Computational Approaches to the Assessment of Motor Competence in Children: Prediction, Interpretation and Markerless Measurement



Colm O'Donaghue, BSc, MSc

A thesis submitted for the award of Doctor of Philosophy (PhD)

School of Health and Human Performance
Dublin City University
Dublin
Ireland

Supervisors: Dr. Stephen Behan, Dr. Michael Scriney & Prof. Sarahjane Belton

Submitted to Dublin City University, August 2026

Authors Declaration:

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original and have conformed to the regulations on the use and declaration of Generative AI, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Colm O'Donaghue (Candidate) ID No.: 22267916 Date: 26/08/2026

Template for Reporting of Gen AI Tool Use in Research Theses

Tools: Clearly indicate the tool name and model (or version) you used and the date(s) of use. If you used a tool multiple times, provide the range of dates over which the tool was used and the model/versions used.	***'Google Gemini, 3.5 Flash', used between 01/11/2025 and 01/06/2026***	***'Google Gemini, 3.5 Flash', used between 01/11/2025 and 01/06/2026***	***'Google Gemini, 3.5 Flash', used between 01/11/2025 and 01/06/2026***	***'Claude Sonnet 4.6', used between 01/11/2025 and 01-06-2026***
Outputs: Specify what tool outputs were generated and used by you and for what purpose.	***Google Gemini, 3.5 Flash was used to assist in table design and layout***	***Google Gemini, 3.5 Flash was used to sense-check findings from our research***	***Google Gemini, 3.5 Flash was used to break up the thesis into sections with appropriate formaFng in scenarios where numerous paragraphs were excessively long and needed to be structured more appropriately***	***Claude Sonnet 4.6 was used to aid in the development of the classification system, specifically on the data loading aspects and optimisation and aiding with navigating the Polars and Pandas documentation***
Prompts: Provide a record of the prompts you used here or link to an appendix where the prompts used are listed.	***Gemini was provided with Jamovi outputs of the analyses and asked to provide suitable approaches for styling and presenting tables to demonstrate the findings***	***Gemini was provided with prompts relating to our research design and it provided feedback on the extent to which the findings were aligned with our chosen methodology***	***Gemini was provided with prompts such as "divide this paragraph into sections and space out accordingly"***	***What is the best panda's function to create a sliding window partitioning data across test, participant and trial***
Iterations: Provide some commentary on the number and type of iterations. This can be linked to your report on prompts and outputs above.				**Code was naturally developed in an iterative context with multiple rounds of testing at each stage manually validating the expected results**

Critical Reflection: If Gen AI tools have been used minimally, a very brief critical reflection on its role in the research submitted for				**While AI for writing code is suitable for most "boilerplate" tasks (such as navigating documentation) a human (and their understanding of the data) and wider programming**
examination can be included here. Otherwise, specify the parts of the thesis where critical reflection on the use and role of Gen AI tools is included.				**knowledge is required in order to deliver functioning software.**

Acknowledgements

I would like to acknowledge and thank the following people who have greatly supported me during the completion of my thesis and PhD journey more generally.

To my supervisors, Dr Stephen Behan, Dr Michael Scriney and Professor Sarahjane Belton, thank you for your constant patience, guidance, encouragement.

Stephen, thanks for simply taking a chance on me four years ago. You owed me absolutely nothing yet gave everything, whether it was a meeting, a call or a coffee you always made time for me. It was an absolute honor to have you as my lead supervisor, I'm beyond appreciative of the unbelievable support you've given me on this journey. You always challenged me to be better and get the best out of the work I did and above all else just to back myself. I certainly wouldn't have seen the finish line without your guidance and support.

Michael, thanks for taking me under your wing, I came in with no prior computing experience, yet I finished with a lot which is a testament to your guidance, your style of teaching but above all else your patience. These qualities made many a daunting task seem doable and for that I'll always be grateful to have worked with you.

Sarahjane, although we didn't meet until the very end of my journey, it was amazing to have you on board. You challenged me to become a more critical thinker, writer, researcher and for that I thank you. I simply couldn't have completed this journey without all three of you, I'll always be incredibly proud to say you were my supervisors.

I would like to thank all faculty and students in the Schools of Health and Human Performance and Computing for their unwavering support throughout my PhD. I particularly want to thank the HHP faculty for providing their time during data collection. Special thanks to Simon Devenney for providing access to the biomechanics laboratory. I would like to thank the wider ML Labs community for their support throughout my PhD, it was great to have a network of other students to chat with over the highs and lows of the journey with. In particular, I would like to acknowledge and thank centre manager Carla and CRT coordinator Vicky for the amazing support they provided me.

I would like to thank my friends and family for their unconditional support throughout the past few years. In particular, I'd like to thank my amazing mother Roisin for her unwavering encouragement, support and belief in me throughout the highs and lows of this journey. I simply wouldn't have gotten this far in life without your guidance, motivation, and patience, and the values you have instilled in me. Lastly, I'd like to pay homage to my late grandfather Pól who sadly passed in April. I had always hoped to celebrate this milestone in his company, but it wasn't to be, perhaps in another life we will. Miss you always grandad.

IN ÓMÓS DO

Pól Ó Donnchadh

1941 – 2026

"*Ní bheidh do leithéid ann arís*"

Table of Contents

Authors Declaration:.....ii
Template for Reporting of Gen AI Tool Use in Research Theses.....iii
Acknowledgements.....v
Table of Contents.....vii
List of Appendices.....x
List of Tables.....xi
List of Figures.....xii
Abstract.....xiii
Chapter 1: Introduction to Thesis.....1
1.1 Publications.....2
1.2 Introduction.....4
1.3 Philosophical Position.....10
1.4 Aims and Objectives.....13
1.5 Research Question and Hypothesis.....14
1.6 Thesis Structure.....15
1.7 Statement of Candidate Contribution.....17
1.8 Definition of Terms.....19
1.9 List of Abbreviations.....21
Chapter 2: Literature Review.....22
2.1 Introduction to Literature Review.....23
2.2 Physical Activity in Childhood and Adolescence.....23
2.3 Motor Competence and its Role in Physical Activity.....26
2.4 Actual Motor Competence.....30
2.4.1 Importance of Actual Motor Competence.....30
2.4.2 Measurement of Actual Motor Competence.....30
2.4.3 Current AMC Trends.....34
2.5 Perceived Motor Competence.....38
2.5.1 Importance of Perceived Motor Competence.....38
2.5.2 Measurement of Perceived Motor Competence.....39
2.5.3 Current PMC Trends.....42

2.6 The Association Between AMC and PMC 43
2.7 Analytical Approaches in Motor Competence Research 48
2.7.1 Traditional Statistical Methods 48
2.7.2 Summary and Future Directions 55
2.8 Machine Learning and Data-Driven Methods 56
2.8.1 Applications of Machine Learning in Motor Competence Research 57
2.8.2 Machine Learning Applications in TGMD Assessment 58
2.8.3 Movement Classification and Analysis in PA Research 69
2.8.4 Limitations of Machine Learning 74
2.9 Markerless Motion Capture; Overview and Applications 76
2.9.1 Can Markerless Motion Capture Aid AMC Assessment? 76
2.9.2 Towards Automated Motor Competence Assessment: A Computer Vision and Machine Learning Pipeline 82
2.10 Summary and Research Gaps 83
Link from Chapter 2 to Chapter 3 85
Chapter 3: Streamlining Motor Competence Assessments via a Machine Learning Approach 87
3.1 Abstract 88
3.2 Introduction 89
3.3 Methodology 93
3.4 Results 99
3.5 Discussion 104
Link from Chapter 3 to 4 109
Chapter 4: Modelling Motor Competence Interdependencies Using Post-hoc Model Interpretability and Graph Analytics 110
4.1 Abstract 111
4.2 Introduction 112
4.3 Methodology 114
4.4 Results 121
4.5 Conclusion 126
Chapter 5: Markerless Motion Capture for Motor Competence Assessment in Children: Promises, Limitations, and a Research Agenda 128
5.1 Abstract 129
5.2 Introduction 130
5.3 Position-paper Approach 131

5.4 Limitations of Traditional Motor Skill Assessments 132
5.6 Markerless Motion Capture: A New Paradigm for MC Assessment 135
5.7 Emerging Open-Source Markerless Platforms 138
5.8 Benefits of Markerless Motion Capture for MC Assessment 139
5.9 Challenges in Applying Markerless Motion Capture to Paediatric Populations 144
5.10 Applications in Research, Education, and Clinical Practice 147
Chapter 6: Evaluation of a Markerless 3D Motion Capture System for Locomotor Skill Assessment: A Pilot Study 152
6.1 Abstract 153
6.2 Introduction 154
6.3 Methodology 158
6.4 Results 168
6.5 Discussion 174
Chapter 7: Conclusions and Future Directions for Markerless Motion Capture Research 180
7.1 Overview 181
7.2 Key Findings 181
7.3 Integrated discussion 183
7.4 Theoretical and methodological contributions 185
7.5 Implications for research and practice 185
7.6 Strengths and limitations 186
7.7 Reflexive note 187
7.8 Conclusion 188
Bibliography 189
Appendices 232

List of Appendices

Appendix A DCU REC Formal Letter of Approval **233**

Appendix B TGMD-3 Criterion Operationalisation, Features and Heuristic Rules **234**

Appendix C Descriptive Statistics Across Camera Configurations **237**

Appendix D Intraclass Correlation Coefficient Absolute Agreement Results Across Skills and Camera Configurations **239**

Appendix E Rules Engine Performance by Skill, Camera Configuration and Feature **241**

List of Tables

Table 2.1 Sample of Statistical Models Adopted in Motor Competence Research....**49**
Table 2.2: Short form versions of the original TGMD assessment batteries **59**
Table 2.3: Sample of ML Methods Used for Classification of PA Type and Intensity **69**
Table 2.4: Sample of Measurement Tools and Statistical Methods Used for Objective Locomotor Skill Assessment....**77**
Table 3.1: Brief Synopsis of Features Analysed....**95**
Table 3.2 Example of Splitting Data by Gender Into several Subsets for Testing. **96**
Table 3.3 Top Performing Target Features—Gender Combined Subset....**99**
Table 3.4. Top Performing Target Features—Female-Only Subset....**99**
Table 3.5. Top Performing Target Features—Male-Only Subset....**100**
Table 3.6. Worst Performing Target Features—All Subsets. **100**
Table 4.1. Best Performing Classification Algorithms **118**
Table 4.2. Degree Centrality Results **122**
Table 4.3 Betweenes Centrality Results **122**
Table 4.4 PageRank Results **122**
Table 6.1. TGMD-3 Short Form (Duncan, Martins, et al., 2022) - Locomotor Subset....**159**
Table 6.2 Features Representing TGMD-3 Performance Criteria **163**
Table 6.3 Friedman Repeated-Measures ANOVA Results Across Locomotor Skills **169**
Table 6.4 Mean ICC(2,1) Values Across Camera Configurations **170**
Table 6.5 Mean Heuristics Rules Performance by Skill and Camera Configuration....**171**
Table 6.6 Feature-Level Classification Performance of Gallop Heuristics Across Camera Configurations **172**

List of Figures

Figure 2.1: Developmental mechanisms influencing the physical activity trajectories of children. Reproduced from Stodden et al. (2008)....**28**

Figure 2.2: Research consensus on motor competence and health related variables. Reproduced from Robinson et al. (2015)....**45**

Figure 2.3: Levels of evidence for each pathway, summarised by skill domain. Reproduced from Barnett et al. (2022)....**46**

Figure 2.4: Proposed extension of the Stodden et al. (2008) model. Reproduced from Lima et al. (2022)....**47**

Figure 2.5: Network analysis model. Reproduced from Duncan, Martins, et al. (2022)....**62**

Figure 2.6: Network analysis model. Reproduced from Bandeira et al. (2020)....**63**

Figure 2.7: Exploratory graph analysis and item replication in the EGA dimensions. Reproduced from Bandeira, Lemos, et al. (2025)....**65**

Figure 2.8: Multilayer network model of MC. Reproduced from Bandeira, Estevan, et al. (2025)....**66**

Figure 2.9: Multilayer network emergent patterns of MC. Reproduced from Bandeira, Estevan, et al. (2025)....**67**

Figure 3.1: Distribution difference of Throw AMC feature for Female Subset....**97**

Figure 3.2 Subset Comparison Amongst All Analysed AMC Features....**103**

Figure 3.3 Distribution of Subset MNM and Positive-Class F1 Score....**104**

Figure 4.1 Process for Modelling MC Inter-dependencies....**115**

Figure 4.2 SP-LIME & Graph Analytics Flowchart....**121**

Figure 4.3 12-Year-Old Female's Interdependencies....**125**

Figure 4.4 9-Year-Old Male's Interdependencies....**126**

Figure 6.1 Experimental Design....**160**

Figure 6.2 Calibration of Cameras with ChArUco Board....**162**

Figure 6.3 2D-to-3D Pose Triangulation....**167**

Name of Candidate: Colm O'Donaghue
Thesis Title: Computational Approaches to the Assessment of Motor Competence in Children: Prediction, Interpretation and Markerless Measurement

Abstract

Motor competence (MC), considered key in supporting child development, is associated with physical activity (PA) and broader lifelong health outcomes. However, traditional assessment and analytical approaches remain constrained by labour-intensive procedures that may not fully capture the complexity of MC development. Consequently, there is growing interest in computational approaches to address such limitations.

This thesis examined the application of computational methodologies within MC research across four related studies. Study 1 evaluated machine learning (ML) classification models for predicting actual motor competence based off remaining actual (AMC) and perceived motor competence (PMC) data. Results demonstrated that several MC outcomes could be accurately predicted, highlighting the potential of ML within MC research. Building on these findings, study 2 utilised post-hoc model interpretability and graph analytics to characterise interdependencies between AMC and PMC by age and gender. Findings highlighted influential skills and constructs within MC networks while reinforcing the importance of PMC as an important correlate of broader MC development.

The third study, a position paper, critically evaluated the opportunities, limitations, and future directions of markerless motion-capture technologies for MC assessment. The review noted the potential of computer vision systems to enhance the scalability and objectivity of assessment while considering methodological and practical limitations requiring further investigation. In response, study 4 explored the feasibility of integrating an open-source markerless motion-capture pipeline ('*FreeMoCap*') with a heuristics-based rules engine to automate classification of TGMD-3 Short Form locomotor performance criteria. Findings suggested that classification performance varied between performance criteria but remained relatively stable across camera configurations. While reducing camera quantity influenced kinematic data quality, a three-camera configuration provided a practical balance between data quality, classification performance, and deployment considerations.

Collectively, this thesis demonstrates the value of computational techniques for advancing MC research and provides a foundation for future technology-enhanced assessment and analysis.

Chapter 1: Introduction to Thesis

1.1 Publications

Published:

Streamlining Motor Competence Assessments via a Machine Learning Approach.
MDPI Youth – accepted 25th June 2025, published online 6th July 2025.
Authors: Colm O'Donaghue, Michael Scriney, Sarahjane Belton & Stephen Behan

Modelling Motor Competence Interdependencies: An XAI and Graph Analytics Approach.
2025 IEEE International Workshop on Sport, Technology and Research – published online 3rd December 2025.
Authors: Colm O'Donaghue, Stephen Behan, Sarahjane Belton & Michael Scriney

Under Review:

Markerless Motion Capture for Motor Competence Assessment in Children: Promises, Limitations, and a Research Agenda
Authors: Colm O'Donaghue, Michael Scriney, Sarahjane Belton & Stephen Behan

Evaluation of a Markerless 3D Motion Capture System for Locomotor Skill Assessment: A Pilot Study
Authors: Colm O'Donaghue, Stephen Behan, Sarahjane Belton & Michael Scriney

Oral Presentations:

'Modelling Motor Competence Interdependencies: An XAI and Graph Analytics Approach'
2025 IEEE International Workshop on Sport, Technology and Research – October 2025, Predazzo Ski Jumping Stadium, Predazzo, Italy

Poster Presentations:

'The Accuracy of Several Classification Algorithms to Predict Movement Competence'
DCU Faculty of Engineering and Computing Research Day – April 2023, DCU, Dublin, Ireland

'Using Machine Learning to Solve Ireland's 'Movement Crisis'
Irish Strength & Conditioning Network Conference – May 2023, University of Galway, Galway, Ireland

‘Investigating The Relationship Between Actual and Perceived Movement Competence’
7th Annual All-Ireland Postgraduate (AIPG) Conference in Sport Sciences, Physical Activity & Physical Education – May 2023, MTU, Cork, Ireland

‘Modelling Motor Competence Interdependencies Using Explainable Artificial Intelligence and Graph Analytics’
8th AIPG Conference in Sport Sciences, Physical Activity & Physical Education – May 2024, ATU Galway City, Galway, Ireland

‘Objectifying Motor Control Assessment via Markerless Motion Capture’
DCU Faculty of Engineering and Computing Research Day – October 2024, DCU, Dublin, Ireland

‘Objectifying Motor Control Assessment via Markerless Motion Capture’
Irish Sport and Exercise Sciences Association Conference – May 2025, SETU, Waterford, Ireland

‘Optimising Camera Configuration for Markerless Locomotor Motion Capture’
Irish Sport and Exercise Sciences Association Conference – May 2026, Technological University of the Shannon: Midlands, Athlone, Ireland

1.2 Introduction

Physical inactivity among children and adolescents has emerged as one of the most significant public health challenges of the twenty-first century (Ding et al., 2016). Despite widespread recognition of the physical, psychological and social benefits of regular participation in physical activity (PA) (Eime et al., 2013), global levels of youth PA remain consistently below recommended guidelines (Aubert et al., 2021). Findings arising from the most recent Irish Children's Sport Participation and Physical Activity (CSPPA) indicate that only 23% of children attending primary school and 12% of post-primary children met the physical activity guidelines of at least 60 minutes of daily moderate-to-vigorous physical activity (MVPA) (Woods et al., 2023). Worryingly, these low levels of PA adherence amongst the youth demographic are not unique to Ireland, with findings reporting 81% of youth aged 11 – 17 years globally failed to meet the 2010 World Health Organisation (WHO) recommendation of accumulating at least 60 minutes of MVPA daily in 2016 (Guthold et al., 2020). Consequently, understanding the factors that influence children's participation in PA remains an important priority for researchers, practitioners, and policymakers alike (Hu et al., 2021; Sallis et al., 2000; Tannehill et al., 2015). Given the well-established links between PA and physical, psychological, and social health outcomes (Eime et al., 2013), identifying the factors that promote lifelong engagement in PA is essential for improving population health.

Participation in PA is influenced by a range of individual, social, and environmental factors (Hu et al., 2021). At a societal level, opportunities to engage in PA must be accessible and inclusive. Socioeconomic status can influence participation, with young people living in disadvantaged areas often reporting barriers to engagement in PA (Chinn et al., 1999). Family can also play a pivotal role in promoting and supporting PA engagement, with parents frequently influencing enrolment in sport clubs and other activities (Sallis et al., 2000). Indeed, the 2022 CSPPA report recommended that coaches, parents and volunteers receive appropriate training regarding the importance of PA and effective strategies for encouraging regular participation among children and adolescents (Woods et al., 2023). These recommendations highlight the importance of understanding the broader factors that shape children's engagement in PA. One framework that has received increasing attention in recent years for explaining PA engagement is physical literacy (PL). A growing body of research has explored the extent to which PL supports participation in PA across the lifespan (Cairney,

Dudley, et al., 2019; Cornish et al., 2020). The Australian physical Literacy Framework defines PL as *'lifelong holistic learning acquired and applied in movement and physical activity contexts, integrating physical, psychological, social, and cognitive capabilities that support participation in health-promoting movement throughout life'* (*Sport Australia*, 2019). Similarly, (Whitehead, 2013) conceptualises PL as encompassing physical competence, confidence, motivation, and the knowledge and understanding necessary to value and engage in PA. Collectively, these attributes are believed to support an individual's capacity and desire to participate in PA throughout the lifespan.

Despite widespread adoption of PL within research, policy, and practice (Carl et al., 2023; Corbin, 2016; Dudley et al., 2017), considerable debate remains regarding its conceptualisation and measurement (Barnett et al., 2019, 2023; Edwards et al., 2018). While many scholars advocate for comprehensive approaches that capture the multidimensional nature of PL (Britton, Belton, et al., 2023; Cairney, Clark, et al., 2019; Cairney, Dudley, et al., 2019), others argue that attempts to operationalise and quantify PL may conflict with its original philosophical foundations (Barnett et al., 2023). Consequently, researchers continue to grapple with questions surrounding how PL should be defined and assessed (Grauduszus et al., 2023; Liu & Chen, 2021). Nevertheless, there is broad agreement that understanding the relationships between the various domains of PL is important for explaining PA participation and informing intervention development (Cairney, Clark, et al., 2019; Cairney, Dudley, et al., 2019).

Within the broad construct of PL, actual motor competence (AMC) has emerged as one of the most extensively studied and measurable domains (Barnett et al., 2022; Cattuzzo et al., 2016; Haga, 2008; Hill et al., 2024; Lopes et al., 2020; Robinson et al., 2015; Stodden et al., 2008). AMC refers to an individual's ability to execute a range of movement skills and is widely recognised as an important determinant of PA participation (Logan et al., 2018). As one of the first large-scale studies of PL conducted in Ireland, Behan et al. (2019) assessed multiple dimensions of PL across a broad age range using validated, age-appropriate measures. Within this framework, particular emphasis was placed on AMC due to its central role in movement development and PA engagement (Stodden et al., 2008). AMC is commonly assessed through the evaluation of fundamental movement skills (FMS), which represent the foundational movement patterns, such as running, kicking, jumping, and throwing, that underpin participation in PA, sport, and recreation (Logan et al., 2018). The development of FMS during

childhood is considered particularly important due to their foundational role in facilitating lifelong engagement in PA and sport (Logan et al., 2015). Early childhood (2–5 years) has been identified as a critical period for the acquisition of FMS and the development of overall AMC (Foulkes et al., 2015; Morley et al., 2015). Children are said to be capable of achieving proficiency in these foundational skills by approximately seven years of age, enabling participation in more specialised sport- and locomotor-related movements (Bolger et al., 2021).

Contrary to the belief that such skills develop naturally through maturation alone, evidence suggests that FMS require instruction, practice, and reinforcement (Clark, 2007; Cools et al., 2009; Haywood & Getchell, 2019; Stodden et al., 2008). Successful development of FMS is particularly important given the central role AMC is believed to play within children's developmental trajectories (Stodden et al., 2008). Children possessing greater levels of AMC are more likely to participate in PA throughout childhood and adolescence, while those with lower levels of AMC may be less inclined to engage due to difficulties performing the skills required for participation (Barnett et al., 2009; Hardy et al., 2013; Jaakkola et al., 2016; Logan et al., 2015; Seefeldt & Gould, 1980; Stodden et al., 2008). However, AMC is not thought to operate in isolation. Rather, it exists within a broader developmental system involving perceived motor competence (PMC), health-related fitness (HRF), and PA participation itself (Stodden et al., 2008). The conceptual model proposed by Stodden et al. (2008) provides one of the most influential frameworks for understanding these relationships. The model suggests that AMC, PMC, HRF, and PA interact dynamically across childhood and adolescence, producing either positive or negative developmental trajectories. Children with higher levels of AMC are proposed to develop greater confidence in their movement abilities, engage in more PA, improve their fitness, and subsequently further enhance their motor competence (MC). Conversely, children with lower levels of AMC may experience reduced confidence, lower participation in PA, and poorer fitness outcomes, potentially resulting in a negative developmental spiral (Stodden et al., 2008).

Consequently, understanding the mechanisms underpinning AMC and its relationship with related constructs remains an important area of investigation (Bandeira, Estevan, et al., 2025; Bandeira, Lemos, et al., 2025; Lima et al., 2022). PMC, defined as an individual's perception of their own motor-skill capabilities (Estevan & Barnett, 2018; Harter, 1982), has emerged as a particularly important factor within this developmental framework (Stodden et al., 2008).

Similar to AMC, higher levels of PMC have been associated with increased motivation, persistence, and engagement in movement-related activities, while lower levels have been linked to reduced participation and disengagement from PA (Rudisill et al., 1993; Weiss & Amorose, 2005). Research has consistently demonstrated positive associations between PMC, motivation, and PA participation (Bagøien & Halvari, 2005; Baker & Davison, 2011; Barnett et al., 2008; Carcamo-Oyarzun et al., 2025; De Meester, Maes, et al., 2016; De Meester, Stodden, et al., 2016; Den Uil et al., 2023). Indeed, PMC has been suggested to exert a greater influence on motivation to engage in PA than AMC itself (Bardid et al., 2016; Harter, 1978, 1982; Welk, 1999). Within the model proposed by Stodden et al. (2008), PMC is positioned as a key mediator between AMC and PA, with this relationship strengthening throughout childhood and adolescence.

Higher levels of PMC have also been associated with greater PA participation, improved fitness, and more favourable body composition outcomes (Barnett et al., 2008; Coppens et al., 2021; Den Uil et al., 2023; Jones et al., 2010; Sallis et al., 2000). PMC is influenced by factors such as prior experiences, perceived task difficulty, social reinforcement, and intrinsic motivation (Estevan & Barnett, 2018; Harter, 1978). However, young children often lack cognitive maturity to accurately evaluate these factors and frequently overestimate their motor-skill capabilities (Almeida et al., 2017; Harter, 1978; Morano et al., 2020). As cognitive development progresses during middle childhood, PMC gradually becomes aligned with AMC (Estevan & Barnett, 2018; Harter, 1978; Robinson et al., 2015; Stodden et al., 2008). However, PMC and AMC remain conceptually distinct, and agreement between the two is not always observed (De Meester et al., 2020; Estevan & Barnett, 2018).

Collectively, the literature suggests that AMC represents a central component of children's PL and plays an important role in shaping participation in PA across childhood and adolescence (Cairney, Clark, et al., 2019; Stodden et al., 2008). Furthermore, the developmental relationships between AMC, PMC, HRF, and PA are increasingly recognised as complex, dynamic, and reciprocal rather than linear in nature (Lima et al., 2022; Stodden et al., 2008). Despite growing evidence supporting the importance of these constructs, significant challenges remain regarding both their measurement and interpretation. From a measurement perspective, researchers continue to debate the most appropriate methods for assessing AMC (Barnett et al., 2023; Fransen et al., 2014; Hulteen et al., 2020, 2023; True et al., 2017). Although numerous process oriented and product-oriented assessment tools have been

developed, concerns remain regarding objectivity, assessor burden, scalability, and the ability to capture meaningful aspects of movement competence across diverse populations and settings (Bandeira et al., 2020; Duncan, Martins, et al., 2022; Hulteen et al., 2023; Valentini et al., 2018). Consequently, there remains a need for assessment approaches that can provide objective information regarding movement quality while simultaneously improving efficiency, consistency, and accessibility (Lander et al., 2020).

In addition to measurement challenges, many of the relationships proposed within contemporary MC developmental models may not be fully captured using traditional analytical approaches, which are sensitive to the synergic, dynamic, and nonlinear interactions between MC constructs (Salami et al., 2024). The interactions between AMC, PMC, HRF, and PA are likely characterised by multiple interdependencies and potentially non-linear relationships (Bandeira, Estevan, et al., 2025; Bandeira, Lemos, et al., 2025). While conventional statistical techniques have contributed substantially to understanding these constructs, emerging analytical approaches may offer additional opportunities to identify patterns and relationships that are not readily observable using traditional methods (Bandeira, Estevan, et al., 2025; Bandeira, Lemos, et al., 2025; Britton, Belton, et al., 2023; Duncan et al., 2024; Farrahi & Rostami, 2024; Pilati Filho et al., 2023; Salami et al., 2024). Recent developments in machine learning (ML) (Britton, Onibonoje, et al., 2023; Duncan et al., 2024) and markerless motion capture (Bossavit & Arnedillo-Sánchez, 2022; Chow et al., 2020; Lander et al., 2025) technologies offer potential solutions to both of these challenges. Markerless motion capture systems have the capacity to provide objective movement data without the practical limitations associated with traditional laboratory-based motion analysis systems (Colyer et al., 2018). Similarly, ML approaches may facilitate the identification of complex relationships within MC data while supporting the development of more efficient and scalable assessment methods (Clark et al., 2021; Lander et al., 2020). However, despite growing interest in these technologies, further validation and application within paediatric MC research remain necessary before widespread implementation can be justified (Hulteen et al., 2023; Lander et al., 2020).

Therefore, this thesis seeks to address two interrelated gaps within the literature. First, it explores the application of advanced analytical approaches to enhance understanding of the complex relationships underpinning MC and its associated constructs. Second, it explores the potential of markerless motion capture technology (via '*FreeMoCap*', an open-source pipeline

by Matthis et al. (2022) to support the development of more objective, efficient, and scalable approaches to AMC assessment. In doing so, the thesis aims to contribute to both the methodological advancement of MC assessment and the broader understanding of children's movement development.

1.3 Philosophical Position

Research is guided by a set of assumptions regarding the nature of reality, how knowledge is generated, and the most appropriate means of investigating a phenomenon. These assumptions are commonly described through the concepts of ontology and epistemology. Ontology concerns the nature of reality and what can be said to exist, while epistemology relates to the nature of knowledge and how knowledge can be acquired and justified (Moon & Blackman, 2014). Together, these philosophical positions influence how researchers formulate questions, select methods, interpret findings, and evaluate evidence. My interest in physical activity and motor-skill development is informed by both personal and professional experiences. Participation in sport and PA has been a consistent feature throughout my life and has shaped my interest in understanding the factors that influence children's engagement in movement. Through coaching, teaching, and research experiences, I developed an appreciation of the important role that physical activity can play in supporting health, wellbeing, and development. Recognising these experiences is important, as they inevitably influence how I approach research questions and interpret evidence. Acknowledging these influences represents an important aspect of reflexive practice throughout this thesis.

Throughout my undergraduate and master's studies, and during my early experiences in research, I largely viewed knowledge through a positivist lens, placing considerable emphasis on objective measurement and the identification of observable relationships. However, as my research experience developed, my philosophical position evolved towards pragmatism. Pragmatism rejects strict adherence to any single methodological tradition and instead prioritises the research question when determining the most appropriate approach to inquiry (Giacobbi et al., 2005; Poucher et al., 2020). From this perspective, the value of knowledge lies not only in its explanatory power but also in its capacity to address real-world problems and inform practical action.

Ontologically, this thesis adopts a realist position, acknowledging that an external reality exists independently of the researcher and can be investigated through systematic inquiry. However, consistent with pragmatic thought, it is recognised that our understanding of that reality is always incomplete and subject to refinement as new evidence, methods, and perspectives emerge. Epistemologically, knowledge is viewed as provisional and fallible rather than absolute. Consequently, the aim of research is not necessarily to establish definitive truths,

but to develop increasingly useful and robust explanations that improve understanding and support decision-making (Giacobbi et al., 2005; Poucher et al., 2020). This philosophical position shares considerable overlap with post-positivist and critical realist perspectives, both of which acknowledge an external reality while recognising that knowledge of that reality remains partial and fallible. What locates the present study within pragmatism is the criterion by which knowledge is judged. Rather than prioritising the pursuit of increasingly accurate representations of reality as an end in itself, pragmatism emphasises the practical usefulness of knowledge and its capacity to address real-world challenges (Giacobbi et al., 2005; Poucher et al., 2020). Accordingly, the methods employed throughout this thesis were selected not only for their ability to generate robust evidence, but also for their potential to improve the understanding and assessment of MC.

A pragmatic paradigm was considered particularly appropriate for this thesis because of its dual focus on understanding MC and improving how it is assessed. While pragmatism is often associated with mixed-methods research, its relevance extends beyond methodological flexibility. In the present thesis, pragmatism facilitated the selection of analytical approaches based on their suitability for addressing specific research questions rather than adherence to a particular methodological tradition. Consequently, traditional statistical techniques, ML approaches, graph-based analyses, and computer vision methodologies were employed where they offered the greatest potential to advance understanding or address practical challenges within the field. This philosophical position influenced several methodological decisions throughout the thesis.

In Chapters 3 and 4, ML and graph-based approaches were employed to explore complex relationships within MC data that may not be readily identifiable through conventional analytical methods. In Chapters 5 and 6, markerless motion capture technology was investigated as a potential solution to the practical limitations associated with existing assessment approaches, including issues relating to time, cost, objectivity, and scalability. Across all studies, methodological decisions were guided by a commitment to identifying evidence-based approaches capable of advancing both the understanding and assessment of MC. At the same time, I remain aware that my professional background in sport, PA, and motor-skill development may predispose me towards viewing emerging technologies as potentially valuable tools for addressing contemporary challenges in assessment and practice.

To minimise the influence of this potential bias, the studies presented throughout this thesis sought to evaluate new approaches against established methods, report limitations alongside strengths, and critically examine findings irrespective of whether they aligned with initial expectations. Consequently, the conclusions presented should be viewed as contributions to an evolving body of evidence rather than definitive solutions, with further validation and refinement required as the field continues to develop.

1.4 Aims and Objectives

Primary aims of research:

1.4.1 Evaluate the ability of several machine learning classification models in predicting actual motor competence assessment outcomes from existing actual and perceived motor competence assessment data. (Chapter 3)

1.4.2 Model and evaluate the interdependencies between actual and perceived motor competence using post-hoc model interpretability analyses and graph analytics techniques (Chapter 4).

1.4.3 Explore the theoretical rationale, opportunities, and challenges associated with the application of markerless motion capture technologies within motor competence research (Chapter 5).

1.4.4 Evaluate the feasibility of using an open-source markerless motion-capture pipeline to support automated classification of TGMD-3 Short Form locomotor performance criteria. (Chapter 6)

Secondary objectives of research:

1. Explore the contribution of perceived motor competence variables to the prediction of actual motor competence outcomes (Chapter 3).

2. Examine the differences in graph analytics measures across gender and age (Chapter 4).

3. Uncover the minimal viable camera configuration required for accurate classification by an open-source markerless motion capture system (Chapter 6)

1.5 Research Question and Hypothesis

1.5.1 Can machine learning classification models accurately predict individual actual motor competence assessment outcomes using data derived from other actual and perceived motor competence assessments?

Hypothesis: It is hypothesised that the predictive capability of machine learning classification models will differ across actual motor competence assessments, with some skills demonstrating sufficient predictive accuracy.

1.5.2 Do machine learning classification models demonstrate differing predictive performance across male, female, and gender-combined motor competence data subsets?

Hypothesis: It is hypothesised that the predictive performance of machine learning classification models will be improved when trained on gender-specific datasets compared with a gender-combined dataset.

1.5.3 Can post-hoc model interpretability analyses and graph analytics identify meaningful age- and gender-specific interdependencies within motor competence data?

Hypothesis: It is hypothesised that age- and gender-specific motor competence networks will exhibit distinct structures, with both actual and perceived motor competence variables acting as influential nodes within these networks.

1.5.4 What is the minimal viable camera configuration required for accurate classification by an open source markerless motion capture system?

Hypothesis: It is hypothesised that reducing the camera configuration from four to three cameras will not significantly affect kinematic data quality; however, reducing the configuration from four to two cameras will result in a significant decline in data quality.

1.6 Thesis Structure

Chapter 1: Introduction to Thesis

This chapter gives a broad overview of the thesis, including the aims and objectives of the study.

Chapter 2: Literature Review

This chapter gives a critical evaluation of the previous literature in areas such as: physical activity, fundamental movement skills, actual motor competence, perceived motor competence and fundamental movement skill assessment. The chapter examines both traditional and emerging assessment approaches, including machine learning and markerless motion capture technologies, highlighting current limitations and identifying key gaps that underpin the research objectives of this thesis.

Chapter 3: Streamlining Motor Competence Assessments via a Machine Learning Approach

This chapter evaluates the predictive capability of machine learning classification models within a motor competence context. It identifies actual motor competence assessments that can be accurately predicted using other actual and perceived motor competence measures and examines differences in predictive performance across gender groups.
RQ; Can machine learning classification models accurately predict individual actual motor competence assessment outcomes using data derived from other actual and perceived motor competence assessments?
RQ; Do machine learning classification models demonstrate differing predictive performance across male, female, and gender-combined motor competence data subsets?

Chapter 4: Modelling Motor Competence Interdependencies Using Post-hoc Model Interpretability and Graph Analytics

This chapter explores age- and gender-specific interdependencies between actual motor competence and perceived motor competence among Irish primary school children. By integrating post-hoc model interpretability and graph analytics, it identifies influential competencies and highlights the important role of perceived motor competence within broader motor competence networks.

RQ; Can post-hoc model interpretability and graph analytics identify meaningful age- and gender-specific interdependencies within motor competence data?

Chapter 5: Markerless Motion Capture for Motor Competence Assessment in Children: Promises, Limitations, and a Research Agenda

This chapter explores the potential of markerless motion capture technologies as an objective, scalable, and accessible approach to motor competence assessment. It examines the promises and limitations of markerless motion capture systems and outlines key research priorities required to support their future adoption within motor competence research and practice.

Chapter 6: Evaluation of a Markerless 3D Motion Capture System for Locomotor Skill Assessment: A Pilot Study

This chapter investigates the feasibility and validity of a computer vision-based markerless motion capture system for motor competence assessment. It examines the impact of reduced camera configurations on kinematic data quality and evaluates the ability of the system to automatically assess locomotor skill proficiency using established assessment criteria.
RQ; What is the minimal viable camera configuration required for accurate classification by an open-source markerless motion capture system?

Chapter 7: Conclusions and Future Directions for Markerless Motion Capture Research

This chapter provides an overview of the thesis, presenting the main findings, strength and limitations of the study, directions for future research, and potential options to advance markerless motion capture research.

1.7 Statement of Candidate Contribution

Declaration

In accordance with Regulation 11.2.2, this section sets out my individual contribution to each of the four studies presented in this thesis and to the pre-existing dataset on which two of them draw. Contributions are described using the Contributor Roles Taxonomy (CRediT). I am the first author of all four studies and was responsible in each case for the design and execution of the analysis, the development of all code, the interpretation of results and the preparation of the original manuscript. My supervisors contributed to the conception of the programme of work, provided domain expertise in motor competence assessment and in computer science respectively, provided access to data and to equipment, and revised each manuscript critically. Initials used below are CO'D for Colm O'Donaghue, MS for Michael Scriney, SB for Stephen Behan and SJB for Sarahjane Belton.

Role matrix

Study 1 is Streamlining Motor Competence Assessments via a Machine Learning Approach, published in Youth. Study 2 is Modelling Motor Competence Interdependencies, published in the proceedings of the 2025 IEEE International Workshop on Sport, Technology and Research. Study 3 is Markerless Motion Capture for Motor Competence Assessment in Children, currently under review. Study 4 is Evaluation of a Markerless 3D Motion Capture System for Locomotor Skill Assessment, currently under review.

CRediT role	Study 1 (Ch. 3)	Study 2 (Ch. 4)	Study 3 (Ch. 5)	Study 4 (Ch. 6)
Conceptualisation	CO'D, MS, SB, SJB	CO'D, MS, SB, SJB	CO'D, MS, SB, SJB	CO'D, MS, SB, SJB
Methodology	CO'D, MS	CO'D, MS	CO'D	CO'D, MS, SB
Software	CO'D	CO'D	Not applicable	CO'D
Validation	CO'D, MS	CO'D, MS	Not applicable	CO'D, SB
Formal analysis	CO'D	CO'D	Not applicable	CO'D
Investigation	Not applicable (secondary data)	Not applicable (secondary data)	CO'D	CO'D
Resources	SB, SJB (dataset access); MS (computing)	SB, SJB (dataset access); MS (computing)	Not applicable	MS (equipment and computing)
Data curation	CO'D	CO'D	Not applicable	CO'D
Writing, original draft	CO'D	CO'D	CO'D	CO'D
Writing, review and editing	CO'D, MS, SB, SJB	CO'D, MS, SB, SJB	CO'D, MS, SB, SJB	CO'D, MS, SB, SJB
Visualisation	CO'D	CO'D	CO'D	CO'D
Supervision	MS, SB, SJB	MS, SB, SJB	MS, SB, SJB	MS, SB, SJB

CRediT role	Study 1 (Ch. 3)	Study 2 (Ch. 4)	Study 3 (Ch. 5)	Study 4 (Ch. 6)
Project administration	CO’D	CO’D	CO’D	CO’D

Where a role is recorded as not applicable, the study did not involve that activity. Studies 1 and 2 are secondary analyses and therefore involved no investigation in the CRediT sense, since no new data were generated. Study 3 is a position paper and involved no software, formal analysis or data curation.

1.8 Definition of Terms

Actual Motor Competence: An individual's ability to perform a wide range of motor acts, including the movement coordination and control underlying a particular motor outcome that are necessary to manage everyday tasks (Robinson et al., 2015).

Children: Used to describe those under 18 years of age (Lansdown & Vaghri, 2022).

Computer Vision: Uses image and pattern mappings in order to find solutions (Esposito & Malerba, 2001), considers an image as an array of pixels and is said to automate the monitoring, inspection, and surveillance tasks (Bond et al., 2019).

Fundamental Movement Skills: Fundamental movement skills are described as the basic building blocks of more advanced complex movements that support successful engagement in sport and PA (Logan et al., 2018).

Locomotor subtest: In the context of FMS, the locomotor subtest refers to the motor skills that require fluid coordinated movements of the body as the child moves through space (Gallahue & Ozmun, 2006).

Machine Learning: A subfield of artificial intelligence concerned with producing mathematical models that identify and reason about dependencies between the variables describing a phenomenon (Imam et al., 2024).

Mastery/Near Mastery: in the context of FMS, mastery is defined as correct performance of all components of a skill. Near mastery is defined as correct performance of all components but one (Van Beurden et al., 2002).

Non-Communicable Disease: A non-communicable disease is defined as chronic health conditions that are not infectious and cannot be transmitted between individuals (Maphumulo et al., 2025).

Object Control subtest: In the context of FMS, object control subtest assesses motor skills that demonstrate efficient movements which involve an object, such as throwing, striking, and catching movements (Ulrich, 2013).

Physical Activity: Physical activity is widely defined as “any bodily movement produced by skeletal muscles that results in energy expenditure” (Caspersen et al., 1985, p. 126).

Perceived Motor Competence: How someone perceives his/her ability to competently perform a motor task (Harter & Pike, 1984).

Physical Inactivity: Physical inactivity is defined as the failure to meet the recommendations on physical activity for health (Boisgontier & Iversen, 2020).

Physical Literacy: ‘Lifelong holistic learning acquired and applied in movement and physical activity contexts, integrating physical, psychological, social, and cognitive capabilities that support participation in health-promoting movement throughout life’ (*Sport Australia*, 2019).

Sedentary Behaviour: Sedentary behaviour is defined as any behaviour that incurs ≤1.5 METs and includes behaviours such as sitting, television watching and lying down (Pate et al., 2008).

1.9 List of Abbreviations

Actual Motor Competence = AMC

BMI = Body mass index

BOT-2 = Bruininks Oserestky Test of Motor Proficiency, 2nd edition

CI = Confidence interval

CSPPA = Children's Sport Participation and Physical Activity report

CVD = Cardiovascular Disease

DCUREC = Dublin City University Research Ethics Committee

FMS = Fundamental movement skills

HRF = Health related fitness

KTK = Korperkoodinations fur Kinder

M = Mean

MNM = Mastery/Near-Mastery

MWBW = Moving Well-Being Well

MVPA = Moderate to vigorous intensity physical activity

N = Number

PA = Physical activity

PE = Physical education

PMC = Perceived Motor Competence

SD = Standard deviation

TGMD = Test of Gross Motor Development

Y-PATH = Youth Physical Activity Towards Health

Chapter 2: Literature Review

2.1 Introduction to Literature Review

This chapter provides a critical synthesis of the theoretical and empirical literature that underpins the thesis. It is structured to develop a coherent argument that moves from the identification of a public health issue to the justification of a methodological response. The chapter begins by examining physical activity (PA) in youth, outlining the epidemiological context and the consistent evidence demonstrating declining participation across childhood and adolescence. It then focuses on actual motor competence (AMC), reviewing its role as a key determinant of PA engagement, and evaluating the tools commonly used to assess it. Perceived motor competence (PMC) is subsequently considered, with particular attention given to its interaction with AMC and its influence on PA behaviour, especially during the transition into adolescence.

Building on this theoretical foundation, the chapter critically evaluates the dominant statistical approaches within motor competence (MC) research. While these approaches are methodologically robust, they have largely produced incremental and often repetitive findings, with limited capacity to account for the complex and non-linear nature of motor-skill development. In response to these limitations, machine learning (ML) is introduced as a promising alternative, and its application within PA and MC research is reviewed. The chapter concludes by examining the emergence of markerless motion capture technologies and their potential to integrate with ML approaches to enable scalable and objective assessment of AMC in real-world contexts. Collectively, this review establishes the rationale for the current thesis, positioning it within a field that is increasingly recognising the need to advance both measurement practices and analytical approaches.

2.2 Physical Activity in Childhood and Adolescence

Physical activity (PA) is widely defined as “any bodily movement produced by skeletal muscles that results in energy expenditure” (Caspersen et al., 1985, p. 126), a definition adopted in similar terms by the World Health Organization (WHO, 2018) and which underpins current Irish PA guidelines and health promotion strategies (Murtagh et al., 2023). PA can be accumulated through various forms, from active travel and free play through to organised recreational and competitive sport. For children and adolescents, the WHO advises an average of at least 60 minutes of moderate to vigorous intensity PA (MVPA) daily, incorporating muscle and bone strengthening activities across the week (Bull et al., 2020).

This guidance has been adopted by numerous countries and by the European Union (Parrish et al., 2020), with current Irish guidelines mirroring this international consensus (Murtagh et al., 2023). At the outset, it is important to distinguish PA from sedentary behaviour. Sedentary behaviour refers to waking behaviours involving low energy expenditure while sitting, reclining or lying down; it is addressed by separate guidelines (Murtagh et al., 2023) and is independently associated with adverse health outcomes in youth, including increased body weight and poorer metabolic profiles (Ekelund et al., 2006; Marshall et al., 2004). The two constructs are therefore not interchangeable, as a child may meet the MVPA guidelines whilst still accumulating substantial sitting time, and the terms are not used interchangeably within this thesis.

Regular PA engagement during childhood and adolescence is associated with a broad range of physical, psychological and cognitive health benefits. Evidence from systematic reviews and meta-analyses indicates that greater levels of MVPA are associated with improved cardiorespiratory fitness, muscular fitness, bone health, body composition, and cardiometabolic health markers such as blood pressure and insulin sensitivity (Hupin et al., 2015; Janssen & LeBlanc, 2010; Poitras et al., 2016). These relationships appear to follow a dose-response pattern, with greater volumes and intensities of activity conferring larger benefits, and with vigorous intensity activity potentially conferring even greater benefit (Janssen & LeBlanc, 2010). Importantly, even modest amounts of PA appear to confer measurable benefit among higher-risk groups, such as children who are overweight or obese (Janssen & LeBlanc, 2010), whilst across the intensity spectrum it is MVPA that shows the most consistent favourable associations with health indicators when PA is measured objectively (Poitras et al., 2016).

Regular participation in MVPA has also been linked with lower levels of adiposity, including reduced waist circumference and BMI z-scores (Hay et al., 2012; Poitras et al., 2016). The benefits extend beyond physical health. PA participation is associated with reduced symptoms of depression and anxiety (Paluska & Schwenk, 2000), enhanced mental wellbeing, and improved cognitive and academic performance among children and adolescents (Li et al., 2023; Singh et al., 2023; Warburton, 2006). The strength of this evidence does vary across outcomes, and much of the underpinning research in youth remains observational in design (Poitras et al., 2016). Nonetheless, the consistency of these associations across multiple health domains positions PA as a fundamental component of healthy development and

explains why promoting lifelong PA engagement remains a major public health priority. Childhood warrants particular attention within this priority. Non-communicable diseases account for an estimated 40 million deaths annually worldwide and 82% of deaths annually in Ireland (WHO, 2025), and physical inactivity is recognised as a leading modifiable risk factor, with estimates suggesting over 5 million deaths per year could be prevented if population levels of PA were increased (Salvo et al., 2026).

In adults, the scale of the relationship is well established, with meta-analytic evidence including over 30 million participants demonstrating a clear inverse dose-response relationship between PA and all-cause mortality (Garcia et al., 2023) and umbrella review evidence reinforcing the protective role of PA against all-cause, cardiovascular and cancer mortality (Li et al., 2025). Importantly, the precursors of these outcomes are established early in life. Cardiovascular risk factors that emerge during childhood and adolescence frequently persist into adulthood and are associated with poorer cardiovascular health in later life (Juonala et al., 2013; Raitakari & Juonala, 2003), and recent evidence indicates a rising prevalence of such risk factors among adolescents (Arshad et al., 2024). Behaviour itself also tracks across the lifespan, as physically inactive children are more likely to become inactive adults (Martínez-Andrés et al., 2020). Given that greater PA engagement in youth is associated with lower blood pressure, improved glycaemic control and healthier weight status (Barone Gibbs et al., 2021; Loh et al., 2020; Swift et al., 2014), promoting PA during childhood represents an important target for efforts to reduce cardiometabolic risk across the lifespan.

Despite these well-established benefits, a substantial gap exists between recommended and achieved PA levels. Globally, findings report 81% of youth aged 11 to 17 years are estimated not to meet the WHO PA guidelines (Guthold et al., 2020), whilst a large-scale European study categorised more than two thirds of European youth as insufficiently active (Steene-Johannessen et al., 2020). The economic burden is correspondingly large, with physical inactivity estimated to have accounted for $53.8 billion in global healthcare costs in 2013 alone (Ding et al., 2016). Irish data reflect the same pattern. The most recent Children's Sport Participation and Physical Activity (CSPPA) study reported that only 23% of children attending Irish primary schools and 12% of Irish post-primary children self-reported meeting the PA guideline of at least 60 minutes of daily MVPA (Woods et al., 2023), albeit a slight improvement on the 17% and 10% reported in 2018 (Woods et al., 2018). Such prevalence estimates should be interpreted with a degree of caution, as they rest largely on self-report and the accurate

measurement of PA remains methodologically challenging (Hidding et al., 2018). Even allowing for measurement error, however, the consistency of these findings across instruments and settings indicates that the substantial majority of children, in Ireland and internationally, are insufficiently active. Compounding these low absolute levels, PA declines as children age. Pooled accelerometer data from 27,637 participants indicated that, cross-sectionally, PA was approximately 4.2% lower with every additional year of age after the age of five (Cooper et al., 2015), whilst systematic review evidence drawn from longitudinal studies confirms that MVPA declines progressively across childhood and adolescence rather than at any single point (Farooq et al., 2020). Longitudinal accelerometer data similarly show weekday MVPA decreasing by approximately 2.2 minutes per day, per year, between the ages of six and eleven, with sedentary time increasing concurrently (Jago et al., 2020). This decline is not evenly distributed by gender. Girls consistently report lower PA participation than boys (Hallal et al., 2012; Ricardo et al., 2022), demonstrate a steeper age-related reduction in vigorous activity (Corder et al., 2016), and have been noted to fall below recommended PA levels at a younger age than boys (13.1 versus 14.7 years) (Nader et al., 2008). Declines are also particularly pronounced around the transition from primary to secondary school, which in Ireland typically occurs at around twelve years of age (De Meester et al., 2014; Marks et al., 2015). Taken together, the evidence indicates that most children are insufficiently active, that activity erodes further with age, and that girls and early adolescents are at particular risk of disengagement. It is therefore imperative to consider modifiable factors that may help counteract this decline and support sustained PA engagement across childhood and into adolescence. One such factor is motor competence, to which the following section turns.

2.3 Motor Competence and its Role in Physical Activity

Motor competence (MC) is broadly defined as an individual's ability to execute a range of goal-directed motor skills (Robinson et al., 2015), encompassing three broad sub-domains: locomotor skills, object control skills and stability skills (Goodway et al., 2013) each of which is examined in detail in Section 2.4. Within the construct, a distinction is commonly drawn between actual motor competence (AMC), referring to objectively assessed motor skill proficiency, and perceived motor competence (PMC), referring to a child's perception of that proficiency (De Meester et al., 2020). In the last decade there has been increased recognition that AMC development is a key factor in promoting PA engagement (Barnett et al., 2009; Burton et al., 2023; Logan et al., 2015; Robinson et al., 2015; Stodden et al., 2008), alongside

a variety of other health benefits such as healthy weight status (Cattuzzo et al., 2016; D'Hondt et al., 2013) and psychological wellbeing (Rose et al., 2015). Childhood is considered a critical period in this regard (Koolwijk et al., 2024). It is during childhood that the foundational skills enabling successful participation in various forms of PA are acquired (Clark & Metcalfe, 2002), and establishing proficiency early is said to contribute to the development of an active lifestyle (Haugen & Johansen, 2018; Holfelder & Schott, 2014). Before examining the empirical evidence for these claims, it is necessary to set out the theoretical model within which they are typically framed.

The conceptual spine of this literature is the developmental model proposed by Stodden et al. (2008) (see Figure 2.1), which hypothesises the relationships among AMC, PA, PMC, health-related fitness and weight status to be key components of a positive health trajectory for children and adolescents. Central to the model is the proposal that the relationship between AMC and PA is both reciprocal and developmentally dynamic. In early childhood, PA is proposed to drive the development of AMC, as young children accumulate the movement experiences from which skills emerge. As children progress through middle and later childhood, the direction of influence is hypothesised to reverse, with AMC increasingly determining the extent to which children are equipped to participate in the games, sports and other activities through which PA is accrued. PMC occupies a mediating role within this dynamic. Young children tend to hold inflated perceptions of their abilities, meaning that early PA participation is theorised to be relatively independent of AMC; however, as self-perceptions become more accurate with age, the relationship between AMC and PA strengthens and becomes increasingly mediated by psychological factors.

The model's best-known prediction follows from these dynamics: children possessing high levels of AMC proficiency and positive PMC are proposed to enter a "positive spiral of engagement" in which competence, perception and participation reinforce one another, whereas children with poor AMC proficiency and low PMC risk being "drawn into a negative spiral of disengagement", resulting in physical inactivity and the negative health consequences associated with it (Stodden et al., 2008, p. 296). It should be emphasised that the model is a hypothesised synthesis rather than an established theory, and the empirical testing of its pathways, along with the reconceptualisations this has prompted, is considered in Section 2.6. What follows here is the evidence for its load-bearing assumption: that AMC and PA are

positively associated.

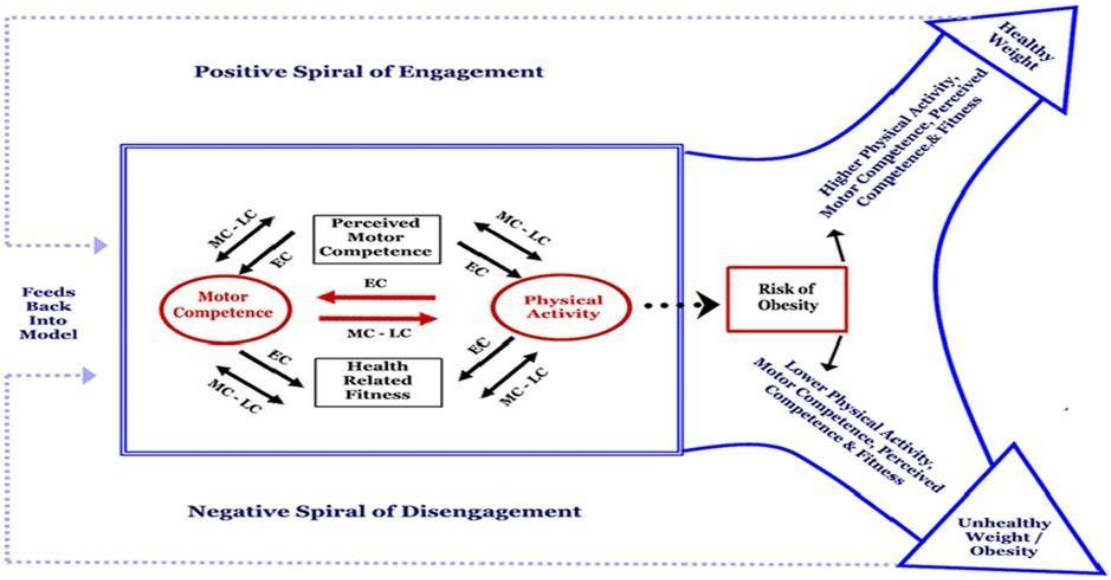


Figure 2.1: Developmental mechanisms influencing the physical activity trajectories of children. Reproduced from Stodden et al. (2008)

Early review-level evidence broadly supported the assumption of AMC and PA being positively related to one another, whilst also exposing its evidential limits. Lubans et al. (2010), in a systematic review of the health benefits associated with motor skill proficiency in youth, concluded that the evidence confirmed a positive cross-sectional relationship between general AMC proficiency and PA among children and adolescents. The evidential base of the review was thin in places, however, including only two longitudinal and two experimental studies, and the authors acknowledged the need for further longitudinal research exploring the relationship. Four years later, Holfelder & Schott (2014) reached a similar verdict, noting strong cross-sectional evidence for a positive relationship between AMC and organised PA whilst cautioning that a cause-and-effect relationship was suspected but had not yet been established, a reservation echoed in earlier work (Barnett et al., 2011; Kambas et al., 2012; Okely et al., 2001; Stodden et al., 2008). Their review included only eight longitudinal studies, and the imbalance between cross-sectional and longitudinal designs that both reviews identified has remained a defining feature of this literature since. Two further reviews published in 2015 sharpened the picture without resolving it. Logan et al. (2015) reported low to moderate positive relationships between AMC and PA in early childhood, middle to late childhood, and adolescence, although one included study found a moderate negative relationship. The authors concluded by returning to the question posed by Fulton et al. (2001, p. 123): "...are

children physically active because they have good motor skills or do they have good motor skills because they are physically active?", noting that many factors likely shape how the relationship changes over time and that research with stronger designs would be required to identify them and determine their magnitude. Robinson et al. (2015) likewise concluded that the literature supported a positive relationship between AMC and PA across childhood, with low to moderate associations evident from early through middle childhood but found insufficient evidence to determine whether the association strengthens during adolescence, as the Stodden et al. (2008) model predicts. Notably, Robinson et al. (2015) also identified measurement as a structural obstacle: until consensus is reached on AMC and PA measurement methodology, and measures are used consistently across the literature, changes in the strength of association across time cannot be meaningfully examined.

Longitudinal and predictive evidence, although scarcer, points in the same direction. Children with stronger object control skills in childhood have been shown to be more physically active in adolescence (Barnett et al., 2009), and poor AMC during childhood has been reported to negatively affect PA adherence as children transition into adolescence and adulthood (Barnett et al., 2009; Lloyd et al., 2014). More recent work has positioned AMC as a predictor of both PA volume (McIntyre et al., 2018) and PA intensity (Gu et al., 2017). Such findings are consistent with the developmental logic of the Stodden et al. (2008) model, although they remain too few in number to establish the causal pathway the model proposes.

The most recent evidence is more contested. Burton et al. (2023) concluded that current findings support the positive associations identified in earlier reviews (Holfelder & Schott, 2014; Logan et al., 2015; Lubans et al., 2010; Robinson et al., 2015), with Den Uil et al. (2023) reporting similarly supportive findings. In direct contrast, Barnett et al. (2022), in a large systematic review of mediation, longitudinal and experimental evidence, found no supporting evidence for an association between AMC and PA in adolescence, with Schmutz et al. (2020) and Coe et al. (2024) also questioning the strength of the relationship. Two features of the evidence base explain how reviews published within a year of one another can reach opposing conclusions. First, the literature remains overwhelmingly cross-sectional: nearly nine in ten of the studies located by De Meester et al. (2020) employed cross-sectional designs, which inflate the appearance of association through the proximal measurement of variables and cannot speak to the direction of influence. Second, there is a lack of consistency in assessment approaches. Studies differ in the AMC instruments employed, in whether PA is

self-reported or measured via accelerometry, and in the skill domains assessed, and each of these differences moderates the associations observed (Logan et al., 2015; Robinson et al., 2015). On balance, the evidence for the bridge between AMC and PA is positive but indeterminate. The weight of review-level evidence supports a low to moderate positive association across childhood, the longitudinal evidence is directionally consistent but sparse, and the strongest recent challenge rests substantially on the absence of adequately designed studies rather than on contradictory findings. Whilst the literature remains undetermined as to whether AMC is the most important determinant of PA participation or whether other factors are more influential, two conclusions can be drawn with some confidence. The first is that AMC matters enough, both theoretically and empirically, to warrant sustained research attention. The second is that progress on the questions left open by this literature depends directly upon the quality, consistency and scalability of AMC measurement, and upon the analytical approaches applied to the resulting data; these are the central concerns of this thesis. The following section therefore examines AMC in greater detail, beginning with its importance and measurement.

2.4 Actual Motor Competence

2.4.1 Importance of Actual Motor Competence

Actual Motor Competence (AMC) encompasses the mastery and ability to perform various fundamental movement skills (FMS) such as coordination and balance, leading to quality movement (De Meester et al., 2020). FMS are considered to be the ‘building blocks’ of more advanced, complex movements (Logan et al., 2018) and are classified into three distinct skill categories: locomotor skills (locomotion of the body e.g., running, jumping), object-control (manipulative skills e.g., ball-catching, kicking) and stability skills such as balancing (Goodway et al., 2013).

2.4.2 Measurement of Actual Motor Competence

Due to the increased interest and evidence surrounding associations between AMC and health, it is imperative that assessments used to quantify MC are accurate and informative (Logan et al., 2017). AMC Measurement in youth can be either process-based or product-based. Process-oriented assessments evaluate how a movement is performed and describe

qualitative movement patterns (Logan et al., 2017). In contrast, a product-oriented assessment evaluates the outcome of a movement, which is typically identified as a quantitative score (e.g., speed, distance or number of successful attempts). Process oriented assessments include the Test of Gross Motor Development-2nd edition (TGMD-2; Ulrich & Sanford, 2000), 3rd edition (TGMD-3; Ulrich, 2013) and Get Skilled; Get Active (GSGA; (Ryde, 2000).

The TGMD-2 assessment battery is one of the most frequently used process-oriented assessments of youth AMC (Barnett et al., 2008; Belton et al., 2014; Logan et al., 2017; O' Brien et al., 2013, 2016). The TGMD-3 (Ulrich, 2013) is a revision of the second version of the TGMD-2 (Ulrich & Sanford, 2000), comprising 12 skills from the TGMD-2 plus one extra object-control skill, the one-handed strike (Ulrich, 2013). Thus, the TGMD-3 contains 6 locomotor skills and 7 object control skills. The TGMD assessment batteries evaluate children's FMS based on the presence or absence of 3–5 performance criteria for each of the skills. Scores across two trials are then summed to provide a raw score for that particular skill. The GSGA skills are measured similarly to that of the TGMD batteries with the GSGA comprising 12 skills (catch, overhand throw, kick, forehand strike, sprint run, leap, dodge, vertical jump, hop, side gallop, skip, and static balance), and uses a similar scoring process to the TGMD-3, where each skill is broken down into its movement components (Ryde, 2000). Participants are required to perform five trials for each skill and the presence or absence of each skill component is noted by the tester. If the participant displays a skill component in four of the five trials then this component is marked as present (Ryde, 2000). Process-oriented assessments differ in their complexity to administer due to the number of skills included, number of performance criteria for each skill and whether or not performance can be observed and recorded live or whether video analysis is required for accurate scoring (Logan et al., 2017).

Examples of product-based assessment tools widely used in research are the Körperkoordinationstest für Kinder (KTK) (Kiphard & Schilling, 1974) and the Movement Assessment Battery for Children (MABC) (Henderson et al., 1992). The KTK (Kiphard & Schilling, 1974) has been validated for children aged 5 – 14 years (Cools et al., 2009; Vandorpe et al., 2011). The KTK consists of four subtests (walking backwards, moving sideways, hopping for height and jumping sideways) and outcome measures are taken for each (Vandorpe et al., 2011). The KTK has been used in previous research to objectively assess AMC among the youth demographic (D'Hondt et al., 2013; Vandendriessche et al.,

2012; Vandorpe et al., 2011). There are limitations associated with the KTK however, most notably, an overall gross AMC score is awarded as opposed to breaking AMC down into its three distinct components (locomotor, object-control, stability). Another key limitation of the KTK is the relatively old normative data (Cools et al., 2009). Another product-based AMC assessment is the MABC-2 (Henderson et al., 1992) which is suitable for children and youth, aged between three to sixteen and is used to identify impairments in FMS across three skill categories: manual dexterity, aiming and catching, and balance. It has been used across several age-ranges (Fisher et al., 2005; Gísladóttir et al., 2014; Livesey et al., 2011; Vedul-Kjelsås et al., 2012) and the tasks change based on the participant's age (Brown & Lalor, 2009). Normative scores are provided for each age category (3-6 years, 7-10 years, 11-16 years) and individuals at risk of movement difficulties are identified using a traffic light system. The MABC-2 can be considered advantageous when compared against the KTK, since it assesses all three components of FMS, and that an assessor can focus on just one component without having to complete the full assessment battery. The biggest limitation of the MABC-2, is the fact that it does not identify movement quality – i.e. if the components of a skill are correctly executed during their execution, however, this issue is evident within all product-based assessments.

Previously, researchers have combined product and process-based assessment techniques to measure AMC proficiency (Rodrigues et al., 2016; Stodden et al., 2009; Stodden et al., 2013). Using such an approach, participants were asked to complete a skill (e.g., kicking a ball) and the process of how the participant executed the components of the skill was assessed, with the outcome of the skill also being measured (e.g. the speed of the ball). Such a mixed-approach can be useful in overcoming a potential ceiling effect for older children and adolescents in some process-based AMC assessments (Logan et al., 2017), but disparities between gender tend to be highlighted when outcome-based measures are used, particularly in adolescence due to physiological changes that occur across puberty (True et al., 2017).

Another form of FMS assessment is the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) and the Bruininks-Oseretsky Test Second Edition (BOT-2) (Bruininks & Bruininks, 2005). These tests are different to the other assessments since they assess both fine and gross motor skill development and are used primarily to identify those with mild or moderate motor coordination difficulties. The BOT-2 is suitable for those aged between 4 and 21 years

of age and the complete test consists of eight sub domains containing 53 items of assessment, although there is a short form version (Bruininks & Bruininks, 2005). There are limitations associated with the BOT-2 however, for instance, the complete assessment takes nearly an hour to complete with the short form lasting between 15 and 20 minutes (Cools et al., 2009) making it impractical for assessing large numbers, particularly in a school setting (Peerlings, 2007). The BOT-2 also does not assess as many skills as other assessments; thus, it does not provide an assessor with a full picture of a participant's AMC proficiency.

Another limitation, which is not just true for the TGMD-3 but for all the aforementioned measurement tools, is that traditionally all the assessments were graded using a pen and paper method (Ulrich, 2013; Yun & Ulrich, 2002). Such a method is a time-consuming process when time required entering the collected data into a database for analysis is considered (Behan et al., 2019). This time constraint can subsequently have a negative knock-on effect by significantly impacting the amount of data collected, not to mention the risk of human error in the data entry process (Bandeira et al., 2020; Bisi et al., 2017; Chow et al., 2020; Duncan, Martins, et al., 2022; Valentini et al., 2021). Despite such limitations, there are a wide variety of assessment tools utilised by practitioners that can assess AMC proficiency accurately and reliably (Logan et al., 2012). Regardless of the assessment measure utilised, the results from such assessments show alarmingly low levels of AMC proficiency reported worldwide (Barnett et al., 2009; Behan et al., 2019; Lubans et al., 2010; Mukherjee et al., 2017).

In summary, from the variety of assessment tools currently available to assessors it appears logical to utilise the latest edition of the TGMD, the TGMD-3, as the primary assessment tool due to the addition of an extra skill. Previously, a meta-analysis recommended the TGMD-2 for use in research (Logan et al., 2012) with similar recommendations also provided for the updated TGMD-3 across a diverse range of participants (Maeng et al., 2016; Magistro et al., 2018; Rintala et al., 2017; Ulrich, 2017; Webster & Ulrich, 2017). The TGMD-3 is regarded as the gold standard of AMC assessment tools and can be considered the most appropriate assessment tool available in a mainstream setting. The TGMD-3 does, however, have several limitations. For instance, it does not assess balance, (Rudd et al., 2015), which, alongside locomotor and object control skills, represents a key AMC component (Logan et al., 2015; Rudd et al., 2015). Furthermore, the absence of explicit quantitative thresholds within some performance criteria can introduce subjectivity into assessor judgements, while variation in score-aggregation approaches may limit comparability across studies (Ng et al., 2025). While

acknowledging these limitations, the prominence of the TGMD-3 in contemporary research has enabled the accumulation of a substantial and comparable evidence base. This, in turn, provides a foundation for identifying consistent patterns in AMC across age groups, gender, and settings. The following section builds on this foundation by examining current trends in AMC reported in the literature.

2.4.3 Current AMC Trends

In recent years, poor AMC proficiency levels have been demonstrated amongst the global youth demographic. For instance, prior research demonstrated across a thirteen-year period that Australian children's proficiency levels remained low with prevalence rarely above 50% (Hardy et al., 2013). Given such stark findings, it is paramount that the skills underpinning such poor proficiency levels are explored further. Each FMS skill comprises several behavioural components that must be present for successful skill completion and are described as performance criteria (Ulrich & Sanford, 2000). Participants must be proficient within all the criteria of a specific FMS skill to be deemed competent. For instance, to be considered competent at kicking in the TGMD-3, participants are required to meet the following criteria: 1) rapid continuous approach to the ball, 2) Child takes an elongated stride or leap just prior to ball contact, 3) non-kicking foot placed close to the ball, and 4) kicks the ball with the instep of the preferred foot (not the toes) (Ulrich, 2013).

When characterising AMC competency, the terms "mastery" and "near-mastery" are frequently applied within the literature. 'Mastery' of process-oriented locomotor and object control skills is the correct performance of all the skill components associated with each skill over both the trials. 'Near Mastery' is when an individual performs all but one of the components correctly over both trials (Van Beurden et al., 2002). Participants that do not fulfil the criteria to achieve mastery or near-mastery (MNM) status are classified as poor (Booth et al., 1999). Such criteria guidelines can prove useful for researchers who have used process-based AMC assessments and can use MNM as a comparative tool. Gallahue & Ozmun (2006) proposed that children have the potential to master most of the FMS by six years of age, while others contradict this view, believing all FMS can be mastered by ten years of age (Hardy et al., 2010). Considering such estimates, it would not be unreasonable to suggest that children should have mastered most of the basic motor skills as they are finishing primary school. Considering the youngest cohort (two- to six-year-olds), an AMC mastery rate of 0% was previously reported among

Australian preschool children (n = 425) (Hardy et al., 2010). Worryingly, this dearth of competency is evident within more recent research. Brian et al. (2018) noted that whilst a group of Belgian preschool children (n = 170) produced superior levels of AMC competency than a group of US preschool children (n = 156) both groups scored below what might be expected of them at this age, especially for object control skills. 32.4% of the Belgian sample and 57% of the US sample scored below average for object control skills compared to 25.33% of the TGMD-2 reference group (Ulrich & Sanford, 2000). Indeed, the aforementioned findings are unsurprising however given that children are only expected to achieve mastery across most of the FMS by age six (Gallahue & Ozmun, 2006).

Whilst AMC competency might not be expected at the preschool level, it is plausible that AMC competency levels may gradually increase as children advance through primary school. Prior research would contradict this sentiment with numerous studies noting children to not have achieved their projected levels of AMC competency (Barnett et al., 2009; Kelly et al., 2019; Lubans et al., 2010; Mukherjee et al., 2017; Okely & Booth, 2004). Furthermore, Australian children (n = 1288, six and nine years of age) did not produce proficiency levels above 35% mastery in any of the six FMS skills assessed (Okely & Booth, 2004). With the exception of balance, the aforementioned study also noted that prevalence of near mastery did not exceed 50% across the remaining FMS skills (Okely & Booth, 2004). Such findings fall in line with prior Australian research which noted within primary school children (n = 1045) that less than half of participants (47%) achieved MNM in FMS (Van Beurden et al., 2002).

Indeed, more recent literature notes that this competency dearth is still evident. Bardid et al. (2016) noted Belgian children (n = 1614, age three to eight) to rate poorly against the normative scores of the TGMD-2 (Ulrich & Sanford, 2000) whilst Mukherjee et al. (2017) reported children in Singapore were rated ’average’ or ‘below average’ for locomotor skills and ‘poor’ or ‘below average’ for object control skills. Recent research by Lawson et al. (2021), which investigated AMC competency among English primary school children (n = 219, aged 7–10 years) noted less than 40% of participants achieved mastery in each FMS, with 27% not achieving mastery in any of the skills with such findings aligning with prior British research (Bryant et al., 2014; Duncan et al., 2020; Foulkes et al., 2015; Goodway et al., 2010). Within an Irish context, several studies have assessed the AMC competency of Irish primary school children (Behan et al., 2019; Bolger et al., 2018; Kelly et al., 2019; O’ Brien et al., 2016). Such research confirmed that Irish children’s levels of AMC competency are similar to children

worldwide with age and gender differences evident (Bolger et al., 2018). Behan et al. (2019) highlighted several skills with low levels of MNM, such as; balance at 41.8% MNM, underhand roll at 37.7% MNM, vertical jump at 34% MNM, two-handed strike at 30.5% MNM, one-handed strike at 28.6% MNM, gallop at 19.8% MNM and overhand throw at 16%. Similar findings were also noted by Kelly et al. (2019) with lower levels of MNM noted across several skills such as two-handed strike and vertical jump, both at 22% MNM. The difference in MNM scores between Behan et al. (2019) and Kelly et al. (2019) may be attributed to the inclusion by Behan et al. (2019) of a wider age range (five to twelve years) and the large number of schools who participated in their study. Additionally, the findings of Kelly et al. (2019) had limitations, due to the lack of generalisability in the convenience sample consisting of only three primary schools (n = 414). Research has consistently reported that males produce superior levels of object control skill proficiency when compared to girls (Bardid et al., 2016; Barnett et al., 2010; Behan et al., 2019; Bolger et al., 2018; Butterfield et al., 2012; Eather et al., 2018; Kelly et al., 2019; Okely & Booth, 2004; Van Beurden et al., 2002). Behan et al. (2019) noted that boys scored significantly higher in object control skills, with the exception of the five-year-old group and corroborated the findings of Bolger et al. (2018) which noted Irish boys to perform better in object control skills.

In turn, gender differences have not been shown to be as pronounced in terms of locomotor skills. Previous studies have found significant gender differences in locomotor skills, with females achieving higher proficiency in the locomotor skills (Behan et al., 2019; Cliff et al., 2012; Hardy et al., 2010) whereas other studies found no significant differences between males and females (Bardid et al., 2016; Kelly et al., 2019). Behan et al. (2019) reported that whilst male participants significantly outscored the female participants across most ages, it was noted that effect size increased among the older age groups, maintaining a moderate effect size across the ages of ten to twelve. This may represent a pivotal turning point in girls object control development, as O'Brien et al. (2016) also noted female adolescents to have significantly less mastery of the object control skills. The 2022 CSSPA report (Woods et al., 2023) noted that significantly more Irish girls participate in activities such as dance (29%) and gymnastics (20%) compared to 4.5% for Irish boys. Such statistics could perhaps explain why girls develop their locomotor and balance skills at a faster rate than boys. In turn, the 2022 CSSPA report noted Irish boys to produce superior rates of extra-curricular sport participation, predominantly made up of Gaelic games, soccer and rugby all of which involve object control skills (Woods et al., 2023). Prior research has suggested object control skills to be a predictor

of PA (Barnett et al., 2009). Given that the most popular sports in Ireland revolve around object control skills (Woods et al., 2010, 2018, 2023), it seems key to place an emphasis upon such skills among primary school aged children, particularly the female youth demographic. Independently of the underlying reasons, gender differences in AMC competency are a key consideration. Furthermore, analysing the individual components (locomotor, object control and stability) is key for determining the disparities between gender. Opposing disparities could cancel each other out for instance if males scored higher in object control and lower in locomotor, and females scored higher in locomotor and lower in object control (Fisher et al., 2005). This could result in both males and females holding similar AMC competencies which is an important consideration given that any gender disparities would not be apparent if only the overall AMC competency is reported (Fisher et al., 2005).

When AMC competency is considered among the adolescent population, previous research shows that they have not achieved the anticipated competency levels for their age (Booth et al., 1999; Hardy et al., 2010, 2013). Recent Irish research found that Irish adolescents showed generally poor AMC competency, lower than two previous samples of Irish adolescents (Lester et al., 2017; O'Brien et al., 2018). Indeed, such findings were consistent with previous studies of Irish adolescents that reported low levels of AMC competency (Duncan, Martins, et al., 2022; Lester et al., 2017; O'Brien et al., 2016).

In recent years, researchers have shifted their focus from cross-sectional approaches towards more longitudinal designs. Barnett et al. (2010) noted low levels of AMC competency among ten-year-olds yet noted improvements by age sixteen, with 80% of males acquiring mastery in five skills, and 80% of females achieving mastery in three skills. Recently, a large-scale US study by Chen et al. (2024) sought to assess the changes in U.S. children's AMC competencies using the data of the TGMD from 1985 to 2019. Overall, the study showed that US children increased their AMC competency through ages three to seven with a decrease/plateau at ages eight to ten before the year 2000. However, children showed a slight decline/no change in AMC competency at ages three to six years and a significant decline at ages seven to ten years, after 2000. Given that children's AMC levels are shown to decrease with age, it is important to consider additional factors that may help counteract this decline. In line with the developmental model outlined in Section 2.3, individuals' perceptions of their own MC have been identified as a key mediating variable in the relationship between AMC and PA (Barnett et al., 2008; Coe et al., 2024; De Meester et al., 2020; Gu et al., 2017). While variable-

centered studies have demonstrated that children with higher AMC tend to be more physically active than their less competent peers (Lubans et al., 2010; Okely et al., 2001), it has been argued that PMC may be an even more influential determinant of PA behaviour (Bardid et al., 2016; Harter, 1978, 1982; Welk, 1999). The following section therefore focuses on PMC and its role in shaping children's engagement in PA.

2.5 Perceived Motor Competence

While AMC is considered imperative in driving PA engagement, it has also been suggested that the perception of competence otherwise known as perceived motor competence (PMC) can be considered equally as important (Stodden et al., 2008). Additionally, PMC can support other psychological health outcomes that are critical for promoting a positive environment of social interaction and acceptance in childhood, which also are important to promote PA (Reunamo et al., 2014). PMC has been defined as a person's belief in their capacity to master a task (Feltz, 1988) whilst Harter's Competence Motivation theory defines PMC as an individual's perception of their physical abilities to perform successfully in specific domains (Harter, 1982).

2.5.1 Importance of Perceived Motor Competence

Research has consistently highlighted a positive association between PMC and PA engagement (Bailey et al., 2012; Baker & Davison, 2011; Crocker et al., 2000; Davison et al., 2010; De Meester, Stodden, et al., 2016; Ekelund et al., 2012; Hills et al., 2011; Kriemler et al., 2010; Raudsepp et al., 2002; Warburton, 2006; Zhang et al., 2015) with it theorized that children with higher PMC would be more motivated to engage in PA and, thus, further develop their AMC proficiency (Harter, 1999; Stodden et al., 2008). Harter (1999) suggested that PMC adjusts throughout a child's development with younger children demonstrating a limited ability to perceive their AMC proficiency resulting in inflated levels of PMC. Simply put, a young child may demonstrate low AMC proficiency whilst simultaneously perceiving themselves as being highly competent (Harter & Pike, 1984). Indeed, this inflation of PMC may be considered a valuable asset as children will be more likely to persistently engage in activities in which they believe they are already proficient which will in turn drive PA and AMC development (Stodden et al., 2008). As children age, their ability to accurately perceive their competence will increase over this time (Harter, 1999). During this period, children with a poor skill level will also perceive

themselves as lowly skilled, and this can result in physical inactivity. Such children may realise that they are not as competent as their peers (Goodway & Rudisill, 1997; Weiss & Amorose, 2005) and may feel self-conscious displaying their poor competence in a public setting (Horn & Weiss, 1991; Weiss & Amorose, 2005). Considering the significance of PMC, it is important to examine the methods used to measure it.

2.5.2 Measurement of Perceived Motor Competence

Numerous measurement tools have previously been presented within the literature. Harter & Pike (1984) reported psychometric data supporting a PMC construct being divided into 'physical competence' and 'cognitive competence', and from this designed the Pictorial Scale of Perceived Competence and Acceptance for Young Children to measure physical self-perceptions in children. The scale contained four separate subscales; cognitive competence, physical competence, peer acceptance and maternal acceptance with each subscale containing six items with individual subscales designed for both preschool-kindergarten and first-second grade children. The physical competence scale included questions on competency in skipping, running, hopping, ball bouncing alongside swinging, climbing, jump-roping and shoe-tying. Of note, only climbing, skipping and running were included in both versions of the scale. Each item was scored on a four-point scale, where a score of 4 would be the most competent or accepted and a score of 1 would account for the least competent. Item scores were averaged across the six items for a given subscale, and these four means provided the child's profile of perceived competence and social acceptance. To assess the ability of the scale, children (n = 90) aged 4 – 7 were recruited for primary data collection. Perceived physical competence scores were consistently high across age groups ($M \approx 3.2 – 3.4$), with moderate internal consistency and modest correlations with teacher ratings ($r = .30$), suggesting generally positive but only partially accurate self-perceptions of physical ability in early childhood. A limitation of the physical competence scale was that whilst it assessed typical childhood actions, it did not assess AMC in a comprehensive manner. With this issue in mind further instruments have been proposed to more accurately measure PMC of the most common FMS skills.

In 2005, The Children's Perception of Motor Competence Scale was developed and used in Spain (Pérez & Sanz, 2005). The scale originally proposed 23 skills validated by physical education experts that were representative of typical physical education and movement

sessions for children. The final scale included 22 skills, as, after analysis, the kicking task was removed as the depictions of the skill were deemed too confusing for the children, a notable limitation of the scale. Depictions of all skills were prepared as pictures in a booklet with pictures depicting children successfully or unsuccessfully performing diverse fundamental perceptual motor skill tasks such as running, throwing and catching. Participants (n = 495) who were aged between 4 – 6 years old were required to point their finger or highlight with a pencil to which boy or girl in each scene appeared most like their own ability.

To examine the relationship between AMC and PMC, participant's AMC was measured via the Movement ABC test (Henderson et al., 1992). Results noted that scores on the scales of the Children's Perception of Motor Competence Scale, Perceived Gross Motor Competence, Perceived Fine Motor Competence, and the total score were significantly related with some of the AMC variables. These relationships were low and moderate, ranging from .10 to .31. The global index of the Perceived Motor Competence had significant correlations with all items of the General Motor Competence dimension of the Motor Competence Observation Scale, and with Balance skills and Ballistic skills measures of the Movement ABC Test. These coefficients were between .20 and .10. The estimate of the internal consistency of the global Children's Perception of Motor Competence Scale indicated that responding to this scale was reliable ($\alpha = .81$), that is, responses of these young children were consistent. Whilst results were promising limitations included the geographically restricted sample, cross-sectional design, and modest PMC-AMC associations. Additionally, the reliance on teacher observations introduced potential measurement bias, while subscale reliability and item functioning issues suggested the need for further validation.

Given that such key FMS skills are demonstrated within AMC measurement tools such as the TGMD (Logan et al., 2012) more recent attempts have been made to better align PMC measurement alongside the TGMD with the Pictorial Scale of Perceived Movement Skill Competence (PMSC) for Young Children (Barnett, Ridgers, Zask, et al., 2015) being a prime example. The PMSC scale consists of thirteen skills, akin to that of the TGMD-3, with each skill depicted through two cartoon pictures of a boy or a girl performing that particular skill. One picture will depict a child performing a skill well whilst another picture will depict a child performing the same skill poorly. In completing the PMSC assessment, participants are initially asked to choose the picture they believe represents their ability in that AMC skill. If the chosen picture depicts a child performing an AMC skill well, the participant is then asked if they feel

they are either ‘pretty good’ or ‘really good’ at a particular skill. Conversely, should the participant choose the opposing picture, they would then either answer ‘sort of good’ or ‘not too good’ at a particular AMC skill. A score of four points is awarded for the highest PMC, i.e.., ‘really good’, with just one point awarded for the lowest PMC, i.e., not too good’, with a score of two or three points assigned for the other two answers.

Findings from Barnett, Ridgers, Zask, et al. (2015) noted the PMSC to demonstrate acceptable reliability, with good test–retest reliability (ICC = 0.78) and moderate internal consistency (α = 0.63–0.72), alongside similar values reported in an independent sample (α ≈ 0.70). Further findings from the study noted a significant positive association between actual and perceived object control competence (B = 0.11, p = 0.027), indicating that higher skill proficiency corresponded with higher self-perception. However, neither AMC (p = 0.092) nor PMC (p = 0.827) were significantly associated with moderate-to-vigorous physical activity (MVPA). Descriptively, boys demonstrated higher PMC (20.1 vs 17.5), AMC (33.8 vs 28.4), and MVPA (72.3 vs 58.1 mins/day) compared to girls. Overall, while the study’s participant’s perceptions aligned with their actual abilities, these competencies did not appear to influence PA behaviour at this age.

Whilst the study by Barnett, Ridgers, Zask, et al. (2015) demonstrated notable methodological strengths, there were also several limitations highlighted which should be considered. First, the cross-sectional design precludes causal inference regarding the relationships between AMC, PMC, and PA. Second, the use of a relatively small convenience sample (n = 102) and a low parent-consent rate (~23%) may limit generalisability and introduce potential selection bias. Third, although the pictorial scale showed acceptable reliability (ICC = 0.78; α = 0.63–0.72), the moderate internal consistency suggests some measurement variability, likely reflecting the developmental stage of the participants. Additionally, the focus solely on object control skills excludes locomotor competence, which may also influence PA behaviours. However, these limitations are balanced by key strengths, including the use of validated, process-oriented assessments of MC and objective measurement of PA via accelerometry, enhancing the overall robustness and credibility of the findings. A further consideration is the age range across which pictorial instruments remain appropriate. Scales such as the PMSC rely on cartoon depictions precisely because young children possess limited reading ability and benefit from concrete, visual item formats (Barnett, Ridgers, Zask, et al., 2015; Harter & Pike, 1984). However, as children age, their self-perceptions become more differentiated and

more accurate (Harter, 1999), and a forced choice between two pictured performances may fail to capture the increasingly nuanced self-evaluations of older children and adolescents. Consequently, pictorial scales are best regarded as instruments for early and middle childhood, and findings derived from them in older cohorts should be interpreted with this constraint in mind.

The PMSC measurement tool has been widely utilised by researchers due to its validity and reliability (Barnett, Ridgers, & Salmon, 2015; Barnett, Ridgers, Zask, et al., 2015). The PMSC can be used to accurately assess young children's AMC proficiency and is particularly useful as a comparison tool when using an iteration of the TGMD to measure AMC proficiency (Barnett, Ridgers, & Salmon, 2015). Considering the evident importance placed upon PMC towards PA engagement, it seems prudent to first consider how PMC develops across childhood before examining its association with AMC proficiency.

2.5.3 Current PMC Trends

PMC follows a characteristic developmental trajectory. In early childhood, perceptions of MC are typically inflated relative to actual proficiency, as young children demonstrate a limited capacity to appraise their own abilities accurately (Harter, 1999; Harter & Pike, 1984). This is evident in the consistently high perceived physical competence scores reported among four- to seven-year-olds, which correlate only modestly with external ratings of ability (Harter & Pike, 1984). Such inflation is not necessarily problematic. Indeed, it is theorised to be adaptive, encouraging young children to persist with activities in which they believe themselves proficient, thereby driving both PA engagement and AMC development (Stodden et al., 2008). As children progress through middle childhood, cognitive maturation and increasing social comparison render self-perceptions progressively more accurate (Harter, 1999), and the alignment between AMC and PMC correspondingly strengthens; modest but significant positive associations between actual and perceived object control competence have been reported even among samples of four- to eight-year-olds (Barnett, Ridgers, & Salmon, 2015), suggesting that this calibration begins early. The consequence of growing accuracy is double-edged, however. Children with genuinely poor skills increasingly recognise their limitations, may feel self-conscious displaying poor competence in public settings (Horn & Weiss, 1991; Weiss & Amorose, 2005) and become vulnerable to the withdrawal from PA described in the Stodden et al. (2008) model (Goodway & Rudisill, 1997).

Gender differences in PMC are also evident, although the evidence base is considerably thinner than for AMC. Boys have been reported to hold significantly higher PMC than girls among young British children (Duncan et al., 2018), with descriptively higher PMC also noted among Australian samples (Barnett, Ridgers, & Salmon, 2015). These differences broadly mirror the object control advantage observed for boys in actual competence (Section 2.4.3), although whether perceptions track, exaggerate or lag actual differences remain unclear, given the paucity of studies reporting both constructs by gender. The patterning nonetheless matters for intervention, particularly in light of longitudinal evidence suggesting that the influence of PMC on PA engagement may be stronger for girls than for boys (Jaakkola et al., 2019) and given that girls demonstrate the steepest declines in PA participation across the preadolescent years (Section 2.2). Taken together, current trends indicate that PMC is high but poorly calibrated in early childhood, becomes progressively better calibrated with age, and differs by gender in ways that broadly parallel actual competence. What remains is to characterise precisely how the two constructs relate to one another, which is the focus of the following section.

2.6 The Association Between AMC and PMC

As outlined in Section 2.3, the conceptual model proposed by Stodden et al. (2008) hypothesised dynamic relationships among AMC, PMC, PA, health-related fitness and weight status (see Figure 2.1). The pathways of the model have since been subjected to empirical testing, refinement and reconceptualisation, and this section reviews that evidence with particular attention to the association between AMC and PMC.

Considering such proposals, researchers have sought to measure both AMC and PMC longitudinally to measure its impact, if any, upon PA engagement. In a longitudinal study, Barnett et al. (2009) noted children demonstrating both high AMC and PMC proficiency in object control skills were more active during adolescence than their less proficient peers. In 2015, another study noted a strong correlation to exist between children's PMC of their object control skills to their actual object control skills (Barnett, Ridgers, & Salmon, 2015). More recently, a Finnish study found that the impact of aligned PMC and AMC had different impacts upon PA levels for boys and girls (Jaakkola et al., 2019). While the study found that AMC and PA were directly associated only for the boys' group, they also demonstrated an indirect path

to PA engagement for the girls' group through PMC (Jaakkola et al., 2019). Such findings suggest that PMC may be more critical among girls in driving future PA engagement, particularly in this preadolescence phase when the opinions of their peers begin to garner more weight. The same findings suggested that PE teachers should focus on developing girls PMC in order to promote continued participation in sport and PA.

Further research has sought to investigate the pathways noted within the conceptual model outlined by Stodden et al. (2008). In their review of health-benefits associated with AMC in youth, Lubans et al. (2010) found that PMC and at least one aspect of AMC were associated in each of the studies included in the review. Indeed, most studies investigating this relationship among children and adolescents have reported significant positive correlations (Barnett et al., 2011; LeGear et al., 2012; Toftegaard-Stoeckel et al., 2010), with some studies even reporting r values of 0.40 or higher (Raudsepp et al., 2002; Robinson, 2011). On the contrary, numerous studies have not found a significant relationship between AMC and PMC/physical self-perception (Breslin et al., 2012; Goodway & Rudisill, 1997; Tietjens et al., 2018).

Such conflicting findings have led researchers to critically examine the Stodden et al. (2008) conceptual model in more critical detail and thus present more updated iterations of the model within the literature. In 2015, a narrative review was published to synthesise evidence related to the original model, and a figure was produced to summarise the available evidence at this time (see Figure 2.2) (Robinson et al., 2015). The findings of this review provided convincing evidence that AMC was positively associated with PA, cardiorespiratory fitness, muscular strength, muscular endurance and healthy weight status. Whilst the review did note that evidence indicated AMC to likely be positively associated with PMC, there was an appreciation that it is difficult to truly ascertain strengths of association between PMC and AMC as PMC assessments do not closely align with AMC assessments in terms of particular skill domains (Barnett, Ridgers, Zask, et al., 2015) or even general measures of self-concept (Babic et al., 2014). The review did however note that two Australian studies (Barnett, Ridgers, & Salmon, 2015; Liong et al., 2015) with aligned measures of AMC and PMC in young children, finding positive associations but that further research would be required to see whether the strength of association differed for different skill or activity types and across age.

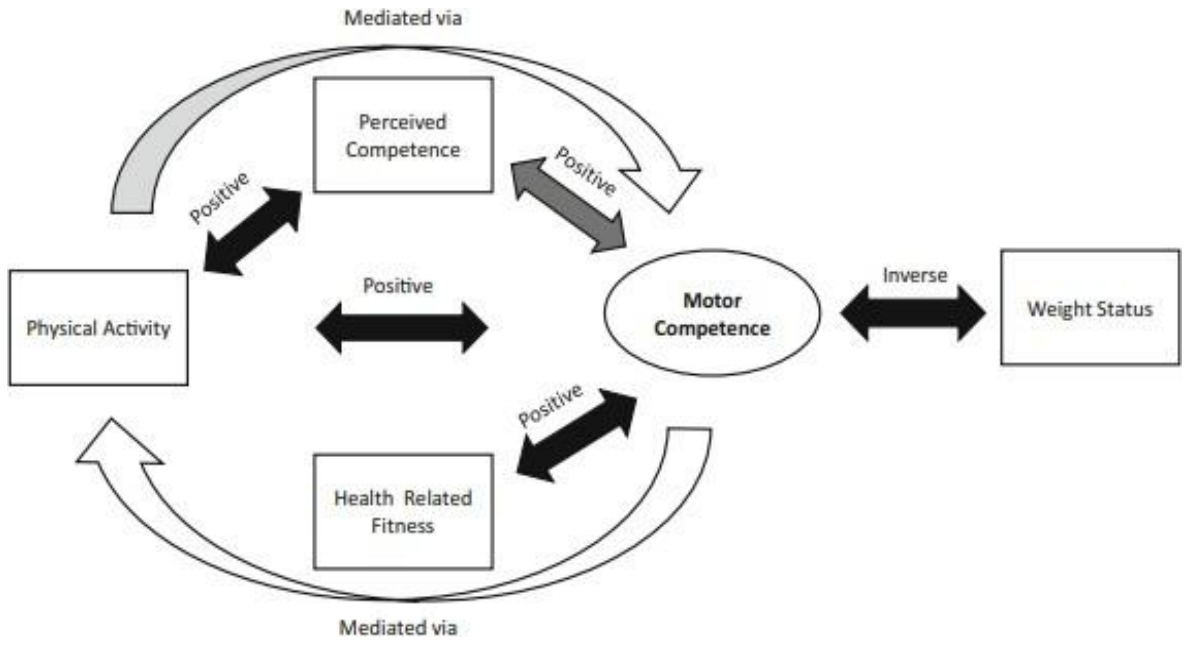


Figure 2.2: Research consensus on motor competence and health related variables. Reproduced from Robinson et al. (2015)

In 2022, a systematic review by Barnett et al. (2022) aimed to systematically compile mediation, longitudinal and experimental evidence in support of the Stodden et al. (2008) conceptual model. The authors produced a figure (see Figure 2.3) to summarise the findings of their review. In terms of the pathway from PMC to AMC and vice versa, the authors concluded that evidence associating PMC and AMC in either direction was insufficient and inconsistent. It was noted that such findings were likely due to the small number of longitudinal/intervention studies conducted in the 5 years prior. De Meester et al. (2020) found that 88.5% of the 87 studies included in their systematic review were cross-sectionaldemonstrating that the PA research field is in need of an abundance of longitudinal studies. A factor hypothesised to be potentially important in enhancing associations between PMC and AMC is alignment between the measure used for assessing both variables. However, the results from the review by Barnett et al. (2022) (albeit based on a small number of studies) suggested that more aligned measures did not moderate the association. Similarly, a review published two years prior by De Meester et al. (2020) did not find that the level of instrument alignment moderated the association between AMC and PMC.

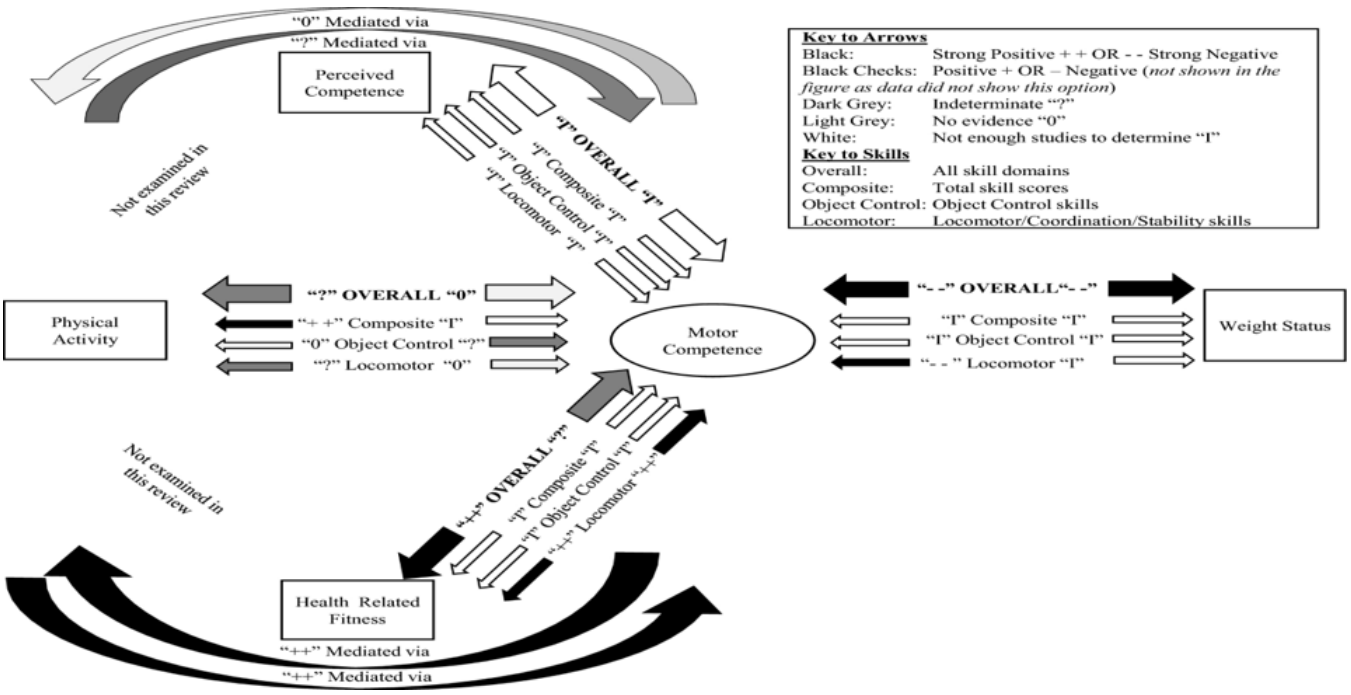


Figure 2.3: Levels of evidence for each pathway, summarised by skill domain. Reproduced from Barnett et al. (2022)

Further interations of the original Stodden et al. (2008) model was produced in 2022 when Lima et al. (2022) proposed an extension to the original model (see Figure 2.4). The extension proposed that variables from the original model would be assumed to be directly associated with weight status. Moreover, PA, AMC, health-related fitness, and weight status would be expected to be directly associated with variables such as metabolic health, cognition and academic performance and mental health. In proposing an extension to the model, Lima et al. (2022) did not propose PMC to be a key contributor towards the new chain of variables. The authors concluded their review by suggesting a need for longitudinal research, particularly experimental research using different doses of interventions to better critically examine the original Stodden et al. (2008) model in detail. Furthermore, they suggested that experimental studies should evaluate the possible causal chain of the associations among AMC, PMC, PA and health-related fitness with metabolic health, mental health, and cognition and academic performance, including as many components from the Stodden et al. (2008) original model as possible.

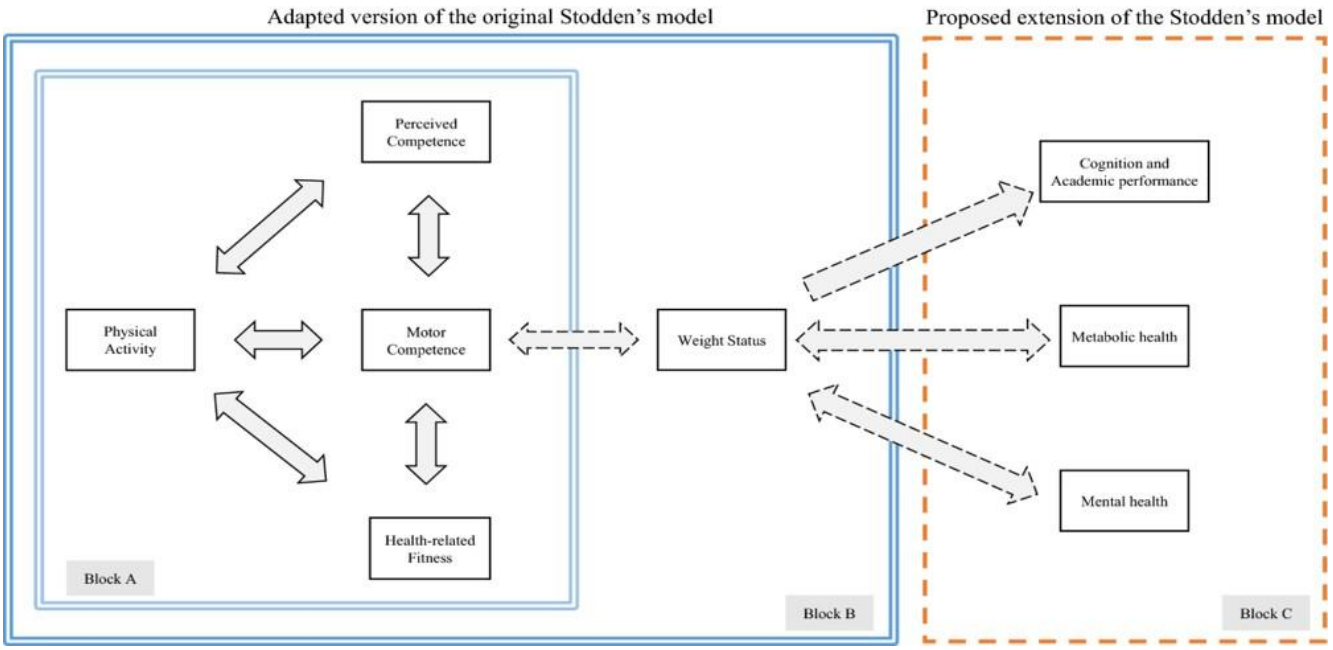


Figure 2.4: Proposed extension of the Stodden et al. (2008) model. Reproduced from Lima et al. (2022)

Such claims for further longitudinal research are warranted when the existing literature is considered. Upon review there have been significantly more cross-sectional studies completed examining the aforementioned pathways (Barnett et al., 2013; Barnett et al., 2011; Behan et al., 2019; Cohen et al., 2014; D'Hondt et al., 2009; Hardy et al., 2010, 2013; Jones et al., 2010; Kemp & Pienaar, 2013; Larouche et al., 2014; LeGear et al., 2012; Lopes et al., 2012; Martins et al., 2024; Okely & Booth, 2004; Okely et al., 2001; Spessato et al., 2013; Vandendriessche et al., 2012) when compared to longitudinal studies (Barnett et al., 2010; D'Hondt et al., 2013; Jaakkola et al., 2013; Maccobb et al., 2005; Viholainen et al., 2006).

Barnett et al. (2022) noted that conclusions on the mediation outcomes for the pathways within the Stodden et al. (2008) model were weakened by the predominantly cross-sectional nature of the available evidence. The authors concluded their review by noting that the reliance upon many cross-sectional studies for evidence creates a bias, as the proximal measurement of variables is likely to contribute to more associations. Thus, to truly test the pathways proposed in the Stodden et al. (2008) model, the field requires more robust longitudinal studies across early childhood and into adolescence that include multiple variables from the model, have

multiple time points and account for potential confounding factors (Barnett et al., 2022). In summary, the literature generally supports the theory of AMC and PMC being (at minimum) loosely associated with one another. However, it must be noted that the strength of this association is ultimately indeterminate given the lack of longitudinal research conducted and therefore, the theory of PMC being the primary influencer upon AMC proficiency and vice versa requires further longitudinal research with longer-term follow ups to draw stronger conclusions (Lima et al., 2022). In addition to these research design limitations, the manner in which these relationships have been statistically analysed may also contribute to this uncertainty. Given recent calls for more rigorous research designs and robust forms of analyses to be employed in the PA research domain (Caldwell et al., 2020; Melby et al., 2022), it is therefore key to consider how these constructs of MC development are operationalised and analysed within the existing literature, and whether current approaches adequately capture these constructs in a sufficiently objective and reliable manner.

2.7 Analytical Approaches in Motor Competence Research

2.7.1 Traditional Statistical Methods

Within research focused on AMC, group-level comparisons have formed a central component of empirical literature. Studies have consistently relied on independent samples t-tests and analyses of variance (ANOVA) to examine differences in AMC proficiency across gender and age (Behan et al., 2019; Bolger et al., 2018; Foulkes et al., 2015; O'Brien et al., 2016). These approaches are well aligned with testing directional, theory driven hypotheses, such as the expectation that boys outperform girls in object control skills, or that AMC improves with age until a developmental plateau is reached. However, their repeated application within largely similar cross-sectional designs has contributed to a body of evidence characterised by consistent and predictable findings. While such consistency strengthens confidence in these findings, it also highlights a tendency within the literature towards describing group differences rather than advancing understanding of the mechanisms underpinning MC development and variation. Across studies, similar patterns emerge, including gender differences favouring boys in object control, age-related improvements in locomotor skills, and overall proficiency levels that fall below normative benchmarks (Behan et al., 2019; O'Brien et al., 2016). As illustrated in Table 2.1, there is a clear convergence in methodological approaches across recent Irish and UK studies, with similar analytical frameworks yielding similarly well-established, but

ultimately constrained, conclusions.

Table 2.1 Sample of Statistical Models Adopted in Motor Competence Research

Study & Authors	**Participants & Statistical Models Used**	**What Was Analysed**	**Results**
Fundamental Movement Skills of Preschool Children in Northwest England (Foulkes et al., 2015)	n = 168 Independent Sample T-Tests & Univariate Analyses of Covariance (ANCOVA) models	Independent Sample T-Tests - Assessed differences in total, locomotor, and object-control scores ANCOVA - Examined gender differences in total and subscale scores, controlling for age, deprivation score and Body Mass Index z score.	No significant differences in either total ($p = .41$) or locomotor ($p = .11$) score were observed between gender. However, a significant difference in object-control score was found ($p = .002$), with boys showing greater competence than girls.

Fundamental movement skill proficiency amongst adolescent youth (O'Brien et al., 2016)	n = 242 Independent Sample T-Tests	Gender differences in individual AMC performance	There was a significant difference in the mean composite FMS score (total object control and locomotor) between genders, with males scoring higher; $t(221) = 2.454$, $p < .05$. An independent t-test showed that male participants obtained an overall higher object control score compared to females; $t(221) = 3.382$, $p < .01$.

Age and gender Differences in Fundamental Movement Skills Among a Cohort of Irish School Children (Bolger et al., 2018)	n = 203 Analysis of Variance (ANOVA) & One Sample T-Tests models	ANOVA - Investigate differences across total FMS scores, locomotor subset scores and object-control scores with respect to age and gender. One Sample T-Tests - Compare mean age and gender specific subset scores with those of a US reference sample included in the TGMD-2 manual	Older children scored significantly higher than younger children in both locomotor ($p < .05$) and object-control score ($p < .05$). Boys scored significantly higher than girls in object-control score ($p < .05$), while girls scored significantly higher in locomotor score ($p < .05$).

Self-Perceived and Actual Motor Competence in Young British Children (Duncan et al., 2018)	n = 258 Independent Sample T-Tests & ANCOVA models	Independent Samples T-Test - Examine differences in PMC between boys and girls. ANCOVA - Examine whether AMC differed by gender and the PMC classifications	Boys obtained significantly higher scores than girls for both the process ($p = .044$) and product ($p = .001$) measures of AMC. Boys had significantly ($p = .04$) higher scores for PMC compared to girls. Compared to children classified as medium and high self-perceived competence, children classified as low self-PMC had lower process ($p = .001$) and product scores (i.e., medium, $p = .009$ and high, $p = .0001$) of AMC.

Moving Well-Being Well: Investigating the maturation of fundamental movement skill proficiency across gender in Irish children aged five to twelve (Behan et al., 2019)	n = 2098 Independent Sample T-Tests & One Way Between-Groups ANOVA models	Independent Sample T-Test - Analyse differences between gender in the locomotor, object control and balance subtest scores, as well as the overall FMS totals One Way Between-Groups ANOVA - Examined the impact of age on overall FMS, locomotor, object control, and balance skills.	Girls scored significantly higher than boys in the locomotor and balance subtests with the boys outperforming the girls in object control skills. Improvements in ability can be seen over time ($F(8,1968)=70.18$, $p<0.001$), with significant increases in FMS proficiency seen up to the age of ten, after which proficiency begins to decline.
Differences in Weight Status and Autonomous Motivation towards Sports among Children with Various Profiles of Motor Competence and Organized Sports	n = 206 Independent Sample T-Tests & One-way ANOVA models	Independent Samples T-Test - Examine whether boys and girls significantly differed from each other in terms of the outcome variables One-way ANOVA - Analyze potential	AMC was moderately correlated with PMC ($r = 0.502$, $p < 0.001$) and weakly correlated with organised sports participation ($r = 0.308$, $p < 0.001$). Superior AMC proficiency was associated with lower

Participation (Coppens et al., 2021)		age-related differences in the cluster and outcome variables.	BMI ($r = -0.292$, $p < 0.001$) and higher autonomous motivation ($r = 0.302$, $p < 0.001$).

The Independent Sample T-Test has proven to be a popular statistical model in comparing group means with numerous authors utilising the method (Behan et al., 2019; Coppens et al., 2021; Duncan et al., 2018; Foulkes et al., 2015; O'Brien et al., 2016). Foulkes et al. (2015) noted their independent samples t-test results showed no significant gender differences in total score ($p = .41$) or locomotor score ($p = .11$). However, a significant difference was found in object-control score ($p = .002$), with boys showing greater competence than girls. O'Brien et al. (2016) noted male participants to obtain an overall higher object control score compared to females. Additionally, no gender difference was noted in the overall locomotor mean score performance, females did however outperform males with significantly higher advanced skill proficiency in the skip.

On the contrary, males displayed superior skill proficiency in the overhand throw, run and horizontal jump. Using the same approach, Duncan et al. (2018) assessed whether any difference existed in PMC between young British boys and girls with significant differences found in favour of boys over girls. Behan et al. (2019) compared overall AMC proficiency for boys and girls across the Irish primary school cycle whilst also comparing the three individual AMC subset scores (i.e., object-control, locomotor and balance skills). The results of Behan et al. (2019) noted significant differences between gender in overall AMC proficiency among eleven- and twelve-year-olds, however on the contrary, girls were noted to produce significantly higher scores in locomotor skills than boys at age, five, eight, nine and ten. Coppens et al. (2021) conducted independent samples t-tests to examine whether boys and girls significantly differed from each other in terms of the outcome variables (i.e., weight status and autonomous motivation towards sports). Their independent sample t-tests analysis revealed no significant differences between boys and girls in weight status, and autonomous motivation towards sports in the total sample. Furthermore, weight status and autonomous motivation towards sports did not differ between age groups, therefore both gender and age were not taken into consideration in the subsequent analyses. Several forms of ANOVAs have

been employed across recent studies to examine whether significant differences exist between the means of two or more groups (Behan et al., 2019; Bolger et al., 2018; Coppens et al., 2021; Foulkes et al., 2015). Foulkes et al. (2015) utilised ANCOVA analysis as they sought to assess group means differences whilst controlling for age, deprivation score, and BMI z-score, and noted that boys still showed significantly greater competence in object-control skills compared to girls ($p = .002$), however, no significant gender differences were noted in total score ($p = .41$) or locomotor score ($p = .11$). Bolger et al. (2018) noted a significant age effect for both locomotor and object-control subset scores whilst 10-year-old children were shown to score higher than their younger counterparts in both locomotor (medium to large effect size) and object-control subset scores (large effect size). A significant main effect for gender was also found, with girls scoring higher than boys in locomotor score (small effect size) while boys scored significantly higher than girls in object-control score (large effect size). Duncan et al. (2018) sought to examine whether AMC (via both TGMD-2 scores and product measures) differed by gender and the PMC classifications and analysed data using 2 (gender) × 3 (high, medium, low PMC) ANCOVA models, controlling for both age and BMI. The ANCOVA analysis revealed significant main effects for both gender and PMC. Boys had higher TGMD-2 scores, and children classified as low perceived competence had poorer composite Z-scores for product AMC compared to those classified as medium and high PMC.

Behan et al. (2019) sought to test the commonly held hypothesis that AMC proficiency improves through maturation (Robinson & Goodway, 2009) by using a one-way between-groups ANOVA with planned comparisons used to explore the impact of age on overall AMC, locomotor, object control, and balance skills. They found statistically significant age effects across all four outcome measures (overall AMC, locomotor, object control, and balance) with large effect sizes throughout except balance, which reached a moderate effect size. Taken collectively, the studies reviewed above highlight the extent to which traditional statistical approaches have produced a consistent, but relatively narrow, evidence base. Across varied national contexts, age groups, and sample sizes, similar conclusions are repeatedly observed: boys tend to demonstrate higher proficiency in object control skills, girls often perform comparably or better in locomotor domains, and AMC generally improves with age until reaching a developmental threshold, beyond which progress stabilises or declines. While these patterns are both theoretically meaningful and practically informative, their consistent replication across predominantly cross-sectional designs reflects a degree of methodological convergence that constrains the potential for new insight. Consequently, although the

evidence base is robust in establishing age- and gender-related differences in AMC, comparatively less attention has been devoted to understanding the developmental, behavioural, and environmental mechanisms that underpin these differences. The implications of this methodological convergence are taken up in the following subsection.

2.7.2 Summary and Future Directions

Traditional statistical approaches have played a key role in advancing MC research, providing robust and interpretable methods for testing clear, theory-driven hypotheses. The patterns they have established, gender differences in object control, age-related improvement in proficiency, and competency levels falling below normative benchmarks, constitute some of the field's most secure knowledge. The limitation identified in this thesis is therefore not that such methods are poor. Rather, it is that the field has tended to re-apply the same designs and the same analyses in a copy-and-paste fashion: cross-sectional samples, group comparisons by gender and age conducted via t-tests and analyses of variance, and conclusions drawn at the group level, as documented in Section 2.7.1 and Table 2.1. The predictable consequence is a body of evidence that is consistent but incremental, in which the replication of established differences has increasingly substituted for the interrogation of mechanisms. This concern is not unique to this thesis. Methodological critiques within the PA and MC literatures have repeatedly called for more rigorous designs and more robust analytical strategies (Caldwell et al., 2020; Melby et al., 2022; Song et al., 2021), with recent work arguing specifically that MC research requires approaches capable of representing the complexity and interdependence of motor-skill development (Bandeira, Estevan, et al., 2025).

Three questions in particular remain poorly served by group-level inferential designs: why certain children struggle to develop competence in specific skill components; how the multiple determinants of AMC interact in combination; and how developmental trajectories vary at the level of the individual child. It is this gap, an inability to capture complexity and heterogeneity within the data rather than any defect in the methods themselves, that provides the rationale for the data-driven analytical approaches examined in the following section.

2.8 Machine Learning and Data-Driven Methods

ML is a subfield of artificial intelligence (AI) concerned with producing mathematical models that identify and reason about dependencies between the variables describing a phenomenon (Imam et al., 2024), and its adoption has expanded rapidly across science, technology and healthcare, driven by new learning algorithms, the growing availability of data and low-cost computation (Jordan & Mitchell, 2015). Within this thesis, three families of analytical approach are distinguished, and the distinction is maintained throughout the chapters that follow.

Supervised learning refers to models trained on labelled examples in order to predict a known outcome, encompassing classification, where the outcome is categorical (for example, classifying a movement skill performance as proficient or not), and regression, where the outcome is continuous. Unsupervised learning refers to methods that identify structure within unlabelled data, such as clustering or dimensionality reduction, without a predefined outcome (Oladipupo, 2010). Data-driven statistical modelling, by contrast, refers to techniques that estimate the relationships or latent structure among a set of variables, including network analysis, exploratory graph analysis and structural equation modelling. Although such methods are data-driven, they are more accurately described as network psychometrics than as machine learning (Epskamp & Fried, 2018), and they are treated as distinct from ML throughout this thesis. Logistic regression occupies an intermediate position, functioning both as a classical inferential method and, within a predictive pipeline, as a supervised classifier (Levy & O'Malley, 2020); where it is used in the latter sense in this thesis, this is stated explicitly.

Taken together, supervised and unsupervised approaches offer a flexible and scalable framework for analysing complex, high-dimensional data. Their advantage over traditional inferential methods is, however, frequently misstated. It does not lie in freedom from statistical assumptions, since ML methods carry assumptions of their own, including the independence of observations and the need for sample sizes commensurate with model complexity (Bashir et al., 2020; Song et al., 2021), although many ML algorithms are less dependent on assumptions of linearity and normality and more robust to multicollinearity among predictor variables (Song et al., 2021). Rather, the advantage lies in flexibility. ML methods allow the data to shape the model rather than imposing a fixed parametric structure from the outset (Imam et al., 2024), and they are accordingly better suited to capturing non-linear

relationships, higher-order interactions and latent structures that fixed parametric models are not designed to detect (Song et al., 2021). As such, these approaches present a promising avenue for addressing many of the methodological limitations previously identified within MC research. The following section examines their application within this domain, with particular attention to how they have been used to enhance the assessment, classification and understanding of MC.

2.8.1 Applications of Machine Learning in Motor Competence Research

Building on the limitations identified in the previous sections, researchers have increasingly turned to ML techniques to capture more complex, predictive relationships within MC datasets. The application of such techniques within this field has a longer history than is often acknowledged, with predictive modelling approaches appearing in the literature more than a decade ago (Barnett et al., 2009; Hardy et al., 2013). Logistic regression (LR) occupies a distinctive position within this context, as it functions both as a traditional inferential method and as a predictive classifier within ML frameworks (Levy & O'Malley, 2020). This dual role made it an accessible entry point for MC researchers beginning to engage with predictive modelling, allowing a shift towards data-driven approaches without requiring a complete departure from the inferential methods that have traditionally dominated the field. For this reason, LR features prominently in the early stages of ML-related work within the MC literature.

Barnett et al. (2009) sought to use LR to examine the probability of participating in some vigorous activity, as a function of AMC proficiency in childhood. LR has been noted to be adopted in the years that followed, such as Hardy et al. (2013), Henrique et al. (2016) and Henrique et al. (2020). Barnett et al. (2009) noted predicted values from the LR model showed that children with good (>10) object control skills would have at least a 20% greater chance of participating in at least some vigorous PA in adolescence, compared with those with poor (<5) object control skills. Hardy et al. (2013) found that after adjusting for fitness, BMI, and socioeconomic status, primary and high school boys were 1.3 times as likely to improve proficiency in the kick, 1.6 times in the vertical jump, and 3.6 times in the catch in 2004 compared with that in 1997. Additionally, girls were 2.0 times in the vertical jump, and 3.6 times in the catch in 2004 compared with that in 1997 (Hardy et al., 2013). Henrique et al. (2016) found that participation in sports at baseline was the strongest predictor of sports participation after two years. Furthermore, each one-point increase in locomotor skills

represented an increase of approximately 21% probability of participating in sports after two years. Weight status was not a significant predictor of sports participation, but it was important for the model fit (Henrique et al., 2016). Henrique et al. (2020) noted that the LR model was statistically significant and revealed that an increase in age and an increase in locomotor skills performance were associated with a reduction in the likelihood of exhibiting central obesity ≥ 0.5. However, gender and object control skills were not associated with the prevalence of central obesity ≥ 0.5.

Taken together, these studies demonstrate that even relatively straight forward ML classifiers can identify predictive relationships between MC and health-related outcomes. Crucially, these relationships extend beyond what can typically be detected using traditional group-level inferential methods, which are not designed to capture complex, multivariate patterns within the data. As such, the application of ML highlights the potential to uncover more informative and practically relevant associations that may otherwise remain obscured within conventional analytical frameworks. Whilst these studies establish the utility of ML for predictive modelling in MC research, more recent work has directed these techniques specifically toward the commonly used TGMD assessment battery, with a focus on enhancing both its predictive validity and its practical efficiency in applied settings.

2.8.2 Machine Learning Applications in TGMD Assessment

There has been a growing need to develop more time-efficient and practical versions of the TGMD assessment batteries without compromising validity. In response, several studies have sought to enhance the original TGMD by developing and validating shorter-form versions (Bandeira et al., 2020; Duncan, Martins, et al., 2022; Valentini et al., 2018), with some also incorporating data-driven statistical modelling in the form of network analysis.

The approaches used in these studies are summarised in Table 2.2. Such studies have incorporated network analysis (NA) to examine the relationship between skills in the TGMD. In this case a network model is depicted by a set of nodes (circles) representing variables (e.g., MC variables) and edges (lines) representing the partial correlation between two nodes conditioned on all other nodes (Epskamp & Fried, 2018). In NA, the role of each skill in the network, assessed from the strength centrality indicator, is equivalent to the factorial load of the skill in the latent model (Christensen & Golino, 2021). The network perspective for

instrument validation provides important information on the relationships between the tested AMC skills and those skills that are most important in different contexts and settings, such as for boys and girls (Duncan, Martins, et al., 2022).

Table 2.2: Short form versions of the original TGMD assessment batteries

Study	**Aim**	**Statistical Approach**	**Study Findings**
The development of a short form of the Test of Gross Motor Development-2 in Brazilian children: Validity and reliability (Valentini et al., 2018)	Develop a short form of the TGMD-2 and to examine its validity, interrater reliability and test–retest reliability.	Exploratory Factor Analysis (EFA) & Confirmatory Factor Analysis (CFA)	The construct validity of the factorial short form indicated that the six skills correlated satisfactorily with each correspondent subtest, locomotor (i.e., run, gallop, and hop) and object control (i.e., strike, kick, and throw).
TGMD-2 short version: evidence of	Analyse the evidence of validity	Confirmatory Factor Analysis and	A two-factor model

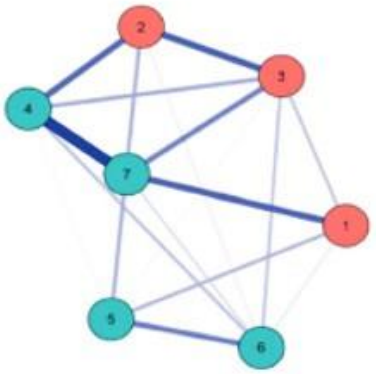
Validity and Associations with gender, age and BMI in preschool children (Bandeira et al., 2020)	and reliability of the TGMD-2 for low-income preschoolers; and to investigate the associations between the final model with gender, age and BMI.	Network Analysis	(locomotion and object control) with adequate values for the six skills (gallop, leap, slide, strike, throw and roll) and presented excellent indexes. The network analysis to associate AMC with gender, age and BMI seems useful to optimise future interventions.

TGMD-3 short version: Evidence of validity and associations with gender in Irish children (Duncan, Martins, et al., 2022)	Examined the internal structure and evidence of validity of the TGMD-3 short-form in primary school aged children.	Confirmatory Factor Analysis and Network Analysis	A two-factor model (locomotion and object control) with adequate values for the seven skills (gallop, hop, jump, two-hand strike, bounce, catch, overhand throw) presented excellent indexes. The use of a network analysis approach confirmed that the skills included in the TGMD-3 short version are similar for boys and girls.

In the case of Bandeira et al. (2020) and Duncan, Martins, et al. (2022) both studies adopted network analysis, a data-driven statistical modelling approach, examples of the NA models from Duncan, Martins, et al. (2022) and Bandeira et al. (2020) are shown in Figures 2.5 and 2.6 respectively. Indeed, the use of NA can be considered useful to discriminate differences between countries, and to determine the priority of intervention proposals, respecting cultural aspects, as previously suggested (Barnett et al., 2016; Hulteen et al., 2018). Whilst the results from such NA highlight the associations between AMC skills, a key limitation of these studies are that the feature weights representing these skills in the NA model were not highlighted. Thus, whilst the relationships between skills are visually demonstrated there is no way of quantifying the magnitude of these relationships. Another key limitation in such studies is the lack of coinciding PMC scores for each specific skills; given the aforementioned discussion upon the potential importance of PMC on AMC proficiency in previous sections, the inclusion

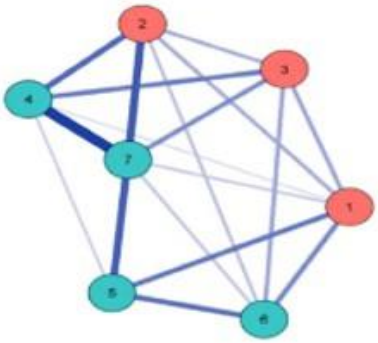
of such variables may have influenced the models produced in these studies.

a)



Orange circles = locomotion skills (1: Gallop; 2: Hop; 3: Horizontal Jump); blue circles = object control skills (4: Catch; 5: Throw; 6: Two Hand Strike; 7: Bounce).

b)



Orange circles = locomotion skills (1: Gallop; 2: Hop; 3: Horizontal Jump); blue circles = object control skills (4: Catch; 5: Throw; 6: Two Hand Strike; 7: Bounce).

Figure 2. Network perspective of fundamental movement skills for a) girls Orange circles = locomotion skills (1: Gallop; 2: Hop; 3: Horizontal Jump); blue circles = object control skills (4: Catch; 5: Throw; 6: Two Hand Strike; 7: Bounce). and b) boys Orange circles = locomotion skills (1: Gallop; 2: Hop; 3: Horizontal Jump); blue circles = object control skills (4: Catch; 5: Throw; 6: Two Hand Strike; 7: Bounce).

Figure 2.5: Network analysis model. Reproduced from Duncan, Martins, et al. (2022)

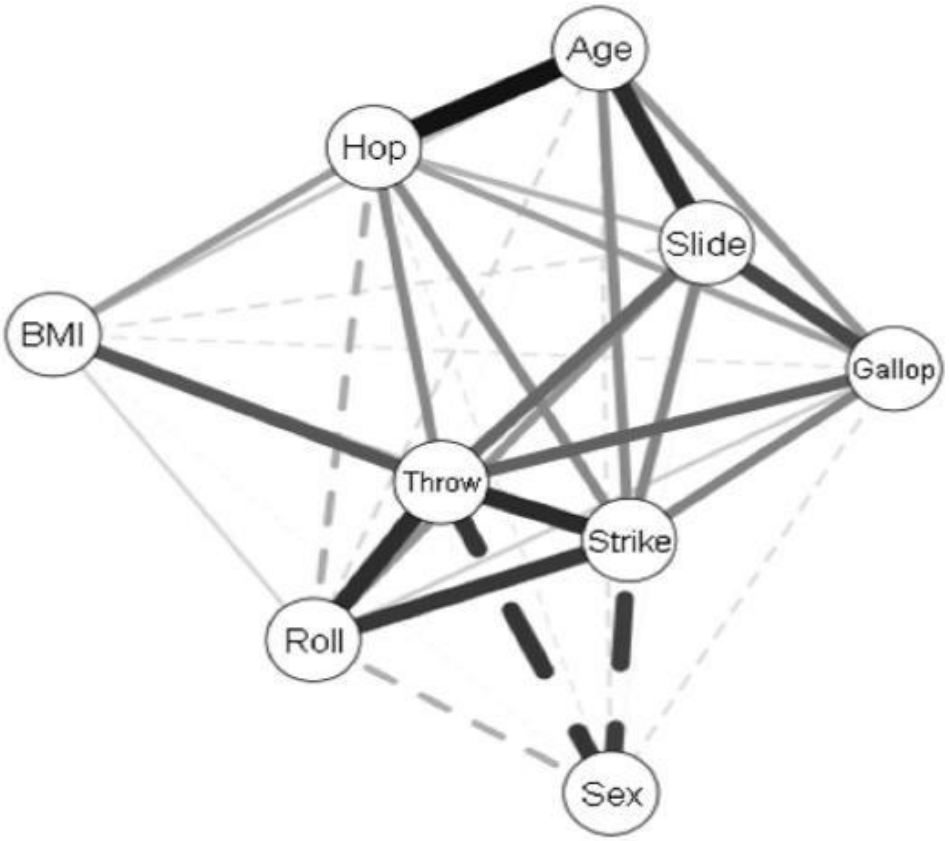


Figure 2.6: Network analysis model. Reproduced from Bandeira et al. (2020)

Recently, literature has begun to appreciate such aforementioned limitations and thus broaden the range of variables which can be modelled within network structures. Earlier studies (Bandeira et al., 2020; Duncan, Martins, et al., 2022) utilizing the TGMD-2 and TGMD-3 predominantly classified FMS into two predefined domains—locomotor and object control—using Confirmatory Factor Analysis (CFA) and correlation-based network models. CFA examines whether data fits a predefined factor structure whilst Exploratory Factor Analysis (EFA) examines which underlying variables naturally emerge from a dataset. While these studies provided valuable insights into skill development, they assumed motor skills to operate within fixed categories, potentially overlooking the complex interconnections between different aspects of MC. Bandeira, Lemos, et al. (2025) recently sought to advance our current understanding of such complex MC interconnections by employing Exploratory Graph Analysis (EGA), a network-based method that allows data-driven identification of MC structures.

Instead of relying on predetermined constructs, the study found a four-dimensional model that better reflects how motor skills interact. Notably, the study's findings (highlighted in Figure 2.7) challenge conventional classifications by revealing that certain object control skills, such as kicking, share movement characteristics with locomotor patterns due to the need for acceleration and deceleration. Additionally, object control skills were divided into process-based (e.g., throwing technique) and product-based (e.g., throwing speed) measures, while balance and coordination tasks formed a distinct category, reinforcing their critical role in overall MC proficiency. This multidimensional approach presented by Bandeira, Lemos, et al. (2025) integrated multiple assessment tools, including TGMD-2 (process-oriented), the Motor Competence Assessment (product-oriented) (Luz et al., 2016) an assessment of locomotor, manipulative and stability skills for preschool children, and Supine-to-Stand (Nesbitt et al., 2016), an assessment of functional movement ability, providing a more holistic view of MC.

By treating MC as a self-organizing system rather than a static construct, the study highlighted the need for assessment tools that capture interconnected networks of skills rather than isolated domains. This perspective has important implications for early childhood MC assessments, suggesting that both process- and product-oriented measures should be combined to achieve a more comprehensive evaluation. The findings from this network-based approach help shift the paradigm in MC research, emphasizing the complexity of skill development and the role of cultural and environmental factors in shaping motor skill proficiency. Unlike previous research that categorized skills into fixed domains (Bandeira et al., 2020; Duncan, Martins, et al., 2022), EGA suggests that MC emerges from dynamic interconnections between skills, which can vary based on context. This new perspective challenges traditional assessment models and supports the adoption of more flexible, data-driven approaches in evaluating and developing motor skills in young children.

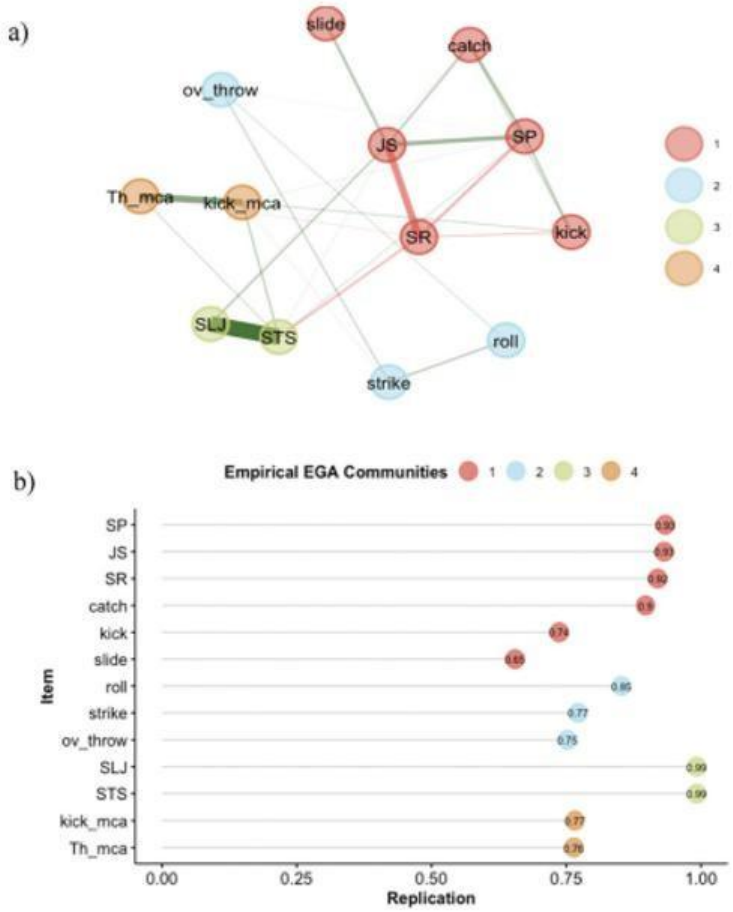


Figure 2. a) dimensionality structure of motor competence using Exploratory Graph Analysis for Model 2; b) item replication in the Model 1EGA dimensions; h_jump = jump; ov_throw = overhand throw; SR = shuttle run; SLJ = standing long jump; Th_mca = throwing; SP = shifting platforms; JS = jumping sideways; STS = supine-to-stand.

Figure 2.7: Exploratory graph analysis and item replication in the EGA dimensions. Reproduced from Bandeira, Lemos, et al. (2025)

However, while Bandeira, Lemos, et al. (2025) provided strong evidence that MC is structured as a network, it did not account for how these relationships evolve over time or the broader influences beyond motor skill tasks themselves. Expanding on these findings, Bandeira, Estevan, et al. (2025) refined and extended this network-based approach by introducing a Multilayer Network Model. The Bandeira, Estevan, et al. (2025) model moves beyond a static depiction of MC, incorporating elements from complex systems theory to show how MC dynamically emerges from multiple interrelated domains. Unlike the authors' prior work, (Bandeira, Lemos, et al., 2025), which primarily examined motor tasks in isolation, this new framework integrates biological, psychological, social, and environmental factors,

emphasizing that the importance of different variables shifts over time and across contexts (highlighted in Figure 2.8). By framing MC as a self-organizing, non-linear system, this study highlights that its development cannot be fully understood without considering the dynamic interactions between various influencing factors.

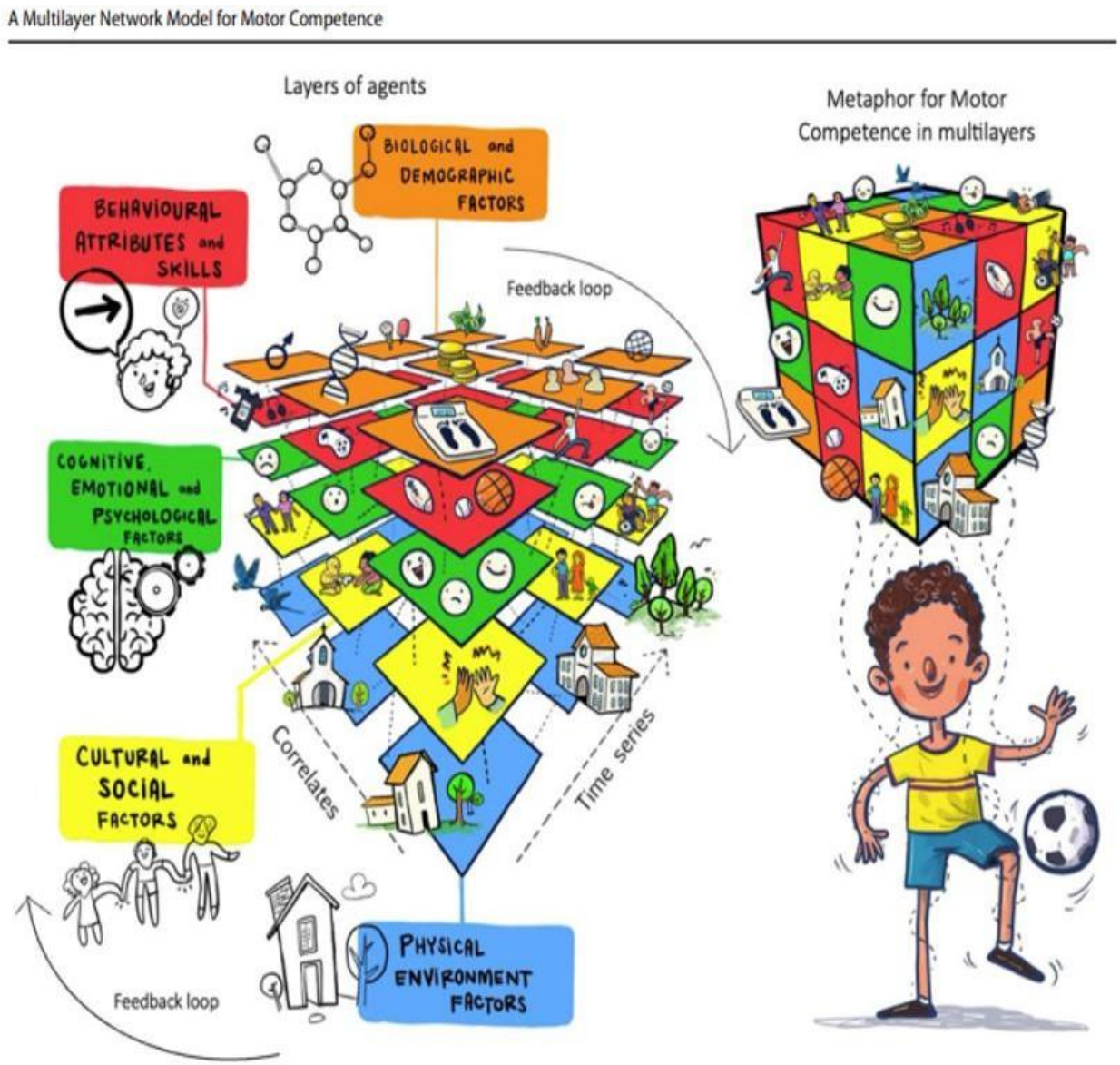


Figure 2.8: Multilayer network model of MC. Reproduced from Bandeira, Estevan, et al. (2025)

This conceptual expansion also brings methodological advancements. While Bandeira, Lemos, et al. (2025) used graph-based modelling to identify MC dimensions, Bandeira,

Estevan, et al. (2025) applied multilayer network science, thus, enabling researchers to examine how multiple factors interact across different layers and time points (highlighted in Figure 2.9).

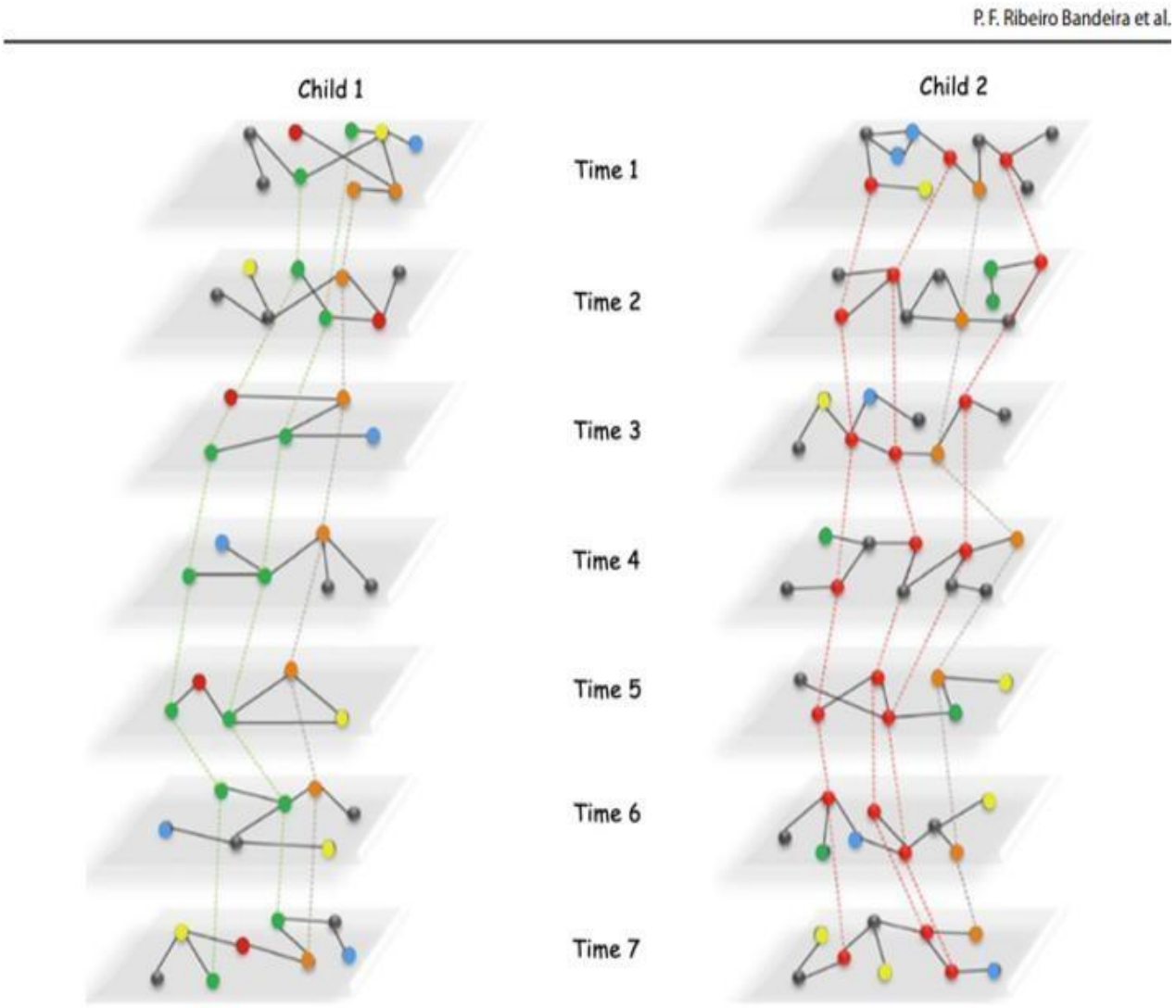


Figure 2.9: Multilayer network emergent patterns of MC. Reproduced from Bandeira, Estevan, et al. (2025)

This deeper, systems-based approach enhances our ability to predict and track changes in MC development, offering a more precise foundation for assessment and intervention. By acknowledging that MC is shaped by varying influences at different life stages, the concepts presented by significantly advances the field, bridging the gap between traditional MC assessments and a more holistic, developmental perspective.

Overall, the suite of NA studies presented in the current literature have highlighted the interdependencies that exist between sub-components of children's MC (i.e., locomotor, object-control, stability skills) revealing their interactions and evolution over time. Importantly however, whilst such analyses have added further insight to the domain, their impact on practitioners remains limited given not all AMC skills discriminate in assessing AMC proficiency and are actually related to sports and PA according to the context/culture where they are assessed (Barnett et al., 2016; Newell, 2020).

Therefore, to truly advance MC research, there are more prudent structural procedures which must be thoroughly examined and, in cases, reimagined if best-practices within MC research are to be advanced. Previous sections in this thesis have discussed MC measurement procedures with popular assessments such as the TGMD-3 (Ulrich, 2013) and BOT-2 (Bruininks & Bruininks, 2005) following an 'observe and score' approach. In isolation such instruments have been shown to be valid and reliable for the assessment of children's MC proficiency, however, when considered at scale, the utilisation of such instruments has been questioned (Bisi et al., 2017; Chow et al., 2020; Duncan, Weldon, et al., 2022). Furthermore, subjectivity in assessor scoring, data precision limitations and inconsistencies in aggregation and interpretation of performance criteria may limit scoring consistency and comparability across studies (Ng et al., 2025).

Given the aforementioned benefits and widespread global implementation of ML across all forms of society, it must be considered if such approaches could be implemented within real world MC assessment procedures. Whilst the implementation of ML into MC research is key it is also imperative to reimagine practical real-world approaches which can be implemented to aid practitioners in assessing MC proficiency. PA research has made extensive use of motion capture technologies, typically classified as optical (camera-based) and non-optical (sensor-based, e.g., inertial units) systems, alongside the growing use of advanced ML analyses. Therefore, the following section shall consider the use of such systems intertwined with ML analyses in PA measurement research, after which we will consider if such approaches could be applicable for implementation in real-world MC assessment.

2.8.3 Movement Classification and Analysis in PA Research

One of the most reliable means of assessing children's PA is through the use of accelerometers to measure movement intensity and frequency (Konstabel et al., 2014). This method provides quantified values for activity intensity over time. However, methodological variations in data gathering and analysis methods used between studies have led to a growing saturation of conflicting cut-points (the quantitative boundaries between PA intensity classes) in the literature (Trost et al., 2011). Therefore, several new research directions have been proposed to attempt to address this issue. One approach is the use of ML techniques for the detection and prediction of activity types and their associated intensity (Ahmadi, Pavey, et al., 2020; Ahmadi, Pfeiffer, et al., 2020; Ahmadi & Trost, 2022; Csizmadia et al., 2022; Fergus et al., 2017; Goodlich et al., 2020; Kwon et al., 2022). Several well-cited papers from recent literature are highlighted below in Table 2.3.

Table 2.3: Sample of ML Methods Used for Classification of PA Type and Intensity

Study / Author	**Aims**	**Measurement Tool**	**ML Model**	**Activities Measured**
Machine Learning Models for Classifying Physical Activity in Free-Living Preschool Children (Ahmadi, Pavey, et al., 2020)	Evaluate the accuracy of ML activity classification models for preschool-aged children	Accelerometers	Random Forest	Five Activity Classes; Sedentary, light activity and games, moderate to vigorous activities and games, walking and running
Physical Activity Classification in Youth Using Raw Accelerometer Data from the Hip (Ahmadi, Pfeiffer, et al., 2020)	Develop and evaluate ML algorithms to predict children's PA category from raw accelerometer data collected at the hip	Accelerometers	Random Forest & Logistic Regression	Sedentary, light household activities and games, moderate-vigorous games and sports, walking and running

Device-based measurement of physical activity in pre-schoolers: Comparison of machine learning and cut point methods (Ahmadi & Trost, 2022)	Compare the accuracy of PA intensity predictions provided by ML classification models and previously published Cut-points for preschool-aged children.	Accelerometers	Random Forest	Five Activity Classes; Sedentary, light activity and games, moderate to vigorous activities and games, walking and running
A machine learning approach to measure and monitor physical activity in Children (Fergus et al., 2017)	Develop a reliable measure of children's free-living PA	Accelerometers	Multilayer Perceptron Neural Network	Resting, drawing, free play, DVD watching, playground activities, Jogging, and walking
Machine learning to quantify habitual physical activity in children with cerebral palsy (Goodlich et al., 2020)	Investigate whether activity-monitors and ML models could provide accurate information about PA performed by children and adolescents with cerebral palsy who use mobility aids for ambulation.	Accelerometers	Decision Tree, Random Forest and Support Vector Machine	Supine Rest, Colouring, Ball Throwing and Catching, Overground Walking with a mobility aid, wheelchair propulsion, cycling on a modified tricycle

Physical Activity among U.S. Preschool-Aged Children: Application of Machine Learning Physical Activity Classification to the 2012 National Health and Nutrition Examination Survey National Youth Fitness Survey (Kwon et al., 2022)	Evaluate PA levels in a representative sample of U.S. preschool-aged children.	Accelerometers	Random Forest	Five Activity Classes - Run, Walk, Other MVPA, Light-intensity PA and Sedentary
Estimating Physical Activity Energy Expenditure with the Kinect Sensor in an Exergaming Environment (Nathan et al., 2015)	Estimate the mechanical work performed by the human body and estimate subsequent metabolic energy using predictive algorithmic models.	Kinect Sensor Camera and Microsoft SDK	Gaussian Process Regression, k-Nearest Neighbour Regression and Linear Regression	Arm swings, standing jumps, body-weight squats, and jumping jacks
Quantitative Contact-Less Estimation of Energy Expenditure from Video and 3D Imagery (Koporec et al., 2018)	Present a comprehensive study on means to extract energy consumption from 2D or 3D image data, without assuming or detecting body pose.	Two Cameras and Two Kinect Sensor Cameras	SVM regression with developed v-RBF grid search optimization	Treadmill running and squash match
Depth-Camera-Based System for Estimating Energy Expenditure of Physical Activities in Gyms (Lin et al., 2019)	Propose a depth-camera-based system for energy expenditure estimation of PA in gyms.	Three Kinect Sensor Cameras	Linear Regression, Random Forest, Multilayer Perceptron, bagging regression tree and support vector regression	

Depth-Camera Based Energy Expenditure Estimation System for Physical Activity Using Posture Classification Algorithm (Lin et al., 2021)	Propose an accurate energy expenditure estimation model based on depth camera data with physical activity classification to solve the limitations in the previous research	Three Kinect Sensor Cameras	Linear regression, multilayer perceptron and convolutional neural network models	Standing, walking and running at various speeds

Classification of PA type has seen extensive use of accelerometer-based measurement across several studies. Ahmadi, Pfeiffer, et al. (2020) utilised accelerometers which were placed on the children's right hip and non-dominant wrist. The accuracy of Random Forest PA classification models was evaluated with the results noting classification models with prediction windows of 10 or 15 seconds provided the accurate recognition of five activity classes representative of young children's movement behaviours. In a similar study, Ahmadi, Pavey, et al. (2020) (using the same accelerometers) developed and evaluated random forest and logistic regression algorithms to predict children's PA class from raw accelerometer data collected at their right hip. Their results found random forest and logistic regression activity classifiers trained on features in the raw acceleration signal from the hip can be used to predict PA class in youth with a high level of accuracy.

Ahmadi & Trost (2022) adopted a similar approach to Ahmadi, Pfeiffer, et al. (2020) with accelerometers placed on the children's right hip and non-dominant wrist to classify various activity classes. Their results noted under true free-living conditions, random forest PA classification models for preschool-aged children exhibited significantly better recognition of PA intensity than cut-point methods. Utilising a supervised ML model (Multilayer Perceptron Neural Network), Fergus et al. (2017) evaluated the use of the model in classifying PA by activity type with the dataset taken from a prior study. The results of the study noted that it was possible to obtain an overall accuracy of 96% with 95% for sensitivity, 99% for specificity and a kappa value of 94% when three and four feature combinations were used. Goodlich et al. (2020), had participants wear accelerometers on their wrist, hip and mid-thigh with three supervised learning algorithms (decision tree, random forest and support vector machine) employed to examine if such algorithms could classify PA type among children and adolescents with cerebral palsy who use mobility aids for ambulation. The results noted the best performing single-placement model was a random forest classifier trained on wrist

features, yielding an overall classification accuracy of 79%, whilst the best performing model developed on a combination of two placements was a random forest classifier trained on wrist and hip features, yielding an overall classification accuracy of 92%. Kwon et al. (2022) sought to estimate daily time spent in MVPA and total PA via a random forest classification algorithm which was applied upon wrist-worn accelerometer data with their results noting that the random forest algorithm estimated that U.S. preschool-aged children engaged in 28 min/day of MVPA and 361 min/day of total PA, on average.

While the aforementioned studies demonstrate the application of accelerometer-derived data and ML for classification of PA type, duration and intensity, subsequent research marked a shift towards markerless, camera-based approaches that removed the constraint of body-worn sensors. Nathan et al. (2015) provided early evidence that Kinect-derived skeletal data and segmental mechanical work could be combined with regression models to estimate energy expenditure across several functional movements. Koporec et al. (2018) extended this approach by demonstrating that energy expenditure could be estimated directly from 2D and 3D imagery using computer vision techniques, including within a realistic sporting environment. Lin et al. (2019) subsequently demonstrated that depth-camera-derived whole-body skeletal data could accurately estimate energy expenditure during treadmill activity without requiring participants to wear monitoring devices.

Finally, Lin et al. (2021) further developed this approach by integrating PA classification with depth-camera data and comparing different predictive models and camera configurations. Together, these studies represent a progression from body-worn accelerometers towards non-intrusive, markerless-based approaches capable of capturing more efficient whole-body movement data. These methodological advancements therefore highlight the increasing scope of ML within PA research, from the analysis of wearable sensor data to the interpretation of increasingly complex, contactless movement data. Despite the many strengths of ML, it is important to highlight its shortcomings and areas which may pose difficulties for stakeholders in our domain who are keen to implement it into research and development. The following subsection shall consider such issues.

2.8.4 Limitations of Machine Learning

As discussed in the prior sections, the predictive capabilities of ML alongside its ability to uncover latent features in noisy and complex datasets enables it to be an appropriate statistical approach regardless of use-case. However, there are many drawbacks which may limit its implementation at scale in AMC and PA research. Firstly, ML models can overfit when they perform well on appropriate training data but poorly on new, unseen data (Yeom et al., 2018). This occurs when a model learns patterns that are too specific to the training set, including noise or random variation. As a result, it fails to generalise on other data. This is typically reflected in a large gap between training performance and test performance. Conversely, underfitting is said to occur when an algorithm lacks sufficient model capacity or sufficient training data to fully learn the true relationship that exists within a given dataset, whether through memorization or not (Bashir et al., 2020). From an MC perspective, an ML model may overfit by learning patterns specific to a particular sample (e.g., one school or setting) and perform poorly on new data. Underfitting may occur if the ML model was too simple to capture key movement patterns, leading to poor performance on both training and unseen data.

Secondly, many academics in the clinical domain have previously raised concerns regarding areas such as model bias (Paulus & Kent, 2020), quality of evaluation and reporting (Dhiman et al., 2021) and censoring (Li et al., 2020). In an MC context, this could manifest as a model trained on a narrow sample (e.g., one age group or distinct cultural setting) producing biased skill classifications, being evaluated using limited or non-representative metrics, or systematically excluding lesser-proficient or failed attempts (e.g., unsuccessful jumps or partial trials), which may bias the model towards higher-performing participants, thereby reducing a model's validity and generalisability. Thirdly, many ML models are inherently 'black boxes' in which they do not explain their predictions in a way that humans can understand (Rudin, 2019). A black box model could be either (i) a function that is too complicated for any human to comprehend, or (ii) a function that is proprietary (i.e., owned by a company which won't share their inner workings) and whilst can be understood can't be accessed. So, by example, a deep learning model could output "proficient" or "non-proficient" from raw movement data yet provide no interpretable insight into which aspects of the movement (e.g., arm trajectory, trunk rotation, timing) drove that proficiency classification. Researchers have attempted to overcome this limitation using ML explainability which in effect results in a black box becoming model-agnostic and thus the cause of the prediction can be pinpointed (Ribeiro

et al., 2016). However, these post-hoc explanations may not fully reflect the model's true internal decision-making process, potentially limiting their interpretability and trustworthiness. Lastly, the computational demands of ML, particularly deep learning models, pose another significant limitation (Saki & Soori, 2026). Training large neural networks is inherently resource intensive, requiring substantial computational power and energy. This presents practical challenges for their deployment within MC assessment contexts. For example, applications such as the automated scoring of fundamental movement skills from classroom or field-based video data often involve processing large volumes of footage and, in some cases, applying explainability techniques in near real time. Such requirements may exceed the computational capacity typically available in school or low-resource settings, thereby limiting the feasibility of widespread implementation.

Recent studies, such as Lander et al. (2025) have shown that this can be done at scale however, this must be considered on a case-by-case basis. Overall, while ML offers clear advantages for MC assessment, including enhanced scalability, greater objectivity, and an improved capacity to model complex relationships and movement patterns, its practical application remains subject to important limitations. In particular, challenges relating to model interpretability and computational demands may constrain its use, especially in applied or resource limited settings. Balancing these strengths and limitations will be key to its effective application in real-world settings.

As was highlighted in section 2.5.2, the use of accelerometers has been prominent in PA event and intensity classification, with their validity and reliability noted in estimating PA type and intensity classification (Barnett, Ridgers, & Salmon, 2015; Bürgi et al., 2011; Crane et al., 2015; De Meester et al., 2011; Jussila et al., 2022; Troiano et al., 2008). However, several notable limitations associated with its implementation has been raised in the literature (Breau et al., 2022). Such limitations are considered below. Firstly, it has been noted that such apparatus can be impractical, intrusive and uncomfortable for participants which in turn may lead to a negative effect on performance (Koporec et al., 2018; Lin et al., 2021). Secondly, accelerometers only quantify the acceleration of the body segment to which they are attached; for example, arm placement of a monitor would not capture movements of the trunk which, in turn, may lead to an underestimation of whole-body PA (Cliff et al., 2009). Additionally, these devices cannot account for increased energy expenditure associated with carrying a load, or

PA completed at an incline, and also have limited capabilities to measure nonweight-bearing activities (i.e., cycling) (Cliff et al., 2009; Pate et al., 2010). Beyond measurements purposes, there are additional factors that limit it's use, particularly among children. In comparison to adults, children have been reported to be prone to either removing accelerometers from their bodies or not adhering to using them in the first place for a multitude of reasons such as appearance, comfort, or if the device becomes wet (Audrey et al., 2013; Scott et al., 2018). It can therefore be concluded that youth populations warrant opposing conditions of compliance in comparison to adult populations (Audrey et al., 2013; Belton et al., 2013). Prior literature has highlighted further feasibility issues such as the number of sensors needed (Lander et al., 2022), calibration, sensitivity of placement on bodies and inability of practitioners to retrieve and utilise the data (Swain et al., 2023). Considering these limitations, both in terms of measurement, comfort and feasibility, markerless video-based solutions may provide a more feasible solution (Lander et al., 2025). The following section shall therefore introduce markerless technologies and it's application across related domains.

2.9 Markerless Motion Capture; Overview and Applications

Markerless motion capture has become a prevalent measurement tool within PA research, where camera and depth-sensor systems combined with ML models have been applied to tasks ranging from activity classification to the estimation of energy expenditure during exercise (Koporec et al., 2018; Lin et al., 2021; Nathan et al., 2015). Energy expenditure estimation, however, lies outside the scope of this thesis, which is concerned with whether such technologies can support the assessment of children's AMC. This section therefore provides a brief survey of the markerless motion capture landscape as it pertains to AMC assessment. A detailed critical synthesis of this literature, including its limitations and a research agenda, is presented in Chapter 5 and is not duplicated here.

2.9.1 Can Markerless Motion Capture Aid AMC Assessment?

The case for technological support in AMC assessment is well established. The timely and ongoing examination of AMC development is hindered by the availability of trained professionals, time and cost (Macy et al., 2014), and the administration burden of observe-and-score batteries such as the TGMD renders them inefficient for screening large populations

(Bisi et al., 2017). In response, the field of Human-Computer Interaction has produced sensors and algorithms capable of recognising human movement patterns through natural interactions with digital applications, collectively termed Motion-Based Technology (MBT) (Bossavit & Arnedillo-Sánchez, 2023). It is therefore imperative to consider whether such approaches can improve access to, and support professionals in, children's motor development monitoring (Bossavit & Arnedillo-Sánchez, 2022).

Early instrumented approaches relied upon wearable inertial measurement units (IMUs). An instrumented version of the TGMD-2 locomotor subtest based on IMUs was demonstrated to constitute a valid approach with high usability, reliability and objectivity (Bisi et al., 2017), removing the need for video recording or experienced raters throughout the assessment (Chow et al., 2020). Two limitations, however, constrain the IMU approach at scale: the placement of multiple sensors on each participant is inefficient and time-consuming when administered to a large population (Zuvela et al., 2011), and the approach could not provide a complete evaluation against all TGMD-2 performance criteria (Bisi et al., 2017). Markerless systems have therefore been proposed as the more scalable alternative (Chow et al., 2020). Table 2.4 summarises the measurement tools and statistical methods that have been applied to objective locomotor skill assessment in this literature.

Table 2.4: Sample of Measurement Tools and Statistical Methods Used for Objective Locomotor Skill Assessment

Study	**Aim**	**Measurement Tool**
Assessing locomotor skills development in childhood using wearable inertial sensor devices: the running paradigm	Investigate the feasibility of using inertial sensors to quantify developmental differences in the running pattern of young children.	IMUs
(Masci et al., 2013)		

Quantitative assessment of developmental levels in overarm throwing using wearable inertial sensing technology (Grimpampi et al., 2016)	Propose an in-field quantitative approach to identify different developmental levels in overarm throwing.	IMU's
Objective assessment of movement competence in children using wearable sensors: An instrumented version of the TGMD-2 locomotor subtest (Bisi et al., 2017)	Facilitate and support the TGMD-2 by designing, implementing and validating an instrumented version of the TGMD-2 locomotor subtest based on IMUs to quantify AMC in children rapidly and objectively	IMU's
Assessing Standing Long Jump Developmental Levels Using an Inertial Measurement Unit (Sgrò et al., 2017)	Sought to use an IMU for the quantitative process-oriented assessment of standing long jump skills in children	IMU's
Assessing Vertical Jump Developmental Levels in Childhood Using a Low-Cost Motion Capture Approach (Sgrò et al., 2015)	Utilise Microsoft Kinect to assess vertical jumping skill and predict developmental levels in 9- to 12-yr.-old children.	Microsoft Kinect

Hop.Skip.Jump.Games: The effect of “principled” exergameplay on children's locomotor skill acquisition (McGann et al., 2020)	What impact would an extended period of exergameplay have on user locomotor skills when deployed with a “principled” human-in-the-loop personalisation process.	Microsoft Kinect
A Video-Based Classification System for Assessing Locomotor Skills in Children (Chow et al., 2020)	Evaluate the accuracy of a newly-developed video-based classification system with a marker-less sensor to assess children’s locomotor skills.	Microsoft Kinect
Using motion capture technology to assess locomotor development in children (Bossavit & Arnedillo-Sánchez, 2022)	To what extent and how accurately can MotorSense recognise and analyse the execution of locomotor skills, Is the current accuracy of MotorSense adequate to support locomotor development assessment	Microsoft Kinect
Agreement of a Portable Motion Capture System to Analyze Movement Skills in Children (Lerma & Gulgin, 2023)	Assess the agree-ment of joint angle measurements obtained from the Microsoft Kinect V2 with AssessLink software to theVicon MX13 in children performing a vertical jumpand land (JL) and forward leap (Leap).	Vicon MX13 Motion Capture System Microsoft Kinect V2

A Machine Learning Model to Automatic Assessment of Gross Motor Development in Children using Posenet (Pilati Filho et al., 2023)	The work presented an approach to solving the automatic evaluation problem of the first criterion of the TGMD-3 jumping skill.	Web cameras (not explicitly reported)
Leveraging Artificial Intelligence to Automate Motor Competence Assessment (Lander et al., 2025)	This study developed and evaluated a video-based machine learning model to automate motor competence assessment and created a user-friendly web platform for teachers and coaches.	Single Webcam

The available systems trade accuracy against practicality, and the comparison criterion matters when interpreting validity claims. Marker-based optoelectronic systems such as the Vicon are regarded as the gold standard for kinematic measurement, but carry high cost, time-intensive setup, reliance on specialised expertise, and intrusive marker placement dependent on a clear line of sight (Van Der Kruk & Reijne, 2018). The Microsoft Kinect V2, an affordable (<€250) markerless alternative, captures full-body joint positions non-intrusively and requires no controlled environment (Bossavit & Arnedillo-Sánchez, 2022; Cai et al., 2019). Its measurement properties are mixed. Test-retest reliability is high across various movements and age groups (Cai et al., 2019; Wochatz et al., 2019), but agreement with the Vicon criterion is weaker: joint angle measurements during children's vertical jump and forward leap demonstrated only poor to moderate correlation and agreement (Lerma & Gulgin, 2023). Kinect-based assessment is accordingly recommended for selective use, according to body segment, plane of view, and the complexity and velocity of the movement (Lerma & Gulgin, 2023). Open-source platforms provide a further means of reducing the financial and technical barriers associated with motion capture (Van Der Kruk & Reijne, 2018). '*FreeMoCap'*, for example, uses synchronised consumer-grade cameras and computer-vision methods to reconstruct three-dimensional skeletal movement without reflective markers or proprietary

hardware. Matthis et al. (2022) demonstrated its application during a complex whole-body task involving three-ball juggling while balancing on an unstable platform, illustrating its capacity to capture dynamic and coordinated movement. Although this application did not involve children or AMC assessment, it demonstrated the potential of '*FreeMoCap'* as an accessible platform for recording complex human movement. Applied to AMC assessment specifically, the evidence for markerless systems is promising. A video-based classification system integrating the markerless Kinect V2 quantified TGMD-2 locomotor performance criteria from parameters such as joint angles and time spent off the ground, enabling an inexperienced rater to assess children's locomotor skills instantly without an experienced evaluator present (Chow et al., 2020), whilst heuristic, rule-based frameworks have enabled the real-time recognition of graded, age-related locomotor skills without reliance on large training datasets (Bossavit & Arnedillo-Sánchez, 2022). Kinect-driven exergames have likewise discriminated proficient from poor performances against TGMD-2 criteria (McGann et al., 2020).

Most recently, a deep-learning platform using a single side-view camera automated the assessment of seven motor skills in children aged 4 to 12 years with an average detection accuracy exceeding 83% (86.9% for object control skills and 88% for locomotor skills), although accuracy varied across skills and criteria, mirroring the difficulties experienced by human assessors (Lander et al., 2025; Ward et al., 2020). Review-level evidence counsels caution, however. Technology-based quantification of gross MC has proven effective but remains restricted to specific motor skills, with the retrieved studies preliminary and exploratory in nature, sample sizes small, and no systematic evaluation at scale nor concrete reporting of time and financial costs (Clark et al., 2021). Of 164 MBT studies reviewed by Bossavit & Arnedillo-Sánchez (2023), only around 16% addressed the screening or assessment of motor skills in young children, and the skills detected account for only approximately half of those considered essential to motor development. The landscape, in short, contains validated components but no complete, scalable solution, and it is within this gap that the contribution of this thesis is positioned.

2.9.2 Towards Automated Motor Competence Assessment: A Computer Vision and Machine Learning Pipeline

The primary contribution of this PhD thesis is a ML pipeline built around computer vision, comprising classification and interpretability elements alongside more traditional computing techniques such as heuristic algorithms (Kokash, 2005). It builds directly upon the recent work of Chow et al. (2020), Bossavit & Arnedillo-Sánchez (2023), Pilati Filho et al. (2023) and Lander et al. (2025), which has produced a clear rationale and template for this approach. Whilst this recent work is impactful, with real-world application, two enhancements distinguish the present contribution.

First, there is a lack of uniformity in experimental design for the use of computer vision devices in this domain. Whilst the most appropriate devices, and the number and positioning of those devices, have been examined (Bossavit & Arnedillo-Sánchez, 2023), this consideration is overlooked in more recent work (Lander et al., 2025; Pilati Filho et al., 2023). This thesis therefore seeks to provide a clear roadmap for practitioners on the quality, quantity and positioning of web cameras when gathering data on children's motor skill performance.

Second, whilst recent contributions can accurately classify performance against TGMD-3 performance criteria (Lander et al., 2025; Pilati Filho et al., 2023), further consideration is required regarding the classification of distinct performance criteria and the interpretability of the resulting outputs.. Applications of ML in this use case must therefore place a strong emphasis on post-hoc model-interpretability techniques as they may provide insight into the features contributing to ML predictions, where sufficient data is available. When data is limited, handcrafted heuristic algorithms may serve as a viable alternative for operationalising and classifying individual performance criteria (Bossavit & Arnedillo-Sánchez, 2022). The pipeline proposed in this thesis is accordingly conceived as open source: accessible, understandable and usable by all associated stakeholders, with the aim of ultimately enhancing children's opportunities to become more physically active from childhood through to adulthood.

2.10 Summary and Research Gaps

This review has established the central importance of actual motor competence (AMC) during childhood, particularly its positive association with physical activity (PA) participation as children progress into adolescence (Barnett et al., 2009; Lubans et al., 2010), alongside the associated health benefits (Cattuzzo et al., 2016; Lubans et al., 2010). The relationship between perceived motor competence (PMC) and AMC has also been considered (Crane et al., 2015; Robinson, 2011; Stodden et al., 2008). Although the literature does not point to a definitive or uniform association between these constructs (Barnett et al., 2022; De Meester et al., 2020), the evidence suggests that each exerts at least a partial positive influence on the other. In terms of measurement, the TGMD assessment batteries (Ulrich, 2013; Ulrich & Sanford, 2000) remain the most widely used tools for AMC assessment, while the PMSC for Young Children (Barnett et al., 2015) represents the PMC measure most closely aligned with TGMD constructs.

Despite the established reliability and validity of the TGMD, its full administration is time and resource intensive, limiting its feasibility in large-scale, school-based contexts (Barnett et al., 2013). In response, several studies have proposed shorter assessment batteries that retain acceptable validity while reducing administration demands (Bandeira et al., 2020; Duncan, Martins, et al., 2022; Valentini et al., 2018). These approaches typically employ network analysis to identify the most informative TGMD items, providing a structured basis for item reduction. However, these shortened batteries have been developed without incorporating PMC variables, and the reporting of item-level associations has been inconsistent. In addition, not all TGMD skills contribute equally to the assessment of AMC, and their relevance may vary across cultural and contextual settings (Barnett et al., 2016; Newell, 2020).

A further limitation across the motor competence literature, examined in detail in Section 2.7, is the field's tendency to re-apply a narrow set of research designs and inferential techniques (Behan et al., 2019; Bolger et al., 2018; Coppens et al., 2021; Duncan et al., 2018; Foulkes et al., 2015; O'Brien et al., 2016), yielding findings that are consistent but increasingly incremental (Bandeira, Estevan, et al., 2025; Caldwell et al., 2020; Melby et al., 2022; Song et al., 2021). This methodological convergence constrains the field's ability to develop a more detailed understanding of the mechanisms underlying AMC development. In contrast, the broader PA literature has seen a notable increase in the use of machine learning (ML) methods, including applications in activity detection and intensity classification using

accelerometry (Ahmadi & Trost, 2022; Kwon et al., 2022), and more recently through Markerless motion capture technologies (Koporec et al., 2018; Lin et al., 2021). Within AMC research, objective measurement approaches have similarly progressed, moving from inertial measurement units (Bisi et al., 2017; Masci et al., 2013; Sgrò et al., 2017) towards less intrusive markerless systems. Open-source platforms such as '*FreeMoCap'* have further demonstrated that complex three-dimensional human movement can be reconstructed using consumer-grade cameras (Matthis et al., 2022), although their application and validation within children's MC assessment remain limited. More recent child-focused applications have examined markerless motion capture as a potentially scalable approach to AMC assessment (Chow et al., 2020; Lander et al., 2025; McGann et al., 2020; Sgrò et al., 2015).The adoption of ML methods within MC research has however been relatively slow, despite growing evidence that such approaches offer advantages in handling large datasets and identifying latent patterns and complex relationships that may not be accessible through traditional analyses (Song et al., 2021). Similarly, while markerless motion capture technologies are increasingly used, they have not yet been fully combined with ML approaches in AMC research. This represents both a methodological limitation and a clear opportunity for advancement.

In summary, this thesis seeks to address several key limitations within the existing literature: the continued reliance on traditional inferential methods and the resulting repetition of established findings; the exclusion of PMC variables in the development of short-form TGMD assessments; and the practical challenges associated with large-scale AMC assessment. By integrating ML approaches, including classification and interpretable models, with markerless motion capture technologies, this research aims to generate novel insights and contribute to more scalable, objective, and contextually relevant assessment of AMC among Irish children and adolescents.

Link from Chapter 2 to Chapter 3

The literature reviewed in Chapter 2 demonstrated the central role of both AMC and PMC within children's PA and developmental trajectories (Barnett et al., 2016, 2022; Barnett, Ridgers, & Salmon, 2015; De Meester et al., 2020; Logan et al., 2015; Robinson et al., 2015; Stodden et al., 2008). Research consistently demonstrates that higher MC proficiency levels are associated with increased PA participation, improved health-related fitness, and more favourable developmental outcomes throughout childhood and adolescence (Hardy et al., 2013; Logan et al., 2015; Robinson et al., 2015). Similarly, PMC has emerged as an important correlate of movement behaviour, with evidence suggesting that children's perceptions of their own competence may influence both motivation and participation in PA (Barnett et al., 2008; Barnett, Ridgers, & Salmon, 2015; Crane et al., 2015; Stodden et al., 2008).

Despite the importance of these constructs, much of the existing literature has relied upon traditional analytical approaches such as correlations, regressions, t-tests, and analysis of variance (Behan et al., 2019; Bolger et al., 2018; Coppens et al., 2021; Duncan et al., 2018; Foulkes et al., 2015; O' Brien et al., 2016). While these methods have provided valuable insights, they are typically designed to test predefined hypotheses and may be limited in their ability to identify complex, multidimensional relationships within large datasets (Farrahi & Rostami, 2024). Consequently, much of the literature has repeatedly demonstrated broadly similar findings, including age-related improvements in MC, gender differences in object-control skills, and positive associations between AMC, PMC, physical activity, and health-related fitness (Barnett et al., 2022).

Recent advances in machine learning provide an opportunity to complement traditional analytical approaches by identifying patterns and predictive relationships that may not be readily observable using conventional statistical techniques (Britton, Onibonoje, et al., 2023; Farrahi & Rostami, 2024; Ley et al., 2022). Rather than focusing solely on explanation, machine learning methods can be used to examine the extent to which existing variables can predict future outcomes and identify the relative importance of individual predictors within complex datasets (Alsareii et al., 2022; Britton, Onibonoje, et al., 2023; Duncan et al., 2024).

Accordingly, the first empirical study presented in Chapter 3 explores the application of machine learning classification models within a MC context. Specifically, the study investigates whether AMC assessment outcomes can be predicted using data derived from existing AMC and PMC assessments, while also examining differences in predictive performance across male, female, and gender-combined datasets. In doing so, the study represents the first step in this thesis' broader objective of investigating how emerging analytical and technological approaches may contribute to advancing MC research.

Chapter 3: Streamlining Motor Competence Assessments via a Machine Learning Approach

Manuscript submitted as: Streamlining Motor Competence Assessments via a Machine Learning Approach.
MDPI Youth – accepted 25th June 2025, published online 6th July 2025.
Authors: Colm O'Donaghue, Michael Scriney, Sarahjane Belton & Stephen Behan

3.1 Abstract

Strong competencies in actual motor competence (AMC) and perceived motor competence (PMC) support lifelong physical activity. However, assessing MC is time-consuming, requiring multiple AMC and PMC evaluations. Streamlining these assessments would improve efficiency at a national level. This study used machine learning (ML) classification to (1) identify AMC assessments that can be accurately predicted in an Irish context using other AMC and PMC assessments, and (2) examine prediction accuracy differences between genders. AMC was measured using the Test of Gross Motor Development (3rd Edition) and the Victorian Fundamental Motor Skills Manual, while PMC was assessed with the Pictorial Scale of Perceived Movement Skill Competence. Five ML classification models were trained and tested on an Irish MC dataset (n = 2,071, mean age 9.2 ± 2.04) to predict distinct AMC assessment outcomes. The highest prediction accuracies (>85%) were found for the Catch (female and gender-combined subsets) and Bounce (male subset) AMC assessments. These assessments could potentially be removed from the current Irish testing battery for their respective gender groups. Our findings highlight the effectiveness of ML classification in optimising Irish MC assessment procedures, reducing redundancy, and enhancing efficiency.

Keywords: Actual motor competence; perceived motor competence; youth; machine learning; prediction; classification

3.2 Introduction

Physical activity (PA) is essential for youth development, as it plays a critical role in reducing the risks of obesity, cardiovascular disease, and mental health issues (Eime et al., 2013; Katzmarzyk et al., 2019). PA has also been linked to improved cognitive performance, social skills, and overall well-being in children and adolescents (Erickson et al., 2019). Despite these well-documented benefits, global trends indicate a significant decline in PA levels among youth, with approximately 81% of young individuals failing to meet the World Health Organization's recommended 60 min of moderate-to-vigorous PA per day (Guthold et al., 2020). This decline has been attributed to various factors, including increased screen time, urbanisation, and a lack of access to structured physical education and recreational facilities (Hesketh et al., 2017). Consequently, there is an urgent need for effective interventions that can enhance PA participation among youth and counteract these negative trends.

One key avenue for intervention lies in the assessment and development of motor competence (MC). MC is widely recognised as a fundamental determinant of PA engagement (Lima et al., 2017), as individuals with higher MC levels are more likely to participate in and sustain an active lifestyle (Estevan et al., 2023). The reciprocal relationship between MC and PA suggests that individuals with greater proficiency in movement skills are more confident and motivated to engage in various physical activities, thereby promoting a positive cycle of MC and participation (Stodden et al., 2008). Ultimately, the lack of an adequate foundation of MC may be linked to a hypothetical 'proficiency barrier' (Seefeldt & Gould, 1980) where low-MC individuals may not demonstrate health-enhancing levels of PA and health-related physical fitness later in life (Stodden et al., 2009). Such poorly proficient individuals may also be at greater risk for obesity through childhood and adolescence (Robinson et al., 2015). Therefore, it is key for practitioners to understand the components that comprise MC.

MC is composed of two distinct but interrelated components: actual motor competence (AMC) and perceived motor competence (PMC). AMC is associated with fundamental movement skills (FMS), which are the building blocks of more advanced, complex movements required to participate in games, sports or other context specific PA (Logan et al., 2018). FMS is made up of three components, namely locomotor skills (running and jumping, etc.), object control skills (throwing and kicking, etc.), and stability skills (i.e., balance) (Goodway et al., 2010). PMC, on the other hand, relates to an individual's self-perception of their motor abilities and

has been shown to influence PA participation and motivation (Estevan & Barnett, 2018). Children who perceive themselves as competent are more likely to develop the confidence and self-efficacy required to engage consistently in PA across a range of settings (Peers et al., 2020). PMC has also been shown to exert a significant influence on future PA behaviour, potentially surpassing the influence of AMC alone (Bardid et al., 2016). Research demonstrates a robust association between high levels of PMC and higher levels of PA, underscoring the critical roles that both AMC and PMC play in developing youth physical literacy (Barnett et al., 2008; Stodden et al., 2008). Given this established relationship, both AMC and PMC should be accorded equal consideration in the assessment of MC among children and adolescents. Recent international studies further support this approach, consistently identifying significant correlations between PMC and AMC (Den Uil et al., 2023; Niemistö et al., 2023). Findings indicate that children displaying low PMC during early childhood face increased risks of persistently low PMC and limited AMC progression into middle childhood (Niemistö et al., 2023). Accordingly, prioritising the development of children's MC proficiency from an early age is essential.

Unfortunately, research worldwide consistently reports poor levels of AMC among children and adolescents, with persistent deficits observed through adolescence (Chen et al., 2024; Hardy et al., 2013). Furthermore, research has identified notable gender differences in AMC proficiency, further emphasising the need for targeted intervention strategies. Studies have reported that boys tend to score significantly higher in object control skills, whereas girls typically outperform boys in locomotor and stability skills and balance-related movements (Behan et al., 2019; Zheng et al., 2022). Thus, there is a clear need for assessment and intervention to help all children enhance their AMC trajectory.

Given the well-documented benefits of MC development in promoting active lifestyles, there is a clear need for reliable and efficient assessment tools that can guide appropriate interventions. While there are a wide range of AMC measures, PMC is generally assessed using a reliable and valid self-report pictorial scale developed by Barnett et al. (2015) and enables alignment between the measurement of PMC and AMC competence in young children (Barnett et al., 2015). In the assessment of AMC, a wide range of FMS assessments are currently used to evaluate children's overall AMC proficiency, with the Test of Gross Motor Development (TGMD, 2nd and 3rd editions) (Ulrich, 2000; Ulrich, 2013) being the most

commonly used in related research (Behan et al., 2019; Duncan et al., 2020). As an assessment of FMS, the TGMD-3 can be used to identify developmental delay in relation to gross motor performance, to evaluate programmes aimed at enhancing AMC through intervention, and to assess changes as a function of increasing age, experience, instruction, or intervention (Ulrich, 2000; Ulrich, 2013). However, while reliable and valid as a measure of AMC, employing the TGMD-3 requires a considerable amount of time to administer and interpret (Barnett et al., 2013). In addition to a considerable time burden on participants and trained staff, specific space and equipment are required for assessment, which can make it logistically difficult to administer in school settings (Tamplain et al., 2020). As a result, researchers have been exploring alternative strategies, such as confirmatory factor analysis and machine learning (ML) approaches, to streamline assessment procedures while maintaining accuracy and reliability (Bandeira et al., 2020; Duncan, Martins, et al., 2022; Valentini et al., 2018). These efforts aim to identify key assessments that can predict overall MC, allowing practitioners to optimise evaluation processes and ensure that children receive the necessary support to enhance their motor development.

The integration of ML in MC research offers a promising avenue for refining assessment protocols. ML involves the development of algorithms and models that enable computers to learn and make predictions or decisions independently, without the need for specific instructions from the end-user (Yaqoob et al., 2023). By leveraging data-driven models, ML techniques can analyse complex patterns in movement proficiency and predict performance outcomes based on a combination of AMC and PMC assessments (Yaqoob et al., 2023). This approach has the potential to reduce the number of assessments required while maintaining predictive accuracy, ultimately improving the feasibility of large-scale MC evaluations (Barnett et al., 2013; Duncan, Martins, et al., 2022). Furthermore, ML driven assessment tools can facilitate the identification of at-risk children who may benefit from targeted interventions, ensuring that resources are allocated efficiently and effectively in educational and community sports settings (Bardid et al., 2019). For instance, recent MC research (Britton, Onibonoje, et al., 2023; Duncan et al., 2024) examined the predictive value of AMC features in relation to children's well-being and technical soccer skill performance respectively. Both studies employed five distinct ML algorithms, initially introducing a baseline algorithm and subsequently introducing more complex algorithms, resulting in superior predictions. Duncan et al. (2024) produced superb predictions of 99% for technical skill proficiency in boys aged

seven to fourteen years who were regularly engaged in grassroots soccer whilst Britton, Onibonoje, et al. (2023) produced strong predictions of 87% for predicting well-being off physical literacy variables among children (n = 1073, mean age 10.86 ± 1.20 years). Thus, to address the ongoing challenges in MC assessment and intervention, it is imperative to explore innovative methodologies that balance efficiency, accuracy, and practicality. By harnessing the power of ML classification models, this study aims to contribute to the development of streamlined assessment tools that can support the promotion of MC and PA participation among children and adolescents. Whilst previous studies have presented original contributions to the TGMD literature base, there are numerous areas in which ML classification analyses could build upon such contributions. Firstly, the ML approach in network analysis has been utilised in prior research (Bandeira et al., 2020; Duncan, Martins, et al., 2022), and whilst it is highly effective at assessing AMC relationship patterns, it is less effective than classification techniques when the aim of a study is predictive modelling (Yaqoob et al., 2023). The scoring criteria associated with the TGMD-3 (Ulrich, 2013) results in participants being awarded a score of either '1' ('Mastery/Near Mastery') or '0' ('Not Mastered'). Such scoring criteria is ideal for ML classification analysis given that the variable's data is binary, thus requiring each variable to be classified into its predefined group. Secondly, whilst Bandeira et al. (2020) and Duncan, Martins, et al. (2022) were capable of reducing their original 13-12 AMC skill models to a final model of 6–7 AMC skills, both studies did not include any PMC variables within their analysis. Indeed, the omission of such PMC variables from the aforementioned analyses is key given the impact a positive AMC–PMC relationship can have on a child's motivation to continuously improve their AMC capabilities (Barnett et al., 2008; Stodden et al., 2008).

Building on these efforts, this study explores the potential of ML classification models as a novel approach to optimising MC evaluation. ML techniques have been increasingly utilised in sports science and health research due to their ability to analyse complex patterns in large datasets and generate predictive models with high accuracy. By applying ML classification methods, this study aims to determine which MC assessments can be accurately predicted based on the results of other AMC and PMC evaluations. If successful, this approach could enable educators, coaches, and health practitioners to implement more efficient assessment protocols, ultimately facilitating broader participation in MC testing and fostering more inclusive physical literacy initiatives for youth populations.

3.3 Methodology

Study Design

Cross-sectional data were collected as part of an Irish AMC study 'Moving Well-Being Well' (Behan et al., 2019). Participants (n = 2,071, 47% girls, ranging from 5–12 years of age, mean age 9.2 ± 2.04) were recruited from 44 schools across twelve counties, encompassing all four provinces in Ireland and Northern Ireland. Data was collected from March to June 2017 across the full range of the Irish primary school cycle of typically developing children aged five to twelve. The subjects were classified per age and gender for data analysis. Ethical approval was obtained from the institutional Research Ethics Committee (DCU/REC/2017/029). The principals of each of the participating schools initially consented to participation. Consent procedures ensured that all individuals provided informed written consent, as well as parental consent, prior to data collection, thus supporting participant autonomy. An age-appropriate child assent form was supplied to participants in the younger classes (5–8 years of age). Additionally, robust anonymisation procedures were employed, including assigning participants a unique numerical code, which not only protected participant confidentiality but also enhanced the integrity and reliability of the dataset by minimising potential biases or risks associated with participant identification.

Procedure

AMC was measured using the TGMD-3 (Ulrich, 2013) alongside an additional locomotor skill test, the vertical jump, from the Victorian Fundamental Motor Skills manual (Walkley et al., 1996). Both have been used in previous research and have a high degree of validity and reliability (Walkley et al., 1996; Cools et al., 2009). In the dataset such AMC skills were labelled as 'Mastery/Near Mastery' (MNM) features in which participants were awarded a score of one if the participant fulfilled the necessary skill criteria (considered as 'Mastery/Near Mastery') whilst a score of zero ('Not Mastered') indicated that they failed to meet such skill criteria (Ulrich, 2013). Further details of these protocols can be found in Behan et al. (2019). PMC was measured using the Pictorial Scale of Perceived Movement Skill Competence (PMSC) for Young Children (Barnett et al., 2015). The PMSC administered fifteen skills, of which seven related to perceptions of various object-control skills and eight related to perceptions of locomotor skills. Each skill item was evaluated from 1 to 4 points leading to a maximum sum score for locomotor skills of 32 points (8 × 4) and 28 points (7 × 4) for perceptions of object

control skills, using a binary two-step choice process to arrive at one of four responses (1 = not that good, 2 = sort of good, 3 = pretty good, 4 = really good) per each skill. The maximum total score a child could receive was 60 points. The higher the child scored, the higher the child's PMC. Previous studies have shown good reliability and validity for assessing young children's PMC with PMSC in different cultures for locomotor (Diao et al., 2018) and object-control skills (Diao et al., 2018; Johnson et al., 2016).

Data Collection

The data collection protocol utilised by the research team is outlined fully in Behan et al. (2019). As such, all research team members completed formal training to ensure a comprehensive understanding of the skill assessment procedure as well as consistency during assessment of the test subjects. To ensure the data was consistent and of high quality, the research team was required to meet a 95% inter-observer agreement on a pre-coded data set, which was pre-coded by the lead researcher. The research team members were blind to the conditions of coding.

Experimental Setup

The original dataset consisted of 2378 rows and 492 columns. The initial cleaning process involved removing null values and removing variables that were deemed surplus for the purpose of the study. Features were only included if they were pure measures of AMC or PMC, for instance 'Actual Kick MC' and 'Perceived Kick MC'. This procedure for removing redundant features was applied uniformly across all gender subsets. Table 3.1 prvoides a brief synopsis on several of the analysed variables, including their data type and a description of the variable. After cleaning the dataset, there were a total of 2,071 rows and 30 columns remaining, which comprised the final dataset.

Table 3.1: Brief Synopsis of Features Analysed

Feature Category	Feature Description	Count (Perceived and Actual Features Combined)
Locomotor Skill	Actual MC—How capable children are to run, jump, skip etc. Full marking criteria are available in TGMD-3 (Ulrich, 2013). Perceived MC—The child's perception of their ability to run, jump, skip etc. Full marking criteria available in (PMSC) for Young Children (Barnett et al., 2015).	16
Object-Control Skill	Actual MC—How capable children are at manipulating objects such as balls (i.e., kicking, catching, throwing). Full marking criteria available in TGMD-3 (Ulrich, 2013). Perceived MC—The child's perception of their ability to kick, catch, throw etc. Full marking criteria available in (PMSC) for Young Children (Barnett et al., 2015).	14

As such, 16 locomotor skill and 14 object-control skill features were analysed, which resulted in a total of 30 variables being included in the analysis. Following this the dataset was split into different subsets of data in which the dataset was split by a specific gender subset. Table 3.2 below gives a brief example of what this splitting process looked like for all dependent variables.

Table 3.2 Example of Splitting Data by Gender Into several Subsets for Testing.

Dependent Variable	Gender Subset	Count
Run_MNM	Female Only	969
Run_MNM	Male Only	1107
Run_MNM	Gender Combined	2076

Before proceeding with data sampling, several crucial feature engineering steps were undertaken to refine the dataset for analysis. The process began by encoding categorical variables through one-hot encoding, converting non-numeric attributes into a machine-readable format. This approach generated binary indicator variables for each categorical feature while discarding the original columns, ensuring that classification models could interpret categorical data without imposing any unintended ordinal relationships. Following this, numerical features such as height and weight were standardised using StandardScaler to maintain consistency across varying scales. Since ML models—particularly those that rely on distance-based computations—are sensitive to differences in feature magnitudes, standardisation ensured that all continuous variables were centered at zero with a standard deviation of one. This prevented certain features from disproportionately influencing the model due to differences in scale.

Additionally, a feature selection process was implemented to remove irrelevant, redundant, or null-indicator columns that did not add value to the predictive model. A filtering mechanism identified and excluded these unnecessary features, enhancing both computational efficiency and model interpretability. After these transformations were finalised, the cleaned dataset was stored in Parquet format, preserving its structured and optimised form for the subsequent data sampling phase. Numerous ML methods have proposed that the target classes (i.e., dependent variables) hold similar distributions (Mohammed et al., 2020). However, this assumption does not hold true in several domains, for instance diagnosis of illness (Krawczyk et al., 2016), as most of the instances are labelled with one class, while few instances are labelled as the other. As a result, the models gravitate towards the majority class, thus eliminating the minority class. This reflects on the model's performance, as these models will perform poorly when the datasets are imbalanced. This is known as a class imbalance problem. Indeed, one of the common approaches in counteracting this issue is to incorporate resampling methods to balance the dataset.

As such, a class imbalance issue existed in the current study's dataset, which resulted in several resampling methods being incorporated, with those methods being: undersampling (the process of decreasing the amount of majority target instances or samples), oversampling (the process of increasing the amount of minority class instances or samples by producing new instances or repeating some instances (Mohammed et al., 2020)), no sampling, and SMOTE (the process of analysing the minority samples and integrating new synthetic minority samples according to the minority samples, in order to add them to the data set (Shi et al., 2021)). These resampling methods resulted in the creation of four instances of the dataset for each individual dependent variable. Figure 3.1 below shows an example of the distribution difference for one of these imbalanced dependent variables.

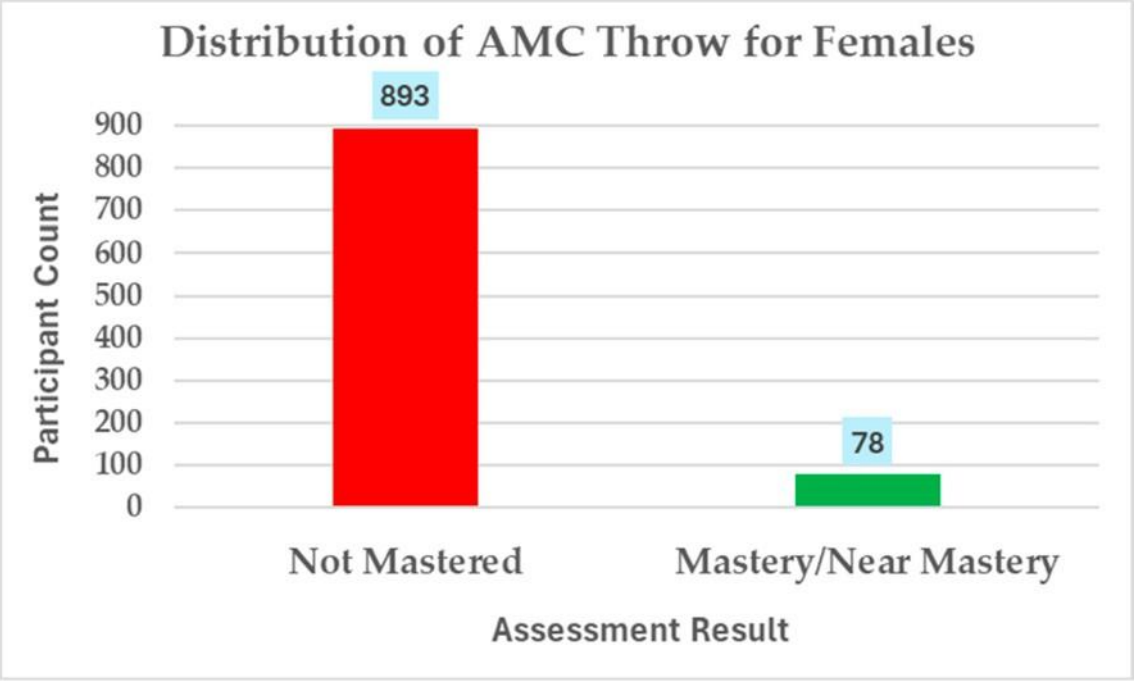


Figure 3.1: Distribution difference of Throw AMC feature for Female Subset

Five classification methods were chosen as candidates for predicting the outcome of individual AMC assessments: decision tree (DT), gradient boosting decision tree (GBDT), k-nearest neighbour (KNN), logistic regression (LR), and XGBoost (XGB). Detailed descriptions of such classifiers are provided within the following references: (Alzubi et al., 2018; Carmona et al., 2022; Yoon, 2021). As a byproduct of the aforementioned class imbalances, the classification models employed were not stratified according to participant age due to a notable variation in age-related MNM proficiency. Most notably, among six-year-old males, for throwing, only 2

MNM cases were noted in comparison to 80 non-mastery cases with seven-year-old females reporting similar distributions in the gallop; 18 MNM cases to 110 non-mastery cases. Furthermore, regardless of class distributions, sample sizes were low across several age groups, with five-year-old males and females most notably producing > 20 overall samples for each skill. Thus, age stratification would have reduced the subsequent samples for model curation and, for some target features, produced insufficient positive class observations to support stable train-test splitting, cross-validation and resampling. As the current dataset is subsequently examined later in Chapter 4, such age-related variations and their implications for subgroup selection are considered further in that chapter, where age stratification serves a distinct analytical purpose.

The metrics used to validate the classifiers were: precision (percentage of observations labelled as positive that are actually positive), recall (percentage of positive observations classified as positive), and F1 (harmonic mean of precision and recall). As such, these metrics were calculated from four distinct counts: true positives (TP), a count of the correctly predicted positive class values (Mastery/Near Mastery); true negatives (TN), a count of the correctly predicted negative class values (Non-Mastery); false positives (FP), which count the occurrences where the actual class is false but is predicted as true; and false negatives (FN), which count the occurrences where the actual class is true but is predicted as false. The sum of these four counts is the total number of observations (Britton, Onibonoje, et al., 2023).

Additionally, the F1 metric was utilised as it is a recognised metric enabling comparison between classifiers trained using differing methodologies, which was present in our case with an imbalanced distribution existing between non-mastery and mastery–near-mastery values. We performed hyperparameter optimisation using GridSearch to determine the optimal model configuration for each method, data subset, and target AMC assessment. The data was split using an 80:20 ratio for training and testing. For each of the 12 AMC target outcomes, five classification algorithms were evaluated under four sampling conditions across gender-combined, female-only and male-only subsets. This resulted in 240 model configurations per subset and 720 predictive classification models overall.

3.4 Results

Results for the top performing target features in terms of highest predictive accuracies (F1 score) amongst each subset alongside their corresponding resampling method, precision, and recall scores are shown in Table 3.3, Table 3.4 and Table 3.5. In turn, results for the worst performing target features are shown in Table 3.6.

Table 3.3 Top Performing Target Features—Gender Combined Subset.

Feature	Classifier	Sampling	Precision	Recall	F_1 Score
Catch MNM	XGB	SMOTE	0.83	0.88	0.85
Run MNM	KNN	SMOTE	0.89	0.79	0.83
Slide MNM	XGB	Unsampled	0.79	0.87	0.83
Bounce MNM	XGB	Undersampled	0.80	0.80	0.80
Hop MNM	XGB	SMOTE	0.75	0.85	0.80
Kick MNM	DT	Oversampled	0.74	0.66	0.70
HJ MNM	LR	SMOTE	0.67	0.68	0.68
Skip MNM	DT	Unsampled	0.66	0.64	0.65
Roll MNM	DT	Unsampled	0.68	0.58	0.63
VJ MNM	DT	Unsampled	0.68	0.53	0.60
Gallop MNM	XGB	Unsampled	0.57	0.44	0.50
Throw MNM	XGB	SMOTE	0.49	0.43	0.46

Table 3.4. Top Performing Target Features—Female-Only Subset.

Feature	Classifier	Sampling	Precision	Recall	F_1 Score
Catch MNM	KNN	Unsampled	0.85	0.88	0.87
Hop MNM	KNN	SMOTE	0.93	0.78	0.85
Run MNM	XGB	Unsampled	0.80	0.91	0.85
Slide MNM	XGB	Unsampled	0.83	0.85	0.84
Kick MNM	DT	Unsampled	1.00	0.62	0.77
Bounce MNM	XGB	Unsampled	0.72	0.73	0.72
Gallop MNM	XGB	Unsampled	1.00	0.57	0.72
HJ MNM	XGB	Undersampled	0.71	0.69	0.70
Throw MNM	LR	Unsampled	1.00	0.54	0.70
Skip MNM	XGB	SMOTE	0.66	0.68	0.67
VJ MNM	XGB	SMOTE	0.61	0.62	0.61
Roll MNM	GB	Oversampled	0.56	0.55	0.56

Table 3.5. Top Performing Target Features—Male-Only Subset.

Feature	Classifier	Sampling	Precision	Recall	F_1 Score
Bounce MNM	XGB	Unsampled	0.84	0.89	0.86
Run MNM	KNN	Unsampled	0.78	0.88	0.83
Catch MNM	KNN	SMOTE	0.88	0.77	0.82
Slide MNM	KNN	Oversampled	0.83	0.80	0.82
Kick MNM	KNN	SMOTE	0.88	0.76	0.81
Hop MNM	KNN	SMOTE	0.82	0.73	0.77
HJ MNM	XGB	Undersampled	0.69	0.69	0.69
Skip MNM	XGB	SMOTE	0.69	0.68	0.68
VJ MNM	DT	Unsampled	0.75	0.58	0.66
Roll MNM	DT	SMOTE	0.58	0.65	0.62
Throw MNM	DT	Unsampled	0.71	0.50	0.59
Gallop MNM	KNN	Unsampled	0.33	0.30	0.31

Table 3.6. Worst Performing Target Features—All Subsets.

Feature	Subset	Classifier	Sampling	Precision	Recall	F_1 Score
Gallop MNM	All Subsets	DT	Unsampled	0.00	0.00	0.00
Throw MNM	Female, GC	Numerous*	Numerous*	0.00	0.00	0.00
Roll MNM	Female	DT	Unsampled	0.50	0.02	0.03
Kick MNM	Female	DT	Unsampled	1.00	0.08	0.14
VJ MNM	GC	DT	Unsampled	0.68	0.16	0.25
Run MNM	GC	KNN	SMOTE	0.89	0.21	0.34
Hop MNM	Female	KNN	SMOTE	0.93	0.21	0.35
Catch MNM	Male	KNN	SMOTE	0.88	0.25	0.39
Skip MNM	Female	KNN	SMOTE	0.66	0.38	0.48
Slide MNM	Male	KNN	SMOTE	0.82	0.36	0.50
HJ MNM	GC	DT	Undersampled	0.65	0.47	0.55
Bounce MNM	Female	KNN	Oversampled Undersampled	0.68	0.59	0.63

Numerous* = Female-only: DT–Unsampled; GBDT–Oversampled; GBDT–SMOTE; KNN–Unsampled; XGB–Unsampled | Gender-combined: DT–Unsampled

Results from the gender combined subset showed that the Catch MNM feature produced the highest F1 score at 85%, whilst the Throw MNM feature produced the weakest F1 score at 46%, resulting in a 39% discrepancy between the best and worst performing features. As such, the three best performing features (Slide, Run, and Catch) ranged from 83–85% respectively. The XGB classifier proved to be the most prominent classifier, appearing on six occasions,

whilst the SMOTE and Unsampled resampling methods were most prominent, appearing on five occasions each. Results from the female-only subset showed that the Catch MNM feature produced the highest F1 score at 87%, whilst the Roll MNM feature produced the weakest F1 score at 56%, resulting in a 31% discrepancy between the best and worst performing features. As such, the three best performing features (Run, Hop, and Catch) ranged from 85–87%, respectively. The XGB classifier proved to be the most prominent classifier, appearing on seven occasions, whilst the Unsampled resampling method was most prominent, appearing on seven occasions.

Results from the male-only subset showed that the Bounce MNM feature produced the highest F1 score at 86%, whilst the Gallop MNM feature produced the weakest F1 score at 31%, resulting in a 55% discrepancy in accuracy between the best and worst performing features. As such, the three best-performing features (Catch, Run, and Bounce) ranged from 82–86%, respectively. The KNN classifier proved to be the most prominent classifier, appearing on six occasions, whilst the Unsampled and SMOTE resampling methods were most prominent, appearing on five occasions each. To gauge the optimism of our best performing classification mode, table 3.6 presents the poorest positive-class results observed for each target across the male-only, female- only and gender-combined subsets. Performance varied substantially, with F1-scores ranging from 0% for Gallop MNM and Throw MNM to 63% for Bounce MNM. The particularly low recall observed for several targets indicates that some classifier–sampling combinations identified only a small proportion of participants demonstrating mastery or near-mastery.

When resampling methods were considered, Unsampled data and SMOTE were associated with the best-performing classifiers. These findings suggest that the classification models employed were able to learn effectively without any alterations to the dataset (i.e., using the original, unsampled data), while the use of synthetic samples generated through SMOTE enhanced the models' predictive performance. Given the consistent success of both Unsampled and SMOTE approaches in our results, it could be inferred that low number of instances where individuals received a mark of not mastered resulted in overfitting and underfitting for the Oversampling and Undersampling approaches, respectively, in the context of the current dataset. The analysis of F1 Score performance across gender subsets revealed relatively small differences in standard deviation values, indicating consistent model

performance within each gender subset. The male-only subset exhibited the highest variability with a standard deviation of 0.15, compared to 0.10 for the female-only subset, while the gender-combined subset showed an intermediate value of 0.13. On average, the standard deviation across all subsets was 0.12, reflecting generally low variation in F1 Score performance. Similarly, the 95% confidence interval was consistent at 0.02 across all groups, highlighting a high degree of precision in the F1 Score estimates. These findings suggest that while the male-only group demonstrated slightly more variability, the model's overall performance remained stable and reliable across gender-based subsets.

Figure 3.2 provides a visual overview of the variation in F1 Scores amongst the three distinct subsets. Considering the results from an overall perspective, most of the features performed quite well, with several features producing F1 scores > 70%. Interestingly, the Gallop MNM feature produced the largest subset-to-subset discrepancy of 41% in favour of females versus males, whilst the male-only subset produced the largest accuracy discrepancy between the most- and least-accurate features at 55%. When the gender combined subset is compared against both the male and female subsets, the latter subsets are shown to essentially mirror and at times exceed the gender combined predictive power (i.e., Run MNM; GC Subset F1 Score—83%, Male Subset F1 Score—83%, Female Subset F1 Score—85%). Figure 3.3 contextualises the positive-class F1-scores according to the prevalence of MNM cases within each held-out test set. A prevalence-informed baseline is provided to indicate whether model performance exceeded that expected from the underlying class distribution, while the accompanying MNM and non-MNM counts provide context for interpreting the robustness of these estimates. In particular, high F1-scores derived from relatively few positive cases should be interpreted cautiously, as they may be more sensitive to sampling variation. Conversely, poor F1-scores observed where larger numbers of positive cases are available provide stronger evidence of limited predictive performance.

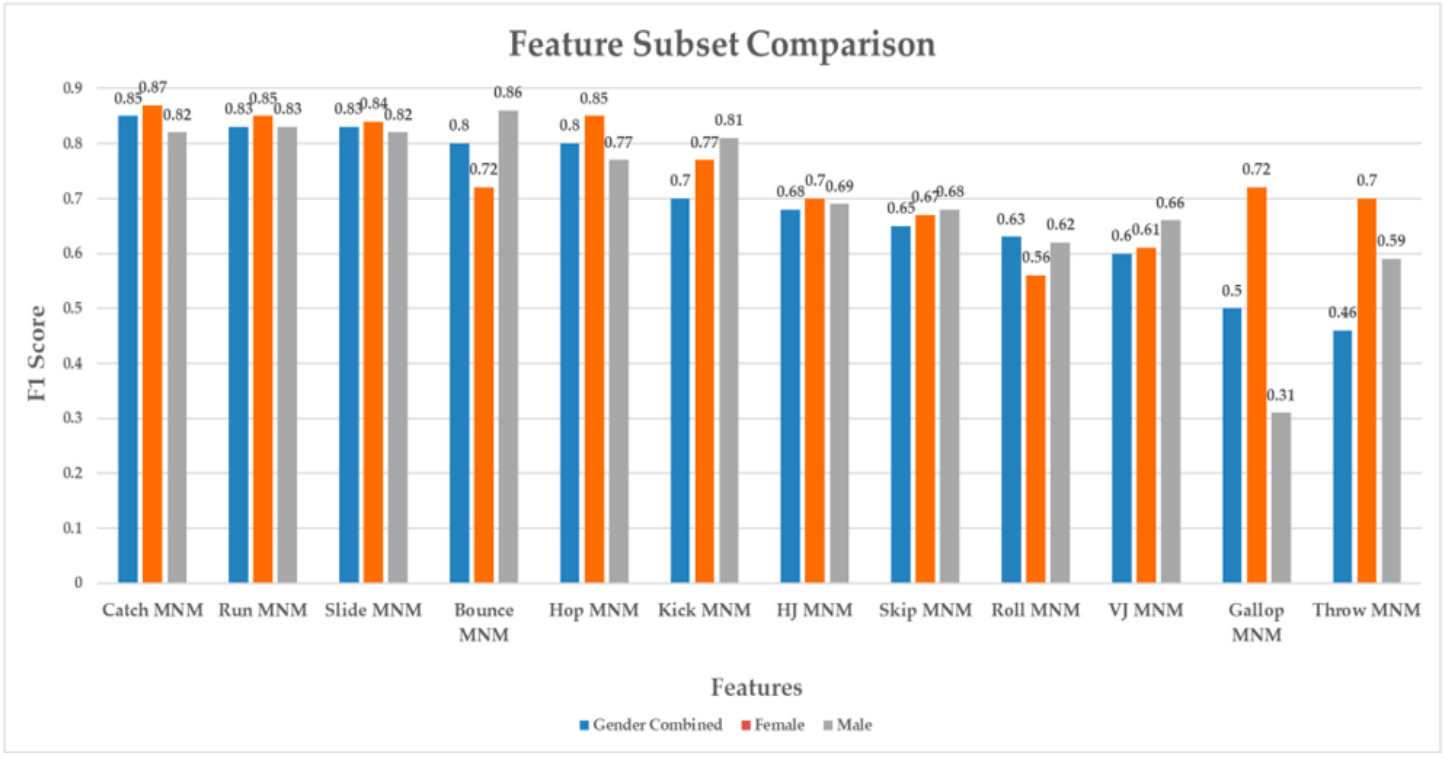


Figure 3.2 Subset Comparison Amongst All Analysed AMC Features

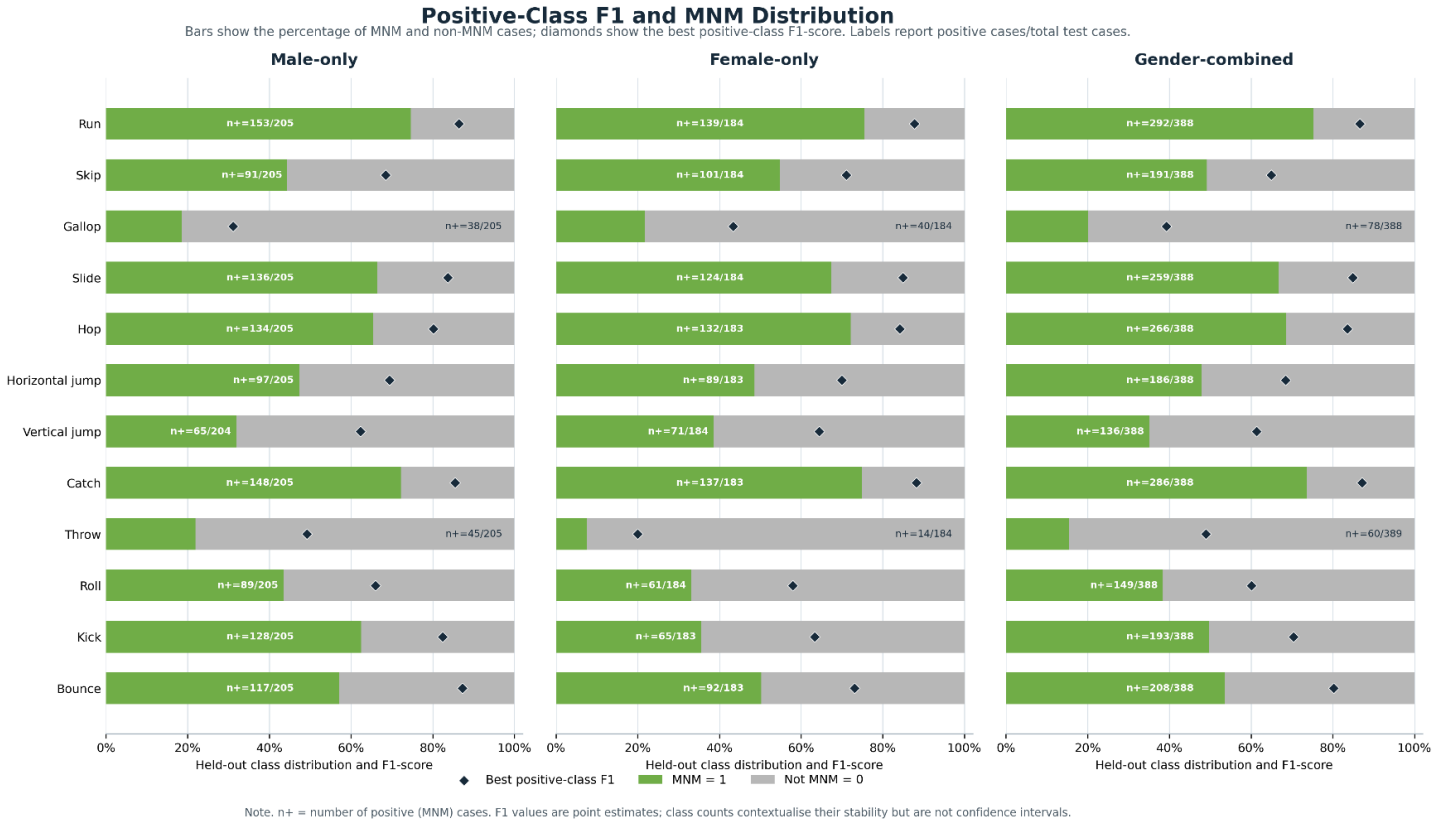


Figure 3.3 Distribution of Subset MNM and Positive-Class F1 Score

3.5 Discussion

The present study aimed to determine which MC assessments can be reliably predicted based on AMC and PMC evaluations, with a particular focus on gender-based differences in prediction accuracy. Our findings suggest that ML offers a promising avenue for streamlining MC assessments by identifying assessments that can be inferred from existing test results with high predictive accuracy. This not only optimises efficiency in testing but also reduces redundancy, potentially allowing practitioners to administer fewer assessments while maintaining robust evaluation standards. Optimising AMC assessment batteries is therefore crucial, as regularly evaluating the most relevant skills for specific subgroups of children can help identify more realistic MC proficiency profiles. This is important because difficulties in mastering MC may hinder success in many sports, potentially leading to subsequent physical inactivity beyond early childhood (Stodden et al., 2008).

The Catch assessment emerged as the most accurately predicted test in the female-only and gender-combined subsets, whereas the Bounce assessment was the most accurately predicted in the male-only subset. These findings align with prior research indicating that object-control skills, which are highly practiced and reinforced in structured and unstructured physical activities, tend to exhibit strong predictive patterns (Behan et al., 2019). The high predictability of these assessments suggests that they may be candidates for removal in future testing batteries, enabling a more targeted approach to MC assessment. However, before such changes are implemented, further validation in diverse populations and across different cultural contexts is warranted.

The contextualisation of these findings within the Irish sports landscape is essential for understanding their broader implications. Sports participation trends in Ireland indicate a strong emphasis on object-control skills, particularly through popular indigenous sports such as Gaelic football and hurling, as well as widely played international sports like soccer, rugby, and basketball (Woods et al., 2023). These sports provide ample opportunities for children to develop object-control skills, contributing to the observed high levels of mastery in skills like Catch and Bounce (Behan et al., 2019). As a result, it is not surprising that these skills are among the most predictable by ML models. Nevertheless, the transferability of these findings to other contexts requires careful consideration, as different regions and cultures emphasise

varying movement skills within their youth sports structures (Newell, 2020). Taking the kick as an example, it could be assumed that in cultures where soccer is the most prevalent sport (i.e., Brazil), AMC proficiency for kicking would likely be superior compared to cultural contexts where soccer is played less (Duncan, Martins, et al., 2022). Indeed, the cultural differences in sport and movement environments between countries may contribute to the reason why the skills may vary in reconceptualisations of the original TGMD model. It can therefore be concluded that children from different countries and cultures display variations in their AMC proficiency (Hulteen et al., 2018). Accordingly, the specificities of each country and culture must be considered by researchers in their validation models, which can be beneficial in discriminating variations between countries, and in determining the priority of intervention proposals, while respecting cultural aspects, as previously suggested (Hulteen et al., 2018; Barnett et al., 2016).

In practical terms, these results have several implications for MC assessment practices. First, the ability of ML models to predict certain MC outcomes with high accuracy suggests that assessment protocols could be streamlined and adapted for gender-specific applications. The removal of the Catch AMC assessment for females and the Bounce AMC assessment for males, based on their high predictive accuracy, could lead to a more efficient assessment process while preserving the integrity of overall MC evaluation. This streamlining could benefit practitioners, including teachers, sports coaches, and MC assessors, by allowing them to focus on evaluating skills that require direct observation due to lower mastery rates or higher variability in proficiency (Woods et al., 2023). Equally, we acknowledge that while the primary focus of the current study was to explore ways in which the burden of practical assessments could be reduced, we appreciate that removing some skills from the assessment battery may leave a practitioner with gaps of nuanced motor skill competence. We do however believe that the current standard testing protocols are impractical and time-intensive, and we feel that any efforts to lighten the burden should be strongly considered. That said, practitioners should be aware that omitting certain skills may result in a lack of data in specific skill competencies that are important in their domain or context (i.e., Gaelic football or soccer). Our findings also align with previous studies that have sought to optimise MC testing batteries through confirmatory factor analysis and network analysis (Bandeira et al., 2020; Duncan, Martins, et al., 2022; Valentini et al., 2018). Unlike these studies, which primarily focused on AMC skills alone, our analysis incorporated PMC features, providing a more holistic understanding of the

relationships between MC skills. The inclusion of PMC is particularly relevant given its role in influencing PA engagement, as children who perceive themselves as competent are more likely to participate in and persist with PA (Stodden et al., 2008). Future research should explore the relative influence of AMC versus PMC in ML prediction models to better understand the interplay between PMC and AMC. Another important consideration is the variation in prediction accuracy across different MC skills. While the Catch and Bounce assessments were highly predictable, other skills, such as Gallop and Throw, exhibited poorer predictive accuracy. These findings are consistent with prior research indicating that mastery rates vary considerably across different MC skills, with certain locomotor and object-control skills being more challenging to develop (Behan et al., 2019). In practical terms, this suggests that future assessment batteries should allocate more time to evaluating skills with lower mastery rates, as they are less likely to be reliably predicted through ML models.

Moreover, our findings provide insights into gender-based differences in MC skill prediction. The female subset produced superior predictions for Catch and Throw AMC assessments compared to males, whereas males showed higher predictive accuracy for Bounce AMC. These results align with existing literature suggesting that gender differences in MC development are influenced by participation patterns and movement skill practice opportunities (Bolger et al., 2018). While these findings suggest potential pathways for gender-specific assessment refinement, further research is needed to explore the underlying factors contributing to these predictive discrepancies, including sociocultural influences, access to PA opportunities, and inherent biomechanical differences.

Limitations and Future Work

A number of limitations must be acknowledged in this study. Firstly, only one dataset was analysed, which may limit the generalisability of the findings. The F1 Scores for each AMC feature might have been improved with the integration of additional datasets that capture a broader range of socio-economic variables. Socio-economic factors, such as access to PA opportunities, parental support, and school resources, can significantly influence MC development. By incorporating such features in future analyses, it may be possible to refine prediction accuracy and provide a more comprehensive understanding of how socio-economic disparities impact youth motor skill acquisition. Our future work will look to build on this by incorporating additional datasets and exploring a range of ensemble learning techniques, such

as Random Forest, Gradient Boosting, and Stacking, to enhance predictive performance. These methods have shown promise in previous studies by combining the strengths of multiple models to reduce overfitting and improve classification accuracy. By leveraging ensemble learning, we anticipate achieving more robust F1 Scores, thereby refining the predictive capability of MC assessments and ensuring greater generalisability across diverse populations. Another limitation of the study was the lack of age-related data, which could have given more insight into the predictions made by the models, allowing for comparison across the age span for all subsets. Our future work will seek to incorporate this going forward. Finally, whilst several classification models were utilised in our analysis, there were several that were not included (i.e., Random Forest, Naive Bayes, AdaBoost etc.). These models were excluded primarily due to computational constraints and the necessity of balancing model complexity with interpretability. However, their inclusion in future studies could provide valuable insights into classification performance, particularly in identifying non-linear patterns and interactions within the dataset. Random Forest, for example, could enhance robustness by aggregating multiple decision trees, while Naive Bayes might offer advantages in handling categorical data efficiently. AdaBoost, with its ability to improve weak classifiers, could further refine prediction accuracy. Future work will explore these models to assess their potential contributions to optimising MC assessments. Our future work will seek to add to this suite of models, eventually including more complex models (i.e., artificial neural networks).

Conclusion

In summary, our study demonstrates the potential for ML to optimise MC assessment procedures by reducing redundancy and improving efficiency. However, caution is warranted when considering assessment streamlining, as the removal of certain tests must be accompanied by rigorous validation across different populations and settings. Future research should focus on longitudinal applications of ML-driven assessments to explore the long-term impact of reduced assessment batteries on MC development and PA engagement. Additionally, further investigation into the role of PMC in MC prediction models is necessary to refine our understanding of how self-perceptions of competence influence motor skill acquisition and performance. Strengthening MC assessment procedures is critical to equipping children with the foundational skills necessary for lifelong PA engagement. This is

particularly important given evidence that early childhood marks a crucial period in which children may follow either a positive developmental trajectory, characterised by increasing PA participation and improved health outcomes, or a negative one, marked by declining engagement and adverse health effects (Stodden et al., 2008). By addressing existing limitations in assessment methodologies, we can foster a more robust, equitable, and evidence-informed approach to MC evaluation, ultimately benefiting both practitioners and youth populations.

Link from Chapter 3 to 4

The findings presented in Chapter 3 demonstrated that machine learning classification models can successfully predict several AMC assessment results from other AMC and PMC assessment results. Predictive performance varied across individual MC measures, suggesting that some skills may share stronger underlying relationships with other competencies than others. Furthermore, differences in model performance across male, female, and gender-combined datasets indicated that the structure of these relationships may not be uniform across populations. While these findings provide evidence that meaningful predictive relationships exist within MC data, ML classification models alone offer limited insight into the nature of these relationships (Carmona et al., 2022; Hassija et al., 2024; Rudin, 2019). A model may successfully predict a movement skill outcome, yet the mechanisms underpinning that prediction often remain unclear. Consequently, prediction represents only one stage in developing a deeper understanding of MC systems. The literature review highlighted that AMC and PMC do not develop in isolation but rather exist within a broader network of interacting factors (Barnett et al., 2008; De Meester et al., 2020; Stodden et al., 2008). Previous research has traditionally examined these relationships through correlation and regression-based approaches, often focusing on individual pairwise associations (Barnett et al., 2008; Crane et al., 2015; Duncan et al., 2018). While informative, such approaches may struggle to capture the complexity of interconnected developmental systems where multiple variables influence one another simultaneously (Bandeira, Estevan, et al., 2025; Bandeira, Lemos, et al., 2025).

Accordingly, the next stage of this thesis shifts from prediction towards interpretation. Chapter 4 adopts a post-hoc model interpretability and graph analytics approach to investigate age- and gender-specific interdependencies within motor competence data. By modelling AMC and PMC variables as interconnected networks, the study seeks to identify influential competencies, uncover hidden patterns of association, and provide greater insight into the relationships that may underpin the predictive outcomes observed in Chapter 3. In doing so, Chapter 4 moves beyond the question of whether MC outcomes can be predicted and instead seeks to understand how these competencies interact within the broader motor competence landscape.

Chapter 4:
Modelling Motor Competence Interdependencies Using Post-hoc Model Interpretability and Graph Analytics

Original Manuscript submitted as; 'Modelling Motor Competence Interdependencies: An XAI and Graph Analytics Approach'.
Presented on 30th October, 2025, at 2025 IEEE International Workshop on Sport, Technology and Research, Trento, Italy – published online 3rd December, 2025.
Authors: Colm O'Donaghue, Stephen Behan, Sarahjane Belton & Michael Scriney

4.1 Abstract

Motor competence (MC), comprising actual (AMC) and perceived (PMC) motor competence, is a notable contributor of lifelong physical activity (PA) engagement and health. Children with stronger AMC and PMC are more likely to sustain PA engagement, while low proficiency increases risks of inactivity and associated health issues. Accurately identifying and developing MC among children is therefore essential. This study explores age- and gender-specific interdependencies between AMC and PMC using post-hoc model interpretability (MI) and graph analytics. A machine learning pipeline—integrating classification models, interpretability techniques, and graph-based centrality measures—was applied to data from 1,124 children (ages 9–12). Results revealed distinct skill networks, highlighting core competencies and AMC–PMC relationships across demographics. These insights can guide targeted, time-efficient interventions and assessment strategies. Practical applications include supporting Physical Education teachers and sports coaches in designing sessions that nurture children's motor development. By combining MI with graph analytics, this work advances MC research and offers evidence-based strategies for educators, coaches, and health professionals.

Key Words: Motor Competence, Classification, Post-hoc Model Interpretability, Graph Analytics

4.2 Introduction

According to recent global data, 81% of adolescents fail to meet daily physical activity (PA) guidelines of at least 60 minutes of moderate-to-vigorous PA (Guthold et al., 2020). Physical inactivity is subsequently a significant public health concern, contributing to chronic diseases such as cancer and diabetes (Lee et al., 2012). Motor competence (MC), comprising actual motor competence (AMC) and perceived motor competence (PMC), is a key determinant of PA participation and lifelong engagement in various activities (Stodden et al., 2008). AMC encapsulates proficiency in locomotor, object-control, and stability skills (Donnelly et al., 2016), while PMC encapsulates an individual's perception of their abilities (Estevan and Barnett, 2018). Paired together, these components influence not only PA participation but also broader physical, cognitive, and social development (Lima et al., 2022).

AMC proficiency has been connected with both PA volume and intensity (Gu et al., 2017; McIntyre et al., 2018), while PMC is theorised to mediate this relationship by motivating children to participate in activities that further refine their skills (Barnett et al., 2022; De Meester et al., 2020). However, evidence on the AMC–PMC relationship remains inconclusive, partly due to a lack of longitudinal and intervention-focused studies (Barnett et al., 2022). This uncertainty has subsequently limited the development of effective strategies to foster sustained PA participation (Ortega-Benavent et al., 2024). Despite its importance, global AMC proficiency levels remain low (Bolger et al., 2021). Gender disparities are evident, particularly in object-control skills, where boys typically outperform girls (Barnett et al., 2016). Socioeconomic and environmental factors, such as access to structured physical education (PE) and opportunities for active play, also shape these disparities (Bellows et al., 2013). Schools therefore provide a critical setting for interventions (Lundvall, 2015), but existing programs face barriers related to time, resources, and teacher expertise, often producing only modest, short-term benefits (Van Sluijs & Kriemler, 2016). Reliable and scalable MC assessment is therefore essential to guide individualised support.

The Test of Gross Motor Development-3 (TGMD-3) is a widely used testing battery to assess locomotor and object-control skills (Ulrich, 2020; Duncan, Martins, et al., 2022), however, its implementation is time-intensive, requiring specialist training and specific equipment (Barnett et al., 2013). To address feasibility issues, recent research has explored short-form assessments and network analysis to identify core AMC skills (Duncan, Martins, et al., 2022;

Bandeira et al., 2020). While promising, these approaches typically omit PMC and age-specific data, despite evidence suggesting that AMC and PMC together are stronger predictors of PA than AMC alone (Barnett et al. 2022). Furthermore, skill development fluctuates across age groups; for example, while running is mastered early, complex skills such as overhand throwing remain challenging at later ages, with AMC proficiency plateauing around age ten (Behan et al., 2019). Addressing these limitations demands more robust analytical approaches capable of modeling the inter-dependencies among AMC, PMC, and PA across demographics. By integrating post-hoc model interpretability (MI) with graph analytics, our current work identifies critical age- and gender-specific competencies and provides evidence-based insights to support PE teachers and sports coaches in designing more individualised assessments and interventions.

Contribution and Paper Structure

We propose a novel methodology for identifying key MC inter-dependencies by integrating MI with graph analytics. By employing a robust ML pipeline, we uncover age-and gender-specific motor skill relationships that can inform targeted MC interventions, PE lesson planning and health promotion initiatives. The following will be presented in the sections to follow; related work, methodology, results, discussion and conclusion

Related Work

ML Classification in PA Related Research

Machine learning (ML) classification has been extensively used in PA research, particularly with accelerometer data (Narayanan et al., 2020), though less so in AMC studies. Prior work has shown that ML can classify PA type, duration, and intensity with high accuracy, often exceeding 90% (Ahmadi, Pavey, et al., 2020; Alsareii et al., 2022; Bonomi et al., 2009). Recent studies have leveraged wearable devices and optimization techniques to measure AMC in schoolchildren (Letts et al., 2024; Sulla-Torres et al., 2024), highlighting the potential of ML to provide scalable and precise alternatives to traditional observational techniques.

Model Interpretability in PA Related Research

MI techniques have recently emerged to enhance trust and transparency in ML for health and PA. Techniques such as Shapley Additive exPlanations (SHAP) and Local Interpretable Model-Agnostic Explanation (LIME) have been used to identify psychological, behavioral, and biometric predictors across domains including exercise behavior (Kuan et al., 2023), cardiology (Loh et al., 2020), diabetes prediction (Khokhar et al., 2025), and mortality risk (Graham & Farrell, 2025). Similar approaches have also uncovered novel digital biomarkers from accelerometer data for predicting physical function in older adults (Fan et al.,

2024). Such findings highlight the value of XAI for uncovering actionable insights from complex PA datasets.

Graph Analytics in PA Related Research

Network analysis (NA) has proven effective in modeling AMC inter-dependencies. Bandeira et al. (2020) demonstrated gender-, age-, and BMI-related differences in preschoolers, while Duncan, Martins, et al. (2022) validated a shortened TGMD-3 using NA, identifying seven core AMC skills central to MC proficiency. Building upon such work, Salami et al. (2024) more recently noted that chronological age, rather than biological maturation, was the strongest predictor of MC in prepubertal children, with centrality measures highlighting key AMC skills (e.g., gallop, slide, dribble, catch, run) that structure broader networks of proficiency. Similarly, Bandeira, Lemos, et al. (2025) adopted exploratory graph analysis to conceptualize MC as a network of interconnected constructs, identifying four stable dimensions that integrate both process- and product-oriented assessments. Together, these findings highlight graph analytics as an effective technique for revealing developmental inter-dependencies, pinpointing essential skills, and informing the design of more efficient and comprehensive MC assessment batteries.

4.3 Methodology

The approach implemented in the current study is demonstrated in Figure 4.1, highlighting the eight distinct phases which were undertaken. The sections that follow will discuss this approach in greater detail.

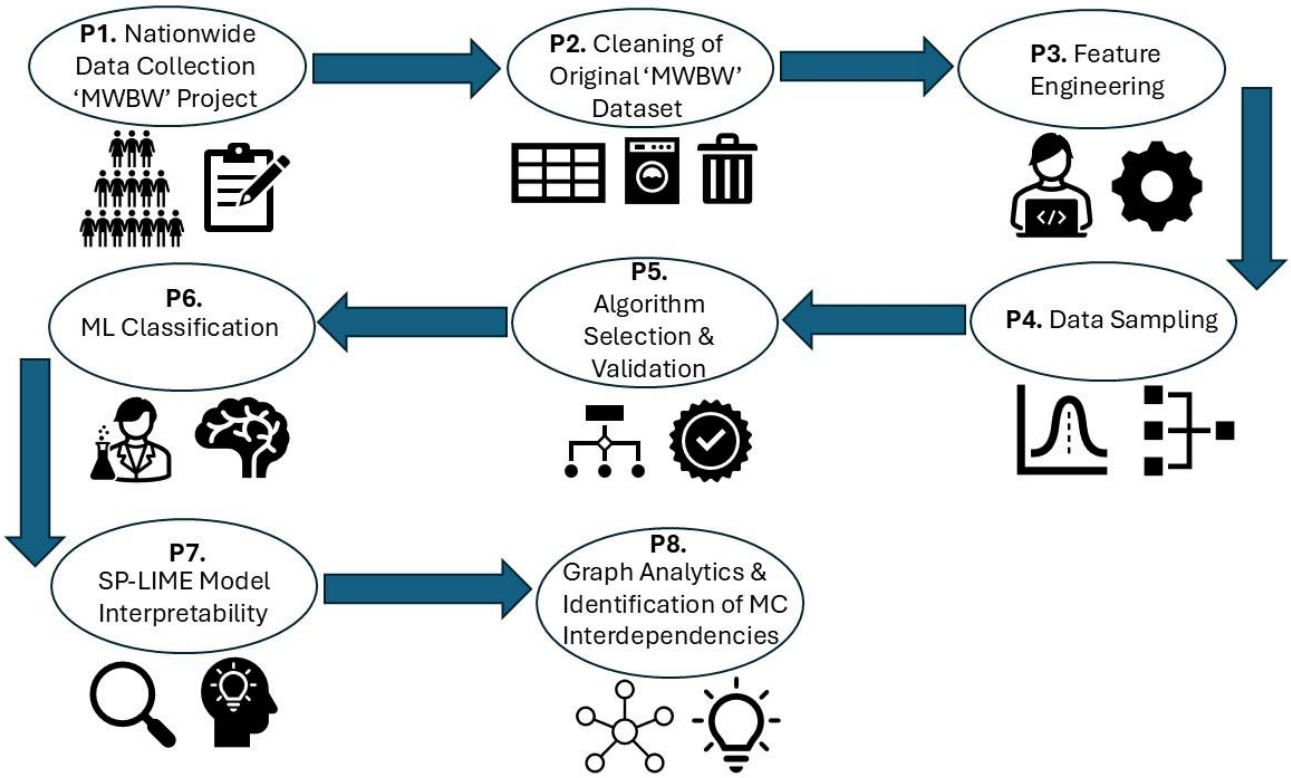


Figure 4.1 Process for Modelling MC Inter-dependencies

Data Collection & Analysis

Cross-sectional AMC data were collected as part of an Irish AMC study 'Moving-Well-Being Well' (MWBW) (Behan et al., 2019). Participants (n = 1,124, ranging from 9-12 years of age) were recruited from 44 schools across twelve counties, encompassing all four provinces in Ireland and Northern Ireland. Full details on ethical and data collection procedures, analysis and synthesis of this AMC data can be found via (Behan et al., 2019). Additionally, participant's PMC was assessed upon completion of AMC assessments using the Pictorial Scale of Perceived Movement Skill Competence (PMSC) for Young Children (Barnett et al., 2015). The PMSC assessed 15 AMC skills, comprising seven object-control and eight locomotor skill perceptions. Further information regarding PMSC assessment procedures is reported in Barnett et al. (2015).

Dataset Cleaning

The original 'MWBW' dataset consisted of 2,378 rows and 492 columns. The initial cleaning process involved handling Null values and removing variables which were deemed surplus for the purpose of the study. Given that the current study sought to uncover age-centered interdependencies, the notable variation in MNM proficiency from chapter 3 was greatly considered when deciding upon which subgroups to examine. From the original dataset, the 5–8-year-old cohorts produced substantial non-mastery cases with limited MNM cases demonstrated, therefore limiting the reliability of comparisons to the older age-groups. Conversely, the 13-year-old cohort was further limited by sample size whilst producing greater variability in class distributions. Therefore, the 9–12-year-old cohorts represented a sample in which mastery was sufficiently developed to enable a greater representation of MNM proficiency across skills, while still retaining meaningful variability between participants and across both age and gender. Consequently, by limiting the analysis to this range it reduced the influence of extreme class sparsity and small subgroup sizes on subsequent comparisons.

Given the aforementioned considerations a robust cleaning procedure was undertaken to filter the dataset accordingly to ensure the bespoke sample was included for analysis with this resulting in a total of 1,124 rows and 30 columns remaining which comprised the final dataset. The final dataset was therefore comprised of 582 males and 542 females. The remaining variables included 16 locomotor skill and 14 object-control skill variables were analysed which resulted in a total of 30 variables being included in the analysis. It could be argued that the implementation of Principal Component Analysis (PCA) for tasks such as multicollinearity mitigation and dimensionality reduction would have streamlined this cleaning process. However, with a key aim of the current study being interpretability of the classification model's predictions, PCA would have resulted in a loss of interpretability as the original features would have been transformed into abstract components, making it difficult to explain what each principal component meant in terms of distinct MC features.

Feature Engineering

Before proceeding with data sampling, several key feature engineering procedures were completed to refine the dataset for analysis. The first step involved encoding categorical variables using one-hot encoding to transform non-numeric attributes into a machine-readable

format. This resulted in the generation of binary indicator variables for each categorical feature while removing the original columns. This ensured the classification models could effectively process categorical data without introducing unintended ordinal relationships.

Next, numerical features such as weight and height were standardized using ‘StandardScaler‘ to ensure uniformity across different scales. Since ML models, particularly those relying on distance-based calculations, can be sensitive to variations in feature magnitude, standardization ensured that all continuous variables had a mean of zero and a standard deviation of one. This step helped prevent any single feature from disproportionately influencing the model due to its scale. Additionally, feature selection was performed to remove unnecessary columns that were either redundant, unrelated to the target variable, or introduced as null indicators. A filtering process identified these extraneous features, which were subsequently excluded from the dataset to improve computational efficiency and model interpretability. Once these transformations were complete, the cleaned and structured dataset was saved in Parquet format, preserving its optimized structure for the next phase of analysis: data sampling.

Data Sampling

After preprocessing, the dataset was split into gender-specific subsets. Each AMC assessment was treated as a classification task with binary labels: "Mastery/Near-Mastery" (MNM = 1) or "Not Mastered" (0) (Ulrich 2020). To address class imbalance (e.g., skewed distributions in certain skills), four resampling strategies were applied: under-sampling, oversampling, unsampled, and SMOTE, producing four balanced variants of the dataset for analysis.

Algorithm Selection and Validation

In the current study our approach was to employ several distinct classification algorithms, several of which were employed in recent related research (Jiabin et al., 2025). Initially, we began with a baseline classification algorithm before gradually implementing more complex algorithms as the analysis progressed. The classification algorithms employed were decision tree (DT), gradient boosting decision tree (GBDT), k-nearest neighbour (KNN), logistic regression (LR) and XGBoost (XGB). The metrics used to validate the classifiers were:

precision (percentage of observations labeled as positive that are actually positive), recall (percentage of positive observations classified as positive) and F1 (weighted average of precision and recall). As such, these metrics were calculated from four distinct counts: True positives (TP), a count of the correctly predicted positive values; true negatives (TN), a count of the correctly predicted negative values; false positives (FP), which count the occurrences where the actual class is false but is predicted as true; and false negatives (FN) , which count the occurrences of where the actual class is true but is predicted as false. The sum of these four counts is the total number of observations.

Additionally, the F1-metric was utilised as it is a recognised metric enabling comparison between classifiers trained using differing methodologies, which was present in our case with an imbalanced distribution existing between non-mastery and mastery-near-mastery values.

The classification algorithms were trained using an 80:20 train-test split, hyperparameter optimization was performed using GridSearch to fine-tune performance with K-Fold Cross Validation also performed for further model validation.

ML Classification

Our classification objective was to examine whether outcomes of specific AMC assessments could be accurately predicted from the remaining AMC and PMC results. For each of the 12 AMC target outcomes, five classification algorithms were evaluated under four sampling conditions across gender-combined, female-only and male-only subsets. This resulted in 240 model configurations per subset and 720 predictive classification models overall. This process involved splitting the dataset into distinct gender subsets—Male Only (M), Female Only (F), and Gender Combined (GC)—to analyze potential differences in prediction accuracy across groups. Each AMC assessment served as a dependent variable, and multiple ML classifiers were applied to each subset. The top five best performing classification algorithms (F1-score) are highlighted in Table 4.1.

Table 4.1. Best Performing Classification Algorithms

Rank	Feature	Classifier	Sampling	Subset	F_1-Score
1st	Catch	KNN	Unsampled	Female	0.87
2nd	Bounce	XGB	Unsampled	Male	0.86
3rd	Run	XGB	Unsampled	Female	0.85
4th	Hop	KNN	SMOTE	Female	0.85
5th	Slide	XGB	Unsampled	Gender Combined	0.84

Model Interpretability; SP-LIME

To interpret the predictions produced by the best-performing classification models, Submodular Pick Local Interpretable Model-Agnostic Explanations (SP-LIME) was employed (Ribeiro et al., 2016). At a basic level, LIME helps interpret an individual prediction from a classifier by understanding the model's behaviour around that observation using a simpler and locally interpretable surrogate model. SP-LIME extends this analysis by choosing a set of representative and comparatively non-redundant local explanations, thereby enabling greater insight into the wider behaviour of the original model.

For each selected classification model, a tabular LIME explainer was initialised in classification mode using the corresponding training predictors, predictor names and class labels. The relevant classification model's predicted class probabilities were supplied through its `predict_proba` function. The SP-LIME submodular-pick procedure was configured to select ten representative instances for each classification model. For each selected instance, the five highest-ranked explanatory features and their locally estimated contributions were retained. This produced 50 feature contributions for each competency-specific model.

Both positive and negative contributions were recorded. A positive contribution inferred that the feature supported the locally predicted class, whereas a negative contribution inferred that it reduced support for that prediction. No further coefficient-magnitude threshold was applied after the SP-LIME phase of the analysis; therefore, the restriction to the five highest-ranked features within each interpretation accounted for the feature-selection threshold. These interpretations should be considered as descriptions of the classifiers' prediction behaviour rather than direct evidence that the identified features causally determined MC proficiency.

The SP-LIME explanations identified several prominent contributors to the model predictions. Among 12-year-old females, one- and two-handed striking, rolling, kicking and running were prominent in the interpretability of catching performance. Among nine-year-old males, vertical jumping, two-handed striking and perceived motor competence measures relating to galloping, two-handed striking and horizontal jumping contributed prominently to predictions of bouncing proficiency. These findings provided an initial indication of potentially key relationships between MC measures and motivated their subsequent examination through graph analytics.

Graph Construction

The SP-LIME outputs were subsequently transformed into a graph representation of the dependencies uncovered by the classification models. Each retained interpretation was initially represented as a set of relationships between the competency predicted by the classifier and its five highest-ranked explanatory features. As the same interpretable feature could appear across several of the ten selected instances, repeated competency–feature contributions were combined by summing their positive and negative values. Retaining the direction of the contributions meant that positive and negative values for the same relationship could offset one another during aggregation. The resulting value represented the combined contribution of that feature across the selected SP-LIME interpretations. Every resulting competency–feature pair was retained, and no further numerical threshold was applied.

Each relationship was represented in Neo4j as a directed `Depends_on` edge from the predicted competency to the interpretable feature. Simply put, an edge from competency A to feature B indicated that the model's predictions of competency A depended, at least in part, on feature B within the SP-LIME interpretability. The combined feature contribution was preserved as a property of the corresponding relationship. A separate subgraph was first constructed for each competency-specific classifier. Nodes with identical names and repeated relationships were then merged, integrating the competency-specific subgraphs into one unified network.

Graph Analytics

To uncover potential interdependencies identified across MC measures, three graph-analytic algorithms were applied to the unified network: degree centrality (DC), betweenness centrality (BC) and PageRank (PR). The measures were calculated using projections of the same unified graph rather than through comparisons between multiple independent networks. Degree centrality was calculated using a reversed directed projection so that features were evaluated according to the competencies with which they were associated (Barthélemy, 2004). Betweenness centrality was calculated using an undirected projection, allowing the analysis to identify nodes positioned along paths connecting otherwise distinct areas of the network (Barthélemy, 2004). PageRank was calculated using the original directed projection to identify features receiving dependencies from influential competency nodes (Page, 1998). No

additional normalisation was undertaken to adjust for differences in network size or density as separate networks of differing sizes were not compared. The directed, reversed and undirected projections represented alternative computational views of the same underlying set of nodes and relationships. Figure 4.2 highlights the approach taken in the current study.

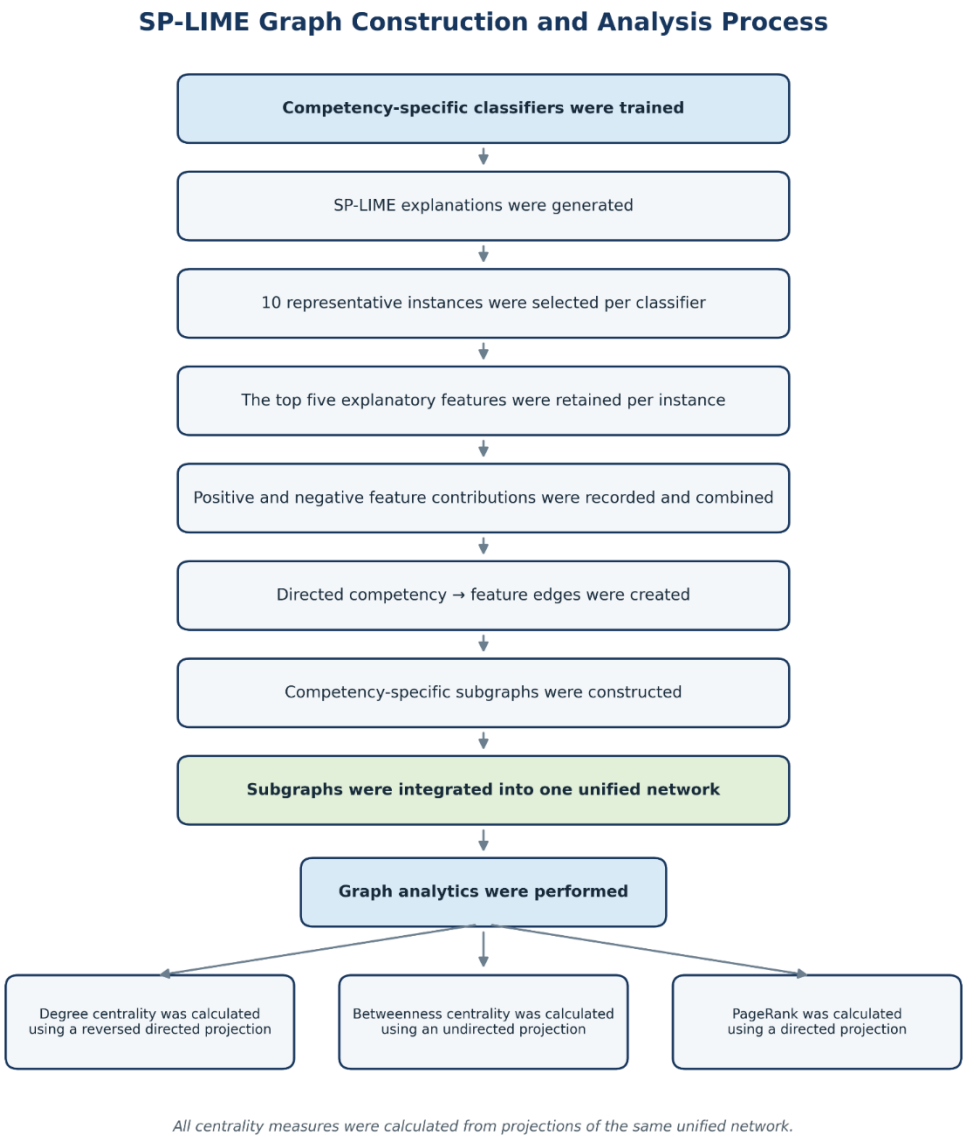


Figure 4.2 SP-LIME & Graph Analytics Flowchart

4.4 Results

Graph analytics results are listed in full within Tables 4.2, 4.3 and 4.4.

Table 4.2. Degree Centrality Results

Feature	Edges	Age	Gender
roll_total	4	12	Male
hj_total	4	9	Female
vj_MNM	3	9	Male
1hs_total	3	10	Male
1hs_MNM	3	10	Male
skip_total	3	10	Female
hj_total	3	10	Female
gallop_total	3	11	Male
gallop_total	3	11	Female
roll_total	3	12	Female

Table 4.3 Betweenes Centrality Results

Feature	Score	Age	Gender
Catch_MNM	335.00	9	Female
Kick_MNM	314.50	9	Male
Slide_MNM	267.83	11	Female
Hop_MNM	164.50	12	Female
Hop_MNM	148.07	10	Female
Slide_MNM	139.36	10	Male
Run_MNM	121.17	11	Male
Kick_MNM	109.24	12	Male

Table 4.4 PageRank Results

Feature	Score	Age	Gender
roll_total	0.23	12	Male
hj_total	0.22	9	Female
bounce_MNM	0.21	11	Male
bounce_total	0.21	11	Male
gallop_total	0.21	11	Male
hj_total	0.21	10	Female
skip_total	0.21	10	Female
roll_total	0.21	12	Female
vj_MNM	0.20	9	Male
1hs_MNM	0.20	10	Male
1hs_total	0.20	10	Male
gallop_total	0.20	11	Female

Degree Centrality Results

DC results highlighted rolling (12-year-old males) and horizontal jumping (9-year-old females) as the most central skills, aligning with prior work by Duncan, Martins, et al. (2022) which identified jumping and catching as pivotal in AMC development. Similar findings were noted

by Salami et al. (2024) with gallop, slide, dribble, catch, and run as central skills in prepubertal networks. Bandeira, Lemos, et al. (2025) further showed, through Exploratory Graph Analysis, that such core nodes map onto stable MC dimensions, supporting the integration of both process- and product-oriented assessments.

Betweeness Centrality Results

BC results revealed catching and kicking as influential bridging skills among male and female nine-year-olds, consistent with earlier findings of their role in linking locomotor and object-control development (Bandeira et al., 2020). Salami et al. (2024) further supported this sentiment showing that such bridging nodes differ by gender, with girls' networks structured around gallop/slide and boys around dribble/catch/run. Such skill-specific differences highlight the contextual nature of central connectors. Bandeira, Lemos, et al. (2025) recently highlighted that bridging roles may also reflect underlying latent dimensions of MC, making them critical in shaping efficient assessment batteries.

PageRank Results

PR results emphasized rolling and horizontal jumping as most influential, echoing Duncan, Martins, et al. (2022) who identified similar foundational skills in Irish children. Salami et al. (2024) likewise demonstrated the prominence of locomotor and object-control tasks as network hubs, while Bandeira, Lemos, et al. (2025) confirmed that these hubs cluster within four robust MC dimensions. The NA demonstrated by Bandeira et al. (2020) showed that children who performed well in galloping and hopping also tended to excel in other AMC skills, suggesting these skills play a foundational role in early motor development. Given these overlapping results, early-stage motor development programs should emphasize these high-ranking skills to maximize long-term competence gains.

Graph Inter-dependencies

To further examine the existing inter-dependencies between MC variables, graph analytics were conducted to model the interactions across all age-grades for both genders. Examples of such graphs are presented in Figures 4.3 and 4.4. Graph analytics provide a nuanced view of MC inter-dependencies across different age groups and genders. As evidenced, in Figure

4.3, by the presence of edges between the nodes 'Kick_MNM', 'PCKick_MNM', and 'PCSkipping_MNM', twelve-year-old female's actual kicking proficiency appeared to have been influenced by numerous perceived abilities, underscoring the influence of PMC in AMC development. Unlike Duncan, Martins, et al. (2022), Bandeira et al. (2020), Bandeira, Lemos, et al. (2025) and Salami et al. (2024) whose graphs focused solely on direct AMC relationships, our findings would suggest that PMC inter-dependencies play a crucial role in overall MC development. Recent research by Niemistö et al. (2023) noting PMC-AMC profiles in early childhood to successfully predict subsequent PMC-AMC profiles in middle childhood further validates these findings.

Similarly Figure 4.3, illustrates that running proficiency among nine-year-old males depended on sliding, catching and skipping proficiencies meaning that targeted interventions on these skills would directly improve running ability. Our graph analytics (Figure 4.4) which demonstrate the influential role of object-control proficiency (specifically kicking) upon numerous MC competencies among nine-year old males can be said to be in line with the findings of Behan et al. (2019) which noted this age-grade as the most proficient in object-control skills across five - twelve-year-old males. Interestingly, it can also be noted among nine-year-old males that both actual bounce and slide proficiency depended on numerous measures of PMC. While Duncan, Martins, et al. (2022) identified bounce and catch as central nodes in their NA and Bandeira et al. (2020) highlighted strong links between throwing and striking, -our findings suggest that PMC may be a key missing factor in such related research and should be considered for inclusion by researchers in their future work.

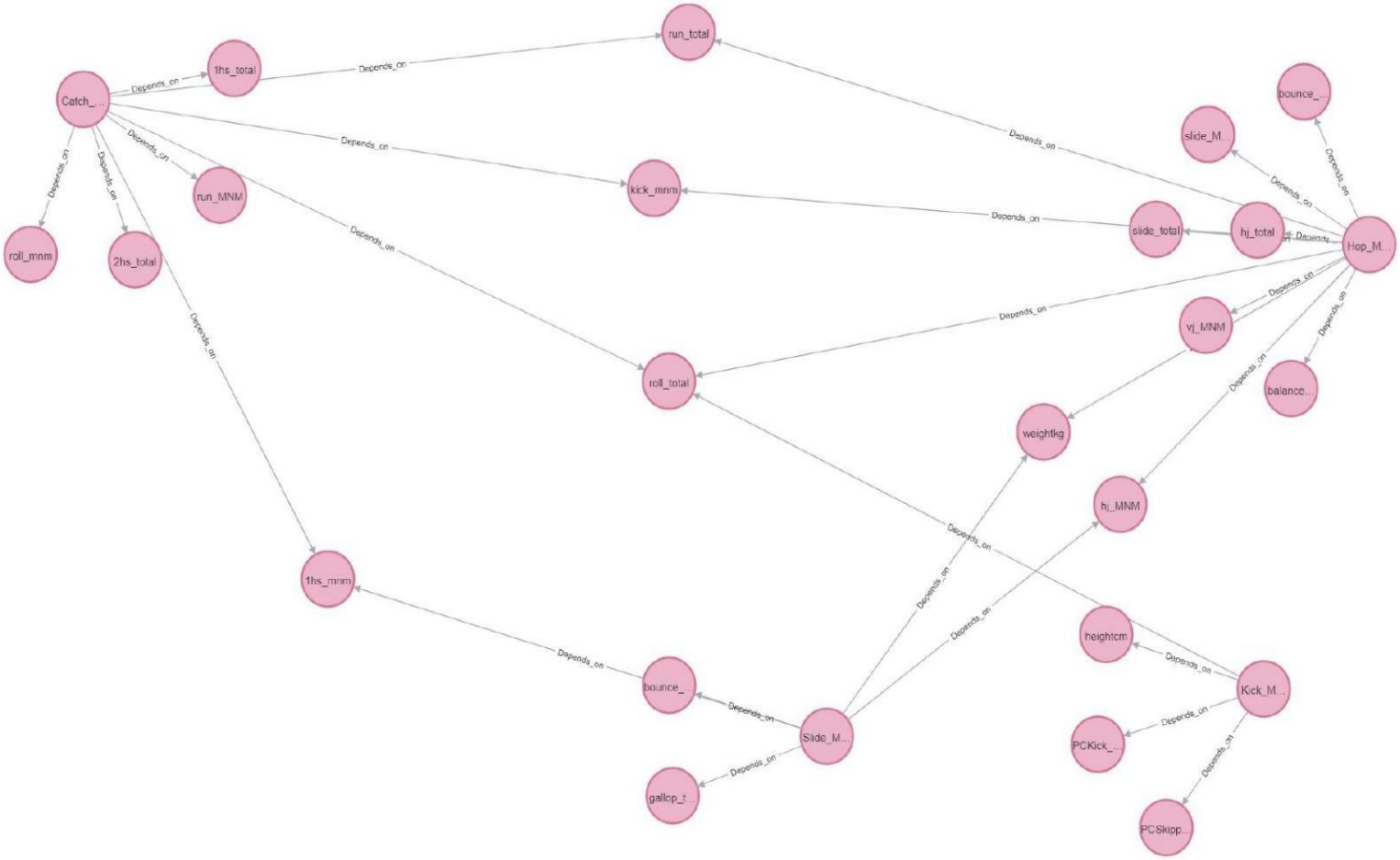


Figure 4.3 12-Year-Old Female's Interdependencies

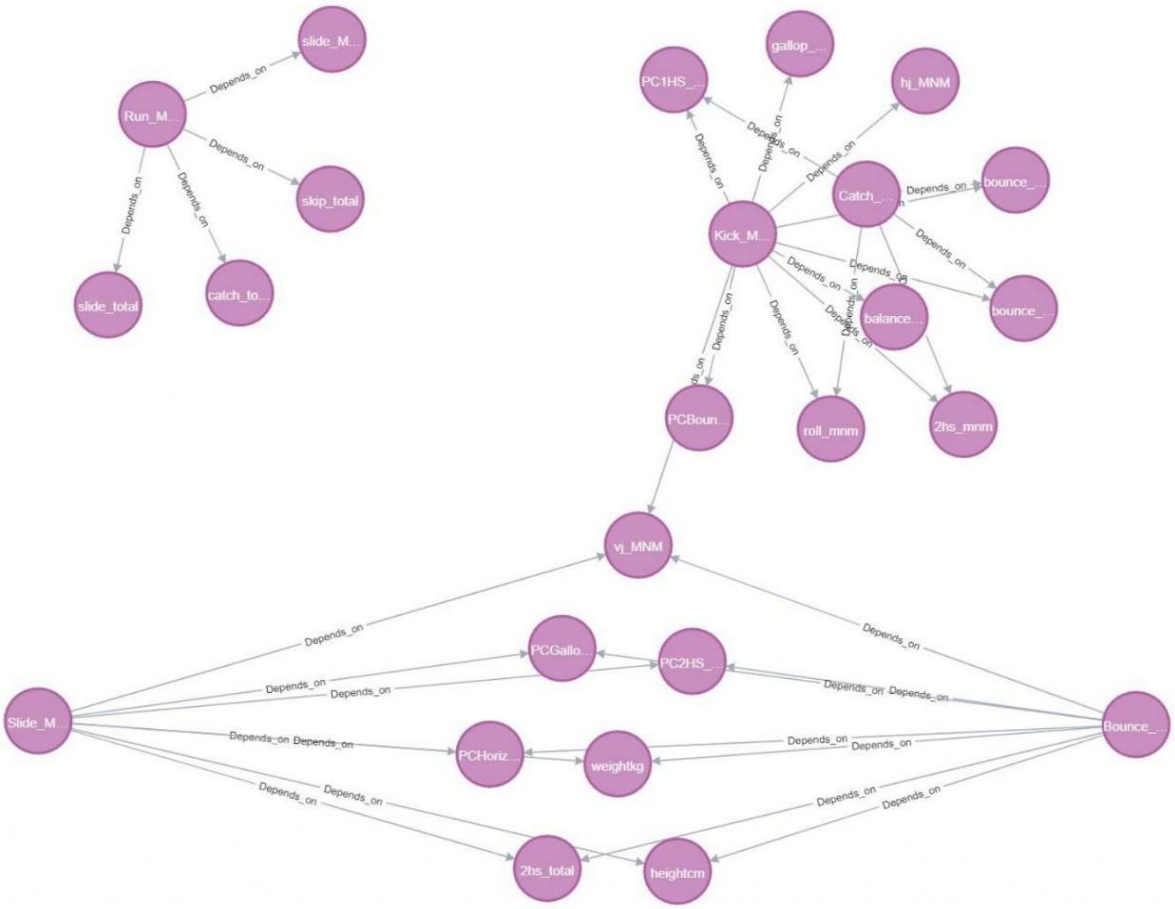


Figure 4.4 9-Year-Old Male's Interdependencies

4.5 Conclusion

Our current study combined MI and graph analytics to uncover key interdependencies between AMC and PMC across age and gender. The findings identified several skills, including catching, kicking, and running, as influential within broader MC networks. Furthermore, the inclusion of PMC measures highlighted important interdependencies with AMC, reinforcing the proposition that children's perceptions of their movement capabilities may play a central role in both MC development and long-term PA engagement (Stodden et al., 2008). Collectively, these findings provide practical guidance for physical education teachers and sports coaches seeking to design more targeted assessment and intervention strategies.

Importantly, further examination of the network structures revealed that several skills contained within the locomotor subset of the TGMD-3 Short Form assessment battery (Duncan, Martins, et al., 2022) occupied influential positions within age- and gender-specific networks.

Given that the TGMD-3 Short Form has previously demonstrated validity and reliability as a measure of MC among Irish children (Duncan, Martins, et al., 2022) alongside it's efficient nature in comparison to the full assessment-battery, such findings therefore reinforced the value of assessing these skills in our future work. However, while identifying influential skills is important, the practical challenge of assessing MC at scale remains. Traditional assessment procedures are often time-intensive, require trained assessors, and can be difficult to implement across large populations (Duncan, Martins, et al., 2022). Consequently, having identified several influential MC constructs, it became pertinent to consider how emerging technologies might facilitate more efficient and scalable assessment procedures. Accordingly, Chapter 5 shifts attention from understanding MC relationships towards examining the potential role of markerless motion-capture technologies in supporting future MC assessment, while critically evaluating the opportunities, challenges, and research directions associated with their implementation.

Chapter 5: Markerless Motion Capture for Motor Competence Assessment in Children: Promises, Limitations, and a Research Agenda

5.1 Abstract

The position paper to follow sets out the methodological position that underpins the empirical work in subsequent chapters of the thesis. It brings forward the argument for the implementation of markerless motion capture as a foundational approach to motor competence assessment in children and provides the rationale for the design choices adopted in the study that follows. Throughout the chapter, the term ‘markerless motion capture’ is used in a system-level rather than an individual sense. It denotes camera-based systems that require no body-worn markers, sensors, or wearables, regardless of who manufactures them. Traditional reflective-marker configurations such as the gold-standard Vicon setup therefore sit outside this definition and are treated as a reference standard against which markerless approaches are benchmarked. The distinction is system-level because several established marker-based vendors now also offer markerless product lines. These qualify as markerless under the operational definition adopted here, even where the same vendor’s marker-based systems do not.

Key Words: Motor Competence; Markerless Motion Capture; Computer Vision; Automated Assessment; Physical Activity; Research Agenda

5.2 Introduction

Motor competence (MC) in childhood underpins the development of fundamental movement skills (FMS), which in turn support lifelong engagement in physical activity (PA) and overall health (Stodden et al., 2008). These skills, spanning locomotor, object-control, and stability domains, are widely recognised as foundational to both physical and psychosocial development (Logan et al., 2018). Despite their importance, evidence consistently indicates that many children demonstrate low levels of MC, therefore, assessment of these skills are critical for identifying developmental delays, informing interventions and quantifying proficiency levels (Bardid et al., 2019; Barnett et al., 2016; Hulteen et al., 2015). Assessments are categorized as either product or process and differ on the number of skills assessed (i.e., individual or subscales), ease of administration and scoring procedures (Logan et al., 2017). Process-oriented assessments assess the performance of a movement and describe qualitative movement patterns whilst product-oriented assessment assesses movement outcomes, which is typically identified as a quantitative score (e.g., speed, distance or number of successful attempts). In spite of an abundance of readily available valid and reliable assessments, a key challenge for researchers and practitioners lie in assessing MC accurately, objectively, and at scale.

Traditional approaches to MC assessment, such as product and process-oriented assessments while widely used, are limited by logistical constraints such as sufficient time for assessor training and specific space and equipment which can result in limited deployment by practitioners such as teachers and coaches (Duncan, Weldon, et al., 2022; Lander et al., 2016, 2020; Lander et al., 2015; Ng et al., 2025). Established observational methods such as the process-oriented 'Test of Gross Motor Development-3rd Edition' (TGMD-3) (Ulrich, 2013), product-oriented 'Körperkoordinationstest für Kinder (KTK)' (Kiphard & Schilling, 1974), and the Movement Assessment Battery for Children (MABC-2) (Henderson et al., 1992) have provided valuable frameworks for evaluating MC proficiency. However, these assessments are labour-intensive, reliant on subjective human judgement, and limited in the granularity of data they produce (Duncan, Martins, et al., 2022). As a result, they are difficult to scale beyond small cohorts and often fail to capture the nuanced movement patterns that underpin motor-skill development.

This paper argues that markerless motion capture should be adopted, not as a replacement for process-oriented scoring but as a complementary measurement tool, as it addresses key

limitations of traditional process-based methods by enabling scalable, objective, and data-rich evaluation of proficiency levels. By leveraging advances in computer vision and sensor technology, markerless systems offer the potential to substantially streamline MC assessment across research, education, and clinical contexts. Marker-based systems provide sub-millimetre three-dimensional kinematic data and are treated in this chapter as the reference standard against which markerless approaches are validated (Matsuda et al., 2025). Within the markerless category, the more accessible options include markerless depth cameras (e.g., Microsoft Kinect) (Chow et al., 2020) and open-source platforms (e.g., '*FreeMoCap'*: Matthis et al. (2022) and OpenCap: Uhlrich et al. (2023)) which offer scalable and cost-effective alternatives that can be deployed in real-world environments without the need for specialised laboratories or wearable sensors.

While these more modern approaches present distinct strengths and limitations, their collective emergence signals a paradigm shift in MC assessment. By moving beyond an exclusive reliance on traditional observational methods, markerless motion capture, as a complementary objective form of measurement enables a more comprehensive and ecologically valid appraisal of MC proficiency. The remainder of the chapter is organised around four arguments. The first considers the limitations of established observational MC assessments and the case for moving beyond this sole observational reliance and instead foster greater levels of objectivity in assessment (Lander et al., 2020). The second reviews the principal markerless approaches, from depth cameras to open-source video pipelines, against the reference standard of marker-based capture and asks where the current evidence supports their use and where it does not. The third addresses the methodological and ethical challenges of applying these systems with children, since this is where the strongest claims of the chapter need to be defended. The fourth considers applications in research, education, and clinical practice, and is careful to separate claims that are evidence-supported from those that remain speculative. The chapter closes with a research agenda that the empirical work in the subsequent chapter begins to address.

5.3 Position-paper Approach

This current thesis chapter is presented as a critical narrative position paper rather than a deep systematic review. Its purpose is to develop and defend a methodological position regarding the role of markerless motion capture in children's AMC assessment, while critically

examining the technical, methodological and practical evidence supporting that position. Relevant evidence was identified through targeted searches of Google Scholar and supplemented by relevant literature identified during the broader thesis review. Sources were selected purposively according to their relevance to: (1) optical and non-optical motion-capture systems; (2) validation against reference systems or expert assessment; (3) applications involving children or AMC assessment; and (4) technical, ethical and implementation limitations. Greater emphasis was placed on empirical validation studies, systematic reviews and studies involving paediatric populations. Adult validation studies were included where they provided important technical evidence unavailable in children. As the identification and selection procedures were not designed as a systematic search, the evidence base should be interpreted as a critical and illustrative synthesis rather than an exhaustive account of all available studies.

5.4 Limitations of Traditional Motor Skill Assessments

Established MC assessments such as the TGMD-3, KTK, and MABC-2 are valid, reliable, and have underpinned much of what is currently known about children's motor development, including, in the case of process-oriented batteries, the capacity to characterise proficiency at the level of individual skill components (O' Brien et al., 2015, 2016). The limitations outlined below should therefore be read not as a dismissal of these tools, but as a description of the practical and methodological constraints that currently hinder their use with both precision and at scale:

- **Resource and Scalability Constraints:** Standardised assessments are labour-intensive and time-consuming. For instance, a full TGMD-3 or MABC-2 assessment must be administered on a participant-by-participant basis by a trained assessor, which typically requires 20 - 40 minutes per participant. Scaling this procedure to hundreds of children (e.g. an entire school) becomes impractical (Tamplain et al., 2020). The administration of large-scale assessment batteries requires substantial time, dedicated space (e.g., school halls or testing labs), and equipment, as well as coordination of schedules between numerous stakeholders (e.g., teachers and assessors). Indeed, the time burden of large-scale assessments is often cited as a major impediment in the field (Bardid et al., 2019). Such logistical barriers mean that comprehensive MC assessments are rarely completed beyond small research samples or specialised

clinics, leaving most children unassessed.

- **Observer Bias and Reliability Issues:** Process-oriented assessments rely upon human judgement for accurate grading against the associated performance criteria. Whilst such performance criteria may be clear and concise (for the most part) in assessment manuals such as the TGMD-3 (Ulrich, 2013) and GSGA (Department of Education and Training, 2000), scoring can be subjective. For instance, different examiners may vary in how strictly they apply the criteria in their grading or which aspects of form they notice, leading to inter-rater variability (Ward et al., 2020). Training and certification of assessors can enhance reliability, but in practice eliminating all bias in live observations is difficult. Fatigue or distractions can also result in assessors missing key details or making data entry errors during large-scale data collection. These human factors highlight the requirement for more objective and automated measurement of MC proficiency to better support observational assessment.

- **Limited Scope and Data Granularity:** A further limitation common to MC assessments more generally concerns the level of detail they capture. Each battery samples a finite set of skills and reduces performance to simplified scores, whether the present or absent judgements of process-oriented criteria or the single outcome values of product-oriented measures, which are then typically aggregated into a composite or a small number of domain scores (Hulteen et al., 2023). This aggregation can mask important information about an individual child. The KTK, for example, yields an overall motor coordination score but cannot indicate whether a child's strength lies in locomotor or object-control proficiency, while outcome measures such as jump distance record what was achieved without describing how the movement was produced. As a result, crucial kinematic such as joint angles, movement timing, and inter-limb symmetry is rarely retained. This lack of detail can obscure specific deficits; for instance, a child might meet the minimum standard for a jump in an outcome score yet still use an atypical technique associated with future injury risk. Existing assessments may also be prone to ceiling effects for highly skilled children and insufficiently sensitive to the small, incremental improvements that occur over time.

- **Outdated Norms and Cultural Validity:** Some widely used assessments are showing their age in terms of normative data and content relevance. For instance, the KTK’s norms date back to the 1970s (Kiphard & Schilling, 1974) with concerns raised over

their outdated standards of borderline results that are not sufficiently heterogeneous to distinguish the degree of motor coordination proficiency levels (Antara et al., 2023; Chovanová et al., 2023). Such concerns have prompted contemporary efforts to reestablish age and country specific reference values (Vandorpe et al., 2011), raising questions about current validity. Whilst the TGMD-3 was recently updated, some assessments (i.e., overhead throw or ball strike) may not be equally familiar to children in all cultures, potentially biasing results (Bardid et al., 2019; Valentini et al., 2017). Moreover, these assessments were developed mainly for typically developing children; adaptations for children with disabilities are often ad hoc, which can impede standardisation (Tamplain et al., 2020). Indeed, a more flexible, detailed assessment approach could better accommodate diverse populations and more niche use cases, such as telerehabilitation.

In conclusion, traditional motor-skill assessments, while invaluable for establishing baselines and comparisons, are laborious, subjective and constrained in the granularity and scale they can offer. As our domain aspires to monitor motor-skill development in larger and more diverse populations, and to do so with superior precision, we must confront the aforementioned limitations by adopting more objectivity in research (Lander et al., 2020) to support and work in tandem with traditional observational methods. Encouragingly, modern technology offers a path forward. Markerless motion capture systems hold the potential to automate analysis, providing objective and rich kinematic data at scale. In the sections to follow we shall examine how these technologies can navigate the hurdles of traditional assessments and discuss their contributions and current limitations.

5.5 The Reference Standard: Marker-Based Optical Systems

Systems like Vicon have long been considered the gold standard for motion analysis (Topley & Richards, 2020). Such optical systems utilise multiple infrared cameras to track reflective markers placed on the body, yielding sub-millimetre accuracy in 3D position data (Greene et al., 2010; Motti Ader & Bossavit, 2023). This technology has provided invaluable insights into movement biomechanics and serves as a benchmark for validating new methods. In theory, practitioners could assess children's motor skills using a Vicon system and obtain extremely precise measurements of their proficiency levels. In practice, however, these lab-based systems are impractical for broad motor-skill development assessment (Lerma & Gulgin,

2023). A full optical setup is exceedingly expensive and requires a dedicated, controlled space (Van Der Kruk & Reijne, 2018). In addition, preparing a child for capture is not trivial. Dozens of reflective markers must be affixed to anatomical landmarks, the system requires repeated calibration, and any marker drop-out during testing can compromise the trial (Topley & Richards, 2020; Van Der Kruk & Reijne, 2018; Bossavit & Arnedillo-Sánchez, 2023). Such a process may be feasible for in-depth studies on small samples, but not for testing hundreds of children in a school environment. Thus, while marker-based systems offer unrivalled accuracy, their cost, complexity, and lack of scalability make them an unrealistic solution for regular or large-scale assessments of children's MC proficiency. In turn, recent technological developments have enabled markerless motion capture to serve as a solution to such limitations. Motion capture technology (MCT) has revolutionised biomechanics by enabling precise human movement tracking. In MC research, the appeal of motion capture is clear; it can capture detailed kinematic information (joint angles, velocities, coordination patterns) that observational tests miss. The following section shall consider the potential of MCT for advancing MC assessment in detail.

5.6 Markerless Motion Capture: A New Paradigm for MC Assessment

Until recently, MCT was confined to specialised labs (Müller et al., 2017), however a spectrum of markerless motion capture approaches are now available, from high-end optical systems to consumer-grade depth sensors to open-source camera-based tools; all of which record movement without body-worn markers or suits (Bossavit & Arnedillo-Sánchez, 2023). Below we highlight examples across this spectrum, noting their strengths as well as limitations in relation to motor-skill research.

Markerless Depth Cameras: In the past decade a breakthrough has been the advent of affordable depth-sensing cameras, exemplified by the Microsoft Kinect (Clark et al., 2012). Originally marketed as a gaming accessory, the Kinect provides both colour images and depth maps, allowing it to infer 3D joint positions (Pagliari et al., 2014; Shotton et al., 2011) without any markers or wearables (Clark et al., 2012). Its introduction signalled a paradigm shift, allowing researchers to perform motion capture in a regular classroom or clinic simply by plugging a transportable device into a laptop. The Kinect enabled exciting possibilities in the space of MC assessment. Given it's non-intrusive, child-friendly (children can move naturally

without wires or markers) and highly portable nature, the system offers real-time skeletal tracking, meaning it can give instant feedback or drive interactive “exergame” exercises (Shotton et al., 2011). For example, a Kinect placed in a school gym can track a child’s skipping motion or throwing motion in real time and potentially evaluate performance against desired patterns (McGann et al., 2020). These features make markerless systems powerful for engaging children and scaling up assessments beyond laboratory settings. However, prior to large-scale deployment, validation should occur at both the kinematic level, through comparison with gold-standard criterion systems such as Vicon, and at an assessment battery level, through comparison with established MC process-oriented assessments such as the TGMD-3. Indeed, early studies demonstrated that even non-expert testers could use Kinect-based software to assess fundamental movement skills with reasonable agreement (‘Moderate’ percentage strength of agreement; 41- 60%; Sasyniuk et al., 2007) to expert human ratings, with the percentage agreement between the MCT and the rater for each performance criterion, the percentage agreement levels for all performance criteria ranged from 43.6% to 92.9% (Chow et al., 2020). Using TGMD-3 assessor evaluations as benchmark data, Guo et al. (2026) noted an AI-based system to demonstrate excellent agreement (ICC = 0.94) and a strong correlation with TGMD-3 scores (r = 0.92), while maintaining low mean absolute error values across running (3.8), jumping (4.2), balancing (4.9), and hopping (5.1), indicating that AI-generated results closely matched traditional human evaluations. Markerless depth cameras thus directly address some resource issues of observational tests, offering a cost-effective, rapid alternative for capturing MC data in the field.

Despite their advantages, commercial markerless systems like Kinect come with notable trade-offs. The most discussed limitation is accuracy when compared to expensive marker-based systems. Depth sensors do not yet reach the fine-grained precision of marker-based systems, especially for complex or high-speed movements or for tracking small segments (fingers, toes) (Cremers et al., 2015). The Kinect’s joint estimates typically have errors on the order of a few centimetres, and accuracy diminishes if the person is not directly facing the sensor or if multiple people occlude each other (Dutta, 2012; Pfister et al., 2014). Validation studies have reported mixed results: some find good to excellent test–retest reliability for certain Kinect-derived metrics, for example, Mentiplay et al. (2018) reported ICC values ≥ 0.73 for single leg squats and drop vertical jumps variables whereas others find only moderate agreement with Vicon data (e.g. correlation coefficients in the 0.5–0.7 range for some joint angles) (Lerma & Gulgin, 2023). In one comparison, a Kinect V2 paired with a custom

assessment app showed poor-to-moderate agreement with a Vicon system on children's vertical jump and leap kinematics(Spearman correlation coefficients ranging from −0.787 to 0.543 during the Jump-Land task and from −0.598 to 0.523 during the Leap task), leading the authors to caution that Kinect should not replace lab-based motion analysis for detailed measures (Lerma & Gulgin, 2023). In general, Kinect can differentiate gross differences in skill (e.g. it can accurately classify a poor versus proficient jump) but may miss subtle nuances or gradual improvements over time (Lerma & Gulgin, 2023). The proprietary and fixed nature of such systems can be considered as a further limitation: the original Kinect models have been discontinued (though a newer Azure Kinect exists (Tölgyessy et al., 2021)), and the algorithms are "black boxes" that users cannot adjust (Gottfried et al., 2011). If the device's default tracking isn't well-tuned for young children's movement patterns, there is little recourse. Additionally, depth cameras have known environmental constraints, they struggle outdoors in bright light, have limited range in very large spaces, and only track a limited number of individuals at once before occlusion becomes an issue (Zhang et al., 2012).

In summary, while Kinect and similar markerless cameras have greatly expanded the toolkit for motor-skill research, they are not a full-proof solution. Such systems strike a balance between laboratory-grade data and real-world practicality; however, researchers must be mindful of their accuracy limitations and dependence on proprietary support. The poor to moderate agreement reported by Lerma and Gulgin (2023) between Kinect V2 and Vicon for children's vertical jump and leap kinematics falls in the same domain of fundamental movement skills that this chapter is concerned with. At that level of error, Kinect-class depth sensors cannot be defended for the granular kinematic analyses that are otherwise advertised as the principal advantage of markerless capture (Cai et al., 2019; Wochatz et al., 2019). Depth-sensor systems such as Kinect retain genuine value for high throughput screening and for engaging children in exergame based interventions (McGann et al., 2020), where binary classification (e.g. proficient versus not-yet-proficient) is sufficient to score against process-oriented criteria such as the TGMD-3 (Ulrich, 2013). However, depth-sensor systems should not be advanced as a substitute for marker-based motion capture where the analytic question concerns joint-level kinematics (Lima et al., 2024). For this reason, the stronger claims advanced in this chapter rest not on depth sensors but on a more recent class of markerless tools: open-source, video-based pipelines such as '*FreeMoCap'* and OpenCap. The supporting paediatric evidence base for these platforms, although still emerging, is more credible than that for depth sensors.

5.7 Emerging Open-Source Markerless Platforms

Recently, a suite of markerless motion capture systems has gained momentum: open-source platforms that leverage bespoke video cameras (e.g. smartphones or webcams) and advanced computer vision algorithms to track movement. These systems aim to combine the accessibility of consumer devices with the sophistication of lab analyses, all within freely available software. For instance, '*FreeMoCap'* (Matthis et al., 2022) is developing a user-friendly pipeline, capable of reconstructing 3D motion from multiple video cameras using algorithms like OpenPose for pose detection (Cao et al., 2021). This illustrates the research-grade potential of open-source frameworks, with tasks that once required an expensive motion lab now analysable with minimal cost.

Similarly, OpenCap (Uhlrich et al., 2023) takes advantage of the ubiquity of smartphones. In OpenCap's design, two or more phone cameras record an individual performing movements (e.g. running, jumping) from different angles; the videos are uploaded to a cloud platform where the software computes 3D kinematics and even performs biomechanical simulations to estimate internal forces. OpenCap is fully open-source and operates via a web application so users need no specialised hardware or local software beyond a camera and internet connection. Recent studies have demonstrated OpenCap's viability: in a 2024 study, researchers found OpenCap performed comparable to a reference 10-camera marker-based system for hip, knee, and ankle sagittal plane motion across all movements with very strong similarity (Turner et al., 2024). In an adult sport orthopaedic field trial of 100 participants, Uhlrich et al. (2023) reported that a clinician using OpenCap could obtain key biomechanical measures approximately 25 times faster and at less than 1% of the cost of an equivalent laboratory assessment. It must be acknowledged however that this figure relates to a return to sport workflow with adults rather than paediatric MC assessment, and should therefore be interpreted as an indication of the potential upper limit of cost and time efficiencies achievable through markerless motion capture, rather than as direct evidence supporting paediatric application (Uhlrich et al., 2023).

Open-source markerless systems bring additional benefits in terms of transparency and adaptability. Due to their open-source nature, code and models are readily available meaning researchers can scrutinise the adopted algorithms and subsequently contribute future improvements. This collaborative development model allows the tools to evolve quickly in

response to user feedback and new scientific needs. For example, if an open-source system has issues recognising a particular skill such as ‘galloping’, developers or users can retrain the model or add a module to improve it. Similarly, open platforms can be configured to address privacy concerns, for instance, analyses can be completed offline or anonymising participant’s faces in video data using a medianBlur function within a haar cascade classifier (Cuimei et al., 2017; Viola & Jones, 2001) for sensitive settings, which is harder to control in closed commercial software.

Such an open ethos approach also facilitates reproducible research; researchers can share their motion capture data and analysis pipelines, since the software is accessible to all, making it easier for other groups to verify findings or perform meta-analyses. These are distinct advantages over proprietary systems, which typically operate as black boxes (Bossavit & Arnedillo-Sánchez, 2022). It should be noted, however, that open-source motion capture is still relatively new in our domain and thus remains under continuous validation. Early adopters therefore must exercise a degree of caution when interpreting data until more studies compare these tools against gold-standard measures. While initial results are promising, there are numerous areas requiring improvement which have been highlighted; all MCT are affected by some level of soft tissue artifact, due to the mobility of participant’s skin, muscle mass or clothing (Motti Ader & Bossavit, 2023). Additionally, issues such as signal interference between different equipment, occlusion or reflection as a byproduct of the participants’ natural movements or objects in the capture environment can negatively affect data capture quality (Motti Ader & Bossavit, 2023). Usability is another consideration; at present, configuring a system like ‘*FreeMoCap’* can require some technical skill (e.g. calibrating multiple cameras, managing large data files), though efforts are underway to streamline this (Matthis et al., 2022). In summary, markerless motion capture systems now span a continuum from laboratory-grade (but impractical) optical rigs, through intermediate solutions like Kinect, to cutting-edge open-source platforms. Together, they offer a toolkit that can potentially overcome many limitations of traditional assessments, as we discuss next.

5.8 Benefits of Markerless Motion Capture for MC Assessment

Adopting markerless motion capture in motor development research facilitates opportunity to measure skills more objectively, in greater detail, and at larger scales than previously possible. In particular, these technologies address the core issues of scalability, objectivity, data

richness, and accessibility that have constrained traditional methods:

- **Scalability and Field Deployment:** Markerless systems can dramatically increase the scale and ecological validity of assessments. Markerless setups using portable cameras or sensors allow testing to move outside the lab into schools, playgrounds, and homes. Researchers can effectively take the lab to the field, capturing children's motor skills in natural environments during physical education classes or free play. This is crucial for interpreting real-world motor behaviour. For example, Chow et al. (2020) showed that an inexperienced assessor using a Kinect-based system (aligned with TGMD-2 criteria) would be capable of assessing children's locomotor skills on the spot, with automated scoring that agreed well with expert human ratings (percentage agreement; 43.6% to 92.9%). Similarly, McGann et al. (2020) implemented a Kinect-driven exergame platform in a school setting with children playing motion-controlled games while evaluating a subset of FMS (hop, jump, and skip; based on TGMD-2 criteria) in real time, providing instant feedback on proficiency. These examples support the potential for a future in which every school could have an interactive motor skill station, or every community centre a simple camera setup for screenings – tasks impossible to scale with manual testing or fragile laboratory equipment. With markerless systems capable of efficiently assessing a high volume of participants (especially video-based systems that can be reviewed at a later stage or are capable of running multiple cameras), they enable high-throughput assessment. Indeed, a single researcher using a few cameras might capture data from dozens of children in a single afternoon, something not possible using traditional observational process-oriented batteries such as the TGMD-3, KTK or MABC-2. Markerless systems therefore have great potential to enable significantly superior sample sizes and more frequent monitoring with much reduced researcher burden, which would also improve statistical power to track developmental trends and evaluate interventions.

- **Objectivity and Rich Data:** By relying on sensors and algorithms, motion capture provides inherently objective measurements, reducing the biases of human error during observation (Duncan, Martins, et al., 2022). Process-oriented batteries such as the TGMD-3 and MABC-2 operationalise MC against theoretically derived movement-form criteria. Replacing those criteria with continuous kinematic variables is not a like

for like upgrade, but a different measurement model with its own validity claims to defend (Hulteen et al., 2023). Whether kinematic features map onto, supersede, or simply describe the existing process criteria is an open empirical question, not a settled one. The position taken in this thesis is that finer grained data are not in themselves better measures of MC. Their value depends on whether the kinematic features extracted are theoretically defensible indicators of motor proficiency, and on whether they predict the developmental and health outcomes that MC assessments are intended to inform (Barnett et al., 2022; Stodden et al., 2008). Markerless motion capture is therefore framed in this chapter not as a replacement for process-oriented scoring but as a complementary measurement model whose construct validity must itself be established. Instead of a binary checklist (pass/fail on a few criteria), these systems yield continuous kinematic variables, joint angles, velocities, trajectories, centre-of-mass movement, etc., offering a more nuanced picture of performance. This data richness allows us to analyse how a movement is performed, not just whether it was successful. It should be noted however that research has demonstrated granularity on determining component-level proficiency factors which determine proficiency levels (O' Brien et al., 2015, 2016).

Therefore, we don't suggest that traditional methods aren't capable of providing high-level feedback on proficiency but rather that this process could be significantly streamlined through markerless systems modelled specifically to TGMD performance criteria. For instance, a motion capture analysis might reveal that the child's take-off speed is low, or their knees do not fully extend during the jump; thus pinpointing specific motor deficits that a teacher or therapist can subsequently target over time. While such elements may be identified by assessors using traditional observational methods such as the TGMD, these elements may not be identified in real-time however when factors such as subjectivity in rater judgement and fatigue during real-time data collection are considered (Ng et al., 2025). Additionally, small improvements (e.g. a slightly smoother coordination pattern or increased range of motion) can be detected quantitatively, whereas they might be overlooked in coarse observational scoring when solely using the TGMD in a live-observational capacity. Additionally, because data collection is automated and digital, the risk of human recording error is minimised, eliminating mistakes tallying scores or illegible handwriting on score sheets. It's important to

acknowledge also that algorithms are not infallible, they can misidentify movements, especially if trained mostly on adult data (Froehle et al., 2013; Lanovaz et al., 2017). Unlike human error, algorithmic errors are systematic and can be debugged and corrected over time once they are identified. With dedicated training on child-specific movement datasets, (e.g., Guo et al., 2026; Lander et al., 2025) the accuracy of these systems will continue to improve.

- **Accessibility and Affordability:** Modern motion capture technologies (especially open-source and consumer-grade options) greatly lower the financial and expertise barriers to advanced process-oriented assessments. Traditional laboratory systems have been confined to well-funded institutions; in contrast, a school or community program with a modest budget can set up a few video cameras or a depth sensor and immediately begin capturing motor skill data. Open-source software is typically free to download, and many platforms run on commodity hardware like smartphones or tablets. This drastically reduces costs. For example, '*FreeMoCap'* and OpenCap have demonstrated that high-quality 3D kinematics can be obtained using a minimal cost set-up (Matthis et al., 2022; Uhlrich et al., 2023). The socioeconomic implications are therefore significant. Even under-resourced schools or clinical settings could adopt these tools, helping narrow gaps in access to process-oriented skill evaluation. Due to the inexpensive nature of these systems, assessments can be repeated frequently, which is ideal for tracking longitudinal progress or evaluating the effects of an intervention.

The open-source ethos further promotes accessibility by inviting a global community to contribute and localise the tools. For instance, an open platform can be adapted by users to different languages or cultural contexts, ensuring that the technology benefits diverse populations. The *'FreeMoCap'* project explicitly frames its mission in terms of equity and inclusion, aiming to facilitate people with no domain expertise or institutional backing to conduct high-quality motion analysis (Matthis et al., 2022). In academic terms, widespread adoption of open, low-cost systems could foster greater participation in research from low- and middle-income countries, enhancing the generalisability of findings in motor-skill development.

- **Reproducibility and Collaboration:** An often-underappreciated benefit of open-source markerless systems is the potential for open science and data sharing. Since many of these tools use common data formats and are transparent in their processing, researchers can more easily share motion capture datasets and even processing code. If multiple groups use the same open platform (say OpenCap or '*FreeMoCap'*) to collect data, their outputs are directly comparable and can be aggregated. It should be noted however that process-oriented data such as the TGMD has lent itself well to data sharing enabling impactful large-scale longitudinal research to be carried out between institutions (Chen et al., 2023; Hardy et al., 2013).

  By proposing the uptake of markerless systems, we suggest this would substantially enhance data-sharing between institutions, perhaps incorporating datasets that include a blend of traditional process-oriented and motion capture data. Such an innovation in data-sharing may pave the way for even greater collaborative studies and meta-analyses of motor-skill development that were previously infeasible. This optimism should, however, be tempered by an important caveat. A central strength of established process-oriented assessments such as the TGMD is their stability over time; because the scoring criteria have remained largely unchanged, data collected decades and continents apart remain directly comparable, and this has underpinned much of the large-scale longitudinal and cross-cultural evidence in the field (Chen et al., 2024; Hardy et al.,2013). Markerless systems do not yet offer this assurance. Pose-estimation algorithms, camera hardware, and processing pipelines are evolving rapidly, such that kinematic data captured and analysed today may not be directly comparable with data produced by a later generation of the same tools.

  Unless software versions are documented and controlled, the very pace of innovation that makes these systems attractive risks fragmenting the long term comparability that standardised observational assessments have long provided. It also accelerates methodological improvements: when one laboratory identifies a limitation (for example, a specific skill is consistently tracked poorly), they can communicate this to the developer community, who can then work on a fix or update, creating in effect a 'crowdsourced' improvement loop for measurement instruments. This communal approach can advance the field much faster than relying on slow updates from proprietary vendors. In the future, researchers could establish normative databases of

kinematic features for certain ages, against which any new dataset can be compared – greatly enhancing our ability to interpret the clinical significance of motion data. All of these collaborative benefits hinge on the adoption of open, interoperable systems that allow data to flow freely (Bossavit & Arnedillo-Sánchez, 2022).

In summary, markerless motion capture offers a compelling set of advantages for motor skill development research. It marries the scale and speed needed for public health monitoring with the depth of data needed for rigorous biomechanical insight. By overcoming many limitations of traditional tests (i.e., time investment and subjectivity) these technologies let us reimagine how and where we assess motor skill development. Children could be evaluated playing a game in their school playground with instant feedback, rather than in a traditional observational session with an assessor, typically separated months apart between test and re-test. Researchers can obtain objective trajectories and angles that may (criteria dependent) provide more impactful insight on criteria specific proficiency, enabling more sensitive measures of change. Importantly, these improvements do not come at exorbitant cost. In fact, many solutions are cheaper than the standard kit (such as stopwatches, balls, bats, scoring booklets) once used for process-oriented testing such as the TGMD. Whilst the many benefits of markerless motion capture have been discussed it is important to acknowledge scalability issues of applying this technology among children. The following section shall consider challenges associated with this in more detail.

5.9 Challenges in Applying Markerless Motion Capture to Paediatric Populations

While markerless motion capture systems undoubtedly offer clear advantages in scalability and accessibility, their application in paediatric populations presents distinct methodological challenges that must be acknowledged. Children differ fundamentally from adults in both movement behaviour and physical characteristics (Berard & Vallis, 2006), yet many pose estimation models underpinning markerless systems have been developed and trained predominantly on adult datasets (Clark et al., 2012; Kanko et al., 2021; Pfister et al., 2014; Tipton et al., 2019; Wochatz et al., 2019). As a result, these models may underfit and not generalise effectively to youth populations, raising important questions regarding validity and reliability in child-focused contexts.

From a technical perspective, children's smaller body segment dimensions can make joint detection and tracking more difficult, particularly when using standard-resolution video or depth sensors (Lerma & Gulgin, 2023). In addition, children tend to exhibit faster, more variable, and less predictable movement patterns than adults (Chihak et al., 2014), which can further challenge the temporal stability and accuracy of pose estimation algorithms (Sciortino et al., 2017). Anthropometric differences, such as altered limb proportions and joint relationships across developmental stages, also introduce complexity, as many existing biomechanical models are calibrated to adult body structures (Rogez et al., 2008). Collectively, these factors can lead to reduced precision in joint localisation and movement reconstruction, particularly during dynamic tasks that are common in MC assessments.

Sciortino et al. (2017) reported pose-estimation error rates on children's images that were roughly twice those obtained on adult benchmarks for equivalent algorithms. Lerma and Gulgin (2023) found that Kinect V2 estimates of children's vertical jump and leap kinematics agreed only poorly to moderately with Vicon, with correlations in the 0.5 to 0.7 range for several joint angles and absolute errors of several centimetres in segment positions. Subsequent work in deep-learning pose estimation has confirmed that adult-trained models underperform on paediatric video and that performance degrades further during fast or atypical movement (Groos et al., 2022; McCay et al., 2020). The sources of this kinematic data degradation should not be treated as a single problem; accordingly healthcare practitioners should not assume that AI methodology intended for adult use will perform to the same standard when implemented in a paediatric population without additional adaptation, particularly for children under 2 years of age (Laborie et al., 2026). Careful model evaluation is thus required before clinical implementation. Below we will consider some pose-estimation discrepancies that have been highlighted in the literature.

The first is anthropometric mismatch, with pose-estimation networks trained largely on adult body proportions, they tend to misplace joint centres on shorter limbs and proportionally larger heads, particularly in children younger than around eight years (Rogez et al., 2008). The second is movement-variability mismatch. Paediatric MC tasks routinely produce faster, less stereotyped, and more frequently corrected movement trajectories than the adult locomotion data on which most public pose-estimation datasets are based, and the temporal smoothing strategies developed for adult gait can themselves introduce systematic bias when applied to

paediatric data (Chihak et al., 2014; Sciortino et al., 2017). The third is behavioural compliance, with younger children likely to adhere to instructions for shorter periods of time than adults (Emanuel et al., 2008), which increases the proportion of trials lost to occlusion, framing failures, and protocol drift, particularly in field deployments outside laboratory conditions (Motti Ader & Bossavit, 2023). Lastly, ethics and data governance involving video data that depicts identifiable minors is subject to substantively stricter consent, storage, retention, and re-identification requirements than adult data under the General Data Protection Regulation and analogous frameworks elsewhere (Singh et al., 2023). These constraints can preclude the cloud-based processing pipelines on which several open-source platforms currently rely. The first three are technical problems amenable to dataset and modelling improvements. The fourth is a governance constraint that better algorithms alone will not solve.

A defensible adoption strategy therefore needs an explicit account of what “good enough” looks like for paediatric MC assessment, calibrated to the use case rather than to a single accuracy threshold. For population level screening or for engaging children in skill learning interventions, agreement at the level of broad proficiency categories (e.g. proficient, emerging, or not-yet-proficient) is plausibly sufficient, and the Kinect-class evidence base already supports this (Chow et al., 2020; McGann et al., 2020). For longitudinal tracking of within child change in MC, joint-angle errors of several centimetres are likely to exceed the developmental signal that needs to be detected, and the use of markerless systems for this purpose should be considered premature in the absence of paediatric-specific validation (Lerma & Gulgin, 2023). For clinical decisions concerning individual children, such as identifying motor coordination difficulties or monitoring response to therapy, agreement against a marker-based reference standard, in the same age range and on the same task, is the minimum acceptable evidence base, and remains rare in the published literature (Lerma & Gulgin, 2023).

Rather than negating the potential of markerless motion capture, the limitations discussed highlight a critical direction for future research. There is a clear need for the development and validation of child-specific datasets, models, and evaluation frameworks to ensure that these systems can accurately capture the nuances of paediatric movement (Lander et al., 2025). Addressing these challenges will be essential for realising the full potential of scalable, markerless assessment in childhood motor-skill development. In this sense, the aforementioned limitations should be viewed not as barriers, but as opportunities to refine and adapt emerging technologies to better serve diverse developmental populations.

5.10 Applications in Research, Education, and Clinical Practice

This final section of this chapter considers concrete applications of markerless motion capture across research and practice domains, illustrating how it can modernise motor-skill assessment.

Research and Discovery: The research community stands to benefit immediately from open-source markerless motion capture. With it's application, motion capture scientists and motor-skill behaviour researchers have the potential to collect a significantly greater volume of data, and from more diverse environments, than previously. Instead of being limited to small samples tested in controlled environments as has been consistently noted in prior literature (i.e., Adank et al., 2018; Bolger et al., 2021; Duncan et al., 2018; Hulteen et al., 2018; Sigmundsson & Haga, 2016), studies can deploy markerless capture systems to, for example, measure hundreds of children during school breaks, physical education classes or within small sports clubs, capturing proficiency levels in an ecologically valid setting. This potential has been recently demonstrated by Lander et al. (2025) in which a deep learning model produced accuracy exceeding 83% on average across seven TGMD-3 skills among a sample of over a thousand children. Such automated approaches may reveal swifter and more tangible insights into how motor skills develop and generalise within such large and diverse populations.

Longitudinal tracking studies would also become more feasible, which would be significant given a lack of longitudinal research has been regularly noted as a limitation in MC research (Barnett et al., 2022; Hands, 2008; Lloyd et al., 2014; Robinson et al., 2015; Romero-Martínez et al., 2026). From a practical perspective, a research team could loan families an inexpensive camera kit and have them periodically record their child's movement at home, building rich developmental trajectories. Indeed, this exact scenario has been demonstrated to be a successful approach through the mobile application *'MoveAhead'* which through gamification offers a fast and accurate appraisal of a child's motor-skill proficiency (Herbaut et al., 2023). Furthermore, precise tracking allows investigation of fine-grained learning processes, researchers can measure subtle changes in movement pattern or consistency as a skill is gradually honed, rather than just pre/post outcomes. This level of granularity can help test and refine long-lasting theories (e.g. the mechanics of motor-skill acquisition; Haubenstricker & Seefeldt (1986)) or the proposed links between MC proficiency and physical activity engagement (Stodden et al., 2008) with far greater nuance than before.

Educational and Community Settings: Schools and youth programmes could use markerless motion capture to enhance both assessment and instruction of motor skills. In practice, brief motor skill stations could be integrated into physical education classes, with children performing a small set of FMS (for example, a short run, a throw, and a horizontal jump) in front of a camera setup that scores their performance against age norms in close to real time. This kind of functionality has been demonstrated by *'MoveAhead'* (Herbaut et al., 2023). Because the scoring is automated and standardised, the system could in principle be run by school teachers with minimal training, in contrast to traditional motor-skill assessments that require specialist examiners and one-to-one testing time. The interactive nature of camera-based assessment also tends to be more engaging for children than traditional tests (Kou et al., 2024; McGann et al., 2020; Oppici et al., 2022; Ye et al., 2018) . As McGann et al. (2020) demonstrated, turning motor-skill practice into an exergame keeps children motivated while simultaneously teaching children the requiste movement patterns to complete motor skills competently. The adoption of a Kinect or camera-based game can reward correct movement technique (e.g. only a properly executed jump earns points) and provide cues to improve, effectively blending teaching with evaluation (McGann et al., 2020). However, these exergame approaches are ultimately coaching tools as opposed to assessment tools which must be considered prior to implementation by practitioners. The approach has obvious appeal in resource-limited schools where large class sizes make individual assessment difficult, however, these camera-based systems should be regarded as untested rather than as established. Community sports programmes can use the same kinds of tools to assess and train foundational skills such as running form and jumping mechanics in young athletes, with coaching directed at measured deficits rather than rough observation (Shah & Prajapati, 2025), Overall, markerless motion capture can help educational and community organisations move towards continuous, objective monitoring of motor-skill development, making it a routine part of youth development programs rather than an occasional specialised test.

Clinical and Therapeutic Practice: In clinical domains, such as paediatric physical therapy and rehabilitation, markerless motion capture offers a means to obtain quantitative movement assessments efficiently (Lam et al., 2023). A physiotherapist evaluating a child with developmental coordination disorder (DCD) (Wood et al., 2017) or cerebral palsy (Daliri et al., 2023) could use a camera-based system to capture a short sequence of movements during a session. Provided the validation evidence supports it for the joints and tasks of interest, this could complement qualitative observation and parent reports with objective metrics such as

joint range of motion, balance sway, and gait symmetry, and can support tracking of change over time. The strongest claim that can currently be made is for tracking within-child change in already-diagnosed paediatric conditions, where the validation requirements are typically less stringent than for diagnostic decisions (Temlali et al., 2024). Whilst Temlali et al. (2024) found limited related research, findings suggested that physical literacy could be regarded as an explanatory framework to understand the way in which exergames can improve PA behaviour in adolescents with DCD. Due to the simplicity of markerless systems such measurements can be taken at each visit without much overhead, allowing more frequent and sensitive tracking of progress.

Telehealth applications are also promising: with an open-source software app, a parent could video their child performing certain exercises at home and send the data to a clinician for analysis, reducing the need for travel to appointments (Bossavit & Arnedillo-Sánchez, 2022). This is particularly valuable for families in remote areas or for children who find clinical environments stressful (De Nocker & Toolan, 2023). In rehabilitation, motion capture can document outcomes of therapy with much greater precision than standard observational scales (Bonanno et al., 2023). A therapist can highlight quantitative improvements, outcomes that might not be fully captured by a likert scale (Bonanno et al., 2023). Such objective evidence could hypothetically guide clinical decisions and also help motivate patients and families by visualising improvement (Rivas-García et al., 2026). Early screening is sometimes proposed as a further application, on the basis that markerless motion capture might flag motor difficulties such as DCD or early motor signs of neurological conditions before in-person specialist assessment. The evidence supporting this proposal is, at present, limited; existing systematic reviews (Knippenberg et al., 2017; Lam et al., 2023) indicate emerging diagnostic potential for paediatric movement disorders but stop short of recommending markerless capture as a stand-alone screening tool, and any such use would need diagnostic-accuracy evidence in the relevant age range that does not currently exist (Lam et al., 2023).

Conclusion

Markerless motion capture technology, encompassing everything from gold-standard marker-based optical systems (i.e., Vicon) to portable depth cameras (i.e., Microsoft-Kinect) and open-source video-based platforms (Lander et al., 2025), presents a significant opportunity for reimagining assessment of MC among paediatric populations. These tools directly address many limitations of traditional MC assessments by enabling scalable, objective, and data-rich

analysis of movement. Early applications across research, education, and clinical settings demonstrate their potential to transform how MC is measured, moving beyond coarse observational scoring toward detailed, evolving representations of movement behaviour. In order to fully achieve this potential however, careful and deliberate progression of research within the field must be achieved. To move from promise to practice, a clear and targeted research agenda must guide the development and adoption of markerless motion capture in paediatric contexts. Four priority areas are detailed below.

Firstly, there is a pressing need for rigorous validation studies in children (Lander et al., 2025). While many systems demonstrate encouraging performance in adult populations, their accuracy, reliability, and sensitivity must be systematically evaluated across developmental stages, movement tasks, and real-world environments. Establishing robust paediatric validation will be essential for ensuring that these technologies can be confidently applied in motor-skill development research and practice.

Secondly, the development of standardised assessment protocols is essential for reliability, repeatability, and validity of conclusions derived within and between assessment protocols (Ng et al., 2025). Current approaches vary considerably in camera configurations, data processing pipelines, and outcome measures, limiting comparability across studies. The establishment of agreed methodological guidelines, including, but not limited to; experimental design, data collection, processing, and reporting of findings would support both reproducibility and the accumulation of comparable datasets across research groups. Crucially, such guidelines should also encompass transparent reporting and version control of the algorithms, models, and software used, so that data captured with one generation of tools can be meaningfully compared with data produced by later iterations, thereby preserving the long-term comparability that standardised observational assessments currently afford.

Thirdly, ethical and data governance considerations must be central to the adoption of video-based motion capture in childhood populations (Singh et al., 2023). The collection of visual data involving children raises important challenges related to informed consent, secure data storage, and privacy protection. Compliance with regulatory frameworks such as GDPR is essential, alongside proactive mitigation of risks associated with re-identification from video or derived skeletal data. Whilst many techniques exist that can be utilised to overcome such hurdles (i.e., face-blur algorithms; Cuimei et al., 2017; Viola & Jones, 2001) developing clear

ethical standards and robust data handling practices is critical to maintaining public trust and safeguarding participant welfare.

Finally, there is a need to establish clear criteria for sufficient validation prior to widespread clinical or educational deployment (Motti Ader & Bossavit, 2023). This includes defining acceptable thresholds for measurement error, reliability, and agreement with benchmark systems, as well as demonstrating sensitivity to meaningful changes in MC proficiency over time. As discussed earlier, collaborations between practitioners will make a consensus view achievable. Without such criteria, the translation of markerless systems into practice risks being premature or inconsistent. Articulating these benchmarks will help guide both technological innovation and evidence-based implementation.

In summary, the discussed priorities aim to provide a roadmap for advancing the motor competence domain beyond proof-of-concept applications toward robust, scalable, and ethically grounded assessment frameworks. Embracing markerless motion capture is not simply a technological shift among MC practitioners, but a methodological one, requiring coordinated efforts in validation, standardisation, and governance. Should these challenges be addressed, such systems have the potential to democratise access to high-quality MC assessment, enabling more inclusive, frequent, and precise monitoring of children's MC development. In doing so, they may play a critical role in supporting healthier developmental trajectories and more equitable opportunities for physical activity and wellbeing across populations.

Chapter 6: Evaluation of a Markerless 3D Motion Capture System for Locomotor Skill Assessment: A Pilot Study

6.1 Abstract

The Test of Gross Motor Development-3 (TGMD-3) is widely regarded as a standardised assessment of motor competence (MC); however, its implementation can be time-intensive and dependent on trained assessors. Markerless motion-capture technologies offer a potential means of automating assessment, but the extent to which low-cost, open-source systems can support TGMD-3 classification remains unclear. Therefore, the purpose of this pilot study was to investigate the feasibility of combining an open-source markerless motion-capture pipeline with handcrafted heuristic rules to automate classification of TGMD-3 short form locomotor performance criteria. A secondary aim was to examine the influence of camera configuration on kinematic reconstruction and classification performance. Adults (n = 10) performed two trials of the hop, gallop, and horizontal jump using a four-camera webcam setup. The resulting recordings were processed using '*FreeMoCap'*, with data subsequently analysed under four-, three-, and two-camera configurations. Variations between configurations were examined using Friedman repeated-measures ANOVAs, with Durbin–Conover post-hoc pairwise comparisons where appropriate. Agreement between configurations was assessed using ICC(2,1), while classification performance was evaluated using accuracy, precision, recall, and F1-scores.

Results indicated that reducing the system from four to three cameras had minimal impact on kinematic data capture quality, whereas larger discrepancies were observed under the two-camera configuration. Despite these differences, overall classification performance remained relatively stable across configurations. Feature-level analyses revealed substantial variability between individual TGMD-3 criteria, with performance criteria linked to discrete kinematic events generally outperforming those reliant on subjective interpretations of movement quality or coordination. These findings demonstrate the feasibility of using low-cost, open-source markerless motion-capture technologies to support automated MC assessment. However, the results also suggest that future advances may depend not only on improvements in motion-capture and computational systems, but also on the continued refinement and operationalisation of TGMD performance criteria into objective and measurable kinematic constructs.

Key Words: Motor Competence; Markerless Motion Capture; Computer Vision

6.2 Introduction

The relationship between motor competence (MC) and physical activity (PA) is now well established, with numerous systematic reviews and meta-analyses reporting a consistent positive association (Burton et al., 2023; Logan et al., 2015). MC is considered an umbrella term for the motor coordination, movement and stability skills needed to effectively engage in PA (Robinson et al., 2015). The acquisition of superior proficiency levels of MC is associated with numerous health benefits, including improved physical fitness, perceived MC and healthy weight status (Barnett et al., 2022). Globally, it has been widely observed that children and adolescents do not achieve age-appropriate proficiency standards (Bolger et al., 2021; Duncan et al., 2020; Hardy et al., 2013). In light of this, assessments have been developed to identify developmental delays, inform interventions and quantify levels of MC proficiency (Bardid et al., 2019; Barnett et al., 2016; Hulteen et al., 2015).

MC assessments are typically presented as either product- or process-oriented. Product scores are quantitative outcomes (e.g., throwing velocity, distance hopped) and don't analyse the process of how the skill is performed. In comparison, process-oriented assessments rely on set performance criteria of which are either present or absent during performance of a skill (Hulteen et al., 2015). Process-oriented assessments therefore provide practitioners with a more nuanced understanding of a participant's skill execution (Burton & Miller, 1998). However, to rate process-oriented assessments, sufficient training is required as scoring can differ depending on the level of an assessor's expertise (Griffiths et al., 2018). Indeed, a multitude of MC assessments exist (Cools et al., 2009), and still more continue to be developed for various purposes and populations. Despite the abundance of available assessments there is minimal evidence to suggest that teachers or coaches assess children's MC proficiency systematically (Duncan, Weldon, et al., 2022; Eddy et al., 2021; Lander et al., 2015). Limited feasibility and utility in the school and youth sport setting is a key barrier to teacher's or coach's use of existing assessment tools, where assessments need to be swift and efficient, only requiring minimal equipment, resources and space (Duncan, Weldon, et al., 2022; Eddy et al., 2021; Lander et al., 2022). Systematically there is a consensus view among researchers that assessment batteries are too time consuming (Hulteen et al., 2020) and therefore approaches to assessment must be re-imagined.

The Tests of Gross Motor Development 3rd edition (TGMD-3) (Ulrich, 2013) has been noted for its application of process-oriented performance criteria to successfully guide impactful age-appropriate MC interventions (Belton et al., 2014; O' Brien et al., 2015, 2016). Despite such successful outcomes, recent research has sought to streamline the assessment battery with (Duncan, Martins, et al., 2022) proposing a modified short form version. Such research demonstrated that a shortened, seven-skill version of the TGMD-3, in comparison to the original thirteen-skill version, is a valid measure of FMS among Irish primary school children, while also highlighting differences in skill patterns between boys and girls. Whilst such findings are promising there remains greater scope for assessment batteries to be streamlined further. Of note, technological advancements have been demonstrated in more recent MC assessment options (Bossavit & Arnedillo-Sánchez, 2022; Chow et al., 2020; Guo et al., 2026; Lander et al., 2025; Lerma & Gulgin, 2023; Maulidiya et al., 2025; McGann et al., 2020). Such research proposes automated assessment through markerless motion capture, an advancement from sensor wear technology, which has been used extensively to detect various aspects of movement proficiency (Clark et al., 2021; Grimpampi et al., 2016; Lander et al., 2020; Swain et al., 2023). Whilst noted as being both valid and reliable for human movement detection (Swain et al., 2023) the same review also highlighted that such a methodology was unlikely to be adopted at scale among practitioners. Feasibility is limited by practical challenges such as the number of sensors required, time for setup and removal, calibration demands, and sensitive placement, alongside limitations in capturing all aspects of performance (e.g., detecting a successful catch) and difficulties for teachers or coaches in accessing and using the data (Barnett et al., 2013; Lander et al., 2022; Swain et al., 2023).

Recent research highlighting the adoption of markerless motion capture (via computer vision) (Bossavit & Arnedillo-Sánchez, 2022; Chow et al., 2020; Guo et al., 2026; Harsted et al., 2019; Lander et al., 2025; Maulidiya et al., 2025) provides a more feasible, cost effective and scalable solution. Whilst high-precision optical systems such as Vicon have long been considered the gold standard for motion analysis (Topley & Richards, 2020) more recently a spectrum of markerless motion capture approaches have emerged with consumer-grade depth sensors and open-source camera-based tools capable of recording movement data without body-worn markers or suits (Bossavit & Arnedillo-Sánchez, 2023). In comparison to the gold-standard approach which is impractical for assessment at scale (Lerma & Gulgin, 2023), markerless systems offer numerous benefits. A full optical setup is exceedingly expensive and requires a

dedicated, controlled space (Van Der Kruk & Reijne, 2018). Setting up a child for capture using non-optical systems such as IMU's can involve attaching dozens of markers and frequent calibrations (Bossavit & Arnedillo-Sánchez, 2023). While this may be feasible for indepth studies on small samples, but not for testing hundreds of children in a school environment.

Although the TGMD-3 is widely regarded as the gold-standard process-oriented assessment of MC (Behan et al., 2019; Belton et al., 2014; Duncan et al., 2020; Eyre et al., 2018; O' Brien et al., 2016), administering the full assessment remains resource intensive (Barnett et al., 2013). Consequently, the TGMD-3 Short Form (Duncan, Martins, et al., 2022) has emerged as a promising alternative, offering a more time-efficient assessment while retaining representation of the broader locomotor and object-control skill domains. The reduced complexity of the TGMD-3 short form also makes it a more practical candidate for automated evaluation. In parallel, increasing attention has been directed towards markerless motion capture technologies, which offer the potential to objectively assess MC proficiency without the need for markers, sensors, or specialist laboratory facilities. Despite these technological advancements, researchers have recently argued that several TGMD performance criteria remain reliant on subjective terminology and lack explicit quantitative thresholds (Lander et al., 2020; Ng et al., 2025). Consequently, the successful automation of MC assessment may depend not only on advances in motion-capture or data-centered technologies, but also on the extent to which existing assessment criteria can be operationalised as objective and measurable movement constructs.

Consequently, combining the TGMD-3 Short Form with markerless motion capture may provide a feasible pathway towards more objective, efficient, and scalable MC assessment. Markerless systems aren't without limitations however and are sensitive to errors arising from motion blur, self-occlusion and temporary loss of joint visibility, particularly during dynamic locomotor movements (Adlou et al., 2025; Needham et al., 2021; Scataglini et al., 2024; Motti Ader & Bossavit, 2023). These issues may be exacerbated when MC skills are captured from a limited number of viewpoints, potentially impacting upon the accuracy of the derived kinematic measures (Yang & Jung, 2025). As TGMD-3 performance criteria are dependent upon the correct identification of process-oriented criteria, any degradation in kinematic data quality may ultimately influence subsequent process-oriented proficiency classifications. Therefore, understanding the influence of camera configuration on both capture quality of

kinematic data and classification accuracy of process-oriented performance criteria represents an important consideration for real-world deployment. Such evidence may provide preliminary evidence to practitioners regarding the feasibility of different camera configurations when designing markerless motion-capture systems for MC assessment. Such work may guide future large-scale validation research in examining how camera configuration influences markerless motion-capture outputs and their potential application within MC assessment.

While previous studies have demonstrated the potential of markerless systems and automated approaches to assess MC proficiency (Chow et al., 2020; Lander et al., 2025), little is known about how camera configuration influences the quality of kinematic data and the resulting classification outcomes. Furthermore, few studies have evaluated whether a fully open-source, multi-camera markerless motion-capture pipeline can support TGMD-style assessment while balancing reproducibility of kinematic data, classification performance, and practical deployment considerations, notably camera configuration and efficiency of data collection. Accordingly, the present study explored the feasibility of integrating '*FreeMoCap'*-derived kinematic data with a heuristics-based rules engine under differing camera configurations. Specifically, the study aimed to;

(1) Examine the feasibility of operationalising TGMD-3 Short Form locomotor performance criteria within an open-source markerless motion capture pipeline, and assess the extent to which its automated decisions agree with expert human scoring in an adult proof-of-concept sample;

(2) Uncover the minimal viable camera configuration required for accurate classification by an open-source markerless motion capture system.

6.3 Methodology

The current study aimed to evaluate the performance of a hand-crafted rules-based approach using bespoke heuristic algorithms (Kokash, 2005) to assess adult's (n = 10) locomotor skill proficiency per TGMD-3 performance criteria (Ulrich, 2013). We sought a heuristic based approach for the current study given their efficient and practical nature to solving complex problems by using rules of thumb rather than exhaustive searches. While heuristics may not guarantee the optimal answer, they significantly reduce computational time and are especially useful in real-world situations where efficiency matters more than perfection (Kokash, 2005). In our use-case, given certain parameters of the TGMD-3 performance criteria are somewhat objective (i.e., *'brief period where both feet are off the surface'* (Ulrich, 2013)) heuristics was deemed as an appropriate technique to effectively discriminate proficiency per the listed criteria (Ulrich, 2013). Data was collected via *'FreeMoCap'* (Matthis et al., 2022), an open-source 3D computer vision system capable of integrating estimations from multiple cameras into estimates of segment locations with consumer cameras. *'FreeMoCap'* was used to discriminate performance in the TGMD-3 short form which has been demonstrated as a valid measure of FMS among Irish school children (Duncan, Martins, et al., 2022). A limitation of the current '*FreeMoCap'* pipeline was its inability to track external objects (e.g., balls). Consequently, despite the research team's intention to assess the full TGMD-3 Short Form, evaluation was restricted to the locomotor subset. Prior to the onset of data collection, ethical approval was received (Appendix A1) for this research (DCUREC/2024/243).

Recruitment

A convenience sample of participants were sourced internally within the research team's university school (n = 10). Whilst only a modest number of participants were recruited, such a sample is in line with prior locomotion-focused markerless motion capture validation studies (Kanko et al., 2021; Lei et al., 2026; Martiš et al., 2024; Turner et al., 2024; Yang & Jung, 2025). Furthermore, as the goal of the pilot-study was to demonstrate evidence of feasibility in a controlled-environment, the participant sample was deemed sufficient to demonstrate sensitivity to performance-criteria data. All participants were required to be of sound physical health and not demonstrate any physical injuries or ailments that would preclude them from achieving full-range physical motion to complete all skills within the locomotor subset of the

TGMD-3 performance criteria (Ulrich, 2013). All participants were extensively briefed on the aims and objectives of the research project, with the risks and benefits of participating also included. All participants were informed of their right to exit the study at any time without question and were given a full overview of the data protection policies in place to protect confidentiality through the course of the research.

Locomotor Subset - TGMD-3 Short Form

Participants performed the locomotor subset of the TGMD-3 short-form. Performance in these skills were evaluated via the process-oriented TGMD-3 (Ulrich, 2013), an internationally recognized motor skill assessment for children aged 3–10 years (Ulrich, 2017; Webster & Ulrich, 2017). All trials were conducted within the biomechanics laboratory of the research team's university. Participants performed all three skills of the locomotor subset of the TGMD-3 short-form. For each locomotor skill included in the TGMD-3 Short Form, performance is assessed against four predefined performance criteria, with each criterion evaluated in sequence. The skills and their associated performance criteria are listed below in Table 6.1. Participants who successfully fulfilled the component criteria for any skill achieved a score of one whilst a score of zero denotes that they failed to meet said criteria (Ulrich, 2013).

Table 6.1. TGMD-3 Short Form (Duncan, Martins, et al., 2022) - Locomotor Subset

Skill	Performance Criteria 1	Performance Criteria 2	Performance Criteria 3	Performance Criteria 4
Hop	Non-hopping leg swings forward in pendular fashion to produce force	Foot of non-hopping leg remains behind hopping leg (does not cross in front of)	Arms flex and swing forward to produce force	Hops four consecutive times on the preferred foot before stopping
Gallop	Arms flexed and swinging forward	A step forward with lead foot followed with the trailing foot landing beside or a little behind the lead foot (not in front of the lead foot)	Brief period where both feet come off the surface	Maintains a rhythmic pattern for four consecutive gallops

Horizontal Jump	Prior to take off both knees are flexed and arms are extended behind the back	Arms extend forcefully forward and upward reaching above the head	Both feet come off the floor together and land together	Both arms are forced downward during landing

Procedures

Experimental design for data collection is demonstrated in Figure 6.1.

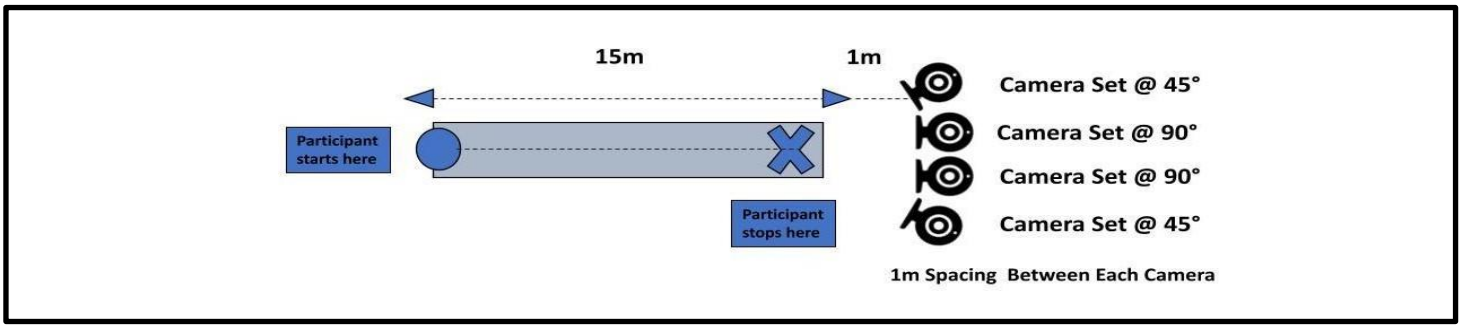


Figure 6.1 Experimental Design

As the *'FreeMoCap'* pipeline was only capable of tracking the kinematic joint data of a single individual at any given time, data collection was conducted on an individual participant basis. The data collection process for each participant therefore took ~10 minutes. Prior to the onset of the data collection participants were given an extensive verbal explanation of the skills they would be required to perform alongside a visual demonstration and one practice trial at each skill. Once each participant had been afforded a trial effort, they were then given two individual trials for each skill.

Participants performed each skill along a marked 15m horizontal walkway (Figure 6.1) which was highlighted to them prior to the onset of data collection. At the end of the walkway lay four mounted USB webcameras ('LOGITECH Pro C920 Full HD Webcam', 1920 x 1080 progressively displayed pixels, sampling rate; 30 frames per second (FPS), all mounted at the same height of 1.6m. All cameras were positioned directly facing participants with two cameras positioned orthogonally (90°) to the movement pathway and supplemented by two oblique cameras positioned at approximately 45°, this configuration is demonstrated in Figure 6.1. The camera arrangement was designed to maximise participant visibility across multiple viewing angles while maintaining a practical and scalable setup using consumer-grade hardware. This

configuration was intended to minimise joint occlusion and improve triangulation accuracy during data capture, where upper- and lower-limb segments may be temporarily obscured from individual camera perspectives. The arrangement was informed by '*FreeMoCap'* developer recommendations (Matthis et al., 2022) which advocate distributing cameras around the capture volume to maximise overlapping fields of view, and by previous markerless motion-capture research demonstrating that reconstruction accuracy is highly dependent on camera viewpoint and participant visibility. For example, Kim et al. (2025) positioned cameras at 45° intervals around participants to evaluate movement from multiple directions and maximise visibility of anatomical landmarks throughout the capture period. Consistent with these recommendations, the use of multiple viewpoints in the present study provided camera redundancy and enhanced spatial coverage of the movement area, thereby supporting robust three-dimensional data capture. In total, 60 locomotor trials were collected (10 participants x 3 skills x 2 trials). Each trial was subsequently processed under three camera configurations (4-, 3-, and 2-camera), resulting in 180 kinematic data captures which were analysed.

Data Processing

Prior to data collection, camera calibration was first conducted. Calibrations were required to establish the spatial relationship between cameras, enabling corresponding joint locations to be identified in each 2D video and subsequent triangulation to estimate participants position in 3-D space, allowing full-body movement trajectories to be reconstructed.

Calibrations were completed in which the lead author held a 7x5 ChArUco board (Figure 6.2) directly in front of all cameras with this captured footage then saved and analysed within '*FreeMoCap'*. A single calibration was sufficient for full data capture of all of locomotor movements across participants and took ~90 seconds to complete. However, a new calibration was required for each separate day of data collection. Video data were captured using the *SkellyCam* application, a companion recording tool developed for the '*FreeMoCap'* pipeline. *SkellyCam* was used to simultaneously record synchronized videos from multiple webcams, ensuring that all camera streams represented the same moment in time. The resulting videos were then imported into '*FreeMoCap'*, where pose estimation, camera calibration, and 3D reconstruction processes were applied to generate full-body kinematic data for subsequent analysis.

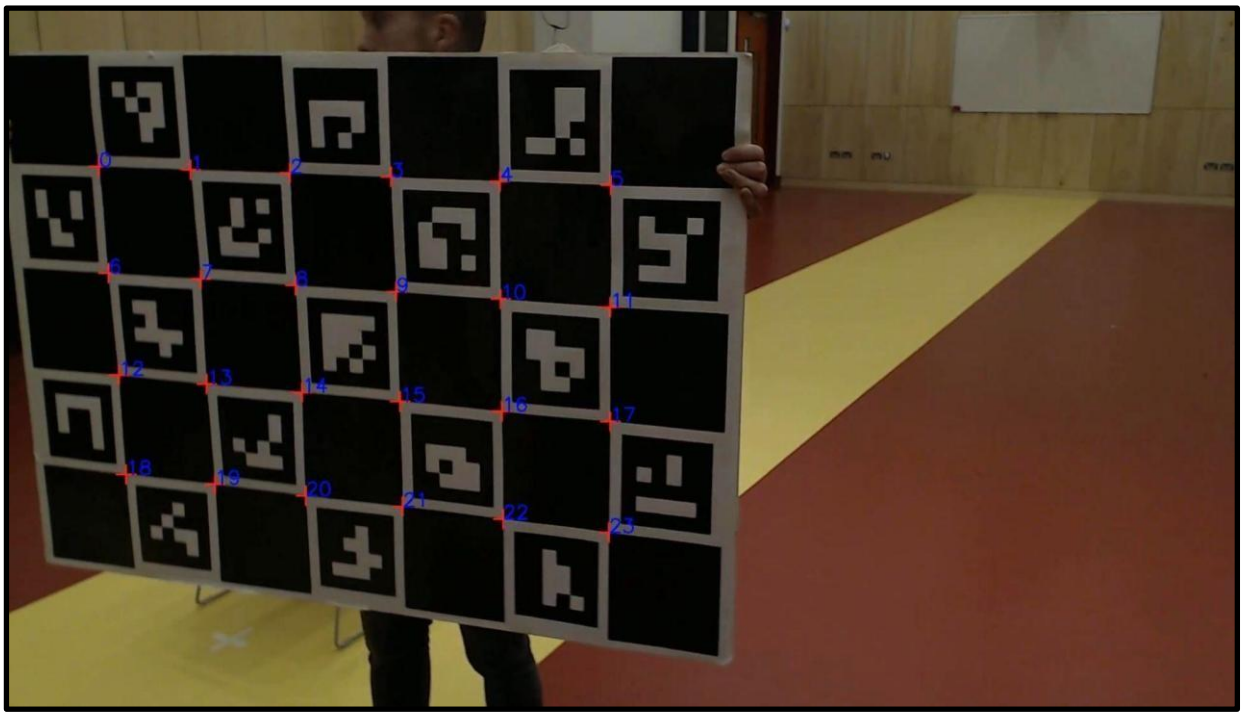


Figure 6.2 Calibration of Cameras with ChArUco Board

Rules-Based Heuristics

Given the current study was of a proof-of-concept nature, heuristics rules-based algorithims were considered appropriate as they represent a transparent and computationally efficient means of operationalising TGMD-3 performance criteria (Kokash, 2005). Algorithms were thus designed to be criterion-specific, reflecting discrete elements of TGMD-3 short form performance criteria (Table 6.1). Using Python, feature engineering was completed in which handcrafted features that encompassed distinct elements of these criteria were created with a list of these featured highlighted in Table 6.2 and a further list of criterion conditions shown in Appendix B. In creating these hand-crafted features, relevant joint-angle thresholds were informed by prior research (Chow et al., 2020), which successfully operationalised TGMD locomotor performance criteria using kinematic measures. The feature-engineering process followed a criterion-driven approach whereby each short-form performance criterion was decomposed into one or more measurable kinematic components. Using reconstructed 3D joint data coordinates, features such as elbow and knee flexion angles, wrist and heel positions, peak movement timing, foot-off-ground events, limb synchronisation, and relative segment

positioning were calculated. These features were subsequently compared against predefined thresholds (Chow et al., 2020) and temporal windows to create objective pass/fail rules for each performance criterion, forming the basis of the heuristics-based classification engine.

Table 6.2 Features Representing TGMD-3 Performance Criteria

Handcrafted Feature	Associated Performance Criteria Definition (Ulrich, 2013)	Example of Analysis
'lift_window'	*"Brief period where both feet are off the surface"* Skills; Gallop & Hop	*Metric; Stepping foot heel rise* *What's Compared; `body_[lead]_heel_z` vs baseline'* *Pass Condition; Heel rises above baseline within 2-second window*
'both_feet_off_surface'	*"Brief period where both feet are off the surface"* Skills; Gallop & Hop	*Metric; Heel rise above floor* *What's Compared; `body_[lr]_heel_z` vs within-window floor'* *Pass Condition; Both heels > 0.02 m above within-window minimum*
'non_dominnant_leg_movement'	*"Non-hopping leg swings forward in pendular fashion to produce force"* Skill; Hop	*Metric; Non-dom heel peak timing* *What's Compared; `argmax(body_[non_dom]_heel_z)` vs `peak_frame`* *Pass Condition; 'Non-dominant heel_z peak within ±10 frames of dominant foot apex'*

‘non_dominant _foot_behind’	*“trailing foot landing beside or a little behind the lead foot (not in front of the lead foot)* Skill; Gallop *“foot on non-hopping leg remains behind hopping leg (does not cross in front of)”* Skill; Hop	*Metric;* *Heel x-position* *What’s Compared;* *`mean(body_[non_dom]_heel_x)` vs `mean(body_[dom]_heel_x)`* *Pass Condition;* *Non-dominant heel mean x < dominant heel mean x (any separation)*
‘elbow_bends’	*“Arms are bent and lifted to about waist level at takeoff.”* Skill; Gallop *“Arms flex and swing forward to produce force”* Skill; Hop	*Metric;* *Mean elbow flexion* *What’s Compared;* *`left/right_elbow_extension_flexion` vs threshold’* *Pass Condition;* *Both elbows > 22.5° flexion within ±5 frames of hop apex*
'arm_bent_levl’	*“Arms are bent and lifted to about waist level at takeoff.”* *Skill; Gallop*	***arm_bent_level** — height* *Metric;* *Elbow height vs hip* *What’s Compared;* *`body_[lr]_elbow_z` vs `body_[lr]_hip_z` ± span* *Pass Condition;* *Both elbows within 40% of hip-to-shoulder span from hip*
‘arm_swing’	*“Arms flex and swing forward to produce force”* *Skill; Hop*	*Metric;* *Wrist peak timing* *What’s Compared;* *`argmax(body_[lr]_wrist_z)` vs `peak_frame`* *Pass Condition;* *Both wrist_z peaks within ±10 frames of dominant foot apex*

‘knees_flexed_ arms_behind’	*“prior to take off both knees are flexed and arms are extended behind the back.”* Skill; Horizontal Jump	***knees_flexed_arms_behind** — knee Metric; Knee flexion angle*
		What’s Compared; `left/right_knee_extension_flexion` vs threshold *Pass Condition; Left and right heel rise within ±5 frames of each other*
‘arms_forward _upward’	*“Arms extend forcefully forward and upward reaching above the head.”* Skill; Horizontal Jump	***arms_forward_upward** — height Metric; Wrist z vs shoulder z* *What’s Compared; `body_[lr]_wrist_z` vs `body_[lr]_shoulder_z`* *Pass Condition; Both wrists above shoulder height at takeoff (any separation)*
‘feet_together’	*“Both feet come off the floor together and land together.”* Skill; Horizontal Jump	***feet_together** — rise Metric; Left/right takeoff sync* *What’s Compared; `left_heel_rise_frame` vs `right_heel_rise_frame`* *Pass Condition; Left and right heel rise within ±5 frames of each other*

‘arms_downward’	*“Both arms are forced downward during landing”* Skill; Horizontal Jump	*Metric;* *Wrist z vs hip z* *What’s Compared;* *`mean(body_[lr]_wrist_z)` vs `mean(body_[lr]_hip_z)`* *Pass Condition;* *Both wrists below hip height in landing phase (any separation)*

Classification Analysis

To assess the accuracy of the deployed heuristic algorithms, classification analysis was carried out using the following metrics; accuracy, precision, recall, and F1-score. Thorough explanations of these metrics can be found by referring back to the methodology section of chapter 3. Importantly, such metrics were derived from comparisons between automated heuristic outputs and manually assigned TGMD-3 ground-truth labels by a trained assessor. The assessor’s qualifications included a BSc in Sports Performance and an MSc in Coaching Science in Sport. The ground-truth scoring procedure involved the assessor providing binary ‘TRUE’ or ‘FALSE’ statements to each handcrafted heuristic rule representing the distinct performance criteria of the TGMD-3 with these summed across both trials for each participant and then used to calculate the subsequent F1 scores given by the heuristic rules engine. Scoring was blinded to the automated output, however, no measure of intra/inter reliability was employed which is deemed a limitation of the analysis.

As precision and recall can provide differing perspectives on performance, the F1-score was also calculated as their harmonic mean, providing an overall measure of agreement between automated and human classifications. Higher values across all metrics indicated stronger correspondence between the heuristics engine and the assessor.

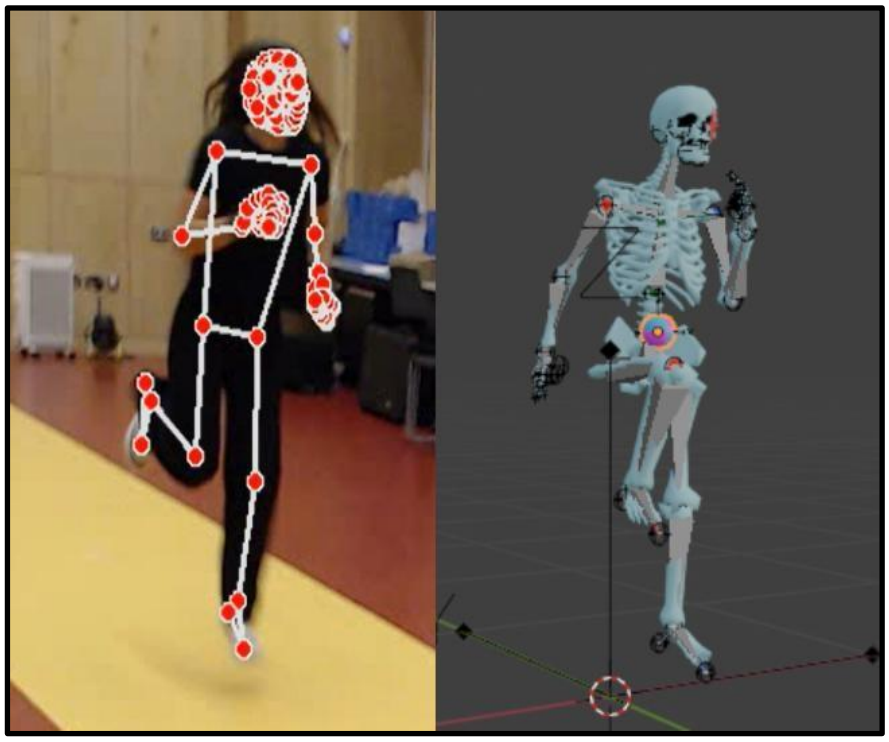


Figure 6.3 2D-to-3D Pose Triangulation

Statistical Analysis

Descriptive statistics (mean, median, standard deviation, minimum, and maximum) across 2-

,3-, and 4-camera configurations for elbow and ankle y-axis variables are presented in Appendix Table A.3. As the data violated assumptions of normality, differences between camera configurations were examined using Friedman repeated-measures analyses of variance (ANOVA). Where significant effects were observed, Durbin-Conover post-hoc pairwise comparisons were performed with appropriate adjustment for multiple comparisons. Effect sizes were calculated using Kendall's coefficient of concordance (W) and interpreted as negligible (<0.10), small (0.10–0.29), moderate (0.30–0.49), or large (≥0.50). Agreement between camera configurations was assessed using Intraclass Correlation Coefficients (ICC; 2,1), calculated using a two-way random-effects model assessing absolute agreement and is

presented in Table 6.4. This approach was selected to quantify the degree of agreement between measurements derived from the reduced camera configurations and those obtained from the 4-camera configuration. ICC values were interpreted as poor (<0.50),

moderate (0.50–0.75), good (0.75–0.90), or excellent (>0.90) (Koo & Li, 2016). While not considered a criterion-standard motion capture configuration, the 4-camera configuration was designated as the de facto comparison condition for agreement analyses as it incorporated the greatest number of viewpoints and therefore provided the greatest opportunity to minimise joint occlusion and maximise the completeness of the reconstructed kinematic data. Analysis was completed using jamovi and python.

6.4 Results

Descriptive statistics for all locomotor skills are reported in Appendix C. Friedman Repeated-Measures ANOVAs for all three locomotor skills are presented in Table 6.3. The data represented in Tables 6.3 and 6.4 are represented through the following variables; left and right ankle, left and right elbow with their corresponding data derived from the y-axis.

Table 6.3 Friedman Repeated-Measures ANOVA Results Across Locomotor Skills
*** = p < .05**

Skill	Variable(s)	χ^2	P	Significant Pairwise Comparison(s)	Kendall's W	95% CI
Hop	Left Ankle (Y-axis)	6.40	.041*	2 vs 4 (p = .011)	0.160	[0.00, 0.64]
Hop	Right Ankle (Y-axis)	5.70	.058	2 vs 4 (p = .017)	0.142	[0.008, 0.571]
Hop	Left Elbow (Y-axis)	7.30	.026*	2 vs 4 (p = .006)	0.182	[0.01, 0.64]
Hop	Right Elbow (Y-axis)	5.70	.058	2 vs 4 (p = .017)	0.142	[0.002, 0.57]
Gallop	All variables (Left/Right Ankle, Left/Right Elbow)	6.40	.041*	2 vs 4 (p = .011)	0.160	[0.00, 0.64]

Horizontal Jump	Left Ankle, Right Ankle, Left Elbow (Y-axis)	6.40	.041*	2 vs 4 (p = .011)	0.160	CI range: [0.00, 0.64] to [0.008, 0.57]
Horizontal Jump	Right Elbow (Y-axis)	10.30	.006*	2 vs 3 (p = .026); 2 vs 4 (p < .001)	0.258	[0.032, 0.73]

Friedman repeated-measures ANOVAs revealed significant effects of camera configuration across the Hop, Gallop, and Horizontal Jump tasks. Significant pairwise differences were consistently observed between the 2-camera and 4-camera configurations, while no significant differences were identified between the 3-camera and 4-camera configurations. Effect sizes were small (Kendall's W = .142–.258), indicating that camera configuration had a modest influence on reconstructed kinematic data outputs. Overall, the findings suggest that reducing camera configuration from four to three cameras had minimal impact on kinematic outputs, whereas further reduction to two cameras resulted in measurable differences in reconstruction performance.

Table 6.4 Mean ICC(2,1) Values Across Camera Configurations

Comparison	Mean ICC(2,1)	Interpretation
3 vs 4 Cameras	0.99	Excellent
2 vs 4 Cameras	0.15	Poor
2 vs 3 Cameras	0.23	Poor

Mean ICC(2,1) values indicated excellent agreement between the 3-camera and 4-camera configurations (ICC = 0.99), suggesting that reducing the system from four cameras to three had minimal impact on kinematic data reconstruction. In contrast, agreement involving the 2-camera configuration was poor, with mean ICC values of 0.15 (2 vs 4 cameras) and 0.23 (2 vs 3 cameras). Examination of the full ICC results (Appendix A.4.) revealed that all variables demonstrated excellent agreement between the 3-camera and 4-camera configurations (ICC = 0.99– 1.00), whereas comparisons involving the 2-camera configuration ranged from moderate agreement for some Hop variables (ICC = 0.51 – 0.69) to poor agreement across most Gallop and Horizontal Jump variables, including several negative ICC values. Collectively, these findings indicate that a 3-camera configuration provides kinematic outputs highly comparable to those obtained

using four cameras, whereas further reduction to two cameras substantially reduces measurement agreement.

Table 6.5 Mean Heuristics Rules Performance by Skill and Camera Configuration

Skill	Cameras	Support True	Support False	Accuracy	Precision	Recall	F1-Score
Hop	2 Cameras	5.00	0.20	0.73	1.00	0.73	0.80
Hop	3 Cameras	2.80	0.20	0.60	0.60	0.60	0.60
Hop	4 Cameras	2.80	0.20	0.60	0.60	0.60	0.60
Gallop	2 Cameras	6.75	1.25	0.50	0.66	0.54	0.51
Gallop	3 Cameras	5.25	0.75	0.54	0.67	0.60	0.59
Gallop	4 Cameras	6.00	1.00	0.50	0.64	0.58	0.56
Horizontal Jump	2 Cameras	3.00	0.00	0.25	0.75	0.25	0.38
Horizontal Jump	3 Cameras	21.75	0.00	0.25	1.00	0.25	0.37
Horizontal Jump	4 Cameras	21.75	0.00	0.27	1.00	0.27	0.39

Heuristics – Mean Level

Mean classification performance varied across skills but remained relatively consistent across camera configurations. Hop demonstrated the strongest overall performance, with F1-scores ranging from 0.60 to 0.80, alongside precision values between 0.60 and 1.00 and recall values between 0.60 and 0.73. Gallop produced moderate classification performance across all camera configurations (F1 = 0.51–0.59), while Horizontal Jump yielded the weakest performance, with F1-scores ranging from 0.37 to 0.39 despite precision values reaching 1.00 in both the 3- and 4-camera configurations. Across all three skills, differences between the 2-, 3-, and 4-camera configurations were modest, suggesting that reductions in camera number had limited influence on overall classification outcomes. Rather, classification performance appeared to be influenced more strongly by the locomotor skill being assessed than by the camera configuration used.

Ground-truth classifications, represented by the Support True and Support False values, provide additional context for these findings. Across all skills, positive classifications substantially outnumbered negative classifications, indicating that participants generally demonstrated the majority of performance criteria according to the traditional observe and grade approach. This pattern was particularly evident for Horizontal Jump, where Support False values were absent across all camera configurations, consequently, Horizontal

Jump classification metrics should be interpreted with caution. Whilst Table 6.5 provides a well-rounded overview of overall classification performance, findings are more notable at an independent skill level when scores associated to the handcrafted features are considered. An example of such findings for the Gallop are demonstrated below in Table 6.6.

Table 6.6 Feature-Level Classification Performance of Gallop Heuristics Across Camera Configurations

Feature	Camera Configuration	Support True	Support False	Accuracy	Precision	Recall	F1-Score
Lift Window	2 Cameras	8	0	1.00	1.00	1.00	1.00
Lift Window	3 Cameras	6	0	1.00	1.00	1.00	1.00
Lift Window	4 Cameras	7	0	1.00	1.00	1.00	1.00
Arm Bent Level	2 Cameras	8	0	0.00	0.00	0.00	0.00
Arm Bent Level	3 Cameras	6	0	0.00	0.00	0.00	0.00
Arm Bent Level	4 Cameras	7	0	0.00	0.00	0.00	0.00
Trail Foot Behind	2 Cameras	6	2	0.38	1.00	0.17	0.29
Trail Foot Behind	3 Cameras	5	1	0.50	1.00	0.40	0.57
Trail Foot Behind	4 Cameras	6	1	0.43	1.00	0.33	0.50
Both Feet Off Surface	2 Cameras	5	3	0.63	0.63	1.00	0.77

Both Feet Off Surface	3 Cameras	4	2	0.67	0.67	1.00	0.80
Both Feet Off Surface	4 Cameras	4	3	0.57	0.57	1.00	0.73

Heuristics – Feature Level

Table 6.6 provides an example of the variations demonstrated in feature-level classification of process-oriented performance criteria for the Gallop. Feature level results for all locomotor skills are demonstrated in Appendix E. While the aggregated skill-level results presented in Table 6.5 suggested minimal differences between camera configurations, feature-level analysis highlight notable fluctuations between individual criteria. Specifically, classification performance ranged from perfect agreement for Lift Window (F1 = 1.00 across all configurations) to complete failure for Arm Bent Level (F1 = 0.00 across all configurations). These findings indicate that overall skill-level metrics may obscure considerable variability in the ability of individual handcrafted features to operationalise specific TGMD-3 performance criteria.

Differences between camera configurations were most apparent for the Trail Foot Behind criterion. Here, the 3-camera configuration produced the strongest performance (F1 = 0.57), compared with the 4-camera (F1 = 0.50) and 2-camera (F1 = 0.29) configurations. Similarly, Both Feet Off Surface achieved its highest performance under the 3-camera configuration (F1 = 0.80), with lower values observed for both the 4-camera (F1 = 0.73) and 2-camera (F1 = 0.77) configurations. In contrast, Lift Window and Arm Bent Level demonstrated identical performance regardless of camera configuration, suggesting that certain criteria were either highly robust or consistently sensitive irrespective of the underlying kinematic data quality.

6.5 Discussion

The aims of the current-pilot study were; (1) Examine the feasibility of operationalising TGMD-3 Short Form locomotor performance criteria within an open source markerless motion capture pipeline and assess the extent to which its automated decisions agree with expert human scoring in an adult proof-of-concept sample and (2) Uncover the minimal viable camera configuration required for accurate classification by an open-source markerless motion capture system. The key contribution of the current pilot study is the presentation of an entirely open-source markerless motion-capture pipeline (via *'FreeMoCap';* Matthis et al. (2022)), with it shown as capable of successfully operationalising TGMD-3 Short Form locomotor criteria (Duncan, Martins, et al., 2022; Ulrich, 2013) and subsequently generate automated classifications (via rules-based heuristics; (Kokash, 2005)) that show meaningful agreement with traditional process-oriented assessor scoring. While fluctuations were observed in classification performance across skills and criteria, the current study's findings provide preliminary evidence that traditional, process-oriented MC assessment can be translated into objective computational rules using low-cost consumer hardware and freely available software. This is noteworthy given that previous studies which have sought to automate TGMD-3 assessment have largely relied on proprietary hardware, bespoke software ecosystems, or large-scale machine learning models (Bossavit & Arnedillo-Sánchez, 2022; Chow et al., 2020; Guo et al., 2026; Lander et al., 2025; Maulidiya et al., 2025; McGann et al., 2020).

The current study contributes to the emerging literature base on markerless motion capture through the development and evaluation of criterion-specific handcrafted features and transparent heuristic rules, transforming qualitative TGMD-3 short form performance criteria into objective kinematic classifications. Notable findings emerged from the feature-level analyses with variability between individual TGMD-3 performance criteria being noted to exceed variability between camera configurations. Although Tables 6.3 and 6.4 demonstrated reduced kinematic agreement when the system was reduced to two cameras, overall classification metrics changed only modestly across camera setups.

In contrast, feature-level results revealed substantial differences between criteria. Features focused upon the Hop and Gallop (i.e., 'lift window', 'non-dominant foot behind', and 'elbow bends') demonstrated strong classification performance, whereas 'arm bent level' and several Horizontal Jump criteria performed poorly regardless of camera configuration. This suggests that the primary challenge was not simply camera quantity,

but the extent to which individual TGMD-3 performance criteria could be translated into objective kinematic rules. These finding align with previous related research with Harsted et al. (2019) and Lerma and Gulgin (2023) both reporting that agreement with criterion-standard systems varied according to the specific form of movement being assessed. Similarly, Chow et al. (2020) demonstrated that markerless locomotor classification was feasible using a Kinect v2-based system, reporting agreement with expert raters ranging from 66.1% to 87.5% across TGMD-2 locomotor skills. However, performance was not uniform across all skills and performance criteria components, indicating that certain performance criteria noted in the hop and leap to reach higher levels of agreement than the horizontal jump, indicating some skills to be more applicable to markerless assessment than others. Chow et al. (2020) further noted certain performance criteria were associated with percentage agreement values of less than 60%; for example, performance criterion 2 and 5 for the hop were associated with percentage agreement values of 43.6% and 55.7%, respectively. A similar pattern emerged in the current study with several handcrafted features including 'elbow_bends' and 'non_dominant_foot_behind' for Hop, and Gallop 'lift_window', demonstraing strong classification performance, other features such as Gallop 'arm_bent_level' and multiple Horizontal Jump criteria consistently underperformed regardless of camera configuration. Across the current study and Chow et al. (2020), performance criteria linked to discrete spatial instances, such as foot position, limb placement, or flight phase, appeared easier to operationalise than those requiring interpretation of coordinated upper-limb actions, movement quality, or more subjective aspects of technique. Collectively, these findings suggest that the feasibility of automated motor competence assessment may depend less on the underlying capture technology and more on the extent to which individual performance criteria can be translated into objective and measurable kinematic features.

When compared against the findings of Bossavit and Arnedillo-Sánchez (2022), who also adopted heuristics, the current findings further support the use of handcrafted heuristics for locomotor assessment. Rather than relying on machine-learning models trained on largescale, labelled datasets, Bossavit and Arnedillo-Sánchez (2022) translated qualitative movement characteristics into explicit decision rules based on kinematic features extracted from pose-estimation outputs. Their approach demonstrated that several locomotor skill components could be classified with moderate to high levels of agreement with human raters (i.e., jump forward (91%), jump high (99%), jump sideway (93%), hop (94%) and jog (92%)), highlighting the potential of transparent, interpretable algorithms for automated assessment. However, their findings also showed that

performance varied across skills and criteria, with some movement components proving more difficult to operationalise reliably than others with ‘jump high’ (83.34%) and ‘hop’ (74.58%) producing lower accuracies than other skills. It should be noted however that the study by Bossavit & Arnedillo-Sánchez, (2023) was centered around a stationary integrated assessment application and thus direct comparisons are not possible. The authors did however demonstrate that handcrafted heuristics can provide meaningful classifications when movement criteria can be clearly defined and represented through measurable kinematic features, while also highlighting the challenges associated with criteria that are more subjective or difficult to quantify. The current findings follow more recent AI-driven approaches to MC assessment. Lander et al. (2025) developed a video-based deep learning system for automated TGMD-3 assessment using a large paediatric sample and reported classification accuracies exceeding 83% across multiple motor skills. Similarly, Guo et al. (2026) evaluated the AI-powered KineticAI platform against TGMD-3 ratings in 200 children and reported strong agreement with human evaluators (ICC = 0.94), a strong correlation with TGMD-3 scores (r = 0.92), and relatively low mean absolute errors across running, jumping, hopping, and balancing tasks. Collectively, these studies demonstrate the growing potential of artificial intelligence and machine learning to support automated MC assessment at scale. However, unlike these data-driven approaches, the current study employed handcrafted heuristic rules informed directly by TGMD-3 performance criteria. While classification performance was lower than that reported by Lander et al. (2025) and Guo et al. (2026), the approach employed in the current study offers a high degree of interpretability allowing for direct examination of how individual performance criteria contributed to classification success and failure, which often gets overlooked in machine learning research (Hassija et al., 2024; Rudin, 2019).

Importantly, feature-level analyses of the current study revealed that some TGMD-3 criteria could be operationalised effectively using simple kinematic-derived rules, whereas others remained challenging regardless of camera configuration. Whilst these findings bring forward the benefit of larger-scale datasets and advanced deep-learning approaches, the current study’s feature-level analyses indicate that the primary challenge may not be technological, but conceptual. Specifically, several TGMD-3 performance criteria presented by Ulrich (2013) lack sufficiently objective definitions to enable consistent and robust computational interpretation, regardless of the sophistication of the underlying assessment system. Whilst future approaches to assessment centred upon automation will undoubtedly benefit from larger datasets, advances in deep-learning methodologies, and continued improvements in markerless motion-capture technology,

the authors of the current study argue that technological advancement alone will not resolve existing limitations. Rather, the current study's findings highlighted that several TGMD-3 performance criteria appeared inherently more difficult to operationalise than others, regardless of camera configuration. Process-oriented performance criteria associated with discrete kinematic data in the current study generally demonstrated stronger classification performance than those reliant on more subjective interpretations of movement quality or coordination. This observation aligns with recent commentary provided by Ng et al. (2025), in which the argument was presented that many TGMD performance criteria lack explicit quantitative thresholds and remain dependent upon assessor judgement. Specifically, the authors proposed that the qualitative nature of several performance criteria can result in inconsistent interpretation between assessors and may limit the extent to which such criteria can be reliably translated into automated assessment systems without first establishing clearer operational definitions. Similar concerns have previously been raised by Barnett et al. (2014), who reported weaker agreement at the individual criterion level than at the overall skill level, while Lander et al. (2020) noted the inherent challenges associated with process-oriented assessment and argued for the introduction of more objective measurement approaches.

Comparatively however, the current authors, as previously noted in this study and in previous chapters of this thesis, highlight the strength of prior research using the TGMD-3 performance criteria in its traditional capacity to subsequently inform age-appropriate MC interventions (Belton et al., 2014; O' Brien et al., 2015, 2016). More broadly, ML and traditional analytical approaches serve related but distinct purposes within MC research. Traditional analytical approaches, such as correlation, regression and group-comparison techniques, are particularly valuable where research questions concern predefined relationships (De Meester et al., 2016), group differences (Behan et al., 2019) or hypothesis testing (Bardid et al., 2015), and their outputs are generally straightforward to interpret within established conceptual frameworks (Stodden et al., 2008). However, the capacity of traditional approaches to model large numbers of interacting variables or complex and potentially non-linear relationships may be limited (Salami et al., 2024).

ML approaches instead prioritise predictive performance and pattern identification, allowing complex relationships within multidimensional MC data to be modelled with fewer assumptions regarding their underlying structure (Bandeira, Estevan, et al., 2025; Bandeira, Lemos, et al., 2025). However, increased predictive complexity can come at the expense of interpretability, and strong predictive performance should not be

interpreted as evidence of causal relationships between variables given the current study's cross-sectional nature and limited sample size assessed. The current study demonstrates that computational approaches need not necessarily involve complex predictive models with handcrafted heuristic rules (Kokash, 2005) providing a convenient approach to solving a problem more efficiently yet accurately, as they have the advantage of meeting real-time constraints without relying on largescale datasets (Bossavit & Arnedillo-Sánchez, 2022). To this effect, while ML models can classify elaborated body motions involving several complex variables, they require largescale datasets and must integrate temporal and online information for real-time constraints (Wang et al., 2018). Consequently, the continued advancement of automated MC assessment does not serve to remove traditional approaches from current use but can serve to support rather than outright replace human expert assessment (Bossavit & Arnedillo-Sánchez, 2023). However, to achieve this it may require not only improvements in motion-capture and computational techniques , but also further refinement of the assessment frameworks upon which they are based. Future research should therefore seek to establish objective, measurable kinematic definitions for TGMD performance criteria which would allow for greater standardisation, reproducibility, and compatibility with other emerging automated assessment systems.

The current study's findings must equally be considered alongside it's limitations. First, the study was conducted as a proof-of-concept pilot investigation involving a small sample of adults rather than a paediatric sample for which the TGMD-3 was originally designed. Consequently, the findings presented in this study should not be interpreted as evidence of criterion validity for full-proof automated TGMD-3 assessment in children, but rather as an initial examination of the feasibility of combining markerless motion capture and heuristic-based classification for enhanced MC assessment. Future work should naturally seek to investigate the outlined approach among a paediatric population, ideally a large sample Second, whilst the handcrafted heuristics were developed to operationalise process-oriented TGMD-3 performance criteria, alternative feature definitions and decision thresholds may have produced different classification outcomes.

Future work should consider expanding upon the definitions and thresholds used in the current study for superior classification results. Third, the locomotor subset of the TGMD-3 Short Form was assessed in isolation, as the markerless motion-capture pipeline utilised within this study was unable to reliably track external objects, precluding evaluation of the object-control skills. The inability of this system to track object-control skills is a notable

limitation given its relevance in promoting long term adherence to physical activity, particularly among the female youth demographic (Barnett et al., 2010).Future technological advancements should hopefully enable external objects in the not so-distant future. Failing that, an analysis of all locomotor skills should be completed on the target population. Finally, the study employed a single markerless motion-capture workflow and did not compare outputs against a criterion-standard motion-capture system such as Vicon. Whilst the four-camera configuration served as a de facto comparison condition for internal comparison purposes, future research should evaluate the extent to which the observed kinematic patterns and classification outcomes generalise across alternative markerless systems, camera configurations, and participant populations.

Conclusion

The current pilot study examined the feasibility of combining an open-source markerless motion-capture pipeline with handcrafted heuristic rules to automate classification of TGMD-3 short-form locomotor performance criteria (Duncan, Martins, et al., 2022; Ulrich, 2013). Whilst reductions in camera quantity influenced kinematic data agreement, overall classification performance remained relatively stable across configurations, with the three-camera setup demonstrating comparable performance to the four-camera reference condition. Feature-level analyses revealed that classification success was highly criterion-dependent, with performance criteria linked to discrete kinematic events generally outperforming those reliant on more subjective interpretations of movement quality or coordination.The current findings offer practical implications to practitioners by suggesting that certain performance criteria components of the TGMD-3 short-form battery can be correctly classified using a three-camera configuration. Future work should seek to enhance F1 scores by developing algorithms which are less sensitive to external factors such as motion blur and signal interference between different equipment (Motti Ader & Bossavit, 2023). Whilst such factors can be accounted for, future advances in automated MC assessment may depend not only on improvements in motion-capture and ML/AI solutions, but also on the continued refinement and operationalisation of existing process-oriented assessment criteria. Although further research involving larger paediatric samples, criterion-standard motion-capture systems, and alternative computational approaches is required, the present study provides encouraging preliminary evidence supporting the potential of low-cost, markerless technologies for scalable MC assessment.

Chapter 7: Conclusions and Future Directions for Markerless Motion Capture Research

7.1 Overview

This thesis set out to investigate how emerging computational methods could enhance the assessment and understanding of motor competence (MC) in children. The preceding chapters established the public health importance of MC, its association with physical activity (PA) and with perceived motor competence (PMC), and the practical limitations of the observational assessment batteries on which the field currently relies. Against that background, four studies were undertaken, each addressing a distinct facet of the overarching aim.

The first two studies were analytical. Chapter 3 examined whether machine learning (ML) classification could predict individual MC outcomes from other MC and PMC assessments, with a view to reducing the testing burden of large-scale screening. Chapter 4 extended this work by combining ML, post-hoc model interpretability and graph analytics to model MC and PMC as an interconnected network and to characterise its age and gender specific structure. The remaining two studies concerned measurement technology. Chapter 5 presented a conceptual position paper on markerless motion capture as a complementary approach to MC assessment, and Chapter 6 evaluated, as a proof-of-concept pilot, whether an open-source markerless pipeline could operationalise the locomotor performance criteria of the Test of Gross Motor Development, third edition (TGMD-3).

This chapter does not restate the empirical detail of those studies. Instead it answers the research questions posed in Chapter 1 (Section 7.2), draws the findings into a single argument (Section 7.3), sets out the theoretical and methodological contributions (7.4) and the implications for research and practice (7.5), appraises the strengths and limitations of the programme of work as a whole (7.6), and proposes a research agenda for what should follow (7.7), before offering a concluding statement (7.8) and a brief reflexive note (7.9).

7.2 Key Findings

The thesis was organised around five research questions and their associated hypotheses, set out in Chapter 1. Each is addressed in turn below.

Predictive classification of motor competence (Chapter 3)

The first question asked whether ML classification models could accurately predict individual MC outcomes from data derived from other MC and PMC assessments. The

hypothesis was that predictive capability would differ across assessments, with some skills demonstrating sufficient accuracy. This hypothesis was supported. Drawing on an Irish dataset of 2,098 children aged 5 to 12 years (Behan et al., 2019), several skills, most notably the catch and the bounce, were predicted with high accuracy (F1 scores above 0.85), whereas others were predicted far less reliably. The pattern indicates that certain assessments demonstrated a degree of predictive overlap with the remaining variables in the original dataset. This predictive overlap does not establish however that the corresponding assessments are therefore redundant or interchangeable. External validation and prospective evaluation of shortened assessment batteries would be required before recommending their removal from future iterations of the TGMD-3.

A related question asked whether predictive performance differed across male, female and gender-combined data subsets, with the hypothesis that gender-specific models would outperform a combined model. This was also supported. The most accurately predicted skill differed by subset: the catch was most accurately predicted in the female-only (F1 = 0.87) and combined (F1 = 0.85) datasets, whereas the bounce was most accurately predicted in the male-only dataset (F1 = 0.86). The redundancy of a given assessment is therefore not uniform across gender, which has direct implications for how, and for whom, an assessment battery might be streamlined.

Interdependencies via post-hoc model interpretability and graph analytics (Chapter 4)

The third question asked whether post-hoc model interpretability and graph analytics could identify meaningful age and gender specific interdependencies within MC data, with the hypothesis that age and gender specific networks would exhibit distinct structures and that both MC and PMC variables would act as influential nodes. The hypothesis was supported. Modelling MC as a network of interconnected skills rather than as a set of isolated assessments revealed that some competencies, notably rolling and the horizontal jump, occupied central positions, while others, such as catching and kicking, acted as bridges linking broader areas of motor development. Network structure varied substantially by age and gender; for example, kicking proficiency among twelve-year-old females was associated with PMC variables such as kicking and skipping, whereas among nine-year-old males bouncing and sliding AMC proficiencies was linked to PMC variables such as Horizontal Jump, Gallop and Two-Handed Strike. Crucially, PMC variables were closely intertwined with MC outcomes throughout, indicating that motor competence is shaped not only by demonstrated skill but also by how competent a child believes themselves to be.

Markerless classification and camera configuration (Chapter 6)

The fourth question asked what the minimal viable camera configuration would be required to allow for all three locomotor skills of the TGMD-3 short form to be accurately classified by the markerless pipeline '*FreeMoCap*' (Matthis et al., 2022). The hypothesis that all skills could be accurately classified was only partially supported. Classification was strongly criterion-dependent: the hop was classified well (F1 0.60 to 0.80) and the gallop moderately (F1 0.51 to 0.59), whereas the horizontal jump performed poorly. Importantly, because the adult sample passed almost every criterion, the horizontal jump in particular yielded too few examples of poor performance for its classification metrics to be interpreted as evidence of discrimination. The pipeline therefore demonstrated feasibility for some criteria but not for all, and the shortfall was as much a property of the assessment criteria as of the technology. The findings provided preliminary evidence of close agreement between the three- and four-camera configurations for the movements and outcomes examined, whereas the two-camera configuration generally demonstrated greater deviation.Agreement between the three and four camera configurations was excellent (ICC(2,1)= 0.99), whereas agreement involving the two-camera configuration was poor (ICC(2,1) = 0.15 to 0.23). A three-camera configuration therefore offers a practical compromise between reconstruction quality and deployment burden.

7.3 Integrated discussion

Considered individually, the four studies address separate questions. Considered together, they advance a single argument: that motor competence is a complex, interdependent construct that is better interrogated with data-driven and interpretable methods than with the repeated group comparisons that have dominated the field, and that its assessment can begin to be made more objective and scalable through markerless technology, but only as a complement to expert observation whose value depends on how clearly the underlying criteria are defined. Two threads run across the thesis and together constitute its principal intellectual contribution.

The standing of perceived motor competence

The first thread concerns PMC. The network analysis in Chapter 4 showed that children's perceptions of their own ability were not peripheral but were woven into the structure of

the MC network, and in some subgroups, such as twelve-year-old females, were among the variables most closely associated with skill proficiency. This finding gives empirical weight to the position, advanced conceptually by Stodden et al. (2008), that PMC is an active participant in motor development rather than a downstream correlate of it. For measurement, the implication is that efforts to streamline or automate MC assessment should not treat PMC as an optional extra; the analytical chapters suggest that a fuller account of a child's motor development requires both what they can do and what they believe they can do.

The operationalisation of assessment criteria

The second thread, which emerges most clearly from Chapters 5 and 6 but is consistent with the analytical work, is that the binding constraint on automated assessment may lie not in the capture technology but in the assessment criteria themselves. In Chapter 6, performance criteria tied to discrete, well-defined kinematic events, such as a foot leaving the surface or a limb passing behind another, were operationalised successfully, whereas criteria requiring interpretation of coordinated upper-limb action or overall movement quality were not, regardless of camera configuration. This mirrors the argument set out in Chapter 5 and the concerns raised in the wider literature (Lander et al., 2020; Ng et al., 2025) that several TGMD criteria rely on qualitative judgement and lack explicit quantitative thresholds. The thesis therefore contributes a finding that is conceptual rather than merely technical: the progress of automated MC assessment will depend on refining the assessment frameworks, not only the instruments that read them.

Augmentation rather than replacement

These threads converge on the stance adopted throughout the measurement chapters, namely that computational methods should augment rather than replace established assessment. The TGMD and similar batteries are valid, reliable, and have underpinned much of what is known about children's motor development, including component-level diagnostic work. What the methods examined here add is efficiency, objectivity, scalability and interpretability, and they do so most effectively where the construct being measured is already well defined. Returning to the developmental model that frames the thesis, the contribution to the Stodden et al. (2008) account is therefore methodological: it offers tools for measuring and modelling the MC and PMC constructs, and the relationships among them, with greater nuance than the inferential methods on which the field has traditionally

relied, while making clear that the quality of those measurements remains bounded by the quality of the underlying criteria.

7.4 Theoretical and methodological contributions

The thesis makes several contributions that can be separated into methodological and substantive categories. Methodologically, it applied ML classification, and subsequently model interpretability and graph analytics, to a combined MC and PMC dataset in order to characterise skill interdependencies and their variation by age and gender, an analytical approach that moves beyond the group-level comparisons that dominate the literature. It also produced practical, open-source guidance on camera configuration for markerless MC capture, demonstrating that three cameras reproduce four-camera reconstructions while two do not, and it presented one of the first fully open-source, transparent and reproducible proof-of-concept frameworks for automated TGMD-3 short-form assessment.

Substantively, the thesis provides evidence that existing assessment batteries contain elements of redundancy that vary by gender, that PMC is more deeply embedded in the structure of motor competence than its conventional treatment suggests, and that the principal obstacle to automated assessment is the ambiguity of qualitative performance criteria. Each of these findings has a direct bearing on how the field designs, delivers and interprets MC assessment.

7.5 Implications for research and practice

For researchers, the analytical chapters suggest that assessment batteries can be examined empirically for redundancy and streamlined in a data-driven rather than purely pragmatic manner, and that a network perspective offers a richer account of motor development than the analysis of skills in isolation. The measurement chapters indicate that the most valuable near-term research investment may be in operationalising existing criteria into objective kinematic definitions, since this would benefit every automated approach regardless of its underlying technology.

For teachers, coaches and others delivering assessment in schools and community settings, the thesis points towards a future in which lower-burden, more scalable screening is feasible, while cautioning that the markerless tools examined here are not yet validated for use with children and should at present be regarded as promising but

preliminary. The three-camera finding offers concrete guidance for anyone seeking to balance data quality against practicality in a field deployment. For clinicians, the strongest current application is the tracking of within-child change in already-diagnosed conditions rather than diagnosis itself, and any clinical use would require diagnostic-accuracy evidence in the relevant age range that does not yet exist.

7.6 Strengths and limitations

The programme of work has several strengths. The analytical studies drew on a large, and geographically broad sample of Irish children recruited from 44 schools across 12 counties and all four provinces lending weight to their conclusions. The methods applied were novel to the field and, in the case of the markerless work, were deliberately open-source and reproducible. Throughout, the thesis maintained a critical and balanced stance, benchmarking new approaches against established methods and reporting where they performed poorly as well as where they performed well.

Several limitations should, however, be acknowledged at the level of the programme rather than the individual study. First, schools and participants from the original dataset were not selected using a national probability-sampling frame and thus the sample should not be considered nationally representative. Furthermore, the analytical chapters drew on cross-sectional data , which precludes causal or temporal inference and means that the direction of the relationship between MC and PMC cannot be established from this work. Second, the empirical evaluation of markerless capture in Chapter 6 was a proof-of-concept conducted with a small sample of adults rather than the paediatric population for which the TGMD-3 was designed; because adults pass most criteria, the sample contained few examples of poor performance, which limits what can be concluded about the system's ability to detect non-proficiency.

Third, the thesis relied throughout on a single assessment framework, the TGMD, which, despite its strengths, does not assess balance and carries the criterion-definition limitations discussed above. Fourth, the work was quantitative throughout; it did not incorporate the perspectives of the teachers, coaches and clinicians who would ultimately use these tools, which would be a valuable addition to future feasibility work. Finally, the use of an open-source pipeline, documented camera configuration and explicitly defined heuristic rules supports transparency and may facilitate future reproduction of the analytical workflow. Reproducibility was, however, not directly assessed through study replication by other researchers or by numerous data collections in different settings across separate sessions. Consequently, the findings of chapter 6 demonstrated that

feasibility and transparency of the workflow rather than establishing its reproducibility across settings.

Future directions

The research agenda set out in Chapter 5 identified four priorities for advancing markerless MC assessment, namely rigorous paediatric validation, standardised assessment protocols, robust ethics and data governance, and clear criteria for what constitutes sufficient validation before deployment. The empirical work in Chapter 6 began to address the protocol element through its camera-configuration analysis, and the present thesis allows that agenda to be made more specific.

The most pressing next step is to evaluate the markerless pipeline in a paediatric sample containing genuine variation in proficiency, so that the system's ability to discriminate proficient from non-proficient movement can actually be tested. Such work should describe and report the human ground-truth scoring in full, including inter-rater reliability, document the heuristic rules and thresholds a priori to avoid circularity, and where possible benchmark the system against a criterion-standard motion-capture system. In parallel, and arguably of equal importance, future research should seek to redefine selected TGMD performance criteria as objective, measurable kinematic benchmarks while preserving their construct validity, since the thesis suggests this is where the greatest gains are available. Standardisation efforts should extend to version control of the algorithms, models and software used, so that data captured now remain comparable as the technology evolves, a concern that does not arise for stable observational batteries. Finally, the analytical findings invite longitudinal designs capable of testing the developmental pathways proposed by Stodden et al. (2008), and in particular the mediating role of PMC that the network analysis in Chapter 4 brought into focus.

7.7 Reflexive note

Consistent with the pragmatist position set out in Chapter 1, the methodological choices throughout this thesis were governed less by allegiance to a particular tradition than by the usefulness of a given method for the problem at hand. That orientation, together with a personal enthusiasm for emerging technologies, carries a risk of over-optimism, and I have sought to guard against it by benchmarking new approaches against established methods, by reporting limitations alongside strengths, and by framing the contributions of this thesis as provisional steps that require further validation rather than as finished solutions. The conclusions above should be read in that spirit: as evidence-informed

contributions to an evolving body of work rather than as settled answers.

7.8 Conclusion

The overarching aim of this thesis was to explore how emerging computational approaches could be leveraged to improve the assessment and understanding of motor competence in children. Collectively, the four studies demonstrate that machine learning, post-hoc model interpretability, graph analytics and markerless motion capture each offer distinct contributions to longstanding challenges in MC assessment. The analytical chapters revealed that MC and PMC exist within a complex, interdependent network in which certain assessments carry redundant information and in which perceived competence is more influential than its conventional treatment implies.

The measurement chapters demonstrated, conceptually and then empirically, that low-cost, open-source markerless systems can begin to support objective and scalable assessment, while showing that their success depends on the operationalisation of assessment criteria as much as on the capture technology itself. Taken together, the thesis argues that the future of MC assessment lies not in replacing established batteries such as the TGMD-3 but in augmenting them through computational methods that improve efficiency, objectivity, scalability and interpretability.

Substantial work remains in validation, standardisation and implementation, but the evidence presented here indicates that such a future is achievable, and that by combining domain expertise in motor development with advances in data science and motion capture, researchers and practitioners may be better positioned to identify, monitor and support children's motor development at a scale not previously attainable.

Bibliography

Adank, A. M., H. H. Van Kann, D., A. A. Hoeboer, J. J., I. De Vries, S., P. J. Kremers, S., & B. Vos, S. (2018). Investigating Motor Competence in Association with Sedentary Behavior and Physical Activity in 7- to 11-Year-Old Children. *International Journal of Environmental Research and Public Health*, *15*(11), 2470. https://doi.org/10.3390/ijerph15112470

Adlou, B., Wilburn, C., & Weimar, W. (2025). Motion Capture Technologies for Athletic Performance Enhancement and Injury Risk Assessment: A Review for Multi-Sport Organizations. Sensors, 25(14), 4384. https://doi.org/10.3390/s25144384

Ahmadi, M. N., Pavey, T. G., & Trost, S. G. (2020). Machine Learning Models for Classifying Physical Activity in Free-Living Preschool Children. *Sensors*, *20*(16), 4364. https://doi.org/10.3390/s20164364

Ahmadi, M. N., Pfeiffer, K. A., & Trost, S. G. (2020). Physical Activity Classification in Youth Using Raw Accelerometer Data from the Hip. *Measurement in Physical Education and Exercise Science*, *24*(2), 129–136. https://doi.org/10.1080/1091367X.2020.1716768

Ahmadi, M. N., & Trost, S. G. (2022). Device-based measurement of physical activity in pre-schoolers: Comparison of machine learning and cut point methods. *PLOS ONE*, *17*(4), e0266970. https://doi.org/10.1371/journal.pone.0266970

Almeida, G., Luz, C., Martins, R., & Cordovil, R. (2017). Do Children Accurately Estimate Their Performance of Fundamental Movement Skills? *Journal of Motor Learning and Development*, *5*(2), 193–206. https://doi.org/10.1123/jmld.2016-0030

Alsareii, S. A., Awais, M., Alamri, A. M., AlAsmari, M. Y., Irfan, M., Aslam, N., & Raza, M. (2022). Physical Activity Monitoring and Classification Using Machine Learning Techniques. *Life*, *12*(8), 1103. https://doi.org/10.3390/life12081103

Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series*, *1142*, 012012. https://doi.org/10.1088/1742-6596/1142/1/012012

Antara, R., Welis, W., Masrun, Irawan, R., Tri Mario, D., Alnedral, Umar, & Wąsik, J. (2023). Effects of agility, coordination, and flexibility on dribbling skills in senior high school female field hockey players. *Physical Activity Review*, *11*(2), 42–51. https://doi.org/10.16926/par.2023.11.20

Arshad, M. M., Ramphul, K., Dachepally, R., Almasri, M., Memon, R. A., Sakthivel, H., Zaman, A., Ahmed, R., & Shahid, F. (2024). Five-year trends in risk factors for cardiovascular disease among adolescents in the United States. *Archives of Medical Science – Atherosclerotic Diseases*, *9*(1), 56–59. https://doi.org/10.5114/amsad/185775

Aubert, S., Brazo-Sayavera, J., González, S. A., Janssen, I., Manyanga, T., Oyeyemi, A. L., Picard, P., Sherar, L. B., Turner, E., & Tremblay, M. S. (2021). Global prevalence of physical activity for children and adolescents; inconsistencies, research gaps, and recommendations: A narrative review. *International Journal of Behavioral Nutrition and Physical Activity*, *18*(1),81. https://doi.org/10.1186/s12966-021-01155-2

Audrey, S., Bell, S., Hughes, R., & Campbell, R. (2013). Adolescent perspectives on wearing accelerometers to measure physical activity in population-based trials. *The European Journal of Public Health*, *23*(3), 475–480. https://doi.org/10.1093/eurpub/cks081

Babic, M. J., Morgan, P. J., Plotnikoff, R. C., Lonsdale, C., White, R. L., & Lubans, D. R. (2014). Physical Activity and Physical Self-Concept in Youth: Systematic Review and Meta-Analysis. *Sports Medicine*, *44*(11), 1589–1601. https://doi.org/10.1007/s40279-014-0229-z

Bagøien, T. E., & Halvari, H. (2005). Autonomous Motivation: Involvement in Physical Activity, and Perceived Sport Competence: Structural and Mediator Models. *Perceptual and Motor Skills*, *100*(1), 3–21. https://doi.org/10.2466/pms.100.1.3-21

Bailey, D. P., Boddy, L. M., Savory, L. A., Denton, S. J., & Kerr, C. J. (2012). Associations between cardiorespiratory fitness, physical activity and clustered cardiometabolic risk in children and adolescents: The HAPPY study. *European Journal of Pediatrics*, *171*(9), 1317–1323. https://doi.org/10.1007/s00431-012-1719-3

Baker, B. L., & Davison, K. K. (2011). I Know I Can: A Longitudinal Examination of Precursors and Outcomes of Perceived Athletic Competence Among Adolescent Girls. *Journal of Physical Activity and Health*, *8*(2), 192–199. https://doi.org/10.1123/jpah.8.2.192

Bandeira, P. F. R., Duncan, M., Pessoa, M. L., Soares, Í., da Silva, L., Mota, J., & Martins, C. (2020). TGMD-2 Short Version: Evidence of Validity and Associations With Sex, Age, and BMI in Preschool Children. *Journal of Motor Learning and Development*, *8*(3), 528–543. https://doi.org/10.1123/jmld.2019-0040

Bandeira, P. F. R., Estevan, I., Duncan, M., Lenoir, M., Lemos, L., Romo-Perez, V., Valentini, N., & Martins, C. (2025). A Multilayer Network Model for Motor Competence from the View of the Science of Complexity. *Sports Medicine*, *55*(2), 245–254. https://doi.org/10.1007/s40279-024-02164-4

Bandeira, P. F. R., Lemos, L. F., Estevan, I., Webster, E. K., Clark, C. T., Duncan, M. J., Mota, J. A., & Martins, C. L. (2025). Are we assessing motor competence? Evidence-informed constructs for motor competence in preschoolers through an Exploratory Graph Analysis. *Journal of Sports Sciences*, *43*(1), 109–116. https://doi.org/10.1080/02640414.2024.2414361

Bardid, F., De Meester, A., Tallir, I., Cardon, G., Lenoir, M., & Haerens, L. (2016). Configurations of actual and perceived motor competence among children: Associations with motivation for sports and global self-worth. *Human Movement Science*, *50*, 1–9. https://doi.org/10.1016/j.humov.2016.09.001

Bardid, F., Rudd, J. R., Lenoir, M., Polman, R., & Barnett, L. M. (2015). Cross-cultural comparison of motor competence in children from Australia and Belgium. Frontiers in Psychology, 6. https://doi.org/10.3389/fpsyg.2015.00964

Bardid, F., Vannozzi, G., Logan, S. W., Hardy, L. L., & Barnett, L. M. (2019). A hitchhiker's guide to assessing young people's motor competence: Deciding what method to use. *Journal of Science and Medicine in Sport*, *22*(3), 311–318. https://doi.org/10.1016/j.jsams.2018.08.007

Barnett, L., Hinkley, T., Okely, A. D., & Salmon, J. (2013). Child, family and environmental correlates of children's motor skill proficiency. *Journal of Science and Medicine in Sport*, *16*(4), 332–336. https://doi.org/10.1016/j.jsams.2012.08.011

Barnett, L. M., Dudley, D. A., Telford, R. D., Lubans, D. R., Bryant, A. S., Roberts, W. M., Morgan, P. J., Schranz, N. K., Weissensteiner, J. R., Vella, S. A., Salmon, J., Ziviani, J., OkelyA. D., Wainwright, N., Evans, J. R., & Keegan, R. J. (2019). Guidelines for the Selection of Physical Literacy Measures in Physical Education in Australia. *Journal of Teaching in Physical Education*, *38*(2), 119–125. https://doi.org/10.1123/jtpe.2018-0219

Barnett, L. M., Jerebine, A., Keegan, R., Watson-Mackie, K., Arundell, L., Ridgers, N. D., Salmon, J., & Dudley, D. (2023). Validity, Reliability, and Feasibility of Physical Literacy Assessments Designed for School Children: A Systematic Review. *Sports Medicine*, 1–25. https://doi.org/10.1007/s40279-023-01867-4

Barnett, L. M., Minto, C., Lander, N., & Hardy, L. L. (2014). Interrater reliability assessment using the Test of Gross Motor Development-2. *Journal of Science and Medicine in Sport*, *17*(6), 667–670. https://doi.org/10.1016/j.jsams.2013.09.013

Barnett, L. M., Morgan, P. J., Van Beurden, E., Ball, K., & Lubans, D. R. (2011). A Reverse Pathway? Actual and Perceived Skill Proficiency and Physical Activity. *Medicine & Science in Sports & Exercise*, *43*(5), 898–904. https://doi.org/10.1249/MSS.0b013e3181fdfadd

Barnett, L. M., Morgan, P. J., van Beurden, E., & Beard, J. R. (2008). Perceived sports competence mediates the relationship between childhood motor skill proficiency and adolescent physical activity and fitness: A longitudinal assessment. *International Journal of Behavioral Nutrition and Physical Activity*, *5*(1), 40. https://doi.org/10.1186/1479-5868-5-40

Barnett, L. M., Ridgers, N. D., & Salmon, J. (2015). Associations between young children's perceived and actual ball skill competence and physical activity. *Journal of Science and*

Medicine in Sport, *18*(2), 167–171. https://doi.org/10.1016/j.jsams.2014.03.001

Barnett, L. M., Ridgers, N. D., Zask, A., & Salmon, J. (2015). Face validity and reliability of a pictorial instrument for assessing fundamental movement skill perceived competence in young children. *Journal of Science and Medicine in Sport*, *18*(1), 98–102. https://doi.org/10.1016/j.jsams.2013.12.004

Barnett, L. M., Stodden, D., Cohen, K. E., Smith, J. J., Lubans, D. R., Lenoir, M., Iivonen, S., Miller, A. D., Laukkanen, A., Dudley, D., Lander, N. J., Brown, H., & Morgan, P. J. (2016). Fundamental Movement Skills: An Important Focus. *Journal of Teaching in Physical Education*, *35*(3), 219–225. https://doi.org/10.1123/jtpe.2014-0209

Barnett, L. M., van Beurden, E., Morgan, P. J., Brooks, L. O., & Beard, J. R. (2009). Childhood Motor Skill Proficiency as a Predictor of Adolescent Physical Activity. *Journal of Adolescent Health*, *44*(3), 252–259. https://doi.org/10.1016/j.jadohealth.2008.07.004

Barnett, L. M., Van Beurden, E., Morgan, P. J., Brooks, L. O., & Beard, J. R. (2010). Gender Differences in Motor Skill Proficiency From Childhood to Adolescence: A Longitudinal Study. *Research Quarterly for Exercise and Sport*, *81*(2), 162–170. https://doi.org/10.1080/02701367.2010.10599663

Barnett, L. M., Webster, E. K., Hulteen, R. M., De Meester, A., Valentini, N. C., Lenoir, M., Pesce, C., Getchell, N., Lopes, V. P., Robinson, L. E., Brian, A., & Rodrigues, L. P. (2022). Through the Looking Glass: A Systematic Review of Longitudinal Evidence, Providing New Insight for Motor Competence and Health. *Sports Medicine*, *52*(4), 875–920. https://doi.org/10.1007/s40279-021-01516-8

Barone Gibbs, B., Hivert, M.-F., Jerome, G. J., Kraus, W. E., Rosenkranz, S. K., Schorr, E. N., Spartano, N. L., Lobelo, F., & on behalf of the American Heart Association Council on Lifestyle and Cardiometabolic Health; Council on Cardiovascular and Stroke Nursing; and Council on Clinical Cardiology. (2021). Physical Activity as a Critical Component of First-Line Treatment for Elevated Blood Pressure or Cholesterol: Who, What, and How?: A Scientific Statement From the American Heart Association. *Hypertension*, *78*(2). https://doi.org/10.1161/HYP.0000000000000196

Barthélemy, M. (2004). Betweenness centrality in large complex networks. *The European Physical Journal B*, *38*(2), 163–168. https://doi.org/10.1140/epjb/e2004-00111-4

Bashir, D., Montanez, G. D., Sehra, S., Segura, P. S., & Lauw, J. (2020). *An Information-Theoretic Perspective on Overfitting and Underfitting* (arXiv:2010.06076). arXiv. https://doi.org/10.48550/arXiv.2010.06076

Behan, S., Belton, S., Peers, C., O'Connor, N. E., & Issartel, J. (2019). Moving Well-Being Well: Investigating the maturation of fundamental movement skill proficiency across sex in Irish children aged five to twelve. *Journal of Sports Sciences*, *37*(22), 2604–2612. https://doi.org/10.1080/02640414.2019.1651144

Bellows, L. L., Johnson, S. L., Davies, P. L., Anderson, J., Gavin, W. J., & Boles, R. E. (2013). The Colorado LEAP study: Rationale and design of a study to assess the short term longitudinal effectiveness of a preschool nutrition and physical activity program. *BMC Public Health*, *13*(1), 1146. https://doi.org/10.1186/1471-2458-13-1146

Belton, S., O' Brien, W., Meegan, S., Woods, C., & Issartel, J. (2014). Youth-Physical Activity Towards Health: Evidence and background to the development of the Y-PATH physical activity intervention for adolescents. *BMC Public Health*, *14*(1), 122. https://doi.org/10.1186/1471-2458-14-122

Belton, S., O'Brien, W., Wickel, E. E., & Issartel, J. (2013). Patterns of Noncompliance in Adolescent Field-Based Accelerometer Research. *Journal of Physical Activity and Health*, *10*(8), 1181–1185. https://doi.org/10.1123/jpah.10.8.1181

Berard, J. R., & Vallis, L. A. (2006). Characteristics of single and double obstacle avoidance strategies: A comparison between adults and children. *Experimental Brain Research*, *175*(1), 21–31. https://doi.org/10.1007/s00221-006-0529-0

Bisi, M. C., Pacini Panebianco, G., Polman, R., & Stagni, R. (2017). Objective assessment of movement competence in children using wearable sensors: An instrumented version of the TGMD-2 locomotor subtest. *Gait & Posture*, *56*, 42–48. https://doi.org/10.1016/j.gaitpost.2017.04.025

Boisgontier, M. P., & Iversen, M. D. (2020). Physical Inactivity: A Behavioral Disorder in the Physical Therapist's Scope of Practice. *Physical Therapy*, *100*(5)

Bolger, L. E., Bolger, L. A., O' Neill, C., Coughlan, E., O'Brien, W., Lacey, S., & Burns, C. (2018). Age and Sex Differences in Fundamental Movement Skills Among a Cohort of Irish School Children. *Journal of Motor Learning and Development*, *6*(1), 81–100. https://doi.org/10.1123/jmld.2017-0003

Bolger, L. E., Bolger, L. A., O'Neill, C., Coughlan, E., O'Brien, W., Lacey, S., Burns, C., & Bardid, F. (2021). Global levels of fundamental motor skills in children: A systematic review. *Journal of Sports Sciences*, *39*(7), 717–753. https://doi.org/10.1080/02640414.2020.1841405

Bonanno, M., De Nunzio, A. M., Quartarone, A., Militi, A., Petralito, F., & Calabrò, R. S. (2023). Gait Analysis in Neurorehabilitation: From Research to Clinical Practice. *Bioengineering*, *10*(7), 785. https://doi.org/10.3390/bioengineering10070785

Bond, R., Koene, A., Dix, A., Boger, J., Mulvenna, M. D., Galushka, M., Bradley, B. W., Browne, F., Wang, H., & Wong, A. (2019). *Democratisation of Usable Machine Learning in Computer Vision* (arXiv:1902.06804). arXiv. https://doi.org/10.48550/arXiv.1902.06804

Bonomi, A. G., Goris, A. H. C., Yin, B., & Westerterp, K. R. (2009). Detection of Type, Duration, and Intensity of Physical Activity Using an Accelerometer. *Medicine & Science*

in Sports & Exercise, *41*(9), 1770–1777. https://doi.org/10.1249/MSS.0b013e3181a24536

Booth, M. L., Okely, T., McLellan, L., Phongsavan, P., Macaskill, P., Patterson, J., Wright, J., & Holland, B. (1999). Mastery of fundamental motor skills among new south wales school students: Prevalence and sociodemographic distribution. *Journal of Science and Medicine in Sport*, *2*(2), 93–105. https://doi.org/10.1016/S1440-2440(99)80189-3

Bossavit, B., & Arnedillo-Sánchez, I. (2022). Using motion capture technology to assess locomotor development in children. *DIGITAL HEALTH*, *8*, 205520762211442. https://doi.org/10.1177/20552076221144201

Bossavit, B., & Arnedillo-Sánchez, I. (2023). Motion-based technology to support motor skills screening in developing children: A scoping review. *Computer Methods and Programs in Biomedicine*, *240*, 107715. https://doi.org/10.1016/j.cmpb.2023.107715

Breau, B., Coyle-Asbil, H. J., & Vallis, L. A. (2022). The Use of Accelerometers in Young Children: A Methodological Scoping Review. *Journal for the Measurement of Physical Behaviour*, *5*(3), 185–201. https://doi.org/10.1123/jmpb.2021-0049

Breslin, G., Murphy, M., McKee, D., Delaney, B., & Dempster, M. (2012). The effect of teachers trained in a fundamental movement skills programme on children's self-perceptions and motor competence. *European Physical Education Review*, *18*(1), 114–126. https://doi.org/10.1177/1356336X11430657

Brian, A., Bardid, F., Barnett, L. M., Deconinck, F. J. A., Lenoir, M., & Goodway, J. D. (2018). Actual and Perceived Motor Competence Levels of Belgian and United States Preschool Children. *Journal of Motor Learning and Development*, *6*(s2), S320–S336. https://doi.org/10.1123/jmld.2016-0071

Britton, Ú., Belton, S., Peers, C., Issartel, J., Goss, H., Roantree, M., & Behan, S. (2023). Physical literacy in children: Exploring the construct validity of a multidimensional physical literacy construct. *European Physical Education Review*, *29*(2), 183–198. https://doi.org/10.1177/1356336X221131272

Britton, Ú., Onibonoje, O., Belton, S., Behan, S., Peers, C., Issartel, J., & Roantree, M. (2023). Moving well-being well: Using machine learning to explore the relationship between physical literacy and well-being in children. *Applied Psychology: Health and Well-Being*, *15*(3), 1110–1129. https://doi.org/10.1111/aphw.12429

Brown, T., & Lalor, A. (2009). The Movement Assessment Battery for Children—Second Edition (MABC-2): A Review and Critique. *Physical & Occupational Therapy In Pediatrics*, *29*(1), 86–103. https://doi.org/10.1080/01942630802574908

Bruininks, R. H., & Bruininks, B. D. (2005). *Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2)*. APA PsycTests

Bryant, E. S., Duncan, M. J., & Birch, S. L. (2014). Fundamental movement skills and weight status in British primary school children. *European Journal of Sport Science*, *14*(7), 730–736. https://doi.org/10.1080/17461391.2013.870232

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., … Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955

Bürgi, F., Meyer, U., Granacher, U., Schindler, C., Marques-Vidal, P., Kriemler, S., & Puder, J. J. (2011). Relationship of physical activity with motor skills, aerobic fitness and body fat in preschool children: A cross-sectional and longitudinal study (Ballabeina). *International Journal of Obesity*, *35*(7), 937–944. https://doi.org/10.1038/ijo.2011.54

Burton, A. M., Cowburn, I., Thompson, F., Eisenmann, J. C., Nicholson, B., & Till, K. (2023). Associations Between Motor Competence and Physical Activity, Physical Fitness and Psychosocial Characteristics in Adolescents: A Systematic Review and Meta-analysis. *Sports Medicine*, *53*(11), 2191–2256. https://doi.org/10.1007/s40279-023-01886-1

Butterfield, S. A., Angell, R. M., & Mason, C. A. (2012). Age and Sex Differences in Object Control Skills by Children Ages 5 to 14. *Perceptual and Motor Skills*, *114*(1), 261–274. https://doi.org/10.2466/10.11.25.PMS.114.1.261-274

Cai, L., Ma, Y., Xiong, S., & Zhang, Y. (2019). Validity and Reliability of Upper Limb Functional Assessment Using the Microsoft Kinect V2 Sensor. *Applied Bionics and Biomechanics*, *2019*, 1–14. https://doi.org/10.1155/2019/7175240

Cairney, J., Clark, H., Dudley, D., & Kriellaars, D. (2019). Physical Literacy in Children and Youth—A Construct Validation Study. *Journal of Teaching in Physical Education*, *38*(2), 84–90. https://doi.org/10.1123/jtpe.2018-0270

Cairney, J., Dudley, D., Kwan, M., Bulten, R., & Kriellaars, D. (2019). Physical Literacy, Physical Activity and Health: Toward an Evidence-Informed Conceptual Model. *Sports Medicine*, *49*(3), 371–383. https://doi.org/10.1007/s40279-019-01063-3

Caldwell, H. A. T., Di Cristofaro, N. A., Cairney, J., Bray, S. R., MacDonald, M. J., & Timmons, B. W. (2020). Physical Literacy, Physical Activity, and Health Indicators in School-Age Children. *International Journal of Environmental Research and Public Health*, *17*(15), 5367. https://doi.org/10.3390/ijerph17155367

Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E., & Sheikh, Y. (2021). OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. IEEE Transactions on Pattern

Analysis and Machine Intelligence, 43(1), 172–186. https://doi.org/10.1109/TPAMI.2019.2929257

Carcamo-Oyarzun, J., Herrmann, C., Gerlach, E., Salvo-Garrido, S., & Estevan, I. (2025). Motor competence, motivation and enjoyment in physical education to profile children in relation to physical activity behaviors. *Physical Education and Sport Pedagogy*, *30*(6), 675–690. https://doi.org/10.1080/17408989.2023.2265399

Carl, J., Bryant, A. S., Edwards, L. C., Bartle, G., Birch, J. E., Christodoulides, E., Emeljanovas, A., Fröberg, A., Gandrieau, J., Gilic, B., Van Hilvoorde, I., Holler, P., Iconomescu, T. M., Jaunig, J., Laudanska-Krzeminska, I., Lundvall, S., De Martelaer, K., Martins, J., Mieziene, B., … Elsborg, P. (2023). Physical literacy in Europe: The current state of implementation in research, practice, and policy. *Journal of Exercise Science & Fitness*, *21*(1), 165–176. https://doi.org/10.1016/j.jesf.2022.12.003

Carmona, P., Dwekat, A., & Mardawi, Z. (2022). No more black boxes! Explaining the predictions of a machine learning XGBoost classifier algorithm in business failure. *Research in International Business and Finance*, *61*, 101649. https://doi.org/10.1016/j.ribaf.2022.101649

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical Activity, Exercise, and Physical Fitness: Definitions and Distinctions for Health-Related Research. *Public Health Reports*, *100*(2), 126

Cattuzzo, M. T., Dos Santos Henrique, R., Ré, A. H. N., De Oliveira, I. S., Melo, B. M., De Sousa Moura, M., De Araújo, R. C., & Stodden, D. (2016). Motor competence and health related physical fitness in youth: A systematic review. *Journal of Science and Medicine in Sport*, *19*(2), 123–129. https://doi.org/10.1016/j.jsams.2014.12.004

Chen, Z., Zhu, W., Ulrich, D. A., & Qin, M. (2024). Have the Fundamental Movement Skills of U.S. Children Changed? *Research Quarterly for Exercise and Sport*, 1–10. https://doi.org/10.1080/02701367.2023.2250828

Chinn, D. J., White, M., Harland, J., Drinkwater, C., & Raybould, S. (1999). Barriers to physical activity and socioeconomic position: Implications for health promotion. *Journal of Epidemiology and Community Health*, *53*(3), 191

Chihak, B. J., Grechkin, T. Y., Kearney, J. K., Cremer, J. F., & Plumert, J. M. (2014). How children and adults learn to intercept moving gaps. *Journal of Experimental Child Psychology*, *122*, 134–152. https://doi.org/10.1016/j.jecp.2013.12.006

Chovanová, E., Majherová, M., & Bendíková, E. (2023). Age- and Gender-Specific Levels and Differences in Children's Gross Motor Coordination During Prepuberty. *Physical Activity Review*, *11*(2), 86–93. https://doi.org/10.16926/par.2023.11.24

Chow, D. H. K., Cheng, W. H. W., & Tam, S. S. M. (2020). A Video-Based Classification System for Assessing Locomotor Skills in Children. Journal of Sports Science &

Medicine, 19(3), 585

Christensen, A. P., & Golino, H. (2021). On the equivalency of factor and network loadings. *Behavior Research Methods*, *53*(4), 1563–1580. https://doi.org/10.3758/s13428-020-01500-6

Clark, C. C. T., Bisi, M. C., Duncan, M. J., & Stagni, R. (2021). Technology-based methods for the assessment of fine and gross motor skill in children: A systematic overview of available solutions and future steps for effective in-field use. *Journal of Sports Sciences*, *39*(11), 1236–1276. https://doi.org/10.1080/02640414.2020.1864984

Clark, J. E. (2007). On the Problem of Motor Skill Development. *Journal of Physical Education, Recreation & Dance*, *78*(5), 39–44. https://doi.org/10.1080/07303084.2007.10598023

Clark, J. E., & Metcalfe, J. S. (2002). *THE MOUNTAIN OF MOTOR DEVELOPMENT: A METAPHOR*

Cliff, D. P., Okely, A. D., Morgan, P. J., Jones, R. A., Steele, J. R., & Baur, L. A. (2012). Proficiency Deficiency: Mastery of Fundamental Movement Skills and Skill Components in Overweight and Obese Children. *Obesity*, *20*(5), 1024–1033. https://doi.org/10.1038/oby.2011.241

Cliff, D. P., Reilly, J. J., & Okely, A. D. (2009). Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0–5 years. *Journal of Science and Medicine in Sport*, *12*(5), 557–567. https://doi.org/10.1016/j.jsams.2008.10.008

Coe, D. P., Post, E. M., Fitzhugh, E. C., Fairbrother, J. T., & Webster, E. K. (2024). Associations among Motor Competence, Physical Activity, Perceived Motor Competence, and Aerobic Fitness in 10–15-Year-Old Youth. *Children*, *11*(2), 260. https://doi.org/10.3390/children11020260

Cohen, K. E., Morgan, P. J., Plotnikoff, R. C., Callister, R., & Lubans, D. R. (2014). Fundamental movement skills and physical activity among children living in low-income communities: A cross-sectional study. *International Journal of Behavioral Nutrition and Physical Activity*, *11*(1), 49. https://doi.org/10.1186/1479-5868-11-49

Colyer, S. L., Evans, M., Cosker, D. P., & Salo, A. I. T. (2018). A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System. *Sports Medicine - Open*, *4*(1), 24. https://doi.org/10.1186/s40798-018-0139-y

Cools, W., De Martelaer, K., Samaey, C., & Andries, C. (2009). Movement skill assessment of typically developing preschool children: A review of seven movement skill assessment tools. *Journal of Sports Science & Medicine*, *8*(2), 154

Cooper, A. R., Goodman, A., Page, A. S., Sherar, L. B., Esliger, D. W., Van Sluijs, E. M., Andersen, L. B., Anderssen, S., Cardon, G., Davey, R., Froberg, K., Hallal, P., Janz, K. F., Kordas, K., Kreimler, S., Pate, R. R., Puder, J. J., Reilly, J. J., Salmon, J., … Ekelund, U. (2015). Objectively measured physical activity and sedentary time in youth: The International children's accelerometry database (ICAD). *International Journal of Behavioral Nutrition and Physical Activity*, *12*(1), 113. https://doi.org/10.1186/s12966-015-0274-5

Coppens, E., De Meester, A., Deconinck, F. J. A., De Martelaer, K., Haerens, L., Bardid, F., Lenoir, M., & D'Hondt, E. (2021). Differences in Weight Status and Autonomous Motivation towards Sports among Children with Various Profiles of Motor Competence and Organized Sports Participation. *Children*, *8*(2), 156. https://doi.org/10.3390/children8020156

Corbin, C. B. (2016). Implications of Physical Literacy for Research and Practice: A Commentary. *Research Quarterly for Exercise and Sport*, *87*(1), 14–27. https://doi.org/10.1080/02701367.2016.1124722

Corder, K., Sharp, S. J., Atkin, A. J., Andersen, L. B., Cardon, G., Page, A., Davey, R., Grøntved, A., Hallal, P. C., Janz, K. F., Kordas, K., Kriemler, S., Puder, J. J., Sardinha, L. B., Ekelund, U., & Van Sluijs, E. M. F. (2016). Age-related patterns of vigorous-intensity physical activity in youth: The International Children's Accelerometry Database. *Preventive Medicine Reports*, *4*, 17–22. https://doi.org/10.1016/j.pmedr.2016.05.006

Cornish, K., Fox, G., Fyfe, T., Koopmans, E., Pousette, A., & Pelletier, C. A. (2020). Understanding physical literacy in the context of health: A rapid scoping review. *BMC Public Health*, *20*(1), 1569. https://doi.org/10.1186/s12889-020-09583-8

Crane, J. R., Naylor, P. J., Cook, R., & Temple, V. A. (2015). Do Perceptions of Competence Mediate The Relationship Between Fundamental Motor Skill Proficiency and Physical Activity Levels of Children in Kindergarten? *Journal of Physical Activity and Health*, *12*(7), 954–961. https://doi.org/10.1123/jpah.2013-0398

Cremers, D., Reid, I., Saito, H., & Yang, M.-H. (Eds.). (2015). *Computer Vision -- ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part III* (Vol. 9005). Springer International Publishing. https://doi.org/10.1007/978-3-319-16811-1

Crocker, P. R., Eklund, R. C., & Kowalski, K. C. (2000). Children's physical activity and physical self-perceptions. *Journal of Sports Sciences*, *18*(6), 383–394

Csizmadia, G., Liszkai-Peres, K., Ferdinandy, B., Miklósi, Á., & Konok, V. (2022). Human activity recognition of children with wearable devices using LightGBM machine learning. *Scientific Reports*, *12*(1), 5472. https://doi.org/10.1038/s41598-022-09521-1

Cuimei, L., Zhiliang, Q., Nan, J., & Jianhua, W. (2017). Human face detection algorithm via Haar cascade classifier combined with three additional classifiers. *2017 13th IEEE*

International Conference on Electronic Measurement & Instruments (ICEMI), 483–487. https://doi.org/10.1109/ICEMI.2017.8265863

Daliri, M., Moradi, A., Fatorehchy, S., Bakhshi, E., Moradi, E., & Sabbaghi, S. (2023). Investigating the Effect of Leap Motion on Upper Extremity Rehabilitation in Children with Cerebral Palsy: A Randomized Controlled Trial. *Developmental Neurorehabilitation*, *26*(4), 244–252. https://doi.org/10.1080/17518423.2023.2203210

Davison, K. K., Schmalz, D. L., & Downs, D. S. (2010). Hop, Skip … No! Explaining Adolescent Girls' Disinclination for Physical Activity. *Annals of Behavioral Medicine*, *39*(3), 290–302. https://doi.org/10.1007/s12160-010-9180-x

De Meester, A., Barnett, L. M., Brian, A., Bowe, S. J., Jiménez-Díaz, J., Van Duyse, F., Irwin, J. M., Stodden, D. F., D'Hondt, E., Lenoir, M., & Haerens, L. (2020). The Relationship Between Actual and Perceived Motor Competence in Children, Adolescents and Young Adults: A Systematic Review and Meta-analysis. *Sports Medicine*, *50*(11), 2001–2049. https://doi.org/10.1007/s40279-020-01336-2

De Meester, A., Maes, J., Stodden, D., Cardon, G., Goodway, J., Lenoir, M., & Haerens, L. (2016). Identifying profiles of actual and perceived motor competence among adolescents: Associations with motivation, physical activity, and sports participation. *Journal of Sports Sciences*, *34*(21), 2027–2037. https://doi.org/10.1080/02640414.2016.1149608

De Meester, A., Stodden, D., Brian, A., True, L., Cardon, G., Tallir, I., & Haerens, L. (2016). Associations among Elementary School Children's Actual Motor Competence, Perceived Motor Competence, Physical Activity and BMI: A Cross-Sectional Study. *PLOS ONE*, *11*(10), e0164600. https://doi.org/10.1371/journal.pone.0164600

De Meester, F., De Bourdeaudhuij, I., Deforche, B., Ottevaere, C., & Cardon, G. (2011). Measuring physical activity using accelerometry in 13–15-year-old adolescents: The importance of including non-wear activities. *Public Health Nutrition*, *14*(12), 2124–2133. https://doi.org/10.1017/S1368980011001868

De Meester, F., Van Dyck, D., De Bourdeaudhuij, I., Deforche, B., & Cardon, G. (2014). Changes in physical activity during the transition from primary to secondary school in Belgian children: What is the role of the school environment? *BMC Public Health*, *14*(1), 261. https://doi.org/10.1186/1471-2458-14-261

De Nocker, Y. L., & Toolan, C. K. (2023). Using Telehealth to Provide Interventions for Children with ASD: A Systematic Review. *Review Journal of Autism and Developmental Disorders*, *10*(1), 82–112. https://doi.org/10.1007/s40489-021-00278-3

Den Uil, A. R., Janssen, M., Busch, V., Kat, I. T., & Scholte, R. H. J. (2023). The relationships between children's motor competence, physical activity, perceived motor competence, physical fitness and weight status in relation to age. *PLOS ONE*, *18*(4),

e0278438. https://doi.org/10.1371/journal.pone.0278438

Dhiman, P., Ma, J., Navarro, C. A., Speich, B., Bullock, G., Damen, J. A., Kirtley, S., Hooft, L., Riley, R. D., Van Calster, B., Moons, K. G. M., & Collins, G. S. (2021). Reporting of prognostic clinical prediction models based on machine learning methods in oncology needs to be improved. *Journal of Clinical Epidemiology*, *138*, 60–72. https://doi.org/10.1016/j.jclinepi.2021.06.024

D'Hondt, E., Deforche, B., De Bourdeaudhuij, I., & Lenoir, M. (2009). Relationship between Motor Skill and Body Mass Index in 5- to 10-Year-Old Children. *Adapted Physical Activity Quarterly*, *26*(1), 21–37. https://doi.org/10.1123/apaq.26.1.21

D'Hondt, E., Deforche, B., Gentier, I., De Bourdeaudhuij, I., Vaeyens, R., Philippaerts, R., & Lenoir, M. (2013). A longitudinal analysis of gross motor coordination in overweight and obese children versus normal-weight peers. *International Journal of Obesity*, *37*(1), 61–67. https://doi.org/10.1038/ijo.2012.55

Diao, Y., Dong, C., Barnett, L. M., Estevan, I., Li, J., & Ji, L. (2018). Validity and Reliability of a Pictorial Instrument for Assessing Fundamental Movement Skill Perceived Competence in Chinese Children. *Journal of Motor Learning and Development*, *6*(s2), S223–S238. https://doi.org/10.1123/jmld.2016-0082

Ding, D., Lawson, K. D., Kolbe-Alexander, T. L., Finkelstein, E. A., Katzmarzyk, P. T., Van Mechelen, W., & Pratt, M. (2016). The economic burden of physical inactivity: A global analysis of major non-communicable diseases. *The Lancet*, *388*(10051), 1311–1324. https://doi.org/10.1016/S0140-6736(16)30383-X

Donnelly, F. C., Mueller, S. S., & Gallahue, D. L. (2016). Developmental physical education for all children: Theory into practice. *Human Kinetics*.

Dudley, D., Cairney, J., Wainwright, N., Kriellaars, D., & Mitchell, D. (2017). Critical Considerations for Physical Literacy Policy in Public Health, Recreation, Sport, and Education Agencies. *Quest*, *69*(4), 436–452. https://doi.org/10.1080/00336297.2016.1268967

Duncan, M. J., Eyre, E. L. J., Clarke, N., Hamid, A., & Jing, Y. (2024). Importance of fundamental movement skills to predict technical skills in youth grassroots soccer: A machine learning approach. *International Journal of Sports Science & Coaching*, *19*(3), 1042–1049. https://doi.org/10.1177/17479541231202015

Duncan, M. J., Jones, V., O'Brien, W., Barnett, L. M., & Eyre, E. L. J. (2018). Self-Perceived and Actual Motor Competence in Young British Children. *Perceptual and Motor Skills*, *125*(2), 251–264. https://doi.org/10.1177/0031512517752833

Duncan, M. J., Martins, C., Ribeiro Bandeira, P. F., Issartel, J., Peers, C., Belton, S., O'Connor, N. E., & Behan, S. (2022). TGMD-3 short version: Evidence of validity and associations

with sex in Irish children. *Journal of Sports Sciences*, *40*(2), 138–145. https://doi.org/10.1080/02640414.2021.1978161

Duncan, M. J., Roscoe, C. M., Noon, M., Clark, C. C., O'Brien, W., & Eyre, E. L. (2020). Run, jump, throw and catch: How proficient are children attending English schools at the fundamental motor skills identified as key within the school curriculum? *European Physical Education Review*, *26*(4), 814–826. https://doi.org/10.1177/1356336X19888953

Duncan, M. J., Weldon, A., Barnett, L. M., & Lander, N. (2022). Perceptions and practices of fundamental movement skills in grassroots soccer coaches. International Journal of Sports Science & Coaching, 17(4), 761–771. https://doi.org/10.1177/17479541211073547

Dutta, T. (2012). Evaluation of the Kinect™ sensor for 3-D kinematic measurement in the workplace. *Applied Ergonomics*, *43*(4), 645–649. https://doi.org/10.1016/j.apergo.2011.09.011

Eather, N., Bull, A., Young, M. D., Barnes, A. T., Pollock, E. R., & Morgan, P. J. (2018). Fundamental movement skills: Where do girls fall short? A novel investigation of object-control skill execution in primary-school aged girls. *Preventive Medicine Reports*, *11*, 191–195. https://doi.org/10.1016/j.pmedr.2018.06.005

Eddy, L., Hill, L. J. B., Mon-Williams, M., Preston, N., Daly-Smith, A., Medd, G., & Bingham, D. D. (2021). Fundamental Movement Skills and Their Assessment in Primary Schools from the Perspective of Teachers. *Measurement in Physical Education and Exercise Science*, *25*(3), 236–249. https://doi.org/10.1080/1091367X.2021.1874955

Edwards, L. C., Bryant, A. S., Keegan, R. J., Morgan, K., Cooper, S.-M., & Jones, A. M. (2018). 'Measuring' Physical Literacy and Related Constructs: A Systematic Review of Empirical Findings. *Sports Medicine*, *48*(3), 659–682. https://doi.org/10.1007/s40279-017-0817-9

Eime, R. M., Young, J. A., Harvey, J. T., Charity, M. J., & Payne, W. R. (2013). A systematic review of the psychological and social benefits of participation in sport for children and adolescents: Informing development of a conceptual model of health through sport. *International Journal of Behavioral Nutrition and Physical Activity*, *10*(1), 98. https://doi.org/10.1186/1479-5868-10-98

Ekelund, U., Brage, S., Froberg, K., Harro, M., Anderssen, S. A., Sardinha, L. B., Riddoch, C., & Andersen, L. B. (2006). TV Viewing and Physical Activity Are Independently Associated with Metabolic Risk in Children: The European Youth Heart Study. *PLoS Medicine*, *3*(12), e488. https://doi.org/10.1371/journal.pmed.0030488

Ekelund, U., Luan, J., Sherar, L. B., Esliger, D. W., Griew, P., & Cooper, A. (2012). Moderate to Vigorous Physical Activity and Sedentary Time and Cardiometabolic Risk Factors in Children and Adolescents. *PHYSICAL ACTIVITY*.

Emanuel, M., Jarus, T., & Bart, O. (2008). Effect of Focus of Attention and Age on Motor Acquisition, Retention, and Transfer: A Randomized Trial. *Physical Therapy*, *88*(2), 251–260. https://doi.org/10.2522/ptj.20060174

Epskamp, S., & Fried, E. I. (2018). A Tutorial on Regularized Partial Correlation Networks. *Psychological Methods*, *23*(4), 617–634. https://doi.org/10.1037/met0000167

Erickson, K. I., Hillman, C., Stillman, C. M., Ballard, R. M., Bloodgood, B., Conroy, D. E., Macko, R., Marquez, D. X., Petruzzello, S. J., & Powell, K. E. (2019). Physical Activity, Cognition, and Brain Outcomes: A Review of the 2018 Physical Activity Guidelines. *Medicine & Science in Sports & Exercise*, *51*(6), 1242–1251. https://doi.org/10.1249/MSS.0000000000001936

Esposito, F., & Malerba, D. (2001). Machine learning in computer vision. *Applied Artificial Intelligence*, *15*(8), 693–705. https://doi.org/10.1080/088395101317018546

Estevan, I., & Barnett, L. M. (2018). Considerations Related to the Definition, Measurement and Analysis of Perceived Motor Competence. *Sports Medicine*, *48*(12), 2685–2694. https://doi.org/10.1007/s40279-018-0940-2

Estevan, I., De Meester, A., & Barnett, L. M. (2023). Advancing the understanding in the association between actual/perceived motor competence and health-related factors among children and adolescents. *Psychology of Sport and Exercise*, *67*, 102420. https://doi.org/10.1016/j.psychsport.2023.102420

Fan, L., Zhao, J., Hu, Y., Zhang, J., Wang, X., Wang, F., Wu, M., & Lin, T. (2024). Predicting physical functioning status in older adults: Insights from wrist accelerometer sensors and derived digital biomarkers of physical activity. *Journal of the American Medical Informatics Association*, *31*(11), 2571–2582. https://doi.org/10.1093/jamia/ocae224

Farooq, A., Martin, A., Janssen, X., Wilson, M. G., Gibson, A., Hughes, A., & Reilly, J. J. (2020). Longitudinal changes in moderate-to-vigorous-intensity physical activity in children and adolescents: A systematic review and meta-analysis. *Obesity Reviews*, *21*(1), e12953. https://doi.org/10.1111/obr.12953

Farrahi, V., & Rostami, M. (2024). Machine learning in physical activity, sedentary, and sleep behavior research. *Journal of Activity, Sedentary and Sleep Behaviors*, *3*(1), 5. https://doi.org/10.1186/s44167-024-00045-9

Feltz, D. L. (1988). Self-confidence and sports performance. *Exercise and Sport Sciences Reviews*, *16*(1), 423–458

Fergus, P., Hussain, A. J., Hearty, J., Fairclough, S., Boddy, L., Mackintosh, K., Stratton, G., Ridgers, N., Al-Jumeily, D., Aljaaf, A. J., & Lunn, J. (2017). A machine learning approach to measure and monitor physical activity in children. *Neurocomputing*, *228*,

220–230. https://doi.org/10.1016/j.neucom.2016.10.040

Fisher, A., Reilly, J. J., Kelly, L. A., Montgomery, C., Williamson, A., Paton, J. Y., & Grant, S. (2005). Fundamental Movement Skills and Habitual Physical Activity in Young Children: *Medicine & Science in Sports & Exercise*, *37*(4), 684–688. https://doi.org/10.1249/01.MSS.0000159138.48107.7D

Foulkes, J. D., Knowles, Z., Fairclough, S. J., Stratton, G., O'Dwyer, M., Ridgers, N. D., & Foweather, L. (2015). Fundamental Movement Skills of Preschool Children in Northwest England. *Perceptualand Motor Skills*, *121*(1), 260–283. https://doi.org/10.2466/10.25.PMS.121c14x0

Fransen, J., D'Hondt, E., Bourgois, J., Vaeyens, R., Philippaerts, R. M., & Lenoir, M. (2014). Motor competence assessment in children: Convergent and discriminant validity between the BOT-2 Short Form and KTK testing batteries. *Research in Developmental Disabilities*, *35*(6), 1375–1383. https://doi.org/10.1016/j.ridd.2014.03.011

Fulton, J. E., Burgeson, C. R., Perry, G. R., Sherry, B., Galuska, D. A., Alexander, M. P., Wechsler, H., & Caspersen, C. J. (2001). Assessment of Physical Activity and Sedentary Behavior in Preschool-Age Children: Priorities for Research. *Pediatric Exercise Science*, *13*(2), 113–126. https://doi.org/10.1123/pes.13.2.113

Gallahue, D. L., & Ozmun, J. G. (2006). Understanding Motor Development: Infants,Children, Adolescents, Adults. *6th Ed. International Edition, McgrawHill Companies, 524 P., New York*

Garcia, L., Pearce, M., Abbas, A., Mok, A., Strain, T., Ali, S., Crippa, A., Dempsey, P. C., Golubic, R., Kelly, P., Laird, Y., McNamara, E., Moore, S., De Sa, T. H., Smith, A. D., Wijndaele, K., Woodcock, J., & Brage, S. (2023). Non-occupational physical activity and risk of cardiovascular disease, cancer and mortality outcomes: A dose–response meta-analysis of large prospective studies. *British Journal of Sports Medicine*, *57*(15), 979–989. https://doi.org/10.1136/bjsports-2022-105669

Giacobbi, P. R., Poczwardowski, A., & Hager, P. (2005). A Pragmatic Research Philosophy for Sport and Exercise Psychology. *The Sport Psychologist*, *19*(1), 18–31. https://doi.org/10.1123/tsp.19.1.18

Gísladóttir, Ó., Haga, M., & Sigmundsson, H. (2014). Motor Competence and Physical Fitness in Adolescents. *Pediatric Physical Therapy*, *26*(1), 69–74. https://doi.org/10.1097/PEP.0000000000000006

Goodlich, B. I., Armstrong, E. L., Horan, S. A., Baque, E., Carty, C. P., Ahmadi, M. N., & Trost, S. G. (2020). Machine learning to quantify habitual physical activity in children with cerebral palsy. *Developmental Medicine & Child Neurology*, *62*(9), 1054–1060. https://doi.org/10.1111/dmcn.14560

Goodway, J. D., Robinson, L. E., & Crowe, H. (2010). Gender Differences in Fundamental Motor Skill Development in Disadvantaged Preschoolers From Two Geographical Regions. *Research Quarterly for Exercise and Sport*, *81*(1), 17–24. https://doi.org/10.1080/02701367.2010.10599624

Goodway, J., Ozmun, J., & Gallahue, D. (2013). Motor development in young children. In *Handbook of research on the education of young children* (pp. 103–115). Handbook of research on the education of young children

Goodway, JD., & Rudisill, M. (1997). Perceived Physical Competence and Actual Motor Skill Competence of African American Preschool Children. *Adapted Physical Activity Quarterly*, *14*(4), 314–326. https://doi.org/10.1123/apaq.14.4.314

Gottfried, J. M., Fehr, J., & Garbe, C. S. (2011). Computing Range Flow from Multi-modal Kinect Data. *Advances in Visual Computing*, *6938*, 758–767. https://doi.org/10.1007/978-3-642-24028-7_70

Graham, B., & Farrell, M. (2025). Mortality prediction using data from wearable activity trackers and individual characteristics: An explainable artificial intelligence approach. *Expert Systems with Applications*, *267*, 126195. https://doi.org/10.1016/j.eswa.2024.126195

Grauduszus, M., Wessely, S., Klaudius, M., & Joisten, C. (2023). Definitions and assessments of physical literacy among children and youth: A scoping review. *BMC Public Health*, *23*(1), 1746. https://doi.org/10.1186/s12889-023-16680-x

Greene, B. R., McGrath, D., O'Neill, R., O'Donovan, K. J., Burns, A., & Caulfield, B. (2010). An adaptive gyroscope-based algorithm for temporal gait analysis. *Medical & Biological Engineering & Computing*, *48*(12), 1251–1260. https://doi.org/10.1007/s11517-010-0692-0

Griffiths, A., Toovey, R., Morgan, P. E., & Spittle, A. J. (2018). Psychometric properties of gross motor assessment tools for children: A systematic review. *BMJ Open*, *8*(10), e021734. https://doi.org/10.1136/bmjopen-2018-021734

Grimpampi, E., Masci, I., Pesce, C., & Vannozzi, G. (2016). Quantitative assessment of developmental levels in overarm throwing using wearable inertial sensing technology. *Journal of Sports Sciences*, *34*(18), 1759–1765. https://doi.org/10.1080/02640414.2015.1137341

Groos, D., Adde, L., Støen, R., Ramampiaro, H., & Ihlen, E. A. F. (2022). Towards human-level performance on automatic pose estimation of infant spontaneous movements. *Computerized Medical Imaging and Graphics*, *95*, 102012. https://doi.org/10.1016/j.compmedimag.2021.102012

Gu, X., Thomas, K. T., & Chen, Y.-L. (2017). The Role of Perceived and Actual Motor Competency on Children's Physical Activity and Cardiorespiratory Fitness During Middle

Childhood. *Journal of Teaching in Physical Education*, *36*(4), 388–397. https://doi.org/10.1123/jtpe.2016-0192

Guo, J. X., Zhang, G. H., & Zhang, Y. M. (2026). AI-Powered Assessment of Motor Development: Using Platforms Like KineticAI to Analyze Fundamental Movement Skills in Children. Perceptual and Motor Skills, 133(2), 276–298. https://doi.org/10.1177/00315125251357047

Guthold, R., Stevens, G. A., Riley, L. M., & Bull, F. C. (2020). Global trends in insufficient physical activity among adolescents: A pooled analysis of 298 population-based surveys with 1·6 million participants. *The Lancet Child & Adolescent Health*, *4*(1), 23–35. https://doi.org/10.1016/S2352-4642(19)30323-2

Haga, M. (2008). The relationship between physical fitness and motor competence in children. *Child: Care, Health and Development*, *34*(3), 329–334. https://doi.org/10.1111/j.1365-2214.2008.00814.x

Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., & Ekelund, U. (2012). Global physical activity levels: Surveillance progress, pitfalls, and prospects. *The Lancet*, *380*(9838), 247–257. https://doi.org/10.1016/S0140-6736(12)60646-1

Hands, B. (2008). Changes in motor skill and fitness measures among children with high and low motor competence: A five-year longitudinal study. *Journal of Science and Medicine in Sport*, *11*(2), 155–162. https://doi.org/10.1016/j.jsams.2007.02.012

Hardy, L. L., Barnett, L., Espinel, P., & Okely, A. D. (2013). Thirteen-Year Trends in Child and Adolescent Fundamental Movement Skills: 1997–2010. *Medicine & Science in Sports & Exercise*, *45*(10), 1965–1970. https://doi.org/10.1249/MSS.0b013e318295a9fc

Hardy, L. L., King, L., Farrell, L., Macniven, R., & Howlett, S. (2010). Fundamental movement skills among Australian preschool children. *Journal of Science and Medicine in Sport*, *13*(5), 503–508. https://doi.org/10.1016/j.jsams.2009.05.010

Harsted, S., Holsgaard-Larsen, A., Hestbæk, L., Boyle, E., & Lauridsen, H. H. (2019). Concurrent validity of lower extremity kinematics and jump characteristics captured in pre-school children by a markerless 3D motion capture system. Chiropractic & Manual Therapies, 27(1), 39. https://doi.org/10.1186/s12998-019-0261-z

Harter, S. (1978). Effectance Motivation Reconsidered Toward a Developmental Model. *Human Development*, *21*(1), 34–64. https://doi.org/10.1159/000271574

Harter, S. (1982). *The Perceived Competence Scale for Children*

Harter, S. (1999). The construction of the self: A developmental perspective. *New York: Guilford*

Harter, S., & Pike, R. (1984). The Pictorial Scale of Perceived Competence and Social Acceptance for Young Children. *Child Development*, *55*(6), 1969–1982 https://doi.org/10.2307/1129772

Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., Scardapane, S., Spinelli, I., Mahmud, M., & Hussain, A. (2024). Interpreting Black-Box Models: A Review on Explainable Artificial Intelligence. *Cognitive Computation*, *16*(1), 45–74. https://doi.org/10.1007/s12559-023-10179-8

Haubenstricker, J., & Seefeldt, V. (1986). Acquisition of motor skills during childhood. *Physical Activity and Well-Being*, 41–92

Haugen, T., & Johansen, B. T. (2018). Difference in physical fitness in children with initially high and low gross motor competence: A ten-year follow-up study. *Human Movement Science*, *62*, 143–149. https://doi.org/10.1016/j.humov.2018.10.007

Hay, J., Maximova, K., Durksen, A., Carson, V., Rinaldi, R. L., Torrance, B., Ball, G. D. C., Majumdar, S. R., Plotnikoff, R. C., Veugelers, P., Boulé, N. G., Wozny, P., McCargar, L., Downs, S., Lewanczuk, R., & McGavock, J. (2012). Physical Activity Intensity and Cardiometabolic Risk in Youth. *Archives of Pediatrics & Adolescent Medicine*, *166*(11), 1022. https://doi.org/10.1001/archpediatrics.2012.1028

Haywood, K. M., & Getchell, N. (2019). The synergetic, probabilistic pathways of typical motor development. In *Understanding Motor Behaviour in Developmental Coordination Disorder* (pp. 11–27). Routledge

Henderson, S. E., Sugden, D., & Barnett, A. L. (1992). Movement Assessment Battery for Children-2. *Research in Developmental Disabilities*

Henrique, R. S., Ré, A. H. N., Stodden, D. F., Fransen, J., Campos, C. M. C., Queiroz, D. R., & Cattuzzo, M. T. (2016). Association between sports participation, motor competence and weight status: A longitudinal study. *Journal of Science and Medicine in Sport*, *19*(10), 825–829. https://doi.org/10.1016/j.jsams.2015.12.512

Henrique, R. S., Stodden, D. F., Fransen, J., Feitoza, A. H. P., Ré, A. H. N., Martins, C. M. L., Dos Prazeres, T. M. P., & Cattuzzo, M. T. (2020). Is motor competence associated with the risk of central obesity in preschoolers? *American Journal of Human Biology*, *32*(3), e23364. https://doi.org/10.1002/ajhb.23364

Herbaut, A., Valgalier, A., Foissac, M., McGann, J., & Issartel, J. (2023). Influence of Visual and Haptic Cues on Motor Learning: Insights from Tennis Skill Acquisition in Children. *Journal of Motor Behavior*, *55*(5), 444–452. https://doi.org/10.1080/00222895.2023.2198494

Hesketh, K. R., Lakshman, R., & Van Sluijs, E. M. F. (2017). Barriers and facilitators to young children's physical activity and sedentary behaviour: A systematic review and

synthesis of qualitative literature. *Obesity Reviews*, *18*(9), 987–1017. https://doi.org/10.1111/obr.12562

Hidding, L. M., Chinapaw, Mai. J. M., Van Poppel, M. N. M., Mokkink, L. B., & Altenburg, T. M. (2018). An Updated Systematic Review of Childhood Physical Activity Questionnaires. *Sports Medicine*, *48*(12), 2797–2842. https://doi.org/10.1007/s40279-018-0987-0

Hill, P. J., Mcnarry, M. A., Mackintosh, K. A., Murray, M. A., Pesce, C., Valentini, N. C., Getchell, N., Tomporowski, P. D., Robinson, L. E., & Barnett, L. M. (2024). The Influence of Motor Competence on Broader Aspects of Health: A Systematic Review of the Longitudinal Associations Between Motor Competence and Cognitive and Social-Emotional Outcomes. *Sports Medicine*, *54*(2), 375–427. https://doi.org/10.1007/s40279-023-01939-5

Hills, A. P., Andersen, L. B., & Byrne, N. M. (2011). Physical activity and obesity in children. *British Journal of Sports Medicine*, *45*(11), 866–870. https://doi.org/10.1136/bjsports-2011-090199

Holfelder, B., & Schott, N. (2014). Relationship of fundamental movement skills and physical activity in children and adolescents: A systematic review. *Psychology of Sport and Exercise*, *15*(4), 382–391. https://doi.org/10.1016/j.psychsport.2014.03.005

Horn, T. S., & Weiss, M. R. (1991). A Developmental Analysis of Children's Self-Ability Judgments in the Physical Domain. *Pediatric Exercise Science*, *3*(4), 310–326. https://doi.org/10.1123/pes.3.4.310

Hu, D., Zhou, S., Crowley-McHattan, Z. J., & Liu, Z. (2021). Factors That Influence Participation in Physical Activity in School-Aged Children and Adolescents: A Systematic Review from the Social Ecological Model Perspective. *International Journal of Environmental Research and Public Health*, *18*(6), 3147. https://doi.org/10.3390/ijerph18063147

Hulteen, R. M., Barnett, L. M., Morgan, P. J., Robinson, L. E., Barton, C. J., Wrotniak, B. H., & Lubans, D. R. (2018). Development, content validity and test-retest reliability of the Lifelong Physical Activity Skills Battery in adolescents. *Journal of Sports Sciences*, *36*(20), 2358–2367. https://doi.org/10.1080/02640414.2018.1458392

Hulteen, R. M., Barnett, L. M., True, L., Lander, N. J., Del Pozo Cruz, B., & Lonsdale, C. (2020). Validity and reliability evidence for motor competence assessments in children and adolescents: A systematic review. *Journal of Sports Sciences*, *38*(15), 1717–1798. https://doi.org/10.1080/02640414.2020.1756674

Hulteen, R. M., Lander, N. J., Morgan, P. J., Barnett, L. M., Robertson, S. J., & Lubans, D. R. (2015). Validity and Reliability of Field-Based Measures for Assessing Movement Skill Competency in Lifelong Physical Activities: A Systematic Review. *Sports Medicine*, *45*(10), 1443–1454. https://doi.org/10.1007/s40279-015-0357-0

Hulteen, R. M., Terlizzi, B., Abrams, T. C., Sacko, R. S., De Meester, A., Pesce, C., & Stodden, D. F. (2023). Reinvest to Assess: Advancing Approaches to Motor Competence Measurement Across the Lifespan. *Sports Medicine*, *53*(1), 33–50. https://doi.org/10.1007/s40279-022-01750-8

Hupin, D., Roche, F., Gremeaux, V., Chatard, J.-C., Oriol, M., Gaspoz, J.-M., Barthélémy, J.-C., & Edouard, P. (2015). Even a low-dose of moderate-to-vigorous physical activity reduces mortality by 22% in adults aged ≥60 years: A systematic review and meta-analysis. *British Journal of Sports Medicine*, *49*(19), 1262–1267. https://doi.org/10.1136/bjsports-2014-094306

Imam, F., Musilek, P., & Reformat, M. Z. (2024). Parametric and Nonparametric Machine Learning Techniques for Increasing Power System Reliability: A Review. *Information*, *15*(1),37. https://doi.org/10.3390/info15010037

Jaakkola, T., Huhtiniemi, M., Salin, K., Seppälä, S., Lahti, J., Hakonen, H., & Stodden, D. F. (2019). Motor competence, perceived physical competence, physical fitness, and physical activity within Finnish children. *Scandinavian Journal of Medicine & Science in Sports*, *29*(7), 1013–1021. https://doi.org/10.1111/sms.13412

Jaakkola, T., Washington, T., & Yli-Piipari, S. (2013). The association between motivation in school physical education and self-reported physical activity during Finnish junior high school: A self-determination theory approach. *European Physical Education Review*, *19*(1), 127–141. https://doi.org/10.1177/1356336X12465514

Jaakkola, T., Yli-Piipari, S., Huotari, P., Watt, A., & Liukkonen, J. (2016). Fundamental movement skills and physical fitness as predictors of physical activity: A 6-year follow-up study: Motor skills, fitness, and physical activity. *Scandinavian Journal of Medicine & Science in Sports*, *26*(1), 74–81. https://doi.org/10.1111/sms.12407

Jago, R., Salway, R., Emm-Collison, L., Sebire, S. J., Thompson, J. L., & Lawlor, D. A. (2020). Association of BMI category with change in children's physical activity between ages 6 and 11 years: A longitudinal study. *International Journal of Obesity*, *44*(1), 104–113. https://doi.org/10.1038/s41366-019-0459-0

Janssen, I., & LeBlanc, A. G. (2010). Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *International Journal of Behavioral Nutrition and Physical Activity*, *7*(1), 40. https://doi.org/10.1186/1479-5868-7-40

Jiabin, L., Shanshan, Z., Xiaomei, L., Kaixin, Y., & Ying, Y. (2025). The benefits of physical literacy for human flourishing: A machine learning–based exploration of adolescents. *Applied Psychology: Health and Well-Being*, *17*(1), e12635. https://doi.org/10.1111/aphw.12635

Johnson, T. M., Ridgers, N. D., Hulteen, R. M., Mellecker, R. R., & Barnett, L. M. (2016). Does playing a sports active video game improve young children's ball skill competence? *Journal of Science and Medicine in Sport*, *19*(5), 432–436. https://doi.org/10.1016/j.jsams.2015.05.002

Jones, R. A., Okely, A. D., Caputi, P., & Cliff, D. P. (2010). Perceived and actual competence among overweight and non-overweight children. *Journal of Science and Medicine in Sport*, *13*(6), 589–596. https://doi.org/10.1016/j.jsams.2010.04.002

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

Juonala, M., Viikari, J. S. A., & Raitakari, O. T. (2013). Main findings from the prospective Cardiovascular Risk in Young Finns Study. *Current Opinion in Lipidology*, *24*(1), 57–64. https://doi.org/10.1097/MOL.0b013e32835a7ed4

Jussila, A.-M., Husu, P., Vähä-Ypyä, H., Tokola, K., Kokko, S., Sievänen, H., & Vasankari, T. (2022). Accelerometer-Measured Physical Activity Levels and Patterns Vary in an Age- and Sex-Dependent Fashion among Finnish Children and Adolescents. *International Journal of Environmental Research and Public Health*, *19*(11), 6950. https://doi.org/10.3390/ijerph19116950

Kambas, A., Michalopoulou, M., Fatouros, I. G., Christoforidis, C., Manthou, E., Giannakidou, D., Venetsanou, F., Haberer, E., Chatzinikolaou, A., Gourgoulis, V., & Zimmer, R. (2012). The Relationship Between Motor Proficiency and Pedometer-Determined Physical Activity in Young Children. *Pediatric Exercise Science*, *24*(1), 34–44. https://doi.org/10.1123/pes.24.1.34

Kanko, R. M., Laende, E., Selbie, W. S., & Deluzio, K. J. (2021). Inter-session repeatability of markerless motion capture gait kinematics. Journal of Biomechanics, 121, 110422. https://doi.org/10.1016/j.jbiomech.2021.110422

Katzmarzyk, P. T., Powell, K. E., Jakicic, J. M., Troiano, R. P., Piercy, K., & Tennant, B. (2019). Sedentary Behavior and Health: Update from the 2018 Physical Activity Guidelines Advisory Committee. *Medicine & Science in Sports & Exercise*, *51*(6), 1227–1241. https://doi.org/10.1249/MSS.0000000000001935

Kelly, L., O'Connor, S., Harrison, A. J., & Ní Chéilleachair, N. J. (2019). Does fundamental movement skill proficiency vary by sex, class group or weight status? Evidence from an Irish primary school setting. *Journal of Sports Sciences*, *37*(9), 1055–1063. https://doi.org/10.1080/02640414.2018.1543833

Kemp, C., & Pienaar, A. E. (2013). Relationship between the body composition and motor and physical competence of Grade 1 learners in South Africa. *THE JOURNAL OF SPORTS MEDICINE AND PHYSICAL FITNESS*, *53*(6)

Khokhar, P. B., Pentangelo, V., Palomba, F., & Gravino, C. (2025). *Towards Transparent and Accurate Diabetes Prediction Using Machine Learning and Explainable Artificial Intelligence*

Kim, T., Jung, M.-C., & Mo, S.-M. (2025). Suggestion for Camera Location in Monocular Markerless 3D Motion Capture System: Focused on Accuracy Comparison With Marker-Based System for Upper Limb Joints. IEEE Access, 13, 47605–47616. https://doi.org/10.1109/ACCESS.2025.3547715

Kiphard, E. J., & Schilling, F. (1974). Körperkoordinationstest für Kinder. *Belz Test, Weinham*

Knippenberg, E., Verbrugghe, J., Lamers, I., Palmaers, S., Timmermans, A., & Spooren, A. (2017). Markerless motion capture systems as training device in neurological rehabilitation: A systematic review of their use, application, target population and efficacy. *Journal of NeuroEngineering and Rehabilitation*, *14*(1), 61. https://doi.org/10.1186/s12984-017-0270-x

Kokash, N. (2005). *An introduction to heuristic algorithms*

Konstabel, K., Veidebaum, T., Verbestel, V., Moreno, L. A., Bammann, K., Tornaritis, M., Eiben, G., Molnár, D., Siani, A., Sprengeler, O., Wirsik, N., Ahrens, W., & Pitsiladis, Y. (2014). Objectively measured physical activity in European children: The IDEFICS study. *International Journal of Obesity*, *38*(S2), S135–S143. https://doi.org/10.1038/ijo.2014.144

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

Koolwijk, P., Hoeboer, J., Mombarg, R., Savelsbergh, G. J. P., & De Vries, S. (2024). Fundamental movement skill interventions in young children: A systematic review. *International Journal of Sport and Exercise Psychology*, *22*(7), 1661–1683. https://doi.org/10.1080/1612197X.2023.2210597

Koporec, G., Vučković, G., Milić, R., & Perš, J. (2018). Quantitative Contact-Less Estimation of Energy Expenditure from Video and 3D Imagery. *Sensors*, *18*(8), 2435. https://doi.org/10.3390/s18082435

Kou, R., Zhang, Z., Zhu, F., Tang, Y., & Li, Z. (2024). Effects of Exergaming on executive function and motor ability in children: A systematic review and meta-analysis. *PLOS ONE*, *19*(9), e0309462. https://doi.org/10.1371/journal.pone.0309462

Krawczyk, B., Galar, M., Jeleń, Ł., & Herrera, F. (2016). Evolutionary undersampling boosting for imbalanced classification of breast cancer malignancy. *Applied Soft Computing*, *38*, 714–726. https://doi.org/10.1016/j.asoc.2015.08.060

Kriemler, S., Zahner, L., Schindler, C., Meyer, U., Hartmann, T., Hebestreit, H., Brunner-La Rocca, H. P., Van Mechelen, W., & Puder, J. J. (2010). Effect of school based physical activity programme (KISS) on fitness and adiposity in primary schoolchildren: Cluster randomised controlled trial. *BMJ*, *340*(feb23 1), c785–c785. https://doi.org/10.1136/bmj.c785

Kuan, G., Musa, R. M., Majeed, A. P. A., Kim, Y., Singnoy, C., & Kueh, Y. C. (2023). Predicting physical activity of young adults based on psychological need satisfaction in exercise using explainable decision tree model. In International conference on mechatronics and intelligent robotics (pp. 451-458). Singapore: Springer Nature Singapore.

Kwon, S., O'Brien, M. K., Welch, S. B., & Honegger, K. (2022). Physical Activity among U.S. Preschool-Aged Children: Application of Machine Learning Physical Activity Classification to the 2012 National Health and Nutrition Examination Survey National Youth Fitness Survey. *Children*, *9*(10), 1433. https://doi.org/10.3390/children9101433

Laborie, L. B., Lee, J., Antram, E., Lein, R. K., & Shelmerdine, S. C. (2026). Performance of adult-trained artificial intelligence models in paediatric imaging—A scoping review. *European Radiology*. https://doi.org/10.1007/s00330-026-12354-5

Lam, W. W. T., Tang, Y. M., & Fong, K. N. K. (2023). A systematic review of the applications of markerless motion capture (MMC) technology for clinical measurement in rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, *20*(1), 57. https://doi.org/10.1186/s12984-023-01186-9

Lander, N. J., Barnett, L. M., Brown, H., & Telford, A. (2015). Physical Education Teacher Training in Fundamental Movement Skills Makes a Difference to Instruction and Assessment Practices. *Journal of Teaching in Physical Education*, *34*(3), 548–556. https://doi.org/10.1123/jtpe.2014-0043

Lander, N., Morgan, P. J., Salmon, J., & Barnett, L. M. (2016). Teachers' Perceptions of a Fundamental Movement Skill (FMS) Assessment Battery in a School Setting. *Measurement in Physical Education and Exercise Science*, *20*(1), 50–62. https://doi.org/10.1080/1091367X.2015.1095758

Lander, N., Nahavandi, D., Mohamed, S., Essiet, I., & Barnett, L. M. (2020). Bringing objectivity to motor skill assessment in children. *Journal of Sports Sciences*, *38*(13), 1539–1549. https://doi.org/10.1080/02640414.2020.1747743

Lander, N., Nahavandi, D., Toomey, N. G., Barnett, L. M., & Mohamed, S. (2022). Accuracy vs. Practicality of Inertial Measurement Unit Sensors to Evaluate Motor Competence in Children. *Frontiers in Sports and Active Living*, *4*, 917340. https://doi.org/10.3389/fspor.2022.917340

Lander, N., Zhou, T., Timperio, A., Barnett, L. M., & Zhang, Y. (2025). Leveraging Artificial Intelligence to Automate Motor Competence Assessment. *Journal of Motor Learning and Development*, 1–22. https://doi.org/10.1123/jmld.2025-0002

Lansdown, G., & Vaghri, Z. (2022). Article 1: Definition of a Child. In Z. Vaghri, J. Zermatten, G. Lansdown, & R. Ruggiero (Eds.), *Monitoring State Compliance with the UN Convention on the Rights of the Child* (Vol. 25, pp. 407–412). Springer International Publishing. https://doi.org/10.1007/978-3-030-84647-3_40

Lanovaz, J. L., Oates, A. R., Treen, T. T., Unger, J., & Musselman, K. E. (2017). Validation of a commercial inertial sensor system for spatiotemporal gait measurements in children. *Gait & Posture*, *51*, 14–19. https://doi.org/10.1016/j.gaitpost.2016.09.021

Larouche, R., Boyer, C., Tremblay, M. S., & Longmuir, P. (2014). Physical fitness, motor skill, and physical activity relationships in grade 4 to 6 children. *Applied Physiology, Nutrition, and Metabolism*, *39*(5), 553–559. https://doi.org/10.1139/apnm-2013-0371

Lawson, C., Eyre, E. L. J., Tallis, J., & Duncan, M. J. (2021). Fundamental Movement Skill Proficiency Among British Primary School Children: Analysis at a Behavioral Component

Level. *Perceptual and Motor Skills*, *128*(2), 625–648. https://doi.org/10.1177/0031512521990330

Lee, I.-M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., & Katzmarzyk, P. T. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *The Lancet*, *380*(9838), 219–229. https://doi.org/10.1016/S0140-6736(12)61031-9

LeGear, M., Greyling, L., Sloan, E., Bell, R. I., Williams, B.-L., Naylor, P.-J., & Temple, V. A. (2012). A window of opportunity? Motor skills and perceptions of competence of children in Kindergarten. *International Journal of Behavioral Nutrition and Physical Activity*, *9*(1), 29. https://doi.org/10.1186/1479-5868-9-29

Lei, A. M. K., Salvador Lis, S., Cabral, S., & Veloso, A. P. (2026). Effect of Camera Configurations on Trunk and Lower-Extremity Kinematic Accuracy Using OpenCap with Lab-Based Motion Capture: A Pilot Investigation. Applied Sciences, 16(4), 1842. https://doi.org/10.3390/app16041842

Lerma, N., & Gulgin, H. (2023). Agreement of a Portable Motion Capture System to Analyze Movement Skills in Children. *Measurement in Physical Education and Exercise Science*, *27*(2), 105–113. https://doi.org/10.1080/1091367X.2022.2088290

Lester, D., McGrane, B., Belton, S., Duncan, M., Chambers, F., & O'Brien, W. (2017). The Age-Related Association of Movement in Irish Adolescent Youth. *Sports*, *5*(4), 77. https://doi.org/10.3390/sports5040077

Letts, E., King-Dowling, S., Kwan, M. Y. W., Obeid, J., Cairney, J., & Trost, S. G. (2024). Machine learning derived physical activity in preschool children with developmental coordination disorder. *Developmental Medicine & Child Neurology*, *67*(7), 910–917. https://doi.org/10.1111/dmcn.16186

Levy, J. J., & O'Malley, A. J. (2020). Don't dismiss logistic regression: The case for sensible extraction of interactions in the era of machine learning. *BMC Medical Research Methodology*, *20*(1), 171. https://doi.org/10.1186/s12874-020-01046-3

Ley, C., Martin, R. K., Pareek, A., Groll, A., Seil, R., & Tischer, T. (2022). Machine learning and conventional statistics: Making sense of the differences. *Knee Surgery, Sports Traumatology, Arthroscopy*, *30*(3), 753–757. https://doi.org/10.1007/s00167-022-06896-6

Li, D., Wang, D., Zou, J., Li, C., Qian, H., Yan, J., & He, Y. (2023). Effect of physical activity interventions on children's academic performance: A systematic review and meta-analysis. *European Journal of Pediatrics*, *182*(8), 3587–3601. https://doi.org/10.1007/s00431-023-05009-w

Li, Y., Li, J., Lei, L., Yu, F., Yu, Y., Zhao, Z., & Zhang, L. (2025). Occupational physical activity and risk of all-cause, cardiovascular disease and cancer mortality among adults in the USA: A national cohort study. *BMJ Public Health*, *3*(2), e002164. https://doi.org/10.1136/bmjph-2024-002164

Li, Y., Sperrin, M., Ashcroft, D. M., & Van Staa, T. P. (2020). Consistency of variety of machine learning and statistical models in predicting clinical risks of individual patients: Longitudinal cohort study using cardiovascular disease as exemplar. *BMJ*, m3919. https://doi.org/10.1136/bmj.m3919

Lima, R. A., Drenowatz, C., & Pfeiffer, K. A. (2022). Expansion of Stodden et al.'s Model. *Sports Medicine*, *52*(4), 679–683. https://doi.org/10.1007/s40279-021-01632-5

Lima, R., Pfeiffer, K., Larsen, L., Bugge, A., Moller, N., Anderson, L., & Stodden, D. (2017). Physical activity and motor competence present a positive reciprocal longitudinal relationship across childhood and early adolescence. *Journal of Physical Activity and Health*, *14*(6), 440–447. https://doi.org/10.1123/jpah.2016-0473

Lin, B.-S., Lee, I.-J., Fahn, C.-S., Lee, Y.-F., Chou, W.-J., & Wu, M.-L. (2021). Depth-Camera Based Energy Expenditure Estimation System for Physical Activity Using Posture Classification Algorithm. *Sensors*, *21*(12), 4216. https://doi.org/10.3390/s21124216

Lin, B.-S., Wang, L.-Y., Hwang, Y.-T., Chiang, P.-Y., & Chou, W.-J. (2019). Depth-Camera-Based System for Estimating Energy Expenditure of Physical Activities in Gyms. *IEEE Journal of Biomedical and Health Informatics*, *23*(3), 1086–1095. https://doi.org/10.1109/JBHI.2018.2840834

Liong, G. H. E., Ridgers, N. D., & Barnett, L. M. (2015). Associations between Skill Perceptions and Young Children's Actual Fundamental Movement Skills. *Perceptual and Motor Skills*, *120*(2), 591–603. https://doi.org/10.2466/10.25.PMS.120v18x2

Liu, Y., & Chen, S. (2021). Physical literacy in children and adolescents: Definitions, assessments, and interventions. *European Physical Education Review*, *27*(1), 96–112. https://doi.org/10.1177/1356336X20925502

Livesey, D., Lum Mow, M., Toshack, T., & Zheng, Y. (2011). The relationship between motor performance and peer relations in 9- to 12-year-old children: Motor performance and peer relations. *Child: Care, Health and Development*, *37*(4), 581–588. https://doi.org/10.1111/j.1365-2214.2010.01183.x

Lloyd, M., Saunders, T. J., Bremer, E., & Tremblay, M. S. (2014). Long-Term Importance of Fundamental Motor Skills: A 20-Year Follow-Up Study. *Adapted Physical Activity Quarterly*, *31*(1), 67–78. https://doi.org/10.1123/apaq.2013-0048

Logan, S. W., Barnett, L. M., Goodway, J. D., & Stodden, D. F. (2017). Comparison of performance on process- and product-oriented assessments of fundamental motor skills across childhood. *Journal of Sports Sciences*, *35*(7), 634–641. https://doi.org/10.1080/02640414.2016.1183803

Logan, S. W., Kipling Webster, E., Getchell, N., Pfeiffer, K. A., & Robinson, L. E. (2015). Relationship Between Fundamental Motor Skill Competence and Physical Activity During Childhood and Adolescence: A Systematic Review. *Kinesiology Review*, *4*(4), 416–426. https://doi.org/10.1123/kr.2013-0012

Logan, S. W., Robinson, L. E., Wilson, A. E., & Lucas, W. A. (2012). Getting the fundamentals of movement: A meta-analysis of the effectiveness of motor skill interventions in children. *Child: Care, Health and Development*, *38*(3), 305–315. https://doi.org/10.1111/j.1365-2214.2011.01307.x

Logan, S. W., Ross, S. M., Chee, K., Stodden, D. F., & Robinson, L. E. (2018). Fundamental motor skills: A systematic review of terminology. *Journal of Sports Sciences*, *36*(7), 781–796. https://doi.org/10.1080/02640414.2017.1340660

Loh, R., Stamatakis, E., Folkerts, D., Allgrove, J. E., & Moir, H. J. (2020). Effects of Interrupting Prolonged Sitting with Physical Activity Breaks on Blood Glucose, Insulin and Triacylglycerol Measures: A Systematic Review and Meta-analysis. *Sports Medicine*, *50*(2), 295–330. https://doi.org/10.1007/s40279-019-01183-w

Lopes, L., Santos, R., Coelho-e-Silva, M., Draper, C., Mota, J., Jidovtseff, B., Clark, C., Schmidt, M., Morgan, P., Duncan, M., O'Brien, W., Bentsen, P., D'Hondt, E., Houwen, S., Stratton, G., De Martelaer, K., Scheuer, C., Herrmann, C., García-Hermoso, A., … Agostinis-Sobrinho, C. (2020). A Narrative Review of Motor Competence in Children and Adolescents: What We Know and What We Need to Find Out. International Journal of

Environmental Research and Public Health, 18(1), 18. https://doi.org/10.3390/ijerph18010018

Lopes, L., Santos, R., Pereira, B., & Lopes, V. P. (2012). Associations between sedentary behavior and motor coordination in children. *American Journal of Human Biology*, *24*(6), 746–752. https://doi.org/10.1002/ajhb.22310

Lubans, D. R., Morgan, P. J., Cliff, D. P., Barnett, L. M., & Okely, A. D. (2010). Fundamental Movement Skills in Children and Adolescents: Review of Associated Health Benefits. *Sports Medicine*, *40*(12), 1019–1035. https://doi.org/10.2165/11536850-000000000-00000

Lundvall, S. (2015). Physical literacy in the field of physical education – A challenge and a possibility. *Journal of Sport and Health Science*, *4*(2), 113–118. https://doi.org/10.1016/j.jshs.2015.02.001

Luz, C., Rodrigues, L. P., Almeida, G., & Cordovil, R. (2016). Development and validation of a model of motor competence in children and adolescents. *Journal of Science and Medicine in Sport*, *19*(7), 568–572. https://doi.org/10.1016/j.jsams.2015.07.005

Maccobb, S., Greene, S., Nugent, J. K., & O'mahony, P. (2005). Measurement and Prediction of Motor Proficiency in Children Using the Bayley Infant Scales and the Bruininks-Oseretsky Test. *Physical & Occupational Therapy In Pediatrics*, *25*(1), 59–79. https://doi.org/10.1300/J006v25n01_05

Macy, M., Marks, K., & Towle, A. (2014). Missed, Misused, or Mismanaged: Improving Early Detection Systems to Optimize Child Outcomes. *Topics in Early Childhood Special Education*, *34*(2), 94–105. https://doi.org/10.1177/0271121414525997

Maeng, H. J., Webster, E. K., & Ulrich, D. A. (2016). Reliability for the test of gross motor development-(TGMD-3). *Research Quarterly for Exercise and Sport*, *87*(S2), A38

Magistro, D., Piumatti, G., Carlevaro, F., Sherar, L. B., Esliger, D. W., Bardaglio, G., Magno, F., Zecca, M., & Musella, G. (2018). Measurement invariance of TGMD-3 in children with and without mental and behavioral disorders. *Psychological Assessment*, *30*(11), 1421–1429. https://doi.org/10.1037/pas0000587

Maphumulo, S. Z., Breukelman, G. J., Shaw, B., & Shaw, I. (2025). The prevalence of physiological non-communicable diseases risk factors among Black Africans in peri-urban community In South Africa. *Clinical Epidemiology and Global Health*, *34*, 102111. https://doi.org/10.1016/j.cegh.2025.102111

Marks, J., Barnett, L. M., Strugnell, C., & Allender, S. (2015). Changing from primary to secondary school highlights opportunities for school environment interventions aiming to increase physical activity and reduce sedentary behaviour: A longitudinal cohort study. *International Journal of Behavioral Nutrition and Physical Activity*, *12*(1), 59. https://doi.org/10.1186/s12966-015-0218-0

Marshall, S. J., Biddle, S. J. H., Gorely, T., Cameron, N., & Murdey, I. (2004). Relationships between media use, body fatness and physical activity in children and youth: A meta-analysis. *International Journal of Obesity*, *28*(10), 1238–1246. https://doi.org/10.1038/sj.ijo.0802706

Martínez-Andrés, M., Bartolomé-Gutiérrez, R., Rodríguez-Martín, B., Pardo-Guijarro, M. J., Garrido-Miguel, M., & Martínez-Vizcaíno, V. (2020). Barriers and Facilitators to Leisure Physical Activity in Children: A Qualitative Approach Using the Socio-Ecological Model. *International Journal of Environmental Research and Public Health*, *17*(9), 3033. https://doi.org/10.3390/ijerph17093033

Martins, C. M. D. L., Hardy, L. L., & Duncan, M. J. (2024). Associations between Biological Maturation, Chronological Age, Body Mass Index, Sex, and Motor Competence in Prepubertal Children: A Network Analysis. Children, 11(9), 1143. https://doi.org/10.3390/children11091143

Martins, C., Webster, E. K., Romo-Perez, V., Duncan, M., Lemos, L. F., Staiano, A., Okely, A., Magistro, D., Carlevaro, F., Bardid, F., Magno, F., Nobre, G., Estevan, I., Mota, J., Ning, K., Robinson, L. E., Lenoir, M., Quan, M., Valentini, N., … Barnett, L. M. (2024). Sex differences in 3- to 5-year-old children's motor competence: A pooled cross-sectional analysis of 6241 children. *Scandinavian Journal of Medicine & Science in Sports*, *34*(5), e14651. https://doi.org/10.1111/sms.14651

Martiš, P., Košutzká, Z., & Kranzl, A. (2024). A Step Forward Understanding Directional Limitations in Markerless Smartphone-based Gait Analysis. https://doi.org/10.20944/preprints202403.0523.v1

Masci, I., Vannozzi, G., Bergamini, E., Pesce, C., Getchell, N., & Cappozzo, A. (2013). Assessing locomotor skills development in childhood using wearable inertial sensor devices: The running paradigm. *Gait & Posture*, *37*(4), 570–574. https://doi.org/10.1016/j.gaitpost.2012.09.017

Matsuda, T., Fujino, Y., Morisawa, T., Takahashi, T., Kakegawa, K., Matsumoto, T., Kiyohara, T., Fukushima, H., Higuchi, M., Torimoto, Y., Miwa, M., Fujiwara, T., & Daida, H. (2025). Reliability and Validity Examination of a New Gait Motion Analysis System. *Sensors*, *25*(4), 1076. https://doi.org/10.3390/s25041076

Matthis, J., Cherian, A., & Wirth, T. (2022). The FreeMoCap project-and-gaze/hand coupling during a combined three-ball juggling and balance task. Journal of Vision, 22(14), 4195–4195.

Maulidiya, D., Pujianto, D., Nopiyanto, Y. E., Wibowo, C., Kurniawan, A., & Herlangga, J. (2025). Preliminary Evaluation of an AI-Based Postural Assessment Application for Motor Skill Perception in Early Childhood Physical Education. 2025 International Conference on Computer, Control, Informatics and Its Applications (IC3INA), 138–143. https://doi.org/10.1109/IC3INA68387.2025.11325347

McCay, K. D., Ho, E. S. L., Shum, H. P. H., Fehringer, G., Marcroft, C., & Embleton, N. D. (2020). Abnormal Infant Movements Classification With Deep Learning on Pose-Based Features. *IEEE Access*, *8*, 51582–51592. https://doi.org/10.1109/ACCESS.2020.2980269

McGann, J., Issartel, J., Hederman, L., & Conlan, O. (2020). Hop.Skip.Jump.Games: The effect of "principled" exergameplay on children's locomotor skill acquisition. *British Journal of Educational Technology*, *51*(3), 798–816. https://doi.org/10.1111/bjet.12886

McIntyre, F., Parker, H., Chivers, P., & Hands, B. (2018). Actual competence, rather than perceived competence, is a better predictor of physical activity in children aged 6-9 years. *Journal of Sports Sciences*, *36*(13), 1433–1440. https://doi.org/10.1080/02640414.2017.1390243

Melby, P. S., Nielsen, G., Brønd, J. C., Tremblay, M. S., Bentsen, P., & Elsborg, P. (2022). Associations between children's physical literacy and well-being: Is physical activity a mediator? *BMC Public Health*, *22*(1), 1267. https://doi.org/10.1186/s12889-022-13517-x

Mentiplay, B. F., Hasanki, K., Perraton, L. G., Pua, Y.-H., Charlton, P. C., & Clark, R. A. (2018). Three-dimensional assessment of squats and drop jumps using the Microsoft Xbox One Kinect: Reliability and validity. *Journal of Sports Sciences*, *36*(19), 2202–2209. https://doi.org/10.1080/02640414.2018.1445439

Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. *2020 11th International Conference on Information and Communication Systems (ICICS)*, 243–248. https://doi.org/10.1109/ICICS49469.2020.239556

Moon, K., & Blackman, D. (2014). A Guide to Understanding Social Science Research for Natural Scientists. *Conservation Biology*, *28*(5), 1167–1177. https://doi.org/10.1111/cobi.12326

Morano, M., Bortoli, L., Ruiz, M. C., Campanozzi, A., & Robazza, C. (2020). Actual and perceived motor competence: Are children accurate in their perceptions? *PLOS ONE*, *15*(5), e0233190. https://doi.org/10.1371/journal.pone.0233190

Morley, D., Till, K., Ogilvie, P., & Turner, G. (2015). Influences of gender and socioeconomic status on the motor proficiency of children in the UK. *Human Movement Science*, *44*, 150–156. https://doi.org/10.1016/j.humov.2015.08.022

Motti Ader, L. G., & Bossavit, B. (2023). Understanding User Motion. In J. Vanderdonckt, P. Palanque, & M. Winckler (Eds.), *Handbook of Human Computer Interaction* (pp. 1–29). Springer International Publishing. https://doi.org/10.1007/978-3-319-27648-9_105-1

Mukherjee, S., Ting Jamie, L. C., & Fong, L. H. (2017). Fundamental Motor Skill

Proficiency of 6- to 9-Year-Old Singaporean Children. *Perceptual and Motor Skills*, *124*(3), 584–600. https://doi.org/10.1177/0031512517703005

Müller, B., Ilg, W., Giese, M. A., & Ludolph, N. (2017). Validation of enhanced kinect sensor based motion capturing for gait assessment. *PloS One*, *12*(4). https://doi.org/10.1371/journal.pone.0175813

Murtagh, E., Power, D., Foster, C., Murphy, M. H., Healy, S., Hayes, G., Murphy, N., Woods, C., & Williamson, C. (with ir@ul.ie & University of Limerick). (2023). *Update of the National Physical Activity Guidelines for Ireland—Final Research Report*. University of Limerick. https://doi.org/10.34961/RESEARCHREPOSITORY-UL.24762942

Nader, P. R., Bradley, R. H., Houts, R. M., McRitchie, S. L., & O'Brien, M. (2008). Moderate-to-Vigorous Physical Activity From Ages 9 to 15 Years. *Jama*, *300*(3), 295–305. https://doi.org/10.1001/jama.300.3.295

Narayanan, A., Desai, F., Stewart, T., Duncan, S., & Mackay, L. (2020). Application of Raw Accelerometer Data and Machine-Learning Techniques to Characterize Human Movement Behavior: A Systematic Scoping Review. *Journal of Physical Activity and Health*, *17*(3), 360–383. https://doi.org/10.1123/jpah.2019-0088

Nathan, D., Huynh, D. Q., Rubenson, J., & Rosenberg, M. (2015). Estimating Physical Activity Energy Expenditure with the Kinect Sensor in an Exergaming Environment. *PLOS ONE*, *10*(5), e0127113. https://doi.org/10.1371/journal.pone.0127113

Needham, L., Evans, M., Cosker, D. P., Wade, L., McGuigan, P. M., Bilzon, J. L., & Colyer, S. L. (2021). The accuracy of several pose estimation methods for 3D joint centre localisation. Scientific Reports, 11(1), 20673. https://doi.org/10.1038/s41598-021-00212-x

Nesbitt, D. R., Cattuzzo, M. T., Molina, S., Stodden, D. F., Robinson, L., & Phillips, D. S. (2016). Supine-to-Stand: A Measure of Functional Motor Competence in Preschool Children. *Research Quarterly for Exercise and Sport*, *87*(S2), A35

Newell, K. M. (2020). What are Fundamental Motor Skills and What is Fundamental About Them? *Journal of Motor Learning and Development*, *8*(2), 280–314. https://doi.org/10.1123/jmld.2020-0013

Ng, J. Y. Y., Jiang, S., Chan, C. H. S., & Ha, A. S. (2025). Assessing fundamental movement skills using the Test of Gross Motor Development (TGMD): Challenges and solutions to comparability and standardization. *Sports Medicine and Health Science*, *7*(4), 299–301. https://doi.org/10.1016/j.smhs.2025.02.007

Niemistö, D., Barnett, L. M., Laukkanen, A., Tolvanen, A., & Sääkslahti, A. (2023). Perceived motor competence in early childhood predicts perceived and actual motor competence in middle childhood. *Scandinavian Journal of Medicine & Science in Sports*, *33*(10), 2025–2038. https://doi.org/10.1111/sms.14438

O' Brien, W., Belton, S., & Issartel, J. (2015). Promoting Physical Literacy in Irish Adolescent Youth: The Youth-Physical Activity Towards Health (Y-PATH) Intervention. *MOJ Public Health*, *2*(6). https://doi.org/10.15406/mojph.2015.02.00041

O' Brien, W., Belton, S., & Issartel, J. (2016). Fundamental movement skill proficiency amongst adolescent youth. *Physical Education and Sport Pedagogy*, *21*(6), 557–571. https://doi.org/10.1080/17408989.2015.1017451

O' Brien, W., Issartel, J., & Belton, S. (2013). Evidence for the Efficacy of the Youth-Physical Activity towards Health (Y-PATH) Intervention. *Advances in Physical Education*, *03*(04), 145–153. https://doi.org/10.4236/ape.2013.34024

O'Brien, W., Duncan, M. J., Farmer, O., & Lester, D. (2018). Do Irish Adolescents Have Adequate Functional Movement Skill and Confidence? *Journal of Motor Learning and Development*, *6*(s2), S301–S319. https://doi.org/10.1123/jmld.2016-0067

Okely, A., & Booth, M. (2004). Mastery of fundamental movement skills among children in New South Wales: Prevalence and sociodemographic distribution. *Journal of Science and Medicine in Sport*, *7*(3), 358–372. https://doi.org/10.1016/S1440-2440(04)80031-8

Okely, A. D., Booth, M. L., & Patterson, J. W. (2001). Relationship of Cardiorespiratory Endurance to Fundamental Movement Skill Proficiency among Adolescents. *Pediatric Exercise Science*, *13*(4), 380–391. https://doi.org/10.1123/pes.13.4.380

Oladipupo, T. (2010). Types of Machine Learning Algorithms. In Y. Zhang (Ed.), *New Advances in Machine Learning*. InTech. https://doi.org/10.5772/9385

Oppici, L., Stell, F. M., Utesch, T., Woods, C. T., Foweather, L., & Rudd, J. R. (2022). A Skill Acquisition Perspective on the Impact of Exergaming Technology on Foundational Movement Skill Development in Children 3–12 Years: A Systematic Review and Meta-analysis. *Sports Medicine - Open*, *8*(1), 148. https://doi.org/10.1186/s40798-022-00534-8

Ortega-Benavent, N., Menescardi, C., Cárcamo-Oyarzún, J., & Estevan, I. (2024). Do perceived motor competence and physical literacy mediate the association between actual motor competence and physical activity engagement? *Revista de Psicodidáctica (English Ed.)*, *29*(2), 158–165. https://doi.org/10.1016/j.psicoe.2024.03.001

Page, L. (1998). The pagerank citation ranking: Bringing order to the web. Technical report. *Stanford Digital Library Technologies Project*

Pagliari, D., Menna, F., Roncella, R., Remondino, F., & Pinto, L. (2014). Kinect Fusion improvement

using depth camera calibration. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL–5*, 479–485. https://doi.org/10.5194/isprsarchives-XL-5-479-2014

Paluska, S. A., & Schwenk, T. L. (2000). Physical activity and mental health: Current concepts. *Sports Medicine*, *29*(3), 167–180. https://doi.org/10.2165/00007256-200029030-00003

Parrish, A.-M., Tremblay, M. S., Carson, S., Veldman, S. L. C., Cliff, D., Vella, S., Chong, K. H., Nacher, M., Del Pozo Cruz, B., Ellis, Y., Aubert, S., Spaven, B., Sameeha, M. J., Zhang, Z., & Okely, A. D. (2020). Comparing and assessing physical activity guidelines for children and adolescents: A systematic literature review and analysis. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 16. https://doi.org/10.1186/s12966-020-0914-2

Pate, R. R., O'Neill, J. R., & Lobelo, F. (2008). The Evolving Definition of 'Sedentary'. *Exercise and Sport Sciences Reviews*, *36*(4), 173–178. https://doi.org/10.1097/JES.0b013e3181877d1a

Pate, R. R., O'Neill, J. R., & Mitchell, J. (2010). Measurement of Physical Activity in Preschool Children. *Medicine & Science in Sports & Exercise*, *42*(3), 508–512. https://doi.org/10.1249/MSS.0b013e3181cea116

Paulus, J. K., & Kent, D. M. (2020). Predictably unequal: Understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *Npj Digital Medicine*, *3*(1), 99. https://doi.org/10.1038/s41746-020-0304-9

Peerlings, W. (2007). Bruininks-Oseretsky Test of Motor Proficiency. 2nd edition. (BOT-2). *In Actuele Themata Uit de Psychomotorische Therapie, Johan Simons (Ed.). Acco, Leuven/Leusden*

Peers, C., Issartel, J., Behan, S., O'Connor, N., & Belton, S. (2020). Movement competence: Association with physical self-efficacy and physical activity. *Human Movement Science*, *70*, 102582. https://doi.org/10.1016/j.humov.2020.102582

Pérez, L. M. R., & Sanz, J. L. G. (2005). New Measure of Perceived Motor Competence for Children Ages 4 to 6 Years. *Perceptual and Motor Skills*, *101*(1), 131–148

Pfister, A., West, A. M., Bronner, S., & Noah, J. A. (2014). Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. Journal of Medical Engineering & Technology, 38(5), 274–280. https://doi.org/10.3109/03091902.2014.909540

Pilati Filho, E. L., Spinosa, R. M. de O., & Brancher, J. D. (2023). A Machine Learning Model to Automatic Assessment of Gross Motor Development in Children using Posenet. *Semina: Ciências Exatas e Tecnológicas*, *44*, e48131. https://doi.org/10.5433/1679-0375.2023.v44.48131

Poitras, V. J., Gray, C. E., Borghese, M. M., Carson, V., Chaput, J.-P., Janssen, I.,

Katzmarzyk, P. T., Pate, R. R., Connor Gorber, S., Kho, M. E., Sampson, M., & Tremblay, M.S. (2016). Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. *Applied Physiology, Nutrition, and Metabolism*, *41*(6 (Suppl. 3)), S197–S239. https://doi.org/10.1139/apnm-2015-0663

Poucher, Z. A., Tamminen, K. A., Caron, J. G., & Sweet, S. N. (2020). Thinking through and designing qualitative research studies: A focused mapping review of 30 years of qualitative research in sport psychology. *International Review of Sport and Exercise Psychology*, *13*(1), 163–186. https://doi.org/10.1080/1750984X.2019.1656276

Raitakari, O. T., & Juonala, M. (2003). Cardiovascular Risk Factors in Childhood and Carotid Artery Intima-Media Thickness in Adulthood. *Jama*, *290*(17), 2277–2283.

Raudsepp, L., Liblik, R., & Hannus, A. (2002). Children's and Adolescents' Physical Self-Perceptions as Related to Moderate to Vigorous Physical Activity and Physical Fitness. *Pediatric Exercise Science*, *14*(1), 97–106. https://doi.org/10.1123/pes.14.1.97

Reunamo, J., Hakala, L., Saros, L., Lehto, S., Kyhälä, A.-L., & Valtonen, J. (2014). Children's physical activity in day care and preschool. *Early Years*, *34*(1), 32–48. https://doi.org/10.1080/09575146.2013.843507

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). 'Why Should I Trust You?': Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

Ricardo, L. I. C., Wendt, A., Costa, C. D. S., Mielke, G. I., Brazo-Sayavera, J., Khan, A., Kolbe-Alexander, T. L., & Crochemore-Silva, I. (2022). Gender inequalities in physical activity among adolescents from 64 Global South countries. *Journal of Sport and Health Science*, *11*(4), 509–520. https://doi.org/10.1016/j.jshs.2022.01.007

Rintala, P. O., Sääkslahti, A. K., & Iivonen, S. (2017). Reliability Assessment of Scores From Video-Recorded TGMD-3 Performances. *Journal of Motor Learning and Development*, *5*(1), 59–68. https://doi.org/10.1123/jmld.2016-0007

Rivas-García, M., Vidal-Ramírez, C., Mora-Traverso, M., Toledano-González, A., Rodríguez-Martínez, M. D. C., Molina-Torres, E., Marín-Marín, J.-A., Triviño-Juárez, J.-M., Gea-Mejías, M., & Romero-Ayuso, D. (2026). A multi-stakeholder preliminary needs assessment to inform the co-design of a digital telerehabilitation tool for children and adolescents with ADHD. *Disability and Rehabilitation: Assistive Technology*, 1–21. https://doi.org/10.1080/17483107.2026.2648864

Robinson, L. E. (2011). The relationship between perceived physical competence and fundamental motor skills in preschool children: Fundamental motor skills and perceived competence. *Child: Care, Health and Development*, *37*(4), 589–596.

https://doi.org/10.1111/j.1365-2214.2010.01187.x

Robinson, L. E., & Goodway, J. D. (2009). Instructional Climates in Preschool Children Who Are At-Risk. Part I: Object-Control Skill Development. *Research Quarterly for Exercise and Sport*, *80*(3), 533–542. https://doi.org/10.1080/02701367.2009.10599591

Robinson, L. E., Stodden, D. F., Barnett, L. M., Lopes, V. P., Logan, S. W., Rodrigues, L. P., & D'Hondt, E. (2015). Motor Competence and its Effect on Positive Developmental Trajectories of Health. *Sports Medicine*, *45*(9), 1273–1284. https://doi.org/10.1007/s40279-015-0351-6

Rodrigues, L. P., Stodden, D. F., & Lopes, V. P. (2016). Developmental pathways of change in fitness and motor competence are related to overweight and obesity status at the end of primary school. *Journal of Science and Medicine in Sport*, *19*(1), 87–92. https://doi.org/10.1016/j.jsams.2015.01.002

Rogez, G., Rihan, J., Ramalingam, S., Orrite, C., & Torr, P. H. S. (2008). Randomized trees for human pose detection. *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8. https://doi.org/10.1109/CVPR.2008.4587617

Romero-Martínez, J., Menescardi, C., García-Massó, X., & Estevan, I. (2026). A Four-Year Longitudinal Study of Relationships Among Actual and Perceived Motor Competence and Physical Activity in Spanish Primary School Children. *Research Quarterly for Exercise and Sport*, *97*(2), 295–309. https://doi.org/10.1080/02701367.2025.2587805

Rose, E., Larkin, D., Parker, H., & Hands, B. (2015). Does Motor Competence Affect Self-Perceptions Differently for Adolescent Males and Females? *Sage Open*, *5*(4), 2158244015615922. https://doi.org/10.1177/2158244015615922

Rudd, J. R., Barnett, L. M., Butson, M. L., Farrow, D., Berry, J., & Polman, R. C. J. (2015). Fundamental Movement Skills Are More than Run, Throw and Catch: The Role of Stability Skills. *PLOS ONE*, *10*(10), e0140224. https://doi.org/10.1371/journal.pone.0140224

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

Rudisill, M. E., Mahar, M. T., & Meaney, K. S. (1993). The Relationship between Children's Perceived and Actual Motor Competence. *Perceptual and Motor Skills*, *76*(3), 895–906. https://doi.org/10.2466/pms.1993.76.3.895

Ryde, N. S. W. (2000). Get Skilled: Get Active: A K-6 Resource to Support the Teaching of Fundamental Movement Skills. *NSW Department of Education and Training: Sydney, Australia*, 7–117

Saki, S., & Soori, M. (2026). Artificial intelligence, machine learning and deep learning in advanced transportation systems, a review. *Multimodal Transportation*, *5*(1), 100242. https://doi.org/10.1016/j.multra.2025.100242

Salami, S., Bandeira, P. F. R., Mashhadi, M. R., Shamsipour Dehkordi, P., Luz, L. G. O., Martins, C. M. D. L., Hardy, L. L., & Duncan, M. J. (2024). Associations between Biological Maturation, Chronological Age, Body Mass Index, Sex, and Motor Competence in Prepubertal Children: A Network Analysis. Children, 11(9), 1143. https://doi.org/10.3390/children11091143

Sallis, J. F., Prochaska, J. J., & Taylor, W. C. (2000). A review of correlates of physical activity of children and adolescents: *Medicine & Science in Sports & Exercise*, 963–975. https://doi.org/10.1097/00005768-200005000-00014

Salvo, D., Crochemore-Silva, I., Wendt, A., Tarp, J., Shiroma, E. J., Simpson, R. J., Lee, I.-M., Ekelund, U., Cerin, E., Keita, Y., Bauman, A., Hallal, P. C., Ramirez Varela, A., Pratt, M., Reis, R., Hinckson, E., Ding, D., Kohl, H. W., & Sallis, J. (2026). Physical activity for public health in the 21st century. *Nature Medicine*, *32*(4), 1479–1489. https://doi.org/10.1038/s41591-026-04237-5

Sasyniuk, T. M., Mohtadi, N. G. H., Hollinshead, R. M., Russell, M. L., & Fick, G. H. (2007). The Inter-rater Reliability of Shoulder Arthroscopy. *Arthroscopy*, *23*(9), 971–977. https://doi.org/10.1016/j.arthro.2007.03.005

Scataglini, S., Abts, E., Van Bocxlaer, C., Van Den Bussche, M., Meletani, S., & Truijen, S. (2024). Accuracy, Validity, and Reliability of Markerless Camera-Based 3D Motion Capture Systems versus Marker-Based 3D Motion Capture Systems in Gait Analysis: A Systematic Review and Meta-Analysis. Sensors, 24(11), 3686. https://doi.org/10.3390/s24113686

Schmutz, E. A., Leeger-Aschmann, C. S., Kakebeeke, T. H., Zysset, A. E., Messerli-Bürgy, N., Stülb, K., Arhab, A., Meyer, A. H., Munsch, S., Puder, J. J., Jenni, O. G., & Kriemler, S. (2020). Motor Competence and Physical Activity in Early Childhood: Stability and Relationship. *Frontiers in Public Health*, *8*, 39. https://doi.org/10.3389/fpubh.2020.00039

Sciortino, G., Farinella, G. M., Battiato, S., Leo, M., & Distante, C. (2017). On the Estimation of Children's Poses. In S. Battiato, G. Gallo, R. Schettini, & F. Stanco (Eds.), *Image Analysis and Processing—ICIAP 2017* (Vol. 10485, pp. 410–421). Springer International Publishing. https://doi.org/10.1007/978-3-319-68548-9_38

Scott, J. J., Hansen, V., Morgan, P. J., Plotnikoff, R. C., & Lubans, D. R. (2018). Young people's perceptions of the objective physical activity monitoring process: A qualitative exploration. *Health Education Journal*, *77*(1), 3–14. https://doi.org/10.1177/0017896917734576

Seefeldt, V., & Gould, D. (1980). *Physical and Psychological Effects of Athletic Competition on Children and Youth.*

Sgrò, F., Mango, P., Pignato, S., Schembri, R., Licari, D., & Lipoma, M. (2017). Assessing Standing Long Jump Developmental Levels Using an Inertial Measurement Unit. *Perceptual and Motor Skills*, *124*(1), 21–38. https://doi.org/10.1177/0031512516682649

Sgrò, F., Nicolosi, S., Schembri, R., Pavone, M., & Lipoma, M. (2015). Assessing Vertical Jump Developmental Levels in Childhood Using a Low-Cost Motion Capture Approach. *Perceptual and Motor Skills*, *120*(2), 642–658. https://doi.org/10.2466/10.PMS.120v12x7

Shah, A., & Prajapati, D. (2025). Recent Advances in Computer Vision and Machine Learning for Athletic Performance in Jump Events. *Augmented Human Research*, *10*(1), 12. https://doi.org/10.1007/s41133-025-00087-x

Shi, Z., Chu, Y., Zhang, Y., Wang, Y., & Wei, D.-Q. (2021). Prediction of Blood-Brain Barrier Permeability of Compounds by Fusing Resampling Strategies and eXtreme Gradient Boosting. *IEEE Access*, *9*, 9557–9566. https://doi.org/10.1109/ACCESS.2020.3047852

Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., & Blake, A. (2011). *Real-Time Human Pose Recognition in Parts from Single Depth Images*. 1297–1304. https://doi.org/10.1109/CVPR.2011.5995316

Sigmundsson, H., & Haga, M. (2016). Motor competence is associated with physical fitness in four- to six-year-old preschool children. *European Early Childhood Education Research Journal*, *24*(3), 477–488. https://doi.org/10.1080/1350293X.2016.1164411

Singh, D., Nagaraj, S., Daniel, R., Flood, C., Kulik, D., Flook, R., Goldenberg, A., Brudno, M., & Stedman, I. (2023). The promises and challenges of clinical AI in community paediatric medicine. *Paediatrics & Child Health*, *28*(4), 212–217. https://doi.org/10.1093/pch/pxac080

Song, X., Liu, X., Liu, F., & Wang, C. (2021). Comparison of machine learning and logistic regression models in predicting acute kidney injury: A systematic review and meta-analysis. *International Journalof Medical Informatics*, *151*, 104484. https://doi.org/10.1016/j.ijmedinf.2021.104484

Spessato, B. C., Gabbard, C., Valentini, N., & Rudisill, M. (2013). Gender differences in Brazilian children's fundamental movement skill performance. *Early Child Development and Care*, *183*(7), 916–923. https://doi.org/10.1080/03004430.2012.689761

Sport Australia: The Australian Physical Literacy Framework. (2019). Sport Australia.

Steene-Johannessen, J., Hansen, B. H., Dalene, K. E., Kolle, E., Northstone, K., Møller, N. C., Grøntved, A., Wedderkopp, N., Kriemler, S., Page, A. S., Puder, J. J., Reilly, J. J., Sardinha, L. B., Van Sluijs, E. M. F., Andersen, L. B., Van Der Ploeg, H., Ahrens, W.,

Flexeder, C., Standl, M., … On behalf of the Determinants of Diet and Physical Activity knowledge hub (DEDIPAC); International Children's Accelerometry Database (ICAD) Collaborators, IDEFICS Consortium and HELENA Consortium. (2020). Variations in accelerometry measured physical activity and sedentary time across Europe – harmonized analyses of 47,497 children and adolescents. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 38. https://doi.org/10.1186/s12966-020-00930-x

Stodden, D. F., Goodway, J. D., Langendorfer, S. J., Roberton, M. A., Rudisill, M. E., Garcia, C., & Garcia, L. E. (2008). A Developmental Perspective on the Role of Motor Skill Competence in Physical Activity: An Emergent Relationship. *Quest*, *60*(2), 290–306. https://doi.org/10.1080/00336297.2008.10483582

Stodden, D. F., True, L. K., Langendorfer, S. J., & Gao, Z. (2013). Associations Among Selected Motor Skills and Health-Related Fitness: Indirect Evidence for Seefeldt's Proficiency Barrier in Young Adults? *Research Quarterly for Exercise and Sport*, *84*(3), 397–403. https://doi.org/10.1080/02701367.2013.814910

Stodden, D., Langendorfer, S., & Roberton, M. A. (2009). The Association Between Motor Skill Competence and Physical Fitness in Young Adults. *Research Quarterly for Exercise and Sport*, *80*(2), 223–229. https://doi.org/10.1080/02701367.2009.10599556

Sulla-Torres, J., Calla Gamboa, A., Avendaño Llanque, C., Angulo Osorio, J., & Zúñiga Carnero, M. (2024). Classification of Motor Competence in Schoolchildren Using Wearable Technology and Machine Learning with Hyperparameter Optimization. *Applied Sciences*, *14*(2), 707. https://doi.org/10.3390/app14020707

Swain, T. A., McNarry, M. A., Runacres, A. W. H., & Mackintosh, K. A. (2023). The Role of Multi-Sensor Measurement in the Assessment of Movement Quality: A Systematic Review. *Sports Medicine*, *53*(12), 2477–2504. https://doi.org/10.1007/s40279-023-01905-1

Swift, D. L., Johannsen, N. M., Lavie, C. J., Earnest, C. P., & Church, T. S. (2014). The Role of Exercise and Physical Activity in Weight Loss and Maintenance. *Progress in Cardiovascular Diseases*, *56*(4), 441–447. https://doi.org/10.1016/j.pcad.2013.09.012

Tamplain, P., Webster, E. K., Brian, A., & Valentini, N. C. (2020). Assessment of Motor Development in Childhood: Contemporary Issues, Considerations, and Future Directions.

Journal of Motor Learning and Development, *8*(2), 391–409. https://doi.org/10.1123/jmld.2018-0028

Tannehill, D., MacPhail, A., Walsh, J., & Woods, C. (2015). What young people say about physical activity: The Children's Sport Participation and Physical Activity (CSPPA) study. *Sport, Education and Society*, *20*(4), 442–462. https://doi.org/10.1080/13573322.2013.784863

Temlali, T. Y., Lust, J., Klaperski-van Der Wal, S., & Steenbergen, B. (2024). The effectiveness of exergames in improving physical activity behaviour and physical literacy domains in adolescents with developmental coordination disorder and cerebral palsy: A scoping review. *Child: Care, Health and Development*, *50*(4), e13293. https://doi.org/10.1111/cch.13293

Tietjens, M., Dreiskaemper, D., Utesch, T., Schott, N., Barnett, L. M., & Hinkley, T. (2018). Pictorial Scale of Physical Self-Concept for Younger Children (P-PSC-C): A Feasibility Study. *Journal of Motor Learning and Development*, *6*(s2), S391–S402. https://doi.org/10.1123/jmld.2016-0088

Tipton, C. C., Telfer, S., Cherones, A., Gee, A. O., & Kweon, C. Y. (2019). THE USE OF MICROSOFT KINECT ™ FOR ASSESSING READINESS OF RETURN TO SPORT AND INJURY RISK EXERCISES: A VALIDATION STUDY. *International Journal of Sports Physical Therapy*, *14*(5), 724–730. https://doi.org/10.26603/ijspt20190724

Toftegaard-Stoeckel, J., Groenfeldt, V., & Andersen, L. B. (2010). Children's self-perceived bodily competencies and associations with motor skills, body mass index, teachers' evaluations, and parents' concerns. *Journal of Sports Sciences*, *28*(12), 1369–1375. https://doi.org/10.1080/02640414.2010.510845

Tölgyessy, M., Dekan, M., Chovanec, Ľ., & Hubinský, P. (2021). Evaluation of the Azure Kinect and Its Comparison to Kinect V1 and Kinect V2. *Sensors*, *21*(2), 413. https://doi.org/10.3390/s21020413

Topley, M., & Richards, J. G. (2020). A comparison of currently available optoelectronic motion capture systems. Journal of Biomechanics, 106, 109820. https://doi.org/10.1016/j.jbiomech.2020.109820

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & Mcdowell, M. (2008). Physical Activity in the United States Measured by Accelerometer. *Medicine & Science in Sports & Exercise*, *40*(1), 181–188. https://doi.org/10.1249/mss.0b013e31815a51b3

Trost, S. G., Loprinzi, P. D., Moore, R., & Pfeiffer, K. A. (2011). Comparison of Accelerometer Cut Points for Predicting Activity Intensity in Youth. *Medicine & Science in Sports & Exercise*, *43*(7), 1360–1368. https://doi.org/10.1249/MSS.0b013e318206476e

True, L., Brian, A., Goodway, J., & Stodden, D. (2017). Relationships Between Product- and Process-Oriented Measures of Motor Competence and Perceived Competence. *Journal of Motor Learning and Development*, *5*(2), 319–335. https://doi.org/10.1123/jmld.2016-0042

Turner, J. A., Chaaban, C. R., & Padua, D. A. (2024). Validation of OpenCap: A low-cost markerless motion capture system for lower-extremity kinematics during return-to-sport tasks. *Journal of Biomechanics*, *171*, 112200. https://doi.org/10.1016/j.jbiomech.2024.112200

Uhlrich, S. D., Falisse, A., Kidziński, Ł., Muccini, J., Ko, M., Chaudhari, A. S., Hicks, J. L., & Delp, S. L. (2023). OpenCap: Human movement dynamics from smartphone videos. PLOS Computational Biology, 19(10), e1011462. https://doi.org/10.1371/journal.pcbi.1011462

Ulrich, D. (2013). The test of gross motor development-3 (TGMD-3): Administration, scoring, and international norms. *Spor Bilimleri Dergisi*, *24*(2), 27–33

Ulrich, D. A. (2017). Introduction to the Special Section: Evaluation of the Psychometric Properties of the TGMD-3. *Journal of Motor Learning and Development*, *5*(1), 1–4. https://doi.org/10.1123/jmld.2017-0020

Ulrich, D. A., & Sanford, C. B. (2000). TGMD-2: Test of gross motor development. *Prod-Ed*

Valentini, N. C., Nobre, G. C., Zanella, L. W., Pereira, K. G., Albuquerque, M. R., & Rudisill, M. E. (2021). Test of Gross Motor Development–3 Validity and Reliability: A Screening Form. *Journal of Motor Learning and Development*, *9*(3), 438–455. https://doi.org/10.1123/jmld.2020-0061

Valentini, N. C., Zanella, L. W., & Webster, E. K. (2017). Test of Gross Motor Development—Third Edition: Establishing Content and Construct Validity for Brazilian Children. *Journal of Motor Learning and Development*, *5*(1), 15–28. https://doi.org/10.1123/jmld.2016-0002

Valentini, N., Rudisill, M., Bandeira, P., & Hastie, P. (2018). The development of a short form of the Test of Gross Motor Development-2 in Brazilian children: Validity and reliability. *Child: Care, Health and Development.*, *44*(5), 759–765. https://doi.org/10.1111/cch.12598

Van Beurden, E., Zask, A., Barnett, L. M., & Dietrich, U. C. (2002). Fundamental movement skills—How do primary school children perform? The 'Move it Groove it'program in rural Australia. *Journal of Science and Medicine in Sport*, *5*(3), 244–252

Van Der Kruk, E., & Reijne, M. M. (2018). Accuracy of human motion capture systems for sport applications; state-of-the-art review. *European Journal of Sport Science*, *18*(6), 806–819. https://doi.org/10.1080/17461391.2018.1463397

Van Sluijs, E. M. F., & Kriemler, S. (2016). Reflections on physical activity intervention research in young people – dos, don'ts, and critical thoughts. *International Journal of Behavioral Nutrition and Physical Activity*, *13*(1), 25. https://doi.org/10.1186/s12966-016-0348-z

Vandendriessche, J. B., Vaeyens, R., Vandorpe, B., Lenoir, M., Lefevre, J., & Philippaerts, R. M. (2012). Biological maturation, morphology, fitness, and motor coordination as part of a selection strategy in the search for international youth soccer players (age 15–16 years).

Journal of Sports Sciences, *30*(15), 1695–1703. https://doi.org/10.1080/02640414.2011.652654

Vandorpe, B., Vandendriessche, J., Lefevre, J., Pion, J., Vaeyens, R., Matthys, S., Philippaerts, R., & Lenoir, M. (2011). The KörperkoordinationsTest für Kinder: Reference values and suitability for 6–12-year-old children in Flanders. *Scandinavian Journal of Medicine & Science in Sports*, *21*(3), 378–388. https://doi.org/10.1111/j.1600-0838.2009.01067.x

Vedul-Kjelsås, V., Sigmundsson, H., Stensdotter, A. -K., & Haga, M. (2012). The relationship between motor competence, physical fitness and self-perception in children. *Child: Care, Health and Development*, *38*(3), 394–402. https://doi.org/10.1111/j.1365-2214.2011.01275.x

Viholainen, H., Ahonen, T., Lyytinen, P., Cantell, M., LicSSc, A. T., & Lyytinen, H. (2006). Early motor development and later language and reading skills in children at risk of familial dyslexia. *Developmental Medicine & Child Neurology*, *48*(5), 367–373. https://doi.org/10.1017/S001216220600079X

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, *1*, I-511-I–518. https://doi.org/10.1109/CVPR.2001.990517

Walkley, J., Holland, B. V., Treloar, R., & O'Connor, J. (1996). *Fundamental motor skills: A manual for classroom teachers*

Wang, P., Li, W., Ogunbona, P., Wan, J., & Escalera, S. (2018). RGB-D-based human motion recognition with deep learning: A survey. *Computer Vision and Image Understanding*, *171*, 118–139. https://doi.org/10.1016/j.cviu.2018.04.007

Warburton, D. E. R. (2006). Health benefits of physical activity: The evidence. *Canadian Medical Association Journal*, *174*(6), 801–809. https://doi.org/10.1503/cmaj.051351

Ward, B., Thornton, A., Lay, B., Chen, N., & Rosenberg, M. (2020). Can Proficiency Criteria Be Accurately Identified During Real-Time Fundamental Movement Skill Assessment? *Research Quarterly for Exercise and Sport*, *91*(1), 64–72. https://doi.org/10.1080/02701367.2019.1646852

Webster, E. K., & Ulrich, D. A. (2017). Evaluation of the Psychometric Properties of the Test of Gross Motor Development—Third Edition. *Journal of Motor Learning and Development*, *5*(1), 45–58. https://doi.org/10.1123/jmld.2016-0003

Weiss, M. R., & Amorose, A. J. (2005). Children's Self-Perceptions in the Physical Domain: Between- and Within-Age Variability in Level, Accuracy, and Sources of Perceived Competence. *Journal of Sport and Exercise Psychology*, *27*(2), 226–244.

https://doi.org/10.1123/jsep.27.2.226

Welk, G. J. (1999). The youth physical activity promotion model: A conceptual bridge between theory and practice. *Quest*, *51*(1), 5–23

Whitehead, M. (2013). Definition of physical literacy and clarification of related issues. *Icsspe Bulletin*, *65*(1.2). https://www.icsspe.org/sites/default/files/bulletin65_0.pdf#page=29

WHO. (2018, June 1). *Global action plan on physical activity 2018–2030: More active people for a healthier world*. World Health Organization. https://www.who.int/publications/i/item/9789241514187

WHO. (2025, July 31). *Cardiovascular diseases (CVDs)*. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Wochatz, M., Tilgner, N., Mueller, S., Rabe, S., Eichler, S., John, M., Völler, H., & Mayer, F. (2019). Reliability and validity of the Kinect V2 for the assessment of lower extremity rehabilitation exercises. *Gait & Posture*, *70*, 330–335. https://doi.org/10.1016/j.gaitpost.2019.03.020

Wood, G., Miles, C. A. L., Coyles, G., Alizadehkhaiyat, O., Vine, S. J., Vickers, J. N., & Wilson, M. R. (2017). A randomized controlled trial of a group-based gaze training intervention for children with Developmental Coordination Disorder. *PLOS ONE*, *12*(2), e0171782. https://doi.org/10.1371/journal.pone.0171782

Woods, C. B., Cormac, P., Saunders, J. A., O'Brien, W., Murphy, M. H., Duff, C., Farmer, O., Johnston, A., Connolly, S., & Sarahjane, B. (2018). *The Children's Sport Participation and Physical Activity Study 2018 (CSPPA 2018)*. Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland, Sport Ireland, and Healthy Ireland, Dublin, Ireland and Sport Northern Ireland, Belfast, Northern Ireland

Woods, C. B., Ng, K. W., Britton, U., McClelland, J. F., O'Keeffe, B., Sheikhi, A., McFlynn, P., Murphy, M., Goss, H., Behan, S., Philpott, C., Lester, D., Adamakis, M., Costa, J., Coppinger, T., Connolly, S., Belton, S., & O'Brien, W. (2023). *Children's Sport Participation and Physical Activity Study 2022 (CSPPA 2022)*. Physical Activity for Health Research Centre, Department of Physical Education and Sport Sciences, University of Limerick, Limerick, Ireland, Sport Ireland and Healthy Ireland, Dublin, Ireland and Sport Northern Ireland, Belfast, Northern Ireland. https://doi.org/10.34961/RESEARCHREPOSITORY-UL.23609157

Woods, C. B., Tannehill, D., Quinlan, A., Moyna, N., & Walsh, J. (2010). *The Children's Sport Participation and Physical Activity Study (CSPPA). Research Report No 1*. School of Health and Human Performance, Dublin City University and The Irish Sports Council, Dublin, Ireland. Yaqoob, A., Musheer Aziz, R., & verma, N. K. (2023). Applications and Techniques of Machine Learning in Cancer Classification: A Systematic Review. *Human-Centric Intelligent Systems*, *3*(4), 588–615. https://doi.org/10.1007/s44230-023-00041-3

Yang, D., & Jung, E. S. (2025). Evaluation of camera configurations of OpenPose-based 3D motion capture system for construction tasks. Measurement, 247, 116826. https://doi.org/10.1016/j.measurement.2025.116826

Ye, S., Lee, J. E., Stodden, D. F., & Gao, Z. (2018). Impact of Exergaming on Children's Motor Skill Competence and Health-Related Fitness: A Quasi-Experimental Study. *Journal of Clinical Medicine*, *7*(9), 261. https://doi.org/10.3390/jcm7090261

Yeom, S., Giacomelli, I., Fredrikson, M., & Jha, S. (2018). Privacy Risk in Machine Learning: Analyzing the Connection to Overfitting. *2018 IEEE 31st Computer Security Foundations Symposium (CSF)*, 268–282. https://doi.org/10.1109/CSF.2018.00027

Yoon, J. (2021). Forecasting of Real GDP Growth Using Machine Learning Models: Gradient Boosting and Random Forest Approach. *Computational Economics*, *57*(1), 247–265. https://doi.org/10.1007/s10614-020-10054-w

Yun, J., & Ulrich, D. A. (2002). Estimating measurement validity: A tutorial. *Adapted Physical Activity Quarterly*, *19*(1), 32–47

Zhang, L., Sturm, J., Cremers, D., & Lee, D. (2012). Real-time human motion tracking using multiple depth cameras. *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2389–2395. https://doi.org/10.1109/IROS.2012.6385968

Zhang, T., Thomas, K., & Weiller, K. (2015). Predicting Physical Activity in 10-12 Year Old Children: A Social Ecological Approach. *Journal of Teaching in Physical Education*, *34*(3), 517–536. https://doi.org/10.1123/jtpe.2013-0195

Zheng, Y., Ye, W., Korivi, M., Liu, Y., & Hong, F. (2022). Gender Differences in Fundamental Motor Skills Proficiency in Children Aged 3–6 Years: A Systematic Review and Meta-Analysis. *International Journal of Environmental Research and Public Health*, *19*(14), 8318. https://doi.org/10.3390/ijerph19148318

Zuvela, F., Bozanic, A., & Miletic, D. (2011). POLYGON – A new fundamental movement skills test for 8 year old children: Construction and validation. *Journal of Sports Science & Medicine*, *10*(1), 157

Appendices

Appendix A DCU REC Formal Letter of Approval

Ollscoil Chathair Bhaile Átha Cliath
Dublin City University

Colm O'Donaghue
School of Health and Human Performance

17th January 2025

REC Reference:	**DCUREC/2024/243 (Expedited Review)**
Proposal Title:	**Validating the Use of a Motion Capture System 'Freemocap' for Assessment of Locomotor Skill Proficiency**
Applicant(s):	Colm O'Donaghue, Stephen Behan, Michael Scriney, Sarahjane Belton

Dear Colleague(s),

Thank you for your application to DCU Research Ethics Committee (REC). Further to its review by the committee and resting on the assumption of information accuracy and completeness, DCU REC is pleased to issue ethical approval for this research project. Please include reference to this approval in all materials used to recruit research participants.

Researchers are responsible for ensuring that the research project to which this ethical approval refers is carried out as specifically described in the application form. Should modifications to the research project be required at a later stage, researchers must submit a research amendment application form to REC for approval, prior to the implementation of modifications.

Please note that it is the responsibility of the PI to ensure that any other DCU compliance requirements relevant to the research project, such as those related to data protection, insurance, health and safety, or legal issues, are fully met in advance of initiating the project.

As part of DCU REC's ongoing monitoring process, a research progress report may be required. DCU REC will request this report from the PI as appropriate.

DCU REC wishes you every success in your research.

Yours sincerely,

DCU Research & Innovation

Taighde & Nuálaíocht Tacaíocht
Ollscoil Chathair Bhaile Átha Cliath,
Baile Átha Cliath, Éire

Research & Innovation Support
Dublin City University,
Dublin 9, Ireland

T +353 1 700 8000
F +353 1 700 8002
E research@dcu.ie
www.dcu.ie

Dr Melrona Kirrane
Chairperson
DCU Research Ethics Committee

Note: Please retain this approval letter for future publication purposes. Research students should include this letter as a thesis appendix.

Appendix B TGMD-3 Criterion Operationalisation, Features and Heuristic Rules

Gallop					
Criterion	**Metric**	**What's Compared**	**Pass Condition**	**Configurable Variable**	**Default**
lift_window	Stepping foot heel rise	`body_[lead]_heel_z` vs baseline	Heel rises above baseline within 2-second window	`FPS`	30 fps (window = fps × 2 frames)
arm_bent_level — angle	Mean elbow flexion angle	`left/right_elbow_extension_flexion` vs 90°	Both elbows within ±30° of 90° (i.e. 60°–120°)	`ELBOW_ANGLE_TOLERANCE`	30.0°
arm_bent_level — height	Elbow height vs hip	`body_[lr]_elb ow_z` vs `body_[lr]_hip_z` ± span	Both elbows within 40% of hip-to-shoulder span from hip	`ELBOW_HEIGHT_TOLE RANCE`	0.4 (fraction of span)
arm_bent_level — window	Analysis window	—	±5 frames around step peak	`PEAK_THRESHOLD_WINDOW`	5 frames
trail_foot_behid	Heel x-position	`mean(body_trail_heel_x)` vs `mean(body_lead_heel_x)`	Trail heel mean x < lead heel mean x (any separation)	—	0 m
both_feet_off_surface — lift	Heel rise above floor	`body_[lr]_he el_z` vs within-window floor	Both heels > 0.02 m above within-window minimum	`FLIGHT_LIFT_THRESHOLD`	0.02 m
both_feet_off_surface — duration	Consecutive airborne frames	—	≥ 2 consecutive frames both heels are off surface	`FLIGHT_MIN_FRAMES`	2 frames
Hop					
Criterion	**Metric**	**What's Compared**	**Pass Condition**	**Configurable Variable**	**Default**
detect_lift_window	Dominant foot heel rise	`body_[dom]_heel_z` vs baseline	Heel rises within 2-second window of	`FPS`	30 fps (window = fps × 2 frames)

			hop start		
non_dominant_leg_movement	Non-dom heel peak timing	`argmax(body_[non_dom]_heel_z)` vs `peak_frame`	Non-dominant heel_z peak within ±10 frames of dominant foot apex	`NON_DOM_PEAK_TOLERANCE`	10 frames
non_dominant_foot_behind	Heel x-position	`mean(body_[non_dom]_heel_x)` vs `mean(body_[dom]_heel_x)`	Non-dominant heel mean x < dominant heel mean x (any separation)	—	0 m
elbow_bends	Mean elbow flexion	`left/right_elbow_extension_flexion` vs threshold	Both elbows > 22.5° flexion within ±5 frames of hop apex	`ELBOW_BEND_THRESHOLD`	22.5°
arm_swing — window	Analysis window	—	±5 frames around hop peak	`PEAK_THRESHOLD_WINDOW`	5 frames
arm_swing	Wrist peak timing	`argmax(body_[lr]_wrist_z)` vs `peak_frame`	Both wrist_z peaks within ±10 frames of dominant foot apex	`NON_DOM_PEAK_TOLERANCE`	10 frames
Horizontal Jump					
Criterion	**Metric**	**What's Compared**	**Pass Condition**	**Configurable Variable**	**Default**
knees_flexed_arms_behind — knee	Knee flexion angle	`left/right_knee_extension_flexion` vs threshold	Both knees > 20.0° flexion in windup phase (start → peak)	`KNEE_FLEX_THRESHOLD`	20.0°
knees_flexed_arms_behind — arms	Wrist x vs hip x	`mean(body_[lr]_wrist_x)` vs `mean(body_[lr]_hip_x)`	Mean wrist x < mean hip x in windup phase (any separation)	—	0 m
arms_forward_upward — height	Wrist z vs shoulder z	`body_[lr]_wrist_z` vs `body_[lr]_shoulder_z`	Both wrists above shoulder height at takeoff (any separation)	—	0 m

arms_forward_upward — forward	Wrist x vs shoulder x	`body_[lr]_wrist_x` vs `body_[lr]_shoulder_x`	Both wrists in front of shoulder x at takeoff (any separation)	—	0 m
arms_forward_upward — window	Analysis window	—	±5 frames around jump peak	`PEAK_THRESHOLD_WINDOW`	5 frames
feet_together — rise	Left/right takeoff sync	`left_heel_rise_frame` vs `right_heel_rise_frame`	Left and right heel rise within ±5 frames of each other	`FEET_TOGETHER_TOLERANCE`	5 frames
feet_together — land	Left/right landing sync	`left_heel_land_frame` vs `right_heel_land_frame`	Left and right heel land within ±5 frames of each other	`FEET_TOGETHER_TOLERANCE`	5 frames
arms_downward	**Wrist z vs hip** z	**`mean(body_[lr]_wrist_z)` vs `mean(body_[lr]_hip_z)`**	**Both wrists below hip height in landing phase (any separation)**	**—**	**0 m**

Appendix C Descriptive Statistics Across Camera Configurations

Skill	Variable	Camera	Mean	Median	SD	Min	Max
Hop	Left Elbow Y	2	734	685	218	324	1264
Hop	Left Elbow Y	3	709	710	208	356	1222
Hop	Left Elbow Y	4	716	730	215	368	1222
Hop	Right Elbow Y	2	191	84.8	274	-99.2	728
Hop	Right Elbow Y	3	158	141	204	-115	693
Hop	Right Elbow Y	4	169	170	206	-118	687
Hop	Left Ankle Y	2	512	429	237	94.5	1004
Hop	Left Ankle Y	3	486	471	206	97.3	957
Hop	Left Ankle Y	4	494	494	209	117	954
Hop	Right Ankle Y	2	373	272	256	-20.1	901
Hop	Right Ankle Y	3	346	335	211	-29.4	860
Hop	Right Ankle Y	4	355	354	213	-4.82	857
Gallop	Left Elbow Y	2	723	687	210	389	1208
Gallop	Left Elbow Y	3	705	747	201	293	1165
Gallop	Left Elbow Y	4	716	776	214	264	1164
Gallop	Right Elbow Y	2	203	73.2	298	-92.0	734

Gallop	Right Elbow Y	3	198	190	198	-210	697
Gallop	Right Elbow Y	4	212	216	206	-230	693
Gallop	Left Ankle Y	2	523	447	234	248	1033
Gallop	Left Ankle Y	3	513	530	189	151	984
Gallop	Left Ankle Y	4	526	558	198	127	980
Gallop	Right Ankle Y	2	403	270	269	103	948
Gallop	Right Ankle Y	3	394	395	201	5.35	898
Gallop	Right Ankle Y	4	408	427	209	-18.5	893
Horizontal Jump	Left Elbow Y	2	679	694	51.4	570	739
Horizontal Jump	Left Elbow Y	3	665	684	77.0	500	765
Horizontal Jump	Left Elbow Y	4	670	693	82.6	494	777
Horizontal Jump	Right Elbow Y	2	193	161	91.2	86.9	367
Horizontal Jump	Right Elbow Y	3	176	201	71.7	13.6	241
Horizontal Jump	Right Elbow Y	4	179	206	76.1	6.81	251
Horizontal Jump	Left Ankle Y	2	595	608	55.8	493	690
Horizontal Jump	Left Ankle Y	3	582	611	73.4	443	674
Horizontal Jump	Left Ankle Y	4	587	614	79.5	435	676
Horizontal Jump	Right Ankle Y	2	294	250	94.5	174	469
Horizontal Jump	Right Ankle Y	3	278	288	74.0	108	395
Horizontal Jump	Right Ankle Y	4	283	297	77.7	99.3	386

Mean = arithmetic mean; SD = standard deviation; Min = minimum; Max = maximum.

Appendix D Intraclass Correlation Coefficient Absolute Agreement Results Across Skills and Camera Configurations

Skill	Variable	Comparison	ICC(2,1)	95% CI	Interpretation
Hop	Left Ankle (Y)	3 vs 4	0.995	[0.99, 1.00]	Excellent
Hop	Left Ankle (Y)	2 vs 4	0.513	[0.09, 0.78]	Moderate
Hop	Left Ankle (Y)	2 vs 3	0.584	[0.20, 0.81]	Moderate
Hop	Right Ankle (Y)	3 vs 4	0.995	[0.99, 1.00]	Excellent
Hop	Right Ankle (Y)	2 vs 4	0.436	[-0.01, 0.73]	Poor
Hop	Right Ankle (Y)	2 vs 3	0.511	[0.09, 0.77]	Moderate
Hop	Left Elbow (Y)	3 vs 4	0.995	[0.99, 1.00]	Excellent
Hop	Left Elbow (Y)	2 vs 4	0.632	[0.27, 0.84]	Moderate
Hop	Left Elbow (Y)	2 vs 3	0.689	[0.37, 0.86]	Moderate
Hop	Right Elbow (Y)	3 vs 4	0.994	[0.98, 1.00]	Excellent
Hop	Right Elbow (Y)	2 vs 4	0.274	[-0.20, 0.64]	Poor
Hop	Right Elbow (Y)	2 vs 3	0.356	[-0.10, 0.69]	Poor
Gallop	Left Ankle (Y)	3 vs 4	0.989	[0.97, 1.00]	Excellent
Gallop	Left Ankle (Y)	2 vs 4	0.204	[-0.28, 0.59]	Poor
Gallop	Left Ankle (Y)	2 vs 3	0.316	[-0.16, 0.66]	Poor
Gallop	Left Elbow (Y)	3 vs 4	0.991	[0.98, 1.00]	Excellent
Gallop	Left Elbow (Y)	2 vs 4	0.369	[-0.09, 0.70]	Poor
Gallop	Left Elbow (Y)	2 vs 3	0.461	[0.02, 0.75]	Poor

Gallop	Right Ankle (Y)	3 vs 4	0.989	[0.97, 1.00]	Excellent
Gallop	Right Ankle (Y)	2 vs 4	0.123	[-0.36, 0.54]	Poor
Gallop	Right Ankle (Y)	2 vs 3	0.237	[-0.24, 0.61]	Poor
Gallop	Right Elbow (Y)	3 vs 4	0.988	[0.96, 1.00]	Excellent
Gallop	Right Elbow (Y)	2 vs 4	-0.019	[-0.48, 0.43]	Poor
Gallop	Right Elbow (Y)	2 vs 3	0.095	[-0.38, 0.52]	Poor
Horizontal Jump	Left Ankle (Y)	3 vs 4	0.993	[0.98, 1.00]	Excellent
Horizontal Jump	Left Ankle (Y)	2 vs 4	-0.331	[-0.71, 0.15]	Poor
Horizontal Jump	Left Ankle (Y)	2 vs 3	-0.287	[-0.67, 0.19]	Poor
Horizontal Jump	Left Elbow (Y)	3 vs 4	0.993	[0.97, 1.00]	Excellent
Horizontal Jump	Left Elbow (Y)	2 vs 4	-0.114	[-0.55, 0.35]	Poor
Horizontal Jump	Left Elbow (Y)	2 vs 3	-0.082	[-0.52, 0.37]	Poor
Horizontal Jump	Right Ankle (Y)	3 vs 4	0.992	[0.98, 1.00]	Excellent
Horizontal Jump	Right Ankle (Y)	2 vs 4	-0.501	[-0.81, -0.05]	Poor
Horizontal Jump	Right Ankle (Y)	2 vs 3	-0.412	[-0.76, 0.06]	Poor
Horizontal Jump	Right Elbow (Y)	3 vs 4	0.994	[0.98, 1.00]	Excellent
Horizontal Jump	Right Elbow (Y)	2 vs 4	-0.809	[-0.98, -0.49]	Poor
Horizontal Jump	Right Elbow (Y)	2 vs 3	-0.762	[-0.95, -0.41]	Poor

Appendix E Rules Engine Performance by Skill, Camera Configuration and Feature

Skill	Feature	Camera Configuration	Support True	Support False	Accuracy	Precision	Recall	F1-Score
Hop	Lift Window	2 Cameras	3	0	1.00	1.00	1.00	1.00
Hop	Lift Window	3 Cameras	3	0	1.00	1.00	1.00	1.00
Hop	Lift Window	4 Cameras	3	0	1.00	1.00	1.00	1.00
Hop	Non-Dominant Leg Movement	2 Cameras	3	0	0.33	1.00	0.33	0.50
Hop	Non-Dominant Leg Movement	3 Cameras	3	0	0.00	0.00	0.00	0.00
Hop	Non-Dominant Leg Movement	4 Cameras	3	0	0.00	0.00	0.00	0.00
Hop	Non-Dominant Foot Behind	2 Cameras	3	0	1.00	1.00	1.00	1.00
Hop	Non-Dominant Foot Behind	3 Cameras	3	0	1.00	1.00	1.00	1.00
Hop	Non-Dominant Foot Behind	4 Cameras	3	0	1.00	1.00	1.00	1.00
Hop	Elbow Bends	2 Cameras	2	1	1.00	1.00	1.00	1.00
Hop	Elbow Bends	3 Cameras	2	1	1.00	1.00	1.00	1.00
Hop	Elbow Bends	4 Cameras	2	1	1.00	1.00	1.00	1.00
Hop	Arm Swing	2 Cameras	3	0	0.33	1.00	0.33	0.50
Hop	Arm Swing	3 Cameras	3	0	0.00	0.00	0.00	0.00
Hop	Arm Swing	4 Cameras	3	0	0.00	0.00	0.00	0.00

Gallop	Lift Window	2 Cameras	8	0	1.00	1.00	1.00	1.00
Gallop	Lift Window	3 Cameras	6	0	1.00	1.00	1.00	1.00
Gallop	Lift Window	4 Cameras	7	0	1.00	1.00	1.00	1.00
Gallop	Arm Bent Level	2 Cameras	8	0	0.00	0.00	0.00	0.00
Gallop	Arm Bent Level	3 Cameras	6	0	0.00	0.00	0.00	0.00
Gallop	Arm Bent Level	4 Cameras	7	0	0.00	0.00	0.00	0.00
Gallop	Trail Foot Behind	2 Cameras	6	2	0.38	1.00	0.17	0.29
Gallop	Trail Foot Behind	3 Cameras	5	1	0.50	1.00	0.40	0.57
Gallop	Trail Foot Behind	4 Cameras	6	1	0.43	1.00	0.33	0.50
Gallop	Both Feet Off Surface	2 Cameras	5	3	0.63	0.63	1.00	0.77
Gallop	Both Feet Off Surface	3 Cameras	4	2	0.67	0.67	1.00	0.80
Gallop	Both Feet Off Surface	4 Cameras	4	3	0.57	0.57	1.00	0.73
Horizontal Jump	Knees Flexed, Arms Behind	2 Cameras	3	0	0.00	0.00	0.00	0.00
Horizontal Jump	Knees Flexed, Arms Behind	3 Cameras	23	0	0.09	1.00	0.09	0.16
Horizontal Jump	Knees Flexed, Arms Behind	4 Cameras	23	0	0.09	1.00	0.09	0.16
Horizontal Jump	Arms Forward Upward	2 Cameras	3	0	0.33	1.00	0.33	0.50
Horizontal Jump	Arms Forward Upward	3 Cameras	23	0	0.43	1.00	0.43	0.61

Horizontal Jump	Arms Forward Upward	4 Cameras	23	0	0.43	1.00	0.43	0.61
Horizontal Jump	Feet Together	2 Cameras	3	0	0.33	1.00	0.33	0.50
Horizontal Jump	Feet Together	3 Cameras	23	0	0.35	1.00	0.35	0.52
Horizontal Jump	Feet Together	4 Cameras	23	0	0.43	1.00	0.43	0.61
Horizontal Jump	Arms Downward	2 Cameras	3	0	0.33	1.00	0.33	0.50
Horizontal Jump	Arms Downward	3 Cameras	18	0	0.11	1.00	0.11	0.20
Horizontal Jump	Arms Downward	4 Cameras	18	0	0.11	1.00	0.11	0.20