initial validation of the original DHAQ, concurrent validity, test-retest reliability, and internal consistency results were mixed, but following modifications, test-retest reliability for practice history information within the updated, online DHAQ improved considerably, and the majority of items reached acceptable reliability standards. The improvements in athlete recall from the original DHAQ to the updated, online DHAQ indicated that in order to facilitate accurate, consistent recall, the item structure of the measurement tool is critical. Items relating to practice history information within the original DHAQ appeared to be overly repetitive, resulting in participant confusion, frustration, and boredom, which ultimately led to poor response quality.

Additionally, a number of items within the original DHAQ appeared to be too specific. While athletes could consistently recall general details of their participation in practice activities, as items prompted them to identify information regarding their involvement in very specific types of practice activities, consistency of recall declined. It is likely that athletes could not remember such specific details of their past practice schedules, and/or the composition of practice varied over the course of the year depending upon periodisation cycles, making it

difficult to provide ‗average' values for the typical number of hours engaged in these specific practice activities each week.

As such, the practice history section of the updated DHAQ was modified to reduce repetition, and practice activities were collapsed into fewer, more general categories. These modifications were successful in improving recall, indicating that an appropriate balance of generality and specificity of information was obtained. These results suggest that athletes are able to consistently recall practice history information, providing the questions take into consideration the specificity of the information required.

It is important to note that while the majority of items within the practice history section of the updated, online DHAQ reached acceptable test-retest reliability standards, supervised practice activities were recalled with slightly greater consistency than unsupervised practice activities. As discussed on several occasions throughout this thesis, supervised practice activities are typically scheduled at regular times for fixed durations, becoming part of a habitual routine for the athlete. Subsequently, these activities may be easier to remember than unsupervised practice activities that are more likely to occur spontaneously and for varied durations.

Additionally, consistency of recall was slightly weaker during the teenage years for several practice activities. The complexity of sport involvement during the adolescent period, due to participation in multiple teams, age group categories, and/or other sports simultaneously, probably results in weaker memories of this time, leading to slightly lower consistency of recall. Therefore, while we can be reasonably confident in athlete recall of practice history information (when data are collected via a carefully constructed instrument such as the DHAQ), it is important to be aware that details of participation in unsupervised practice activities, and/or details of participation in practice activities during the teenage years, might be slightly less reliable, so they should be interpreted with care.

These results highlight the complexities of retrospective recall for practice history information. Since consistency of recall varies according to practice type and age, simple assessments of validity and reliability for total cumulative hours of participation in all practice activities combined (as are common in previous investigations), are unlikely to sufficiently reflect the accuracy of the data. Accordingly, this thesis emphasises the importance of collecting detailed practice histories, addressing involvement in specific types of practice activities separately; along with the necessity to conduct in-depth examinations of validity and reliability for all responses, rather than cumulative measures or a small sample of select measures. These findings are particularly pertinent for investigations utilising measurement tools other than the

DHAQ, which have not been subjected to the same scrupulous validity and reliability assessments.

In relation to parent and coach recall of athlete practice history information, the results from the initial validation of the original DHAQ were surprising. It was expected that since parents are usually responsible for overseeing their child‘s involvement in sporting activities and transporting them to and from practice, convergent validity with parents would be good. Instead, convergent validity with parents for practice history information was generally quite poor. It is understandable that convergent validity with parents could be poor for participation in specific practice activities, because parents may not always be aware of the particular activities in which their child is engaging during practice sessions, especially if they do not stay and watch their child train; however, it was anticipated that convergent validity with parents would be good for recall of general involvement in main sport. Although convergent validity with parents was good for attainment of sporting milestones, parental recall of details pertaining to practice scheduling was not as strong. It may be that parents are so busy overseeing the involvement of all their children in a variety of different activities, they cannot clearly remember how often, and for how long each child engaged in each activity. While the reasons cited for poor parental recall are speculative, it is apparent that parents are not an appropriate source of validation for practice history information.

Given the discrepancies between responses, an initial question was whether athlete or parent recall of practice history information would be more correct. Although athletes are likely to be more vested in their sport involvement, and so would be expected to remember details of their practice history more clearly, as mentioned earlier, parents are usually responsible for transporting their child to and from practice, so they could also be expected to have clear memories of their child‘s involvement in sport. Fortunately, the inclusion of coaches as a third validation source was particularly advantageous and was a great strength of this study, as it allowed for triangulation of the data to establish whether athlete or parent recall may be more correct. In the case of practice history information, convergent validity with coaches was significantly stronger than convergent validity with parents. Given the similarity between athlete and coach responses, it could be assumed that athlete responses were more accurate than parents‘, supporting the earlier suggestion that parents may not be appropriate sources of validation for practice history details. Coaches, on the other hand, do appear to be suitable. The proposed examination of convergent validity with parents and coaches for the updated, online DHAQ will provide further insight into the suitability of parents and coaches as appropriate sources of validation for practice history information.

Another factor considered within the DHAQ that was recalled very consistently by both athletes and parents was family characteristics. The only exception was recall of parent occupation, which was found to be poor during the initial validation of the DHAQ. Since this item was intended as a proxy indicator of socioeconomic status, during modification of the questionnaire it was replaced with highest level of education reached to date. These measures were selected over approximation of household income because they have been recommended as more reliable indicators of socioeconomic status during childhood and adolescence (Galobardes, Shaw, Lawlor; Lynch, & Smith, 2006; Hauser, 1994). Test-retest reliability for highest level of education was very good or good for all family members, suggesting that highest level of education is a more suitable item for inclusion within athlete developmental history questionnaires than parental occupation when attempting to obtain socioeconomic status information. Unfortunately, parent confirmation of highest level of education was not obtained, so convergent validity of this information is unknown at this point, but can be confirmed during the proposed validity assessments for the updated, online DHAQ. Note that convergent validity with coaches was not assessed for items relating to participants‘ family members, as it was not expected that coaches would be aware of these details.

While validity and reliability of recall for familial characteristics was strong, recall of familial involvement in sport and physical activity was not as good. Athletes could consistently recall whether family members had participated in various types of sport and physical activity during the time living together, but could not provide more specific details of the frequency of participation in each activity. Similarly, athletes could consistently identify whether family members had participated in competitive sport during any time in their lives, but could not consistently identify the number of sports, types of sports, or the highest level of competition reached. Athletes were, however, consistent in recalling whether family members had ever participated in the same sport as their main sport. Convergent validity with parents during the initial validation of the DHAQ followed a similar trend. Therefore, we can be reasonably confident in athlete and parent recall of familial characteristics and basic details of familial involvement in sport and physical activity, but if required, more precise details of familial sport participation should be obtained from each family member directly.

In the case of another contextual factor associated with athlete development, places of residence, consistency of recall was mixed. Residences where athletes stayed for long durations were generally recalled consistently by athletes and parents, but residences of short duration were not. Similarly, consistency of recall for details relating to the first two towns/cities of residence was stronger than for later residences. In the majority of cases, participants were adults by the time they relocated to their third and subsequent residences. Therefore, in most

instances, these later residences are likely to have been less impactful on athletes‘ development than earlier residences during the childhood and adolescent years. Given studies of sport expertise development typically focus on factors influencing athlete development during childhood and adolescence, it is not of great concern that recall of places of residence information for residences three and above was less consistent than for residences one and two. Likewise, residences of short duration are likely to have been less impactful on athletes‘ development than residence of long duration, so the lower validity and reliability of recall for short duration residences is also not a great concern.

Positively, details of relocations for reasons relating to sport were recalled consistently by both athletes and parents. Therefore, it appears that recall of relocations for reasons relating to sport can be considered trustworthy, even when these relocations were of short duration, occurred during adulthood, and/or followed multiple prior relocations. It is interesting that although athletes and parents consistently recalled relocations for reasons relating to sport, when relocations were not related to sport, the reason cited was highly variable. This finding suggests that sport-based relocations may have been definitively related to athletic development, whereas other relocations were perhaps made for a combination of reasons including family, work, and lifestyle. Together, these results indicate that we can be reasonably confident in athlete recall of details pertaining to places of residence during the childhood and adolescent years, and recall of relocations for reasons relating to sport. Interestingly, this research suggests that recall of place of residence information during the adult years, as well as recall of details of relocations for reasons other than sport, is questionable.

