2021-09

Peer to Peer Learning - Using Structured Video as a Tool to Improve Performance with Middle School Children

Thacker, Anna

UNIVERSITY OF CALGARY

Peer to Peer Learning – Using Structured Video as a Tool to Improve Performance with Middle School Children

by

Anna Maria Thacker

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN KINESIOLOGY

CALGARY, ALBERTA

SEPTEMBER, 2021

Abstract

Motor competence is observable through an individual's fundamental movement skill development. If deficiencies are not detected at an early age, children may experience lifelong problems engaging in physical activity, leading to increased sedentary behaviours. Therefore, it is important to assess children's motor competence using reliable assessment tools. Research has shown that incorporating digital technologies into physical education has the potential to improve motor skill performance and facilitate real-time collaboration with peers. Additionally, multiple studies have concluded that the reciprocal nature and observational aspect of peer-to-peer learning results in students performing at an equal or improved level with less practice time.

Move Improve® is a video analysis tool designed to evaluate an individual's ability to perform a variety of fundamental movement skills, by breaking down physical skills into easily comprehended components. Through video analysis, the primary purpose of this study was to explore the impact that order of performance has on middle school students' performance of fundamental movement skills within a peer-to-peer learning model. Order of performance refers to the order in which a student performed a skill while paired up with a peer. The secondary purpose of this study was to determine the intra-rater and inter-rater reliability of the Move Improve® assessment tool in middle school students when performing two fundamental movement skills.

Eighteen students (eight males, ten females) in grades 5-8 participated in this study. Using Move Improve®, the students completed a standing jump and a hollow body roll in pairs assigned to order of performance (evaluator/performer). Four groups of raters (students, two independent raters, and a three-person consensus panel) evaluated a total of 34 videos which included 18 standing jump videos and 16 hollow-body roll videos. Overall performance scores were calculated using the Move Improve® assessment tool. Reliability between the four groups of raters was determined using intraclass correlation coefficients and percent agreement.

There was a significant difference in standing jump scores, where students who performed second had higher scores than their peers who went first. Although not statistically significant, results for hollow body roll also showed higher mean performance scores for students who went second compared to those who performed first. The intra-rater reliability of the consensus panel was found to be good for both skills. Standing jump demonstrated good inter-rater reliability between all raters, while hollow body roll had poor inter-rater reliability for the student raters compared to the individual raters and the consensus panel. Further analysis into the individual task components demonstrated inconsistencies depending on the rater, limiting the reliability of the skill assessment based on the components.

These findings suggest that the order of performance within a peer-to-peer learning model may have a significant effect on performance scores for some, but not all fundamental movement skills. Reasons for the discrepancy between the performance scores may be due to a combination of skill familiarity and training of observational learning. The consensus panel results from this study are promising because both fundamental movement skills exhibited good intra-rater reliability. The Move Improve® assessment tool can be used to improve learning, but the quality of the content and the level of training can impact the reliability of the tool. Improving training protocols for students and professionals may improve the reliability when using the Move Improve® assessment tool.

Acknowledgements

I would like to express my sincere gratitude to:

- My supervisor, Dr. Larry Katz, for the mentorship, continued motivation and constant support throughout this entire process. His knowledge and expertise inspired me to learn more and made be a better researcher. I am not the same person I was when I started in his lab. Thank you for believing in me, even when I did not.
- My co-supervisor, Dr. Nicole Culos-Reed, for stepping in when Dr. Katz officially retired. I would not have been able to complete this thesis without your support and patience.
- My supervisory and defence committee members, Dr. Sharon Friesen, Dr. Nadine Van-Wyk, and Dr. Ehud Sharlin. Thank you for your advice and for contributing your unique perspectives to this thesis.
- Dr. Tak Fung, for his guidance and knowledge in the realm of statistical analysis. Thank you for answering all my questions and teaching me the processes along the way. It was the translation of this knowledge that allowed me to understand the data obtained.
- Arsalan and Jennifer, I have no words to thank you for all your help. When Dr. Katz suggested adding two research assistants to the project, I could not have imagined how great it would be. You were both incredibly helpful with data collection, statistical analysis, and publication.
- My cohort of graduate students both in the STRL and in other labs in the HPL and Faculty of Kinesiology. It was a pleasure getting to know all of you and I'm glad we were able to share this time together.
- My family, Adriana, David, and Paul, for being my support system throughout this entire process. Adriana, your nutritious homecooked meals were always greatly appreciated. Paul, thank you for helping me edit and finalize this project. David, thanks for helping me to put things into perspective and keeping things lighthearted. Thank you for supporting me along my educational journey.
- Dan, your unconditional love, and support are what made this MSc possible. You are my biggest fan and have remained by my side throughout the long and difficult process of

completing this dissertation. These past years have taught us that we are stronger together, and I can’t wait to see where life takes us.

Dedication

To Squish. Thanks for all your endless supply of love and purrs.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AR	Augmented Reality
FMS	Fundamental Movement Skills
FSS	Fundamental Sport Skills
HBR	Hollow Body Roll
ICC	Intraclass Correlation Coefficient
sMI	Move Improve®
PE	Physical Education
P2P	Peer-to-Peer
SJ	Standing Jump
STRL	Sport Technology and Research Lab
TGMD	Test of Gross Motor Development – First Edition
TGMD-2	Test of Gross Motor Development – Second Edition
VFB	Video Feedback
VR	Virtual Reality

Chapter One: Introduction

1.1 General Background

The importance of developing children's fundamental movement skills (FMS) to facilitate longer-term engagement with physical activity is well established. Goodway et al. (2014) claim that children who are more highly skilled and motor competent will self-select higher levels of physical activity. Failure to master FMS in childhood acts as a barrier to lifelong participation in physical activity, as it decreases the number of physical activity options available to the individual in adolescence and adulthood (Cliff et al., 2012; Goodway et al., 2014; Robinson et al., 2015; Stodden et al., 2009). Unfortunately, many children are late in mastering the FMS that enable them to feel confident about their competence to engage in sports and other physical activities (Booth et al., 1999; Erwin & Castelli, 2008; O'Brien et al., 2016; Okely & Booth, 2004). As such, physical education (PE) teachers may need to reconsider their methods for motivating and engaging students (Wyant & Baek, 2019). This is particularly evident as some teachers report difficulties in engaging students within their classes and physical activity programs (Jenkinson & Benson, 2010). Given adolescents' desires to interact with and thus be influenced by peers, as well as their accessibility to each other in the school environment, the argument for peer-assisted learning in PE is both practical and compelling (Byra & Marks, 1993; Ellis et al., 2009).