In relation to parent recall of residential information, parents were found to be suitable for confirming details of residences only up to the point at which the athlete left the family home. Beyond this time, convergent validity with parents became particularly poor. As coaches are unlikely to be aware of athletes‘ residential information, alternative sources such as real estate documents, bank statements, and/or utility records may therefore be required to validate athlete recall, particularly for residences other than the family home in which participants were raised.

Yet another factor displaying mixed validity and reliability of recall was competition involvement. During validation of the initial DHAQ, the structure of the questionnaire did not adequately cater to the complexities of competition involvement. The instrument failed to take into consideration that there are a number of different formats of competition, competitions do not always take place at regular intervals, and involvement in competition often includes extended periods of time at the venue but not necessarily actively competing. For these reasons, athletes expressed difficulty quantifying the time involved in competition. Subsequently,

substantial changes were made to this section of the questionnaire within the updated, online DHAQ. Although the modifications to the competition history section of the DHAQ allowed participants to discuss their competition involvement more successfully, a number of reliability issues remained.

Details of involvement in occasional and tournament format competition were recalled with relatively good consistency on an age-by-age basis, but reliability of recall for details of participation in regular season competitions was poor across a number of ages. Furthermore, all cumulative measures of participation in competition received poor test-retest reliability ratings. Regular season competition forms a major component of the competition schedule for many sports, so the poor reliability for related items was a concern. It appears that recall of competition involvement can be difficult because athletes are often competing in a range of age group categories, at multiple representative levels, simultaneously. This results in a complex competition schedule that seems to be difficult to recall accurately.

Competition formats and schedules also vary considerably between sports, so some athletes reported having trouble classifying their competition involvement into the three categories of regular season, occasional, or tournament format. Therefore, poor reliability of recall for some aspects of competition involvement is likely related to a combination of persistent questionnaire design problems, as well as limitations of memory recall for complex, and/or irregular competition scheduling. Given the complex and sport specific nature of participation in competition, a more detailed, sport specific measurement tool may be required to accurately determine the consistency with which athletes can recall competition history information. For the time being, however, the present results indicate that we can be reasonably confident in athlete recall of participation in occasional and tournament format competition on an age-by-age basis, but recall of participation in regular season competitions is questionable. Reliability of cumulative measures of participation in each competition format, along with details of involvement in all formats of competition combined were also problematic, so it is recommended that researchers avoid analyses involving these values.

Unfortunately, due to the inadequate design of the original DHAQ, convergent validity with parents and coaches for competition involvement could not be assessed. It could be assumed that parent and coach recall for details of participation in competition would follow the same trend as practice activities, with coaches representing more suitable sources of validation for this information than parents, but further study is required to confirm this supposition. Given the difficulties reported relating to athlete recall of details of participation in competition, additional sources of validation should also be explored, including training diaries, results archives, newspaper clippings, and awards.

Two factors associated with sport expertise development that were recalled with unexpectedly poor consistency were details of participation in organised sports other than main sport, and details of participation in informal playful sporting games. The relatively poor validity and reliability of recall for participation in organised sports other than main sport was particularly surprising given previous investigations have reported acceptable reliability of recall for this information (Baker et al., 2003b; Berry et al., 2008; Memmert et al., 2010; Soberlak & Côté, 2003). It is believed that the more thorough validity and reliability assessments conducted in the present series of studies is likely responsible for the differences in reported results.

For example, total number of sports played, and total cumulative hours of participation in all other organised sports combined are the main measures typically examined for reliability and validity. While the current studies also displayed good validity and reliability of recall for total number of organised sports, close inspection of the types of sports reported revealed incongruent lists between athletes and parents, as well as between the various athlete test occasions. Therefore, the value for total number of sports in which athletes participated, and subsequently total hours of participation in all other organised sports combined, is highly misleading. The limited information provided in previous investigations regarding validation procedures and results makes it unclear whether such checks were carried out; however, the current findings highlight the necessity for rigorous assessments of validity and reliability when collecting data via retrospective recall methods. The findings suggest that the selection of only a few key measures to represent validity and reliability of the complete data set can be deceiving, and either more comprehensive checks, or rigorously validated instruments such as the DHAQ, should be used where possible.

Furthermore, the results indicate that recall of participation in other organised sports is questionable. The data indicated that athletes had sampled several sports during childhood and adolescence, some for only one or two seasons. Therefore, it is possible that athletes and parents did not always mention sports they only participated in for a short period of time, perhaps because their involvement seemed insignificant compared to their main sport. Either participants forgot to mention these short duration sports, or they did not feel that these sports were worth mentioning because their involvement was so brief. This could explain the discrepancies in sports reported between test occasions. Similarly, details of involvement in activities of low significance to participants are likely to be more difficult to remember, so this could explain the sometimes low consistency of recall for details of participation in sports that were reported.

Indeed, the salience of an event was highlighted by both Coughlin (1990) and Dex (1995) as a major factor influencing recall error; plus, investigations of accuracy of recall for

occupational history information have noted positive associations between recall and employment duration, but negative associations between recall and number of jobs held (Coughlin, 1990). Therefore, the observed results relating to recall of participation in other organised sports are perhaps not so surprising, particularly if athletes have engaged in a high level of sport sampling.

Nevertheless, in order to provide accurate advice to athletes, coaches, and parents regarding the role of diversification in the development of sport expertise, the quality of data collected in this area must be improved. Although longitudinal research would alleviate issues related to reliability of recall, such designs are often not practical due to time and financial demands (Brown, Cozby, Kee, Worden, 1999) not to mention the difficulties associated with ensuring participant samples contain sufficient athletes who go on to achieve elite status (Sosniak, 2006). As such, attempts must be made to improve reliability of retrospective recall for participation in other organised sports. While records of participation such as training diaries, enrolment papers, photographs, awards, and/or results archives might be useful to assist and/or validate recall of this information, it is also recommended that researchers re-examine diversification in the context of sport expertise development.

It is proposed that discussions and further research are required to identify clear criteria for identifying which other sports an athlete may have participated in that warrant reporting in studies of athlete development. Potential criteria could relate to overall duration of participation, average hours of practice per week, level of representation reached, and/or elements of transfer. Following the introduction of minimum requirements for reporting involvement in other organised sports when collecting athlete developmental histories, it is expected that reliability of recall for this information will improve, since only sports of relative significance will be addressed. In the meantime, however, it is important to be aware of the limitations associated with athlete and parent recall of participation in organised sports other than main sport, and to interpret related results with caution.

Similarly, researchers should be cautious interpreting results related to participation in informal, playful sporting games. Although a number of studies have investigated associations between participation in informal, playful sporting games (often referred to as deliberate play), and the development of sport expertise (Baker et al., 2003b; Berry et al., 2008; Ford et al., 2009; Memmert et al., 2010; Soberlak & Côté, 2003), only two reported results pertaining to reliability of recall for details of participation in these activities (Memmert et al., 2010; Soberlak & Côté 2003). Both of these studies reported acceptable validity and reliability of recall for total hours of involvement in playful sporting games; however, the majority of investigations have typically examined the role of play among a range of other variables, only reporting validity and

reliability results for a selection of measures, and not usually those related to sporting play. The present series of studies examined validity and reliability of recall for details related to participation in informal, playful sporting games in great depth, and results indicated that recollection of this information by both athletes and parents was exceptionally poor.

Poor recall of participation in informal, playful sporting games is likely related to the spontaneous, irregular, and highly variable nature of these activities. Play is not typically scheduled, the duration of sessions is not usually fixed, and the types of activities in which participants engage can vary not only between sessions, but also within sessions. The ever-changing, unstructured nature of play appears to make details of participation very difficult to recall with consistency. In addition, youth often engage in these types of activities without direct supervision from an adult, further contributing to the poor convergent validity with parents. Play can occur in the back yard, the school yard, the street, or at neighbourhood parks and facilities (Baker & Côté, 2005). Participation in play does not always require resources such as transport, equipment, or financial support, so parents will not always be aware of specific details of their child‘s involvement in play. Therefore, it is unrealistic to expect parents to provide accurate accounts of the total time their children engaged in sporting play during a typical week, or the types of games they typically played.

Although participation in informal, playful sporting games may be an important contributor to sport expertise, we cannot be confident in the accuracy of retrospective recall for this information. As such, to investigate associations between involvement in sporting play and the development of expertise, longitudinal, prospective experimental designs, in which details of participation in sporting games are recorded in activity diaries, may provide the only suitable method of data collection for this information. It is expected that activity diaries would result in more accurate accounts of details such as time engaged in play, games played, who participants played with, and where they played, so this approach should be explored in the future.

A number of additional factors thought to be associated with the development of sport expertise were addressed within the original DHAQ, but excluded from the updated version due to poor validity/reliability of recall, and/or because the factor appeared to be too complex to address within a quantitative multidimensional developmental history questionnaire such as the DHAQ. Factors such as coaching, sport science/sport medicine support services, and adverse events including injuries and illness were acknowledged as influential on athlete development, so were incorporated into the original DHAQ; however, it was found that they were not addressed in sufficient detail to provide a meaningful contribution to the literature regarding the role they play in the development of sport expertise. Additionally, validity and reliability of recall for these items were generally quite weak. Rather than modify and expand these sections

to attempt to both improve validity and reliability of recall and collect more detailed information, a decision was made to remove them completely to keep the DHAQ at a practical length for participants to complete. As such, although validity and reliability of recall for these sections was not strong within the current study, further investigations involving more detailed instruments are required in order to be certain of the consistency with which athletes, parents, and coaches can recall information pertaining to coaching history, access to, and utilisation of, support services, as well as injury and illness history.

Major Contributions

The methodological approach and results of this thesis provide three major contributions to the field of sport expertise development. Notably, these contributions extend beyond athlete development into domains such as public health, medicine, psychology, and memory. The contributions of this thesis to our understanding of the development of sport expertise plus potential applications to additional areas of research are discussed in detail below.

Major Contribution One: The Creation of the Developmental History of Athletes Questionnaire

Many previous investigations have created measurement tools for the collection of athlete developmental history information (Carlson, 1988; Côté, 1999; Côté et al. 2005; Durand Bush & Salmela, 2005; Gibbons et al., 2002; Hodges & Starkes, 1996; Moesch et al., 2011; Oldenziel et al., 2003; Phillips et al., 2010a; Soberlak & Côté, 2003; Ward et al., 2007; Weissensteiner et al., 2009). While the majority of these tools were subjected to reliability and/or validity assessments, consistency of recall was typically only assessed on a small sample of items. One unique feature of the DHAQ is that validity and/or reliability was assessed for all items, as well as for a number of additional measures obtained via calculations involving a combination of responses. Very few, if any, athlete developmental history questionnaires or interview guides have been examined with such scrutiny prior to use, rendering the DHAQ perhaps the most strongly validated athlete developmental history questionnaire available. This can be said about the DHAQ even now in its current format, before the proposed amendments are made and the recommended assessments to confirm validity for the updated, online version of the DHAQ are conducted.

A second unique feature of the DHAQ is its online format. Traditionally, studies of sport expertise development have involved either one-on-one data collection in the form of

qualitative interviews (Baker et al., 2003a; Baker et al., 2003b; Baker et al., 2005; Berry et al., 2008; Bloom, 1985; Carlson, 1988; Côté, 1999; Durand-Bush & Salmela, 2002; Law et al., 2007; Phillips et al., 2010a; Soberlak & Côté, 2003; Weissensteiner et al., 2009), or distribution of a paper-based questionnaire. Paper-based questionnaires can be completed during supervised data collection sessions (Oldenziel et al., 2003; Ward et al., 2007); administered directly by a researcher/representative, but completed unsupervised and returned by hand (Ford & Williams, 2008; Hodges & Starkes, 1996; Memmert et al., 2010; Oldenziel et al., 2003; Starkes et al. 1996); or distributed and returned via mail (Gibbons et al., 2002; Hodges & Starkes, 1996; Oldenziel et al., 2003). Supervised data collection methods and approaches involving some degree of face-to-face interaction with participants limit the potential sample size of the investigation, and mail-based questionnaires incur additional expenses and time considerations.

An online questionnaire circumvents these limitations, allowing for administration to large sample sizes in a cost-effective and timely manner. As outlined within earlier discussions of study phase three, additional benefits of online data collection methods include: semi-automated participant management procedures including informed consent, completion instructions, and distribution of reminder and thank you emails; personalised completion of the questionnaire via the ability to automatically direct participants to the next relevant item; and semi-automated data entry and analysis procedures through direct export of responses to an external database. With the increasing ease of access to the Internet, web-based data collection is becoming a more viable, attractive alternative to face-to-face and mail-based methods, particularly for investigations involving large sample sizes. The current series of studies, plus a recent investigation of patterns of sport specialisation in Danish athletes that also utilised a web-based practice history questionnaire (Moesch et al. 2011), suggest that online data collection methods are suitable for obtaining athlete developmental history information, and may represent a methodological shift within the field of sport expertise development.

Yet another key feature of the DHAQ is the wide-ranging scope of the questionnaire. Addressed within the updated, online DHAQ are participant demographics, the attainment of sporting career milestones, pathways of progression through various representative levels of competition, a comprehensive practice history, details of participation in other organised sports, familial characteristics and participation in sport and physical activity, and places of residence. Therefore, in reference to the four major approaches to research examining the development of sport expertise outlined in the introduction to this thesis, the DHAQ will facilitate investigations of the contributions of practice to sport expertise development, the role of environmental and contextual factors in the development of sport expertise, and multi-factorial studies of sport expertise development as well. It is, however, acknowledged that there are a number of other

factors associated with the development of sport expertise which are not addressed within the DHAQ; for example, genetics (Bouchard, Malina, and Perusse, 1997; Puthucheary et al. 2011; Singer & Janelle, 1999; Tucker & Collins, 2012), psychological skills and attributes (Gould, Dieffenbach, & Moffett, 2002; Gould, Eklund, & Jackson, 1992; Orlick & Partington, 1988; Weissensteiner Abernethy, Farrow, & Gross, 2012), coaching (Bloom, 1985; Côté, Salmela, Trudel, Baria, & Russell, 1995; Deakin & Cobley, 2003; Durand-Bush & Salmela 2002; Oldenziel et al., 2003), and even chance factors (Gagné, 2004), to name just a few. Nevertheless, the DHAQ incorporates a sufficient number of key practice-related and environmental / contextual contributors to the development of sport expertise to conduct in-depth examinations of associations between sport expertise and these factors in isolation, as well as to explore a variety of interactions between these key components. At the same time, the length of the DHAQ does not place an overly onerous time demand on participants. The DHAQ therefore provides a balance between scientific rigour and participant considerations for completion, resulting in an informative, yet practical measurement tool. Considerations such as time and convenience are critical when conducting research in the high performance sport setting, as athletes have demanding practice and competition schedules. This balance was, therefore, paramount in the construction and modification of the DHAQ.

In summary, the DHAQ is emerging as a practical yet comprehensive and robust measurement tool for the collection of athlete developmental histories. Once the validation process is complete, the DHAQ will be able to be utilised for a wide variety of investigations related to the development of sport expertise. Furthermore, with the advantage of established validity and reliability, it is possible the DHAQ could become the measurement standard for athlete developmental history information. A strongly validated, standardised measurement tool such as the DHAQ will eliminate the need for researchers to conduct their own validity and reliability assessments, affording greater time and resources for data collection. It will also allow for direct comparison of results between studies, advancing our understanding of the development of sport expertise more efficiently and effectively.

Major Contribution Two: The Development of a Rigorous Methodology and Taxonomy for Assessing Validity and Reliability of Retrospective Recall

In addition to the construction of the DHAQ, several additional contributions to the field of sport expertise development emerged from the validity and reliability assessments. The first of these contributions was the development of a rigorous methodology and taxonomy for assessing validity and reliability of retrospective recall.

Upon review of the sport expertise development literature, it became apparent that the statistical tests typically used to examine validity and/or reliability of retrospective recall were perhaps, not the most appropriate given the nature of the data. As explained in greater detail during earlier discussions of study phase two, the most common statistical approach adopted for assessing validity and/or reliability of recall for athlete developmental history information was to calculate the Pearson product moment correlation (Baker et al., 2003a; Baker et al., 2003b; Baker et al., 2005; Berry et al., 2008; Côté et al., 2005; Helsen et al., 1998; Hodges & Starkes, 1996; Memmert et al., 2010; Oldenziel et al., 2003; Ward et al., 2007); however, this type of correlation is best suited for assessing the relationship between two different variables (Haggard, 1958). Intraclass correlations, on the other hand, are more suitable for assessing the relationship between two measurements of the same variable (Haggard, 1958). Therefore, when examining validity and/or reliability of recall, the ICC is a more accurate indicator of the strength of the association between multiple responses than the Pearson correlation coefficient.

Despite this, the ICC still does not provide an accurate indication of the absolute similarity of responses (Bland & Altman, 1986; Costa-Santos et al., 2011; Kottner & Dassen, 2008). Plus, ICCs are only suitable for continuous variables (Müller & Büttner, 1994), and many items within the DHAQ involved categorical responses. Thus, PA values were calculated for all items, both categorical and continuous. PA or percent difference values have been utilised to assess validity and/or reliability in several studies of sport expertise development (Baker et al., 2005; Law et al., 2007; MacDonald et al., 2009; Memmert et al., 2010; Soberlak & Côté, 2003), however, the use of these statistics are far less common than the Pearson product moment correlation. Even less common is the use of multiple statistical methods to assess validity and/or reliability in the field of sport expertise [see Ford and Williams (2008) for an exception], despite this approach being recommended in many other domains including nursing, medicine, and sports medicine (Atkinson and Neville, 1998; Kottner & Dassen, 2008; Kottner et al., 2011). While both PAs and ICCs were calculated for continuous variables in the current series of studies, absolute agreement was considered more important than relative consistency; so, when rating validity and reliability, PA values were weighted more heavily than ICCs.

A second observation from the literature on assessment of validity and reliability was the lack of a standard scale for interpreting the strength of ICCs and PA values. For example, in one case, a PA of 70% was considered to be ‗reasonable agreement' (Baker et al., 2005), whereas, in another study, a PA of 72% was described as ‗less reliable' (MacDonald et al., 2009). Similarly, one investigation suggested that reliability was ‗good' if the ICC was .90 or above (Lemmink et al., 2004), while another indicated that an ICC of .65 or above represented

≥good‘ reliability (Costa-Santos et al., 2011). Therefore, it was apparent that strict criteria for interpreting PA and ICC values were required.

Furthermore, since a combination of PA and ICC statistics were deemed to be the most appropriate approach for assessing validity and reliability of recall for continuous variables in the current series of studies, a classification system encompassing both values was also necessary. Hence, comprehensive criteria for classifying PA values, ICCs, and both in combination were established (see chapter related to study phase two and Table 2 for a more detailed description of the proposed criteria and classification systems). This classification system was then utilised to determine the suitability of items for retention in the DHAQ, with items receiving validity/reliability classifications of very good or good considered suitable for retention, while items displaying moderate or poor validity/reliability were recommended for modification or removal.

The establishment and adoption of strict criteria for determining acceptable validity and/or reliability prior to data collection is a unique feature and great strength of this series of studies. Not only were all items within the DHAQ assessed for validity and/or reliability, but all were exposed to the same objective classification criteria to determine suitability for retention in the questionnaire. Of considerable importance, however, is that the methodology and taxonomy developed can be extended beyond the DHAQ and adopted in other studies examining validity and/or reliability of measurement.

It is acknowledged that the range of acceptable variability when assessing validity and reliability may differ depending upon the nature of the measurement, hence, it is suggested that the methodology and taxonomy proposed in this work is most suitable for assessing validity and reliability of recall for athlete developmental history information; however, it is likely that the same approach could be adopted for assessing validity and reliability of recall for lifetime physical activity behaviours in a variety of health-related investigations. It is also possible that the methodology and taxonomy could be used as a template, and adapted for use in other domains requiring alternative ranges of acceptable variability, for example, medicine and clinical psychology. It is, therefore, particularly noteworthy that this thesis has the potential to contribute to fields outside of the primary domain of sport expertise development.

Major Contribution Three: A Detailed Understanding of Consistency of Recall for Athlete Developmental History Information

A further contribution arising from the validity and reliability assessments was an in-depth understanding of consistency of recall for athlete developmental history information.

These findings relate not only to athlete recall, but also to parent and coach recall of athlete developmental history information. Parents and coaches are occasionally used as sources of validation for athlete recall (Baker et al., 2003a; Baker et al., 2003b; Berry et al., 2008; Côté, 1999; Côté et al., 2005; Durand-Bush & Salmela, 2002; Law et al., 2007; Oldenziel et al., 2003; Soberlak & Côté, 2003), hence, it is important to address the suitability of these sources for future investigations. The detailed validity and reliability assessments conducted on both the original and updated, online DHAQ provide considerable insight into the factors athletes, parents, and coaches could recall with acceptable consistency, those for which recall was questionable, and those for which retrospective recall was so poor that alternative data collection methods should be explored.

The results of the current series of studies indicate that we can be reasonably confident in athlete recall of sporting milestones, practice history, representative history, residential history, as well as family characteristics and basic details of familial participation in sport and physical activity. Athlete recall of more specific details of familial participation in sport and physical activity such as frequency of participation in various physical activities, types of sports family members have played, and highest level of competition reached is more questionable, as is detailed competition history information, and details of participation in other organised sports. Furthermore, athlete recall for details of participation in informal, playful sporting games was unacceptable. Alternative methods of data collection for information relating to factors for which athlete recall is questionable or unacceptable should be explored in future investigations.

The results also suggest that parents are suitable sources of validation for sporting milestones, residential history, and familial characteristics and involvement in sport and physical activity. Parents are not, however, suitable sources of validation for practice and competition history information or details of participation in other organised sports or informal, playful sporting games. Coaches appear to be more appropriate sources of validation for practice history details, competition history information, and sporting milestones achieved during the time working with the athlete, but coaches are not expected to be aware of athlete developmental history information concerning factors outside of the practice and competition environment.

These insights into the consistency with which athlete developmental history information can be recalled are the direct result of exposing the DHAQ to thorough validity and reliability procedures prior to utilisation of the measurement tool in examinations of sport expertise development. Future investigations involving retrospective recall of athlete developmental histories either via the DHAQ or alternative measurement instruments should carefully consider these findings during the study design, data collection, and data interpretation

phases of research to ensure trustworthy reporting of results. Furthermore, these findings may also be interesting to researchers and practitioners concerned with autobiographical memory (Rubin, 1986), as they contribute to our understanding of the content of long-term human memory, the temporal organisation of memories for events and participation in extra-curricular activities such as sport, and parental recall for biographical information relating to their child.

Summary

The creation of the DHAQ and the development of a rigorous methodology and taxonomy for assessing validity and reliability of retrospective recall represent major methodological contributions of this thesis to the field of sport expertise development and beyond. Moreover, the contribution of a detailed understanding of consistency of recall for athlete developmental history information extends our theoretical understanding in this area. Given the broad nature of these contributions, and the relevance of the current findings both within and beyond the domain of sport expertise development, a wide range of potential directions for future research emerge from this dissertation. Several key areas for future research are discussed in the next chapter.

Future Directions

This chapter outlines potential areas for future research based upon the results and recommendations of this thesis.

Following implementation of the few suggested modifications to the questionnaire, the DHAQ could be considered suitable for use in hypothesis testing investigations related to the development of sport expertise; however, until the recommended concurrent and convergent validity testing is complete, researchers would be required to conduct their own validity checks with the sample at hand. Sections retained within the DHAQ include participant demographics, sporting career and milestones, representative history, practice history, participation in other organised sports, familial characteristics and participation in sport and physical activity, and residential history. Therefore, the DHAQ could be utilised for a wide range of investigations pertaining to associations between sport expertise and these factors.

For example, the questionnaire could be completed by athletes of varying skill levels to identify factors that differentiate more highly skilled athletes from lesser skilled athletes; or it could be completed by athletes from different sports to identify characteristics of expertise development that may be sport-specific, those that may be common across groups of similar sports, and those that may be generalised across all sports. Furthermore, the questionnaire could be completed by athletes from different communities, cultures, or countries to explore geographical, cultural, and/or societal differences in sport expertise development. Sex-based differences in athlete development pathways could also be examined.

Investigations utilising the DHAQ could be multidimensional, exploring interactions between a number of factors associated with sport expertise development, or one-dimensional, examining a single influence. As such, it is suggested that the DHAQ could be administered in its entirety, or particular sections of the questionnaire could be selected for more specific research questions involving only one or several of the aspects addressed. Not only would segmenting the questionnaire have the advantage of eliminating irrelevant items and reducing participant requirements during data collection, but it would also allow the DHAQ to be utilised for investigations in areas other than sport expertise development. For example, the familial characteristics and participation in sport and physical activity section could be incorporated into studies examining predictors of health and physical activity beliefs and behaviours; the participation in organised sports section could be utilised within investigations of physical literacy in youth; or the practice history section could be adopted for studies exploring contributors to sports injuries[1].

The scope of the questionnaire is therefore not limited to the field of sport expertise development. However, within this field, it is expected the DHAQ will be utilised for large scale

[1] The DHAQ is already beginning to receive international recognition, with research groups from Australia, Canada, Germany, and Iran requesting to utilise the questionnaire for a variety of investigations ranging from studies of talent development to associations between sport involvement and perceptions of body image.

examinations of factors associated with superior sports performance, as well as for examinations of interactions between these factors. For example, relationships between familial involvement in the same sport as the athlete, attainment of sporting milestones, and rate of progression through the various representative levels could be explored to identify whether family history of involvement in the same sport is associated with accelerated athlete development. Similarly, associations between city size and number of organised sports in which athletes participated could be examined to explore diversification as a potential mechanism for the birthplace effect (Carlson, 1998; Côté et al., 2006; MacDonald, Cheung, Côté, & Abernethy, 2009). The multidimensional nature of the questionnaire is a great strength of the DHAQ as it leads to vast potential for use in a wide range of investigations, both within and outside the field of sport expertise development.

As alluded to in the introduction to this thesis, it would also be interesting to examine the predictive validity of the DHAQ through longitudinal studies of sport expertise development. In these studies, athletes identified as demonstrating potential for future success in their sport at the open international level could be administered the questionnaire upon selection to an elite development training program, then tracked over a period of time. The previously collected developmental history profiles of athletes who ultimately succeeded at the open international level could be compared to the profiles of those athletes who were less successful, in order to identify any differentiating factors between the groups that were evident at the earlier stage of their careers. If any measures incorporated within the DHAQ were identified as having strong predictive validity for eventual attainment of sport expertise, this could have major implications for talent identification and athlete selection procedures.

It is important to highlight that the DHAQ should be considered an evolving measurement tool with the capacity to be updated regularly according to advances in our understanding of the development of sport expertise. It is possible that following large scale examinations, certain factors may be deemed inconsequential to sport expertise development, and hence may be removed from the questionnaire. Alternatively, new factors may emerge as critical influences on athlete development, warranting inclusion as a new section within the DHAQ. Ongoing validity and reliability assessments are also recommended to ensure the DHAQ remains a trustworthy tool for the collection of athlete developmental history information. Minor edits may therefore be required following continued validity and reliability procedures. It is expected that any future amendments to the DHAQ will be documented and published to avoid confusion related to reporting of results obtained from different versions of the questionnaire.

Relating to issues of validity and reliability, additional future directions concern continued efforts to collect valid, reliable information pertaining to competition history, participation in informal playful sporting games, and detailed familial sport histories. While recognised as important influences on the development of sport expertise, retrospective recall of these details within the DHAQ was relatively poor, so alternative methods of data collection are required. For the collection of competition history information, retrospective recall may still be a viable method of data collection; however, since it was apparent the DHAQ did not adequately address the complexity of competition involvement, more detailed sport-specific questionnaires should be developed and evaluated.

Given the concerns raised regarding retrospective recall of irregular, unstructured activities such as sporting play, it is suggested that collection of data relating to involvement in these games be restricted to longitudinal, prospective designs in which subjects record details of their participation in real time. Further investigation is therefore required to assess the appropriateness of such designs for the collection of this information.

Lastly, while retrospective recall may be the only suitable approach for collection of detailed familial sporting history information, particularly for parents and older siblings, it appears that rather than relying on athlete recall, direct consultation with family members may be required to obtain accurate data. Since athlete recall of familial sporting history did not reach acceptable standards, experimental designs involving administration of questionnaires or conduct of interviews with all family members should be explored to gain an in-depth understanding of the role familial participation in sport and physical activity may play in the development of sport expertise. Although the DHAQ can collect reliable information related to a range of factors associated with athlete development, trustworthy methods for obtaining data related to a host of additional influences, including those outlined above plus others discussed throughout this thesis, are yet to be established.

Although items relating to participation in other organised sports were retained in the DHAQ, this thesis highlighted the need for researchers and practitioners to enter discussions regarding the nature of diversification in the development of sport expertise. Given that recall of details of participation in other organised sports in the current investigation was problematic, and previous studies in this area could potentially be based on misleading psychometrics, our understanding of the role of diversification in sport expertise development is called in to question. It is therefore essential to determine the level of involvement in organised sports required for participation to be considered beneficial for athlete development, so that associations between participation in other organised sports and expert performance can be re-visited and established.

The final set of future directions applies to the methodology and taxonomy developed for assessing validity and reliability of retrospective recall. This approach and classification system could be adopted for ongoing assessments of validity and reliability of the DHAQ, when considering new items for inclusion in the DHAQ, and/or when evaluating alternative methods of data collection for factors associated with the development of sport expertise that did not meet acceptable validity and reliability standards for retention in the DHAQ within the current series of studies. The methodology and taxonomy could also be utilised to establish validity and reliability of other measurement instruments designed to collect retrospective athlete developmental history data such as alternative practice history questionnaires or multidimensional interview guides; and, as mentioned previously, it may even be adapted for investigations involving retrospective recall in other domains including health and exercise psychology, or medicine. Therefore, the future directions of this research are, again, not restricted to studies involving the DHAQ, or even investigations related to the development of sport expertise.

It is clear that the construction and validation of the DHAQ provides the foundation for a wide range of future directions. These directions are both methodological and experimental; they apply both within and outside the domain of sport expertise development; and they relate to both the use of the DHAQ as well as alternative measurement instruments for the collection of athlete developmental history information. The impact of this thesis is, therefore, wide reaching, with the capacity to influence a large number of future investigations exploring a variety of research questions.

Although outside the scope of this thesis, studies utilising the DHAQ are currently underway, examining skill-, sport-, sex-, and country-based differences in the attainment of sporting milestones, practice history profiles, details of involvement in organised sports other than main sport, and family characteristics and participation in sport and physical activity. Data have been collected from over 650 Australian and Canadian athletes, representing more than 30 sports and three skill levels (elite, pre-elite, and non-elite). As evidenced by the list of publications and awards included elsewhere in this thesis, preliminary results from these investigations have been presented at international conferences; however, due to the detailed nature of the construction and validation of the DHAQ, further discussion of these subsequent studies is not provided here.

Conclusion

This chapter summarises the findings of this thesis, its contributions, and implications for future research.

Identification of the factors associated with sport expertise development is of particular interest to sport scientists, coaches, athletes, parents, and sport administrators, as this knowledge can aid in the creation of supportive, effective practice environments that promote optimal long-term athlete development. While sport expertise is the result of successful interactions between a wide range of influences, this thesis focussed specifically on working towards a comprehensive understanding of the contribution of practice and contextual factors to the development of sport expertise.

A brief review of the literature in this area highlighted a number of limitations in previous research that restricted the application of knowledge to practical settings. The most noticeable of these limitations were inconsistent or limited findings, and poorly validated measurement tools. Inconsistent findings and recommendations relating to primary contributors to sport expertise development and optimal developmental pathways appeared to be related to small, homogenous sample sizes, and the lack of a standardised measurement tool for the collection of athlete developmental history information. Further, validity and reliability assessments employed to confirm trustworthiness of data were found to be insufficient, and at times, involved the adoption of inappropriate statistics to confirm reliability and/or validity.

Thus, it was apparent that a rigorously validated, standardised measurement tool suitable for large scale collection of athlete developmental history information was required to overcome the limitations of previous research, and contribute to our understanding of the practice and environmental factors associated with the development of sport expertise. Therefore, this thesis aimed to construct and validate the DHAQ.

The DHAQ was constructed based upon the interview guides, questionnaires, results, and recommendations of previous research, and included ten sections relating to a variety of demographic, practice, competition, and environmental/contextual variables. A novel, objective, and comprehensive statistical methodology and taxonomy was developed to establish concurrent validity, convergent validity with parents, convergent validity with coaches, test-retest reliability, and internal consistency for all items within the questionnaire. Items were then identified as suitable for retention, requiring modification, or recommended for removal. Following this process, the DHAQ was updated, and converted to an online format to allow large scale distribution in the future.

As the results of the initial validity and reliability study led to the implementation of a number of major changes to the DHAQ, test-retest reliability of the updated, online measurement tool was reassessed to ensure the questionnaire continued to elicit high quality, trustworthy responses. Although test-retest reliability of the instrument increased markedly, several reliability issues remained following modification of the DHAQ, and several further

changes were recommended. Consequently, another round of validity and reliability testing is required to ensure maximum confidence in the trustworthiness of the measurement tool. Despite the need for further testing, at present the DHAQ could still be considered the most robust instrument available for the collection of athlete developmental history information. Therefore, it is suggested that the DHAQ could be utilised within hypothesis testing investigations of sport expertise development, providing appropriate validity and reliability checks are completed by the researchers, for the sample at hand.

The contributions of this thesis are both methodological and theoretical. The rigorous construction and validation of the DHAQ resulted in the provision of an online, quantitative measurement tool for the collection of athlete developmental history information, with known reliability for all items. Furthermore, a strict methodology and taxonomy for assessing validity and reliability of retrospective recall was established that, while most suitable for investigations of sport expertise development, may also be applied to studies of validity and reliability of measurement in other domains. Finally, the validation process led to a detailed understanding of the consistency with which various aspects of athlete developmental history information can be recalled. Such information is invaluable for interpreting results relating to sport expertise development and informing directions for future research in the area.

Most notably, the DHAQ can now start to be utilised in hypothesis testing investigations of sport expertise development, further contributing to our understanding of the factors that uniquely differentiate international level athletes from lesser skilled sport participants. Given the detailed procedures adopted within the current series of studies, it is proposed that the DHAQ be considered the emergent standard measurement tool for collection of athlete developmental history information.

References

Abernethy, B., Côté, J., & Baker, J. (2002). *Expert decision-making in team sports. Report to the Australian Sports Commission.* Brisbane, Australia: University of Queensland.

Ainsworth, B. E., Richardson, M. T., Jacobs Jr., D. R., Leon, A. S., & Sternfeld, B. (1999). Accuracy of recall of occupational physical activity by questionnaire. *Journal of Clinical Epidemiology, 52*(3), 219-227.

Aizpurua, A., & Koutstaal, W. (2010). Autobiographical memory and flexible remembering: Gender differences. *World Academy of Science, Engineering and Technology, 42*, 1617-1623.

Altman, D. G., & Bland, J. M. (1983). Measurement in medicine: The analysis of method comparison studies. *The Statistician, 32*(3), 307-317.

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine, 26*(4), 217-238.

Bahrick, H. P., Hall, L. K., & Berger, S. A. (1996). Accuracy and distortion in memory for high school grades. *Psychological Science, 7*(5), 265-271.

Bailey, R., Collins, D., Ford, P., MacNamara, Á., Toms, M., & Pearce, G. (2010). *Participant development in sport: An academic review.* Leeds, England: Sports Coach UK.

Baker, J. (2003). Early specialization in youth sport: A requirement for adult expertise? *High Ability Studies, 14*(1), 85-94.

Baker, J., & Côte, J. (2006). Shifting training requirements during athlete development: The relationship among deliberate practice, deliberate play, and other sport involvement in the acquisition of sport expertise. In D. Hackfort & G. Tenenbaum (Eds.), *Essential processes for attaining peak performance* (pp. 92-109). Oxford, United Kingdom: Meyer & Meyer Sport.

Baker, J., Côté, J., & Abernethy, B. (2003a). Learning from the experts: Practice activities of expert decision makers in sport. *Research Quarterly for Exercise and Sport, 74*(3), 342-347.

Baker, J., Côté, J., & Abernethy, B. (2003b). Sport-specific practice and the development of expert decision-making in team ball sports. *Journal of Applied Sport Psychology, 15*, 12-25.

Baker, J., Côté, J., & Deakin, J. (2005). Expertise in ultra-endurance triathletes early sport involvement, training structure, and the theory of deliberate practice. *Journal of Applied Sport Psychology, 17*, 64-78.

Baker, J., & Horton, S. (2004). A review of primary and secondary influences on sport expertise. *High Ability Studies, 15*(2), 211-228.

Baker, J., & Logan, A. J. (2007). Developmental contexts and sporting success: Birth date and birthplace effects in national hockey league draftees 2000-2005. *British Journal of Sports Medicine, 41*, 515-517.

Baker, J., Schorer, J., Cobley, S., Schimmer, G., & Wattie, N. (2009). Circumstantial development and athletic excellence: The role of birth date and birth place. *European Journal of Sport Science, 9*(6), 329-339.

Balyi, I., Cardinal, C., Higgs, C., Norris, S., & Way, R. (2005). *Canadian Sport for Life: Long-Term Athlete Development resource paper*. Vancouver, Canada: Canadian Sport Centres.

Barnsley, R. H., Thompson, A. H., & Barnsley, P. E. (1985). Hockey success and birthdate: The RAE. *Canadian Association of Health, Physical Education & Recreation, 51*(8), 23-28.

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports, 19*, 3-11.

Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin, 83*(5), 762-765.

Berry, J., Abernethy, B., & Côté, J. (2008). The contribution of structured activity and deliberate play to the development of expert perceptual and decision-making skill. *Journal of Sport & Exercise Psychology, 30*(6), 685-708.

Bjork, R. A., & Whitten, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology, 6*, 173-189.

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet, 1*(8746), 307-310.

Bloom, B. S. (Ed.). (1985). *Developing talent in young people*. New York: Ballantine Books.

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. *Cognitive Brain Research, 25*(1), 107-116.

Bouchard, C., Malina, R., & Perusse, L. (1997). *Genetics of fitness and physical performance*. Champaign, IL: Human Kinetics.

Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science, 236*(4798), 157-161.

Bray, M. S., Hagberg, J. M., Perusse, L., Rankinen, T., Roth, S. M., Wolfarth, B., Bouchard, C. (2009). The human gene map for performance and health-related fitness phenotypes: The 2006-2007 update. *Medicine & Science in Sports & Exercise, 41*(1), 35-73.

Brown, K. W., Cozby, P. C., Kee, D. W., & Worden, P. E. (1999). Developmental and single-subject designs. *Research methods in human development* (2nd ed., pp. 137-152). Mountain View, CA: Mayfield Publishing Company.

Bruner, M. W., Erickson, K., McFadden, K., & Côté, J. (2009). Tracing the origins of athlete development models in sport: A citation path analysis. *International Review of Sport and Exercise Psychology, 2*(1), 23-37.

Bruner, M. W., Erickson, K., Wilson, B., & Côté, J. (2010). An appraisal of athlete development models through citation network analysis. *Psychology of Sport and Exercise, 11*(2), 133-139.

Bryant, F. B. (2000). Assessing the validity of measurement. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 99-146). Washington, DC: American Psychological Association.

Bullock, N., Gulbin, J. P., Martin, D. T., Ross, A., Holland, T., & Marino, F. (2009). Talent identification and deliberate programming in skeleton: Ice novice to winter Olympian in 14 months. *Journal of Sports Sciences, 27*(4), 397-404.

Burdock, E. I., Fleiss, J. L., & Hardesty, A. S. (1963). A new view of inter-observer agreement. *Personnel Psychology, 16*(4), 373-384.

Carlson, R. (1988). The socialization of elite tennis players in Sweden: An analysis of the players' backgrounds and development. *Sociology of Sport Journal, 5*(3), 241-256.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Thousand Oaks, CA: Sage Publications.

Clark, R. D. (1982). Birth order and eminence: A study of elites in science, literature, sports, acting and business. *International Review of Modern Sociology, 12*(2), 273-289.

Cobley, S. P., Schorer, J., & Baker, J. (2008). Relative age effects in professional German soccer: A historical analysis. *Journal of Sports Sciences, 26*(14), 1531-1538.

Cobley, S., Baker, J., Wattie, N., & McKenna, J. (2009). Annual age-grouping and athlete development. *Sports Medicine, 39*(3), 235-256.

Costa-Santos, C., Bernardes, J., Ayres-de-Campos, D., Costa, A., & Costa, C. (2011). The limits of agreement and the intraclass correlation coefficient may be inconsistent in the interpretation of agreement. *Journal of Clinical Epidemiology, 64*(3), 264-269.

Côté, J. (1999). The influence of the family in the development of talent in sport. *The Sport Psychologist, 13*(4), 395-417.

Côté, J., Baker, J., & Abernethy, B. (2007). Practice and play in the development of sport expertise. In G. Tenenbaum, & R. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 184-202). Hoboken, NJ: John Wiley & Sons, Inc.

Côté, J., Ericsson, A. K., & Law, M. P. (2005). Tracing the development of athletes using retrospective interview methods: A proposed interview and validation procedure for reported information. *Journal of Applied Sport Psychology, 17*, 1-19.

Côté, J., & Fraser-Thomas, J. (2007). Youth involvement in sport. In P. R. E. Crocker (Ed.), *Sport psychology: A Canadian perspective* (pp. 266-294). Toronto, Canada: Pearson Prentice Hall.

Côté, J., & Hay, J. (2002). Children's involvement in sport: A developmental perspective. In J. M. Silva III, & D. E. Stevens (Eds.), *Psychological foundations of sport* (pp. 484-502). Boston: Allyn and Bacon.

Côté, J., Lidor, R., & Hackfort, D. (2009). ISSP position stand: To sample or to specialize? Seven postulates about youth sport activities that lead to continued participation and elite performance. *International Journal of Sport & Exercise Psychology, 7*(1), 7-17.

Côté, J., MacDonald, D. J., Baker, J., & Abernethy, B. (2006). When "where" is more important than "when": Birthplace and birthdate effects on the achievement of sporting expertise. *Journal of Sports Sciences, 24*(10), 1065-1073.

Côté, J., Salmela, J., Trudel, P., Baria, A., & Russell, S. (1995). The coaching model: A grounded assessment of expert gymnastic coaches' knowledge. *Journal of Sport & Exercise Psychology, 17*(1), 1-17.

Coughlin, S. (1990). Recall bias in epidemiologic studies. *Journal of Clinical Epidemiology, 43*(1), 87-91.

Curtis, J. E., & Birch, J. S. (1987). Size of community of origin and recruitment to professional and Olympic hockey in North America. *Sociology of Sport Journal, 4*(3), 229-244.

Deakin, J. M., & Cobley, S. (2003). A search for deliberate practice: An examination of the practice environments in figure skating and volleyball. In J. L. Starkes, & A. K. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 115-135). Champaign, IL: Human Kinetics.

Dex, S. (1995). The reliability of recall data: A literature review. *Bulletin De Méthodologie Sociologique, 49*, 58-89.

Duffy, M. E. (2002). Methodological issues in web-based research. *Journal of Nursing Scholarship, 34*(1), 83-88.

Durand-Bush, N., & Salmela, J. H. (2002). The development and maintenance of expert athletic performance: Perceptions of World and Olympic champions. *Journal of Applied Sport Psychology, 14*(3), 154-171.

Ebel, R. L. (1951). Estimation of the reliability of ratings. *Psychometrika, 16*, 407-424.

Ericsson, K. A., & Delaney, P. F. (1998). Working memory and expert performance. *Working memory and thinking* (pp. 91-111). East Essex, UK: Psychology Press Ltd.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363-406.

Ford, P. R., Ward, P., Hodges, N. J., & Williams, A. M. (2009). The role of deliberate practice and play in career progression in sport: The early engagement hypothesis. *High Ability Studies, 20*(1), 65-75.

Ford, P. R., & Williams, A. M. (2008). The effect of participation in Gaelic football on the development of Irish professional soccer players. *Journal of Sport & Exercise Psychology, 30*(6), 709-722.

Fraser-Thomas, J., Côté, J., & Deakin, J. (2008). Examining adolescent sport dropout and prolonged engagement from a developmental perspective. *Journal of Applied Sport Psychology, 20*(3), 318-333.

Gabbett, T. J. (2002). Physiological characteristics of junior and senior rugby league players. *British Journal of Sports Medicine, 36*(5), 334-339.

Gagné, F. (2004). Transforming gifts into talents: The DMGT as a developmental theory. *High Ability Studies, 15*(2), 119-147.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., & Smith, G. D. (2006). Indicators of socioeconomic position (part 1). *Journal of Epidemiology and Community Health, 60*(1), 7-12.

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences*. London: MacMillion and Co.

Garson, G. D. (2012). *Reliability*. Asheboro, NC: Statistical Associates Publishing.

Gibbons, T., Hill, R., McConnell, A., Forster, T., & Moore, J. (2002). *The path to excellence: A comprehensive view of development of U.S. Olympians who competed from 1984-1998.* United States Olympic Committee.

Gould, D., Dieffenbach, K., & Moffett, A. (2002). Psychological characteristics and their development in Olympic champions. *Journal of Applied Sport Psychology, 14*(3), 172-204.

Gould, D., Eklund, R. C., & Jackson, S. A. (1992). 1988 U.S. Olympic wrestling excellence: I. Mental preparation, precompetitive cognition, and affect. *Sport Psychologist, 6*(4), 358-382.

Gulbin, J. P., Oldenziel, K. E., Weissensteiner, J. R., & Gagné, F. (2010). A look through the rear view mirror: Developmental experiences and insights of high performance athletes. *Talent Development & Excellence, 2*(2), 149-164.

Haggard, E. A. (1958). *Intraclass correlation and the analysis of variance*. New York: The Dryden Press.

Hauser, R. M. (1994). Measuring socioeconomic status in studies of child development. *Child Development, 65*(6), 1541-1545.

Helsen, W. F., Starkes, J. L., & Hodges, N. J. (1998). Team sports and the theory of deliberate practice. *Journal of Sport & Exercise Psychology, 20*, 12-34.

Helsen, W. F., van Winckel, J., & Williams, A. M. (2005). The relative age effect in youth soccer across Europe. *Journal of Sports Sciences, 23*(6), 629-636; 629.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods, 1998*(1), 104-121.

Hoare, D. G. (2000). Predicting success in junior elite basketball players — the contribution of anthropometric and physiological attributes. *Journal of Science and Medicine in Sport, 3*(4), 391-405.

Hodge, T., & Deakin, J. M. (1998). Deliberate practice and expertise in the martial arts: The role of context in motor recall. *Journal of Sport & Exercise Psychology, 20*, 260-279.

Hodges, N. J., Kerr, T., Starkes, J. L., Weir, P. L., & Nananidou, A. (2004). Predicting performance times from deliberate practice hours for triathletes and swimmers: What, when, and where is practice important? *Journal of Experimental Psychology: Applied, 10*(4), 219-237.

Hodges, N. J., & Starkes, J. L. (1996). Wrestling with the nature of expertise: A sport specific test of Ericsson, Krampe and Tesch-Römer's (1993) theory of "deliberate practice". *International Journal of Sport Psychology, 27*, 400-424.

Hopkins, W. G. (2000a). *A new view of statistics: P values and statistical significance.* Retrieved May 15, 2011, from http://www.sportsci.org/resource/stats

Hopkins, W. G. (2000b). Measures of reliability in sports medicine and science. *Sports Medicine, 30*(1), 1-15.

Kay, T. (2000). Sporting excellence: A family affair? *European Physical Education Review, 6*(2), 151-169.

Keogh, J. (1999). The use of physical fitness scores and anthropometric data to predict selection in an elite under 18 Australian rules football team. *Journal of Science and Medicine in Sport, 2*, 125-133.

Keogh, J. W. L., Weber, C. L., & Dalton, C. T. (2003). Evaluation of anthropometric, physiological, and skill-related tests for talent identification in female field hockey. *Canadian Journal of Applied Physiology, 28*(3), 397-409.

Kottner, J., Audigé, L., Brorson, S., Donner, A., Gajewski, B. J., Hróbjartsson, A., Roberts, C., Shoukri, M., Streiner, D. L. (2011) Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Journal of Clinical Epidemiology, 64*(1), 96-106.

Kottner, J., & Dassen, T. (2008). Interpreting interrater reliability coefficients of the Braden scale: A discussion paper. *International Journal of Nursing Studies, 45*(8), 1238-1246.

Krebs, D. E. (1986). Declare your ICC type. *Physical Therapy, 66*(9), 1431.

Landers, D. M. (1979). Birth order in the family and sport participation. In M. L. Krotee (Ed.), *Dimensions of sport sociology* (pp. 140-167). West Point, NY: Leisure Press.

Lauer, L., Gould, D., Roman, N., & Pierce, M. (2010). Parental behaviors that affect junior tennis player development. *Psychology of Sport and Exercise, 11*(6), 487-496.

Law, M. P., Côté, J., & Ericsson, K. A. (2007). Characteristics of expert development in rhythmic gymnastics: A retrospective study. *International Journal of Sport & Exercise Psychology, 5*(1), 82-103.

Lemmink, K. A. P. M., Elferink-Gemser, M. T., & Visscher, C. (2004). Evaluation of the reliability of two field hockey specific sprint and dribble tests in young field hockey players. *British Journal of Sports Medicine, 38*, 138-142; 138.

MacDonald, D. J., Cheung, M., Côté, J., & Abernethy, B. (2009). Place but not date of birth influences the development and emergence of athletic talent in American football. *Journal of Applied Sport Psychology, 21*(1), 80-90.

MacDonald, D. J., Horton, S., Kraemer, K., Weir, P., Deakin, J. M., & Côté, J. (2009). Application and reliability of the retrospective interview procedure to trace physical activity patterns in master athletes and nonactive older adults. *Educational Gerontology, 35*, 1107-1122.

MacDonald, D. J., King, J., Côté, J., & Abernethy, B. (2009). Birthplace effects on the development of female athletic talent. *Journal of Science and Medicine in Sport, 12*(1), 234-237.

Manfreda, K. L., Batagelj, Z., & Vehovar, V. (2002). Design of web survey questionnaires: Three basic experiments. *Journal of Computer-Mediated Communication, 7*(3). Retrieved from http://jcmc.indiana.edu/vol7/issue3/vehovar.html

Memmert, D., Baker, J., & Bertsch, C. (2010). Play and practice in the development of sport-specific creativity in team ball sports. *High Ability Studies, 21*(1), 3-18.

Moesch, K., Elbe, A. -M., Hauge, M. -L. T., & Wikman, J. M. (2011). Late specialization: The key to success in centimeters, grams, or seconds (cgs) sports. *Scandinavian Journal of Medicine & Science in Sports, 21*(6), e282-e290.

Morrow Jr., J. R., & Jackson, A. W. (1993). How "significant" is your reliability? *Research Quarterly for Exercise and Sport, 64*(3), 352-355.

Müller, R., & Büttner, P. (1994). A critical discussion of intraclass correlation coefficients. *Statistics in Medicine, 13*, 2465-2476.

Murray, P. (1999). Fundamental issues in questionnaire design. *Accident and Emergency Nursing, 7*(3), 148-153.

Musch, J., & Grondin, S. (2001). Unequal competition as an impediment to personal development: A review of the relative age effect in sport. *Developmental Review, 21*(2), 147-167.

Oldenziel, K., Gagné, F., & Gulbin, J. (2003). *How do elite athletes develop? A look through the 'rear-view mirror': A preliminary report from the national athlete development survey (NADS).* Canberra, Australia: Australian Sports Commission.

Oppenheim, A. N. (1992). *Questionnaire design, interviewing and attitude measurement*. London: Continuum.

Orlick, T., & Partington, J. (1988). Mental links to excellence. *The Sport Psychologist, 2*(2), 105-130.

Phillips, E., Davids, K., Renshaw, I., & Portus, M. (2010a). The development of fast bowling experts in Australian cricket. *Talent Development & Excellence, 2*(2), 137-148.

Phillips, E., Davids, K., Renshaw, I., & Portus, M. (2010b). Expert performance in sport and the dynamics of talent development. *Sports Medicine, 40*(4), 271-283.

Pienaar, A. E., Spamer, M. J., & Steyn Jr., H. S. (1998). Identifying and developing rugby talent among 10-year-old boys: A practical model. *Journal of Sports Sciences, 16*(8), 691-699.

Puthucheary, Z., Skipworth, J. R. A., Rawal, J., Loosemore, M., Someren, K. V., & Montgomery, H. E. (2011). Genetic influences in sport and physical performance. *Sports Medicine, 41*(10), 845-859.

Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of Clinical Nursing, 16*(2), 234-243.

Reis, J. P., Dubose, K. D., Ainsworth, B. E., Macera, C. A., & Yore, M. M. (2005). Reliability and validity of the occupational physical activity questionnaire. *Medicine and Science in Sports & Exercise, 37*(12), 2075-2083.

Ropponen, A., Levalahti, E., Simonen, R., Videman, T., & Battie, M. C. (2001). Repeatability of lifetime exercise reporting. *Scandinavian Journal of Medicine and Science in Sports, 11*(185-192), 185-192.

Rubin, D. C. (Ed.). (1986). *Autobiographical memory*. Cambridge: Cambridge University Press.

Rubin, D. C. (Ed.). (1995). *Remembering our past: Studies in autobiographical memory.* Cambridge: Cambridge University Press.

Salmela, J. H., & Moraes, L. C. (2003). Development of expertise: The role of coaching, families, and cultural contexts. In J. L. Starkes, & K. A. Ericsson (Eds.), *Expert performance in sports: Advances in research on sport expertise* (pp. 275-293). Champaign, IL: Human Kinetics.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420-428.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61*, 394-403.

Singer, R. N., & Janelle, C. M. (1999). Determining sport expertise: From genes to supremes. *International Journal of Sport Psychology, 30*, 117-150.

Sloane, K. D. (1985). Home influences on talent development. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 439-476). New York: Ballantine Books.

Smith, D.,J. (2003). A framework for understanding the training process leading to elite performance. *Sports Medicine, 33*(15), 1103-1126.

Soberlak, P., & Côté, J. (2003). The developmental activities of elite ice hockey players. *Journal of Applied Sport Psychology, 15*, 41-49.

Sosniak, L. A. (2006). Retrospective interviews in the study of expertise and expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 287-301). Cambridge, England: Cambridge University Press.

Starkes, J. L., Deakin, J. M., Allard, F., Hodges, N. J., & Hayes, A. (1996). Deliberate practice in sports: What is it anyway? In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 81-106). Mahwah, NJ: Lawrence Erlbaum Associates.

Strube, M. J. (2000). Reliability and generalizability theory. In L. G. Grimm, & P. R. Yarnold (Eds.), *Reading and understanding more multivariate statistics* (pp. 23-66). Washington, DC: American Psychological Association.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn and Bacon.

Tucker, R., & Collins, M. (2012). What makes champions? A review of the relative contribution of genes and training to sporting success. *British Journal of Sports Medicine, 46*(8), 555-561.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352-373.

Vaeyens, R., Malina, R. M., Janssens, M., van Renterghem, B., Bourgois, J., Vrijens, J., Philippaerts, R. M. (2006). A multidisciplinary selection model for youth soccer: The Ghent youth soccer project. *British Journal of Sports Medicine, 40*, 928-934; 928.

van Gelder, M. M. H. J., Bretveld, R., W., & Roeleveld, N. (2010). Web-based questionnaires: The future in epidemiology? *American Journal of Epidemiology, 172*(11), 1292-1298.

van Selm, M., & Jankowski, N. (2006). Conducting online surveys. *Quality and Quantity, 40*, 435-456.

Ward, P., Hodges, N. J., Starkes, J. L., & Williams, A. M. (2007). The road to excellence: Deliberate practice and the development of expertise. *High Ability Studies, 18*(2), 119-153.

Wattie, N., Baker, J., Cobley, S., & Montelpare, W. J. (2007). A historical examination of relative age effects in Canadian hockey players. *International Journal of Sport Psychology, 38*, 178-186; 178.

Weissensteiner, J., Abernethy, B., & Farrow, D. (2009). Towards the development of a conceptual model of expertise in cricket batting: A grounded theory approach. *Journal of Applied Sport Psychology, 21*, 276-292.

Weissensteiner, J. R., Abernethy, B., Farrow, D., & Gross, J. (2012). Distinguishing psychological characteristics of expert cricket batsmen. *Journal of Science & Medicine in Sport, 15*(1), 74-79.

Whitten, W. B., & Leonard, J. M. (1981). Directed search through autobiographical memory. *Memory & Cognition, 9*(6), 566-579.

Wiersma, L. D. (2000). Risks and benefits of youth sport specialization: Perspectives and recommendations. *Pediatric Exercise Science, 12*(1), 13-22.

Williams, J. M. G., Barnhofer, T., Crane, C., Hermans, D., Raes, F., Watkins, E., Dalgeish, T. (2007). Autobiographical memory specificity and emotional disorder. *Psychological Bulletin, 133*(1), 122-148.

Wyatt, J. C. (2000). When to use web-based surveys. *Journal of the American Medical Informatics Association, 7*(4), 426-430.

Young, B. W., & Salmela, J. H. (2002). Perceptions of training and deliberate practice of middle distance runners. *International Journal of Sport Psychology, 33*, 167-181.