Peer-assisted learning, also referred to as peer-to-peer learning (P2P), in PE is one potential strategy that could be used to overcome some aspects of the school environment that impede student enjoyment and learning. More specifically, it is helpful in addressing the teachers' difficulty in directly observing and instructing each individual student (Metzler, 2017). The advantage of this reciprocal teaching style is that all students can receive increased and immediate feedback during practice (Chatoupis, 2018; Mosston & Ashworth, 2008; Mohr & Townsend, 2002). However, practice time is reduced, giving incorrect feedback is possible, and conflicts, verbal abuse, or off-task conversations between students may occur (Byra & Marks, 1993; Chatoupis, 2018; Ernst & Byra, 1998; Mosston & Ashworth, 2008). As a result, teachers must ensure that the use of the reciprocal style benefits all the students without inhibiting individual achievement (Hennings et al., 2010). Research shows that because students must observe and evaluate their peers, the reciprocal style of the P2P model enables them to perform at an equal or higher level with half the practice time (Goldberger et al., 1982; Goldberger & SueSee, 2020; Kolovelonis et al., 2011). As a result, each student's opportunity to receive higher amounts of feedback and reinforcements from peers and teachers may result in improved FMS performance in PE (Alstot, 2018; Mohnsen, 2012).

Observational learning is defined as the ability to learn a motor skill by seeing another person perform the skill (Cross et al., 2009). According to research, observational practice can make important and unique contributions to learning, especially when combined with physical practice (Shea et al., 2000). Individuals who observe a skill they are familiar with or have previously attempted will benefit from observational learning (i.e., watching a peer perform the skill, or watching an instructional video of the skill; Shea et al., 2000), as this learning pathway shares neural paths with physical learning networks (i.e., similar, or same skill done during active play; Rizzolatti & Fogassi, 2014). Observing another person perform a desired skill gives a mental blueprint for how the skill should look, as well as segmental timing and sequential movements of various joints to perform the skill (Blandin & Proteau, 2000; Cross et al., 2009). As a result, observational learning has been shown to promote the learning of a wide range of motor skills (Hodges et al., 2007; Lago-Rodrguez et al., 2014; McCullagh et al., 1989; Ste-Marie et al., 2012; Vogt & Thomaschke, 2007).

Research has examined the significance of PE and how, for certain students, PE is either unimportant or has negative connotations (Sykes & McPhail, 2008; Tannehill et al., 2015). Evidence shows that the reasons many students dislike traditional PE programs are because those students have poor physical fitness, lack the necessary FMS, or are often just not interested (Morey & Karp, 1998). Introducing digital technologies as a new set of practices has the potential to change the way that students participate and therefore make sense of PE (Koekoek & van Hilvoorde, 2018; Wyant & Baek, 2019). Indeed, this is consistent with expectations that technology can increase interest and help young people to connect their experiences in PE with their technologically saturated lives (Papastergiou, 2009). Because children are attracted to digital technologies, by integrating them into PE they can become physical activity tools which may be used to create more engaging lesson plans (Casey et al., 2017; Koekoek & van Hilvoorde, 2018; Mohnsen, 2012; Spittle, 2021; Tanaka et al., 2018; Wyant & Baek, 2019).

Assessment in PE is an important aspect of teaching and learning motor skills, and it should be seen as an integral part of instruction (Chng & Lund, 2018; Payne & Issacs, 2020; Veal, 1995). Continuous assessments give a reliable reflection of students' progress, ensure consistency in grading between teachers and students, and generate objective data that can be shared with students, parents, and administrators (Anderson & Goode, 1997; Lund & Tannehill, 2014; Wright & van der Mars, 2004). Furthermore, assessment in a PE environment has demonstrated numerous advantages for both teachers and students. First, teachers can use assessments as a guide for what they want their students to learn (Chng & Lund, 2018; Veal, 1992). As a result, assessment helps to improve students' ability to acquire knowledge (Collier, 2011; Hensley, 1997; Wright & van der Mars, 2004). Second, teachers can use assessments to re-define goals and objectives to best meet the needs of all their students (Chng & Lund, 2018; Melograno, 1997; Smith, 1997). Depending on how the students perform on an assessment, teachers can adjust unit plans to see what critical components (basic movements that are needed for students to perform a skill successfully) really need to be taught (Barnett et al., 2014; Lund, 1997; Wright & van der Mars, 2004). Third, assessment allows students to interpret information regarding their performance by providing feedback on their progress and mastery of the skill (Chng & Lund, 2018; Johnson, 2004; Mosston & Ashworth, 2008; Ward & Lee, 2005). Assessment enables students to develop an understanding of their own strengths and weaknesses, allowing them to learn how to improve. Finally, students develop a cognitive understanding of all aspects and components of each skill being assessed. For example, using peer assessment requires cognitive understanding of the skill components by the student-evaluator (Chng & Lund, 2018; Johnson, 2004; Ward & Lee, 2005). By developing cognitive understanding of skills students gain a better understanding of the components necessary for psychomotor execution.

There are two types of motor skill assessments: product-oriented (outcomes) and process-oriented (skill development) (Foulkes et al., 2015). Process-oriented assessments are highly beneficial when first learning a skill as they provide insight into specific components of the skill that are meeting expectations or may need improvement (Barnett et al., 2014); product-oriented assessments are beneficial for improving outcomes such as speed and accuracy. Physical educators should use authentic and formative process-oriented assessments (Chng & Lund, 2018). Authentic assessment is designed to take place in a real-life setting and emphasizes validity, fairness, and learning development (Green, 2021; Lund, 1997; Panicucci & Hunt, 2002). Formative assessments provide information and feedback to teachers and students about the students' progress toward a learning goal (Chng & Lund, 2018; Veal, 1992). To evaluate movement or motor performance, several process-oriented assessments are used, with most skill assessments being developed and tested in students for reliability and validity (Goodway et al., 2014). Some factors, however, have an impact on the accuracy and reliability of motor skill assessments. As a result, an assessment tool's intra-rater and inter-rater reliability are important factors to consider when determining its objectivity.

1.2 Rationale

As the integration and prevalence of digital technologies in PE continues to rise, more research is needed to evaluate the effectiveness of multimedia-supported teaching on motor skill development. Research has shown that the use of digital technologies, such as the iPad, has the potential to improve motor skill performance (Hung et al., 2018; Weir & Connor, 2009). Furthermore, multiple studies have demonstrated that P2P learning can be utilized in PE to achieve increased performance scores (Chatoupis, 2015; Ernst & Byra, 1998; Goldberger et al.,1982; Goldberger & SueSee, 2020; Kolovelonis et al., 2011). The combination of digital technology, observational learning, and P2P learning provides an ideal solution for the mastery of FMS, as it allows students to receive rapid feedback on their performance. However, the effectiveness of this combination facilitated through mobile applications in PE settings has not been well researched, especially as it relates to determining the impact that order of performance has on middle school students' performance of FMS within a P2P learning model. Additionally, there is a gap in the literature concerning the reliability of digital technologies to measure FMS performance in a middle school PE environment.

1.3 Research Objectives

The primary objective of this research is to explore the impact that order of performance within a peer-to-peer learning model has on performance outcomes of fundamental movement skills in middle school students. A secondary objective is to determine the inter-rater and intra-rater reliability of the application, Move Improve® (MI), in middle school children when performing two fundamental movement skills, a standing jump (SJ) and a hollow body roll (HBR).

1.4 Research Questions

The two research questions being asked in this investigation are:

1) What impact does the order of performance within a peer-to-peer learning model have on the acquisition of fundamental movement skills in middle school students?

2) What is the inter-rater and intra-rater reliability of the MI application when assessing the performance outcomes of middle school students' fundamental movement skills?

1.5 Research Hypotheses

1) It is hypothesized that students who performed second (order 2) will have greater performance scores than students who performed first (order 1).

2) It is hypothesized that the internal reliability of the MI application is reliable.

1.6 Thesis Overview and Chapter Summary

This document is a manuscript-based thesis, composed of five chapters. Chapter one provides an introduction with background information and presents the thesis research objectives. The following chapter contains a literature review that will provide an overview of the background information necessary to understand the research carried out in this thesis. The third and fourth chapters are original manuscripts as follows:

1. Chapter Three: Thacker, A., Khawaja, A., Ho, J., & Katz, L. (2021). Peer to Peer Learning: The Impact of Order of Performance on Learning Fundamental Movement Skills Through Video Performance Analysis with Middle School Children.
2. Chapter Four: Thacker, A., Ho, J., Khawaja, A., & Katz, L. (2021). Intra-rater and Inter-rater Reliability of a Peer-to-Peer Video Performance Analysis Tool to Measure Fundamental Movement Skills with Middle School Children.

The final chapter summarizes the findings of this research while also discussing limitations. General conclusions from both manuscripts are discussed, and possible future work is proposed.

Finally, Appendix A contains the extended methodology which was used to conduct this research.

This thesis is based on a collection of manuscripts, and therefore some redundancy may be present within the sections.

Chapter Two: Literature Review

2.1 Physical Literacy

2.1.1 Defining Physical Literacy

The creation of the term "physical literacy" is important as it relates to an individual's level of engagement, motor competency, and sense of competence and confidence in physical activity. Margaret Whitehead introduced the concept of physical literacy in 1987 as "a multifaceted conceptualisation of the skills required to fully realize potentials through embodied experience" (Giblin et al., 2014, p. 1177; Whitehead, 2010). The International Physical Literacy Association definition for physical literacy of June 2015 states:

"Physical literacy is the motivation, confidence, physical competence, knowledge and understanding to value and take responsibility for engagement in physical activities for life" (International Physical Literacy Association [IPLA], 2015, p. 1).

The definition recognizes the four essential and interconnected elements of physical literacy. These elements include motivation and confidence (affective), physical competence (physical), knowledge and understanding (cognitive), and engagement in physical activities for life (behavioural) (Sport for Life, 2021).

2.1.2 Physical Literacy and Physical Education

These definitions are important in the context of PE as they build onto the existing foundation of the programming of lessons, and onto the way these lessons are offered to students. The main goal of any educational program is to promote the desire to continue being physically active throughout one's lifespan (Roeterts & Jefferies, 2014). Supporting the student's physical, emotional, social, and affective learning is regarded as the valid learning outcome of PE that can fulfil the subject's aim of promoting a physically active life (Kirk, 2013). Programs that are built around the "whole person", encompassing the physical, cognitive, and affective domains (Mandigo et al., 2009), and that increase the physical literacy of an individual through motor proficiency or competency, have direct relations to increasing the cognitive and affective domains by increasing self-efficacy and self-esteem (Beilghe et al., 2012; Castelli et al., 2014).

2.1.3 Physical Competency

Physical competency refers to the skills, patterns, and complexes which a child learns as they grow, allowing them to interact with their environment (Whitehead, 2010). To be competent in one's physical skills, the individual must be trained in the fundamental blocks to create a foundation from which to learn and build more complex patterns moving through childhood, adolescence, and into adulthood (Cliff et al., 2012). Proficiency in motor skills helps children acquire the competence and confidence necessary to participate in future physical activity and sports (Mandigo et al., 2009). With higher levels of self-esteem and self-efficacy, individuals are more motivated to remain physically active throughout their lifespan (Higgs, 2010).

2.1.4 Physical Literacy and Long-Term Health

The combination of the loss of unstructured play, an increased dependency on technology, and the transition from physical labour to office work, has resulted in increased sedentary behavior and a reduction in physical activity levels among the population, and has directly impacted the development of physical literacy. Unstructured and risky play was common a generation ago, and this helped to develop physical literacy and also prepared children for an active lifestyle. However, today we face a situation where movement opportunities and experiences have been removed from our environment (Sport for Life, 2021).

It cannot be assumed that physical literacy will be achieved through the natural maturation process alone; instead, it must be actively developed. Providing an increased opportunity for unstructured and risky play during childhood forms part of the solution. However, it is important to keep in mind that not everybody innately develops the motivation, confidence, physical competence, knowledge and understanding to value, and to take responsibility for, maintaining an active lifestyle. Therefore, we need to find a variety of ways to support people's personal physical literacy journeys throughout their lives (Sport for Life, 2021).

The fact that regular physical activity delivers physical and mental health benefits is well established in the research literature (Burnham, 1998; Cekin, 2015; Lewis & Hennekens, 2016; Roychowdhury, 2020; Stanton et al., 2014; Warburton et al., 2006). Physical literacy is seen as a precursor to actual physical activity since people who lack the skills, confidence, competence, or knowledge to be physically active are less likely to take part in physical activity. Thus, physical literacy can be seen as a direct determinant of health due to its positive influence on physical activity (Sport for Life, 2021).

Cairney et al. (2019) published an evidence-based model showing the connections between physical literacy, physical activity, and health outcomes throughout life. By being healthy, individuals can continue their physical literacy journeys through their lives, thereby increasing both their engagement in physical activity and their sense of well-being. Through its behavioural, cognitive, and affective domains, physical literacy also promotes a positive self-concept and reduces social isolation, which helps to build resiliency. Finally, the combination of cooperative play and participation with others results in a strong psychological and social foundation that an individual can build on.

2.2 Fundamental Movement Skills

2.2.1 Defining Fundamental Movement Skills

From an early age, the development of a set of gross motor skills, called FMS, allows children to move confidently and with control in a wide range of activity settings (Bai et al., 2020). The acquisition of agility, balance, coordination, and laterality often precede the development of FMS, and as a child matures the acquisition of new complex movement skills builds on previously acquired movement skills. FMS are defined as the building blocks for movement and form the foundation for specialized skills required in more complex or sport specific movement patterns (Clark & Metcalfe, 2002; Gallahue et al., 2012; Hardy et al., 2012). Motor development literature divides FMS into three distinct constructs: (1) locomotor, (2) object control or manipulation, and (3) balance or stability (Gallahue et al., 2012; Ulrich, 2000). Mastering all FMS provides the basis for an active lifestyle, whilst a lack of proficiency in motor development acts as a barrier to lifelong participation in PA (Goodway et al., 2014; Robinson et al., 2015; Stodden et al., 2009).

2.2.2 Locomotor Skills

Locomotor skills are used to move the human body through space and are integral to participation in popular sports, games, and activities. Payne & Issacs (2020) defined running, jumping, hopping, skipping, galloping, and sliding as fundamental locomotion skills. Although walking precedes running, this is overlooked in the Payne & Issacs description. If an individual cannot stand upright to walk, the rest of these skills will not be achieved (Gabbard, 2021; Haywood & Getchell, 2019).

By the age of two, children are capable of walking without support, which is soon followed by running. Running provides the foundation for every other movement pattern associated with locomotion and is typically the first skill to be mastered (Payne & Issacs, 2020).

As soon as children can momentarily propel themselves through space, as is required in running, they will also be able to perform some type of jumping and hopping movement (Gabbard, 2021; Haywood & Getchell, 2019). Jumping skills vary according to the direction of the jump (e.g., downwards, upwards, or forwards) and by the number of feet used to perform the jump. The double- and single-foot actions of propelling the body either upward or horizontally are termed jumping and hopping, respectively (Gallahue et al., 2012; Payne & Issacs, 2020).

As strength, balance, and motor coordination improve, combination patterns will emerge, such as galloping, sliding, and skipping (Gabbard, 2021; Haywood & Getchell, 2019). Hopping and walking contribute to the development and mastery of the gallop and slide. Subsequently, the progression of the gallop and run patterns aid in the development of the skip (Gallahue et al., 2012; Payne & Issacs, 2020).

2.2.3 Object-Control Skills

When a child can walk unassisted, their hands become available to explore the environment. On reaching this developmental milestone, a child begins to demonstrate new movement patterns, commonly referred to as object-control skills (Gabbard, 2021; Gallahue et al., 2012; Haywood & Getchell, 2019). Payne & Issacs (2020) define object-control skills as hand-eye and foot-eye coordination skills, unique to each physical literacy environment. These skills include overarm throwing, underhand throwing, one-handed and two-handed catching, and striking things with, and without, an implement.

The manipulation of objects using either a body part (e.g., hand or foot) or a striking implement (e.g., racquets or bats), requires a series of fundamental and complex movement patterns to ensure performance (Gabbard, 2021; Gallahue et al., 2012; Haywood & Getchell, 2019; Payne & Issacs, 2020). Additionally, the coordination between the hand and eye, or foot and eye, as they pertain to participation in activities and sport, are important throughout the lifetime of an individual (Hardy et al., 2012; Mandigo et al., 2009).

2.2.4 Balance and Stability Skills

Balance refers to the body that remains in equilibrium but moves along its horizontal or vertical axis (Gallahue & Cleland-Donnelly, 2007). Alternatively, balance is sometimes referred to as postural control, which facilitates the body's positioning in space for the purposes of stability and orientation (Payne & Issacs, 2020). Postural stability is defined as the ability to maintain, achieve, or restore a specific state of balance, while postural orientation refers to an individual's capacity to maintain an appropriate relationship between their body and the surrounding environment (Horak, 2006).

There are two types of balance discussed in the motor development literature, known as static balance and dynamic balance. Static balance is the ability to maintain a desired body position when the body is stationary (e.g., standing or sitting in a chair), and dynamic balance refers to the ability to maintain postural control while the body is in motion (e.g., walking or ice skating) (Payne & Issacs, 2020; Westcott et al., 1997).

Stability skills are the ability to sense a change in the relationship of the body parts affecting one's balance, and the ability to respond quickly and accurately to those shifts with appropriate compensating movements (Gallahue et al., 2012). Axial movements, such as bending, stretching, twisting, turning, swinging, body inversion, body rolling/log rolls, and landing/stopping are examples of stability skills (Gallahue & Cleland-Donnelly, 2007).

2.2.5 Fundamental Movement Skills in Adolescence

The quality of motor development in early life has a significant impact on the physical activity behaviours and quality of life experienced in later years. Children with high motor skill proficiency are believed to have higher levels of general fitness and sports competence, and consequently, physically literate children are more likely to remain active throughout adolescence and adulthood (Robinson et al., 2015; Stodden et al., 2009). There is also evidence that fostering enjoyment and interest in physical activity during childhood and adolescence contributes to an increased participation within the adult population (Barker et al., 2020; Martens, 1996; Sallis & Patrick, 1994). Finally, Ulrich & Sanford (2000) point out that children who are less proficient than their peers will usually be selected last to participate in group games during PE and recess. The consequence of consistently being chosen last, or not at all, likely has a negative impact on the child's physical self-concept and motivation to be active.

Prior to advancing into the specialized movement phase, children need to have experience in a variety of basic movement patterns (Clark & Metcalfe, 2002; Robinson et al., 2015; Stodden et al., 2009); however, motor proficiency does not develop naturally or through free play (Goodway & Branta, 2003). A variety of factors influence motor skill development, such as the environment, effective instruction, physical growth (increase in body size), and maturation (development of biological systems) (Logan et al., 2012; Martin et al., 2009).

Throughout adolescence, maturation factors such as the development of secondary sex characteristics and growth spurts may contribute to gender differences in motor skill proficiency (Haywood & Getchell, 2019).

FMS are the building blocks that lead to specialized movement sequences needed to engage in a variety of physical activities for children, adolescents, and adults (Clark & Metcalfe, 2002; Gallahue et al., 2012). Studies conducted in Finland showed that 14-year-old students had better FMS when compared to 11-year-old students (Nupponen, 1997; Nupponen & Telama, 1998; Nupponen et al., 1999), demonstrating that FMS proficiency can be improved during adolescence. Okely et al. (2001) investigated the relationship of participation in organized and non-organized physical activity with FMS proficiency among adolescents, showing that the ability to perform FMS was related to participation in organized physical activity. Specifically, adolescents who performed better on the skills assessment spent more time in organized physical activity during a week. This positive relationship between movement skill proficiency and organized physical activity demonstrates the important role that organized physical activity patterns play in the FMS proficiency levels of adolescents.

2.2.6 Proficiency Barrier

Children lacking motor competence often encounter a barrier to later participation in physical activity (Cools et al., 2009; Goodway et al., 2003; Lubans et al., 2010). This proposed "proficiency barrier" suggests that there is a critical threshold of motor skill competence, and that those individuals who do not meet this threshold will be less likely to be involved in physical activity (Clark, 2007; Lubans et al., 2010; Stodden et al., 2013). According to Gallahue and Cleland-Donnelly (2007), FMS mastery enables individuals to incorporate specialized movement skills into their movement repertoire. Thus, if children enter a competitive sport environment that requires specialized movement skills before they have fully mastered their FMS, they will experience a proficiency barrier which can lead to their eventual withdrawal from the sport.

In order to break through the hypothesized proficiency barrier, and to avoid limiting their future participation in physical activity, it is essential that children develop motor proficiency and their FMS (Clark, 2007; Higgs et al., 2008; Logan et al., 2012; Stodden et al., 2013). The ability to become proficient is a result of practice, encouragement, and reinforcement (Logan et al., 2012). Specifically, children must use every opportunity to develop their basic foundational skills. This development will also contribute to health across their lifespan as regular physical activity is associated with maintaining a healthy weight, reduced blood pressure and positive psychosocial benefits (Barnett et al., 2008; Okely et al., 2001). Providing movement opportunities for children to explore and to understand how their bodies can be manipulated through various actions in different environments helps them to become physically literate.

2.2.7 Fundamental Movement Skill Mastery

Developing children's FMS to facilitate long term engagement with physical activity is important. Without FMS in childhood the number of physical activity options available to the individual in adolescence and adulthood decreases due to their inability to develop the specialized movement skills essential for lifelong activity (Goodway et al., 2014; Robinson et al., 2015; Stodden et al., 2009). Goodway et al. (2014) also claim that children who are more highly skilled and motor competent will self-select higher levels of physical activity. Most children are developmentally able to master most of the basic FMS by the age of 6 (Gallahue et al., 2012), followed by more complex skills by the age of 10 or 11 (O'Brien et al., 2016).

Children do not learn FMS through the maturation process alone, but through effective instruction and practice with a teacher or coach (Goodway & Branta, 2003; Logan et al., 2012; Martin et al., 2009). Although children will naturally develop a rudimentary form of fundamental movement pattern, they are more likely to achieve a mature movement pattern if they have ample opportunities to practice the skills (Clark & Metcalfe, 2002; Gallahue et al., 2012). Thus, while the progression through these fundamental stages generally occurs during childhood, there is no guarantee that an individual will reach the mature stage of any FMS during their lifetime. Unfortunately, many children are late in mastering the FMS that would allow them to feel confident about their competence to engage in sports and other physical activities (Booth et al., 1999; Erwin & Castelli, 2008; O'Brien et al., 2016; Okely & Booth, 2004).

Research suggests that most children and adolescents do not perform FMS to their expected developmental capacities. A United States study of 9-12-year-old children reported that half of the assessed students did not demonstrate proficiency in basketball throwing and dribbling motor tasks (Erwin & Castelli, 2008). Similarly, an Australian study involving students aged 9-15-years old found that the prevalence of mastery only exceeded 40% for one skill in one group (i.e., overarm throw, 15-year-old boys) (Booth et al., 1999).

During a PE class, O'Brien et al. (2016) assessed the performance of FMS in a sample of 12-13-year-old Irish adolescents. The study analyzed the FMS at the component level to identify performance weaknesses, and the commonality of these weaknesses across skills. A reliable instrument protocol was used to assess nine FMS: run, skip, horizontal jump, vertical jump, kick, catch, overhand throw, strike, and stationary dribble. The nine FMS were assessed with the components from three established assessment tools (i.e., Test of Gross Motor Development – First Edition [TGMD], Test of Gross Motor Development – Second Edition [TGMD-2], and the

Victorian Fundamental Motor Skills Manual). Overall, 11% of students scored at the advanced skill proficiency level (i.e., mastery or near mastery) for all nine FMS, with only one participant having complete mastery level across all locomotor and object-control skills. This finding indicates that overall skill execution amongst adolescent youth is low. Furthermore, the overall mean composite FMS score (object-control and locomotor) differed significantly between genders, with adolescent males scoring higher. Additionally, there were differences in the number of participants who failed to achieve mastery level across the nine FMS and their associated components. Aside from the run, advanced skill proficiencies in the locomotor subset skills were lower than in the object-control subset skills. The highest skill performance was the catch, while the poorest performance was the vertical jump. Detailed analysis of the vertical jump results showed that a higher proportion of participants failed to execute components 2 (crouching with the knees bent and arms behind the body) and 3 (forcefully swinging the arms upright). Furthermore, many of participants were unable to demonstrate mastery of the components which required coordinating arm and leg movements. This finding suggests that adolescents may have difficulty making a successful transition to more advanced skills within the sport-specific stage. According to O'Brien et al. (2016), focusing on the weakest skill components during PE and outside school hours may prove a valuable strategy in increasing the current FMS levels and the subsequent physical activity levels amongst adolescent youth.

2.2.8 Fundamental Movement Skills and Physical Education

Given the significant influence movement skills have on lifelong physical activity pursuits and the ramifications of inadequate movement skills on physical health (Lubans et al., 2010), the implications for the physical education system are clear. Payne & Issacs (2020) suggest that both the content of PE classes, and how it is taught, have an influence on the development of children's physical skills.

Quality PE is the single best way to ensure that every child in Canada gets to develop FMS (Sport for Life, 2021). All Canadian children have access to PE through the education system. Many children, particularly those from low socioeconomic backgrounds, may not have access to any form of physical instruction until they enter elementary school (Cohen et al., 2014). It is therefore appropriate to consider the PE curriculum as the principal vehicle for the delivery of physical instruction (McKenzie, 2010). In addition, school-based settings offer easily accessible areas for the investigation of large numbers of children. These investigations can assess the current state of movement skill proficiencies and then, based on the findings, suggest appropriate intervention programs.

2.2.9 Physical Health Education Canada

Individuals who are physically literate are more than just skillful movers. Physical Health Education (PHE) Canada (2013) found that physically literate individuals:

- Move with competence in a wide variety of physical activities that benefit the development of the whole person
- Consistently develop the motivation and ability to understand, communicate, apply and analyze different forms of movement
- Demonstrate a variety of movements confidently, competently, creatively and strategically across a wide range of health-related physical activities
- Demonstrate the ability to "read" what is going on around them in an activity setting and react appropriately to those events

- Make healthy, active choices throughout their life span that are both beneficial to and respectful of themselves, others, and their environment
- Master fundamental movement and fundamental sport skills that allow them to read their environment and make appropriate decisions while moving confidently and with control in a wide range of both indoor and outdoor settings

As children and youth develop and grow, their skills slowly evolve and improve. Basic human movements evolve into fundamental movement skills, and eventually some of the fundamental movement skills are adapted to become fundamental sport skills, used within a sport or activity. Fundamental sport skills (FSS) are defined as the rules, regulations, and specialized skills for a particular sport (Higgs et al., 2008).

It can be argued that a small percentage of children will learn physical skills on their own through trial and error. However, Logan et al. (2012) conclude that children will experience greater success if the skills are taught by a qualified individual who understands the skill mechanics and values the importance of developing such skills. Therefore, quality physical education is the best way to ensure that every child in Canada develops FMS and FSS.

2.2.10 Sport for Life

Sport for Life is a Canadian organization that endorses physical literacy. The organization has developed a chart that provides an interactive visual of the characteristics of a physically literate person. The chart shows how a person develops fundamental movement skills that lead to fundamental sport skills in various decision-making situations and in a variety of environments including the ground, water, snow, ice, and air.

Canadian children should learn FMS and FSS in each of the four basic environments described below (Higgs et al., 2008):

- On the ground, as the basis for most games, sports, dance, and physical activities.
- In the water, as the basis for all aquatic activities. The ability to participate safely in the water is an important life skill.
- In the air, as the basis for gymnastics, diving, and other aerial activities. These activities require spatial awareness which is the ability to understand where one's body is in space.
- On snow and ice, as the basis of all winter activities, which often include a great deal of balance and agility.

Exposure to as many environments as possible will increase the transferability of skills to different contexts providing more opportunity for physical development (Higgs et al., 2008). By fostering physical literacy through quality activity programs, the students of today are not only better prepared to lead healthy and active lives, but they are better prepared to help others, are respectful of the environment, and are creative in generating new and innovative ideas (Corlett & Mandigo, 2013). If children can use movement and sport skills to become more educated and confident, then this confidence can transfer to other areas of learning.

2.2.11 Relationship between Physical Literacy, Physical Activity, and Fundamental Movement Skills

The development and mastery of locomotor, object-control, and balance skills are essential for the growth and development of a physically literate child, which can lead to the continuous engagement in physical activity and sport throughout the lifespan. (Gallahue et al., 2012; Hardy et al., 2012; Higgs, 2010). Without the mastery of these skills, children are more likely to be disengaged from sports and physical activities as they transition to adolescence and are also more likely to lead sedentary lifestyles as they mature into adulthood (Cliff et al., 2012; Goodway et al., 2014; Robinson et al., 2015; Stodden et al., 2009). Physical literacy, physical activity, and FMS are interdependent since those individuals who can perform, enjoy, and participate in physical activity in a variety of environments are more likely to have the skills, fitness, motivation, and knowledge to continue this positive lifestyle throughout life.

2.3 Peer to Peer Learning

Traditional PE teaching methods have revolved around the "command" teaching style, where the teacher demonstrates the skill, and the students attempt to replicate the movement individually and practice independently while receiving minimal feedback from the teacher (Chatoupis, 2018; Mosston & Ashworth, 2008). This approach limits the student to reproduce the observed movement through imitation (Pritchard et al., 2008). The "reciprocal" teaching style is an alternative approach where peers are directly involved in the teaching and evaluation process (Jenkinson et al., 2014; Kelly & Katz, 2016; Topping, 2005; Topping & Ehly, 1998; Ward & Lee, 2005). Although still considered a reproductive teaching method, the reciprocal teaching style requires active student participation to achieve learning outcomes (Chatoupis, 2018; Mosston & Ashworth, 2008).

2.3.1 Peer to Peer Learning in Physical Education

Peers are often used in school settings to influence student learning outcomes. These instructional methods are often referred to as peer-assisted learning (Chng & Lund, 2018; Ward & Lee, 2005). In this thesis, peer-assisted learning is referred to as P2P learning. Peer tutoring and peer teaching are common teaching models that incorporate different components of peer-based interaction in PE and school physical activity programmes (Byra, 2006; Chng & Lund, 2018; Metzler, 2017; Ward & Lee, 2005). A common feature of the P2P model is a reciprocal teaching style, in which one student (the learner) performs a skill while the other student (the observer) provides specific feedback on the task. Most importantly, P2P learning is highly transferable across a range of educational contexts (Topping, 2005; Topping & Ehly, 1998), with pedagogy research suggesting it as a best practice in PE (Siedentop & Tannehill, 2000).

Greenwood et al. (1995) identified seven specific advantages of P2P learning over traditional methods, which are listed below. The P2P learning model:

(1) Creates more favorable evaluator/student ratios;

(2) Increases the student's time on task;

(3) Offers students more opportunities to respond to feedback and assessment;

(4) Increases the opportunities and immediacy for error correction;

(5) Enhances student motivation;

(6) Creates opportunities for individualized help;

(7) Provides an environment for students to work collaboratively.

The peer learning method encourages meaningful learning which involves students teaching and learning from one another; there is mutual benefit to all parties involved (Rohrbeck et al., 2003; Topping, 2005; Topping & Ehly, 1998). It involves the sharing of ideas, knowledge, and experiences, as well as a focus on interdependent learning as opposed to independent learning (Boud et al., 1999). In the reciprocal teaching style, the observer provides positive and corrective feedback on the critical skill components of the task immediately following each performance. Finally, at the end of the session, the two students switch roles (Chatoupis, 2018; Mosston & Ashworth, 2008). If structured properly, peer-assisted learning techniques help both the performer and the observer (Rohrbeck et al., 2003; Topping, 2005; Topping & Ehly, 1998). The model's reciprocal nature is essential, as it ensures that students do not hold power over each other due to their position or responsibilities (Chatoupis, 2018; Mosston & Ashworth, 2008).

The reciprocal teaching style has many advantages. First, it ensures that students receive sufficient and immediate feedback during practice (Chatoupis, 2018; Mohr & Townsend, 2002; Mosston & Ashworth, 2008). Siedentop & Tannehill (2000) claim that the increased frequency of feedback by the observer results in a greater number of correct performances by the student. Second, students learn to both give, and accept, peer feedback which enhances socialization skills (Chatoupis, 2018; Mosston & Ashworth, 2008). Creating an atmosphere in which peers feel comfortable with one another is crucial in movement settings since much of what happens in these settings (e.g., games, team sports) requires the interaction of two or more individuals. Finally, students learn to perform, as well as analyze, movements. By evaluating a peer's performance, comparing it against the criteria, drawing conclusions about the performance, and providing appropriate feedback, the observer gains a deeper understanding of the process of learning a skill (Chatoupis, 2018; Mosston & Ashworth, 2008; Siedentop & Tannehill, 2000).

While the reciprocal teaching style has many advantages, it also has some drawbacks. The first, and probably most significant, is that practice time is reduced, because one student is providing feedback to another. As a result, the amount of skill practice that a student would usually complete in another teaching style is reduced by approximately half (Ernst & Byra, 1998). In the self-check style, for example, the learner is concerned with executing the skills rather than assessing a peer. Hence, students in the self-check style have more opportunities to be physically engaged in the task than students in the reciprocal style (Chatoupis, 2018; Mosston & Ashworth, 2008). Secondly, there is a risk of giving inaccurate feedback since each student is responsible for providing task-specific feedback to their partner. This is a critical problem to address since incorrect feedback may be more detrimental to a student's skill performance than no feedback (Ernst & Byra, 1998). Thirdly, the possibility of student conflict is increased, especially if peers have differing opinions (Byra & Marks, 1993; Chatoupis, 2015). Finally, because of the cooperative nature of the reciprocal teaching style, students can engage in small-talk and off-task conversations (Chatoupis, 2018; Mosston & Ashworth, 2008). As with any instructional approach, teachers and researchers must address these problems when using the reciprocal teaching style.

The application of P2P learning in PE has shown a variety of benefits, including:

(1) Improve the percentage of correct performances of motor skills (Houston-Wilson et al., 1997);
(2) Increase scores in motor skill tests (Chatoupis, 2015; Ernst & Byra, 1998; Goldberger & Gerney, 1986; Goldberger et al., 1982; Goldberger & SueSee, 2020);
(3) Increased total time spent doing moderate to vigorous physical activity (Lieberman et al., 2000);
(4) Increased motivation and interest for learning the sport (Yoncalik et al., 2010);
(5) Enhanced social interaction among students (Chatoupis, 2015; Byra & Marks, 1993; Goldberger et al., 1982; Goldberger & SueSee, 2020).

Peer modelling and demonstrations are particularly effective for motor skill learning (Magill & Anderson, 2014); however, demonstrations require consideration of student age, maturity, and model-observer familiarity. Research suggests that watching similar peers (e.g., skill or ability level, age, gender, etc.) is very beneficial to learning and motor skill acquisition (McCullagh & Weiss, 2001). When young children and novices observe similarly skilled peers, they develop the ability to identify where they are making errors, allowing them to make the appropriate changes and to improve. However, when observing an expert model novices have little or no model-observer familiarity, which may lead to a feeling that the skill is unattainable or else an inability to pay attention to the critical movement elements (Jennings et al., 2013).

The combination of a collaborative learning environment with the sharing of common goals (Duran & Monereo, 2005) allows P2P learning to facilitate the development of personal connections between participants (Tinto, 2003). Morgan et al. (2005) discovered that the reciprocal style resulted in more adaptive, cognitive, and affective student responses, such as a greater emphasis on learning, improvement, and enjoyment, when compared to the “command” style. Peer-assisted learning strategies allow students to improve their cognitive knowledge by practicing their reasoning and decision-making skills (Ploetzner et al., 1999), discussing other student’s opinions (Stahl, 2006), reflecting on misconceptions and misunderstandings, and developing mutual understandings (van den Bossche et al., 2006). In conclusion, the P2P learning model requires students to be actively involved in the teaching and assessment processes, which increases communication, student engagement, and cognitive understanding.

Several studies have been conducted to explore the impact of P2P learning on student psychosocial outcomes (Byra, 2006; Byra & Marks, 1993; Chatoupis, 2015; Ernst & Byra, 1998; Goldberger et al., 1982; Mulvihill et al., 2000; Weiss & Stuntz, 2004). Typically, these characteristics have been measured using psychometric methods that evaluate constructs like self-esteem and self-efficacy, or through interviews and surveys designed to evoke students' perceptions. According to Goldberger et al. (1982), students who used the reciprocal style showed more empathy, praise, and motivation than the control group. Working with a peer was a positive experience (Chatoupis, 2015) that increased social interaction among students (Ernst & Byra, 1998). Byra (2006) adds to this by claiming that fifth grade students who used the reciprocal style reported feeling motivated, challenged, and completely engaged in a supportive and enjoyable instructional environment. Byra and Marks (1993) discovered that observers provided more specific feedback to performers if they were friends, demonstrating that pairing by companionship influences peer learning outcomes. Furthermore, when practicing with friends, students who used the reciprocal style reported feeling more at ease receiving feedback (Byra & Marks, 1993). Similarly, Mulvihill et al. (2000) found that children prefer to work in teams, especially if they were able to interact with their friends. This finding suggests that children may take their peers' opinions into consideration when learning new skills and playing games (Macphail et al., 2008). According to Weiss and Stuntz (2004), establishing positive friendships and having social support are strong predictors of positive affective responses and motivational processes for participants in youth sports. It is not surprising, then, that researchers argue that peer influence is significant in how children learn skills and strategies in PE and sport (Byra, 2006; Byra & Marks, 1993; Chatoupis, 2015; Dyson et al., 2004; Ernst & Byra, 1998; Goldberger et al.,1982; Goldberger & SueSee, 2020; Koekoek et al., 2009; Macphail et al., 2004; Taylor et al., 2009; Weiss & Stuntz, 2004).

2.3.2 Variables that Impact Peer to Peer Learning

Due to the nature of peer-assisted learning, a range of variables can be manipulated, resulting in different learning outcomes. Examples include same-age or cross-age tutoring, same gender or mixed-gender tutoring, reciprocal or unidirectional, settings (classroom or outside the classroom), intensity (one session per week or five sessions per week of various durations), and targeted domain (social, cognitive, physical) (Jenkinson et al., 2014).

Peer learning is bi-directional, while peer tutoring implies an unequal partnership due to one individual's position of power and responsibility. Peer tutoring is advantageous because it provides feedback at a higher rate and provides immediate task-related information to the learner (Chatoupis, 2018; Mosston & Ashworth, 2008). The increased amount, immediacy, and content of feedback, as well as the ability to practice tasks, can lead to improved motor performances.

Same-age tutoring creates an opportunity for peers to gain insight into how to communicate and interact with people their own age. They also have a better understanding of their peers' learning challenges, and learning-through-teaching offers a lot of social reinforcement (Topping, 2005; Topping & Ehly, 1998). In contrast, cross-age tutoring is more likely to allow the tutor to demonstrate greater knowledge, removing the competitive aspect between peers and protecting the tutee's self-esteem (Cohen, 1986; Jenkinson et al., 2014). In conclusion, same-age tutoring is often used when the aim is to improve skill performance and technique, while cross-age tutoring is more often used when the desired outcome is knowledge transfer (Jenkinson et al., 2014).

2.3.3 Research Studies Examining the Relationships between Peer-to-Peer Learning, Motor Skill Performance, and Social Relationships

Motor skill performance and social relationships between peers have been examined in several studies in PE in which the reciprocal teaching style was used.

2.3.3.1 Byra & Marks (1993) and Ernst & Byra (1998)

Byra and Marks (1993) investigated the effect of pairing students by ability level (high, low, and mixed) and companionship (friendship and non-acquaintance) on student feedback delivery, as well as comfort level when giving and receiving feedback. The findings showed that if the performers were friends, the observers provided more specific feedback. Students also reported feeling more comfortable receiving feedback from a friend than a non-acquaintance. Pairing by ability level had little effect on the amount or quality of feedback given.

Ernst and Byra (1998) conducted a subsequent study in a junior high school PE setting to explore the effects of pairing learners by skill ability (high, low, and mixed) on student motor skill performance (skill technique and skill outcome) and cognitive performance. Juggling was chosen as the motor skill since it was new for all students. The results found that low-skilled learners significantly improved both their skill outcome and skill technique regardless of whether they were paired with high-ability or low-ability partners. Furthermore, high-skilled learners who were paired with low-skilled learners also improved significantly, but only in terms of skill outcomes. Finally, significant changes in the students' knowledge of juggling skill components were reported by the high-high, high-low, and low-low skilled learners (Ernst & Byra, 1998).

In summary, when students practiced with friends, they received more specific feedback (Byra & Marks, 1993; Ernst & Byra, 1998), whilst pairing by ability level had no effect on the amount of feedback provided (Byra & Marks, 1993) or the performance (Ernst & Byra, 1998).

2.3.3.2 Johnson and Ward (2001)

Johnson and Ward (2001) investigated the effects of peer tutoring on third-grade students’ performance of striking skills, with 11 children participated in a 20-lesson striking intervention. They examined the effects of peer tutoring on the overall number of trials, as well as the number and percentage of correct trials. The researchers also assessed students' ability to accurately evaluate each other's performance. When compared to baseline data, the results showed that although students performed fewer total trials, they had a higher percentage of correct trials. Furthermore, their analysis on gender effects revealed that females, regardless of ability, were able to perform similarly to their male peers. Finally, findings demonstrated that students accurately evaluated their peer’s performance more than 90% of the time.

2.3.3.3 D’Arripe-Longueville et al. (2002)

D’Arripe-Longueville et al. (2002) investigated the extent that low-, average-, and high-skilled performers acted as tutors for their peers for a flip turn in swimming. Tutors showed a flip turn to their tutees, who then had eight minutes to practice. The number of trials, demonstrations, and verbal interactions were all documented during this time. Results showed that the number of attempts and demonstrations were influenced by the interaction with tutors, with male tutees being more likely to have a greater number of attempts than female tutees. There was a significant relationship between tutee achievement and tutor skill level; students paired with low- or average-skill tutors did not perform as well as students paired with high-skill tutors. Interestingly, gender significantly influenced skill level, with high-skilled male tutors showing significantly better results in their tutee outcomes than high-skilled female tutors.

2.3.3.4 Ayvazo & Ward (2009)

Ayvazo and Ward (2009) examined the impact of peer tutoring on sixth-grade students’ volleyball skills. All tutors received training and were assigned to the same team to complete volleyball skills presented on task cards. Peer tutoring began after this training, with tutees and tutors switching roles and teams completing tasks, as necessary. Video analysis showed that the total number of trials and correct trials both increased. This study supports the trend of a greater number of attempts and correct trials when using peer tutoring in striking skills (Johnson & Ward, 2001) and swimming (D'Arripe-Longueville et al., 2002).

2.3.3.5 Ensergueix & Lucile (2010)

Ensergueix and Lucile (2010) evaluated the effect of peer tutor training on motor performance outcomes, specifically by examining whether a peer tutor training program prior to the start of a reciprocal peer tutoring intervention would help students feel adequately prepared to fulfil their role as a tutor. The study revealed significant improvements in motor skill development, as well as the benefit of having the same gender tutor. The results suggest that implementing reciprocal peer tutoring with trained tutors can lead to higher motor performance for tutees when compared to either individual practice time with no tutoring, or else students completing reciprocal peer tutoring with untrained tutors.

2.3.3.6 Chatoupis (2015)

Using the reciprocal teaching style, Chatoupis (2015) investigated the effects of pairing learners by companionship (friend and non-acquaintance) on 8-year-old children's motor skill performance and comfort levels when giving and receiving feedback with their peers. Based on previously discussed research (Byra & Marks, 1993; Mosston & Ashworth, 2008), it was hypothesized that working with friends would be more effective in terms of motor skill development and comfort levels than working with non-acquaintances. The participants (n = 52) were randomly assigned to a treatment group (n = 40) or a control group (n = 12). Prior to the study, learners in the treatment group were paired by companionship into partners who were friends and partners who were not friends. Children in all groups were taught the same dribbling tasks over the course of eight 30-minute sessions. A soccer dribbling test adapted from Keith (1980) was used to evaluate learners' dribbling skill, in which the time taken to complete the test represents the participants' skill outcome score. To measure how the learners perceived working with a partner, the research used a questionnaire using a 7-point semantic differential scale, which was adapted from Byra & Marks (1993) study. The findings demonstrated that learners paired with friends felt more comfortable giving and receiving feedback compared to learners paired with non-acquaintances. Additional results showed that motor skill development was greater in learners paired with friends than learners paired with non-friends or learners in the control group. Both findings supported their hypotheses. These results support previous research (Ernst & Byra, 1998; Goldberger et al., 1982; Kolovelonis et al., 2011) as well as Mosston and Ashworth's (2008) claim that in the reciprocal teaching style individuals learn motor skills by observing the performance, comparing the performance against criteria, and giving appropriate feedback. Furthermore, children paired with a friend not only improved their motor skill performance, but also felt that working with a peer was a positive experience, thus supporting previous similar findings (Byra & Marks, 1993; Mosston & Ashworth, 2008). This information suggests that PE teachers who value outcomes related to social relationships should use the reciprocal style with the appropriate pairing technique. If the goal is to improve motor skill performance, it is recommended that learners be allowed to select their partner.

2.3.3.7 Madou & Iserbyt (2018)

Madou & Iserbyt (2018) investigated the effect of pairing by ability in peer teaching for swimming performance, physical activity, and time-on-task. They hypothesized that mixed-ability pairing would lead to higher performance improvement and that low-ability learners would benefit the most from this configuration. Additionally, they hypothesized that physical activity and time-on-task would be higher for individuals in mixed-ability pairings. University students specializing in Kinesiology participated in four 50-min lessons in front crawl swimming where the reciprocal style of teaching was used. The students were randomly assigned to gender-homogeneous same-ability (low with low and high with high) and mixed-ability (high with low) dyads. Swimming performance was assessed before and after the four lessons. Physical activity and time on-task was coded based on video recordings of all lessons. Although not statistically significant, results showed higher swimming improvement in mixed-ability dyads, especially for low-ability swimmers, which agrees with their hypothesis. These learners also spent the most time swimming and on-task, which also agrees with their hypothesis. Meanwhile, high-ability learners improved their performance the least when paired with other high-ability learners. This dyad pairing also spent the lowest amount of time swimming and were the least on-task. Overall, students spent 37% of lesson time engaged in moderate-to-vigorous physical activity and were on task 82% of the time.

2.3.3.8 Comparing Research Studies

In conclusion, studies in which peer tutors had higher skill knowledge and skill performance ability (D'Arripe-Longueville et al., 2002; Ernst & Byra, 1998; Madou & Iserbyt, 2018) and received training prior to their tutoring roles in a reciprocal peer teaching framework (Ensergueix & Lucile, 2010) showed significant or positive changes in tutees' measured outcomes. Improvements in psychosocial outcomes were also observed in research using same-age and same-gender tutoring (D'Arripe-Longueville et al., 2002; Ensergueix & Lucile, 2010). Additionally, these results highlight the importance of tutor selection and the impact it might have on both primary and secondary study outcomes (Chatoupis, 2015). The findings suggest that tutors with more proficient motor skills and task expertise can help tutees achieve more positive psychosocial outcomes, thereby improving tutee performance (D'Arripe-Longueville et al., 2002; Ensergueix & Lucile, 2010; Ernst & Byra, 1998; Madou & Iserbyt, 2018; Ward & Lee, 2005). This could be explained by tutees having more confidence in their tutors and therefore concentrating more intently on the task at hand.

2.3.4 Peer Assessment in Physical Education

2.3.4.1 Rationale for Peer Assessments in Physical Education

Incorporating peer-assisted learning and evaluation strategies can enhance learning experiences and facilitate performance assessment. Peer assessment is a variation of peer-assisted learning in which students evaluate the performance of their peers (Melograno, 1997). It can be used as a learning activity or a formative assessment (Chng & Lund, 2018). According to Boud et al. (2001), careful planning is required for assessments to be compatible for peer learning; since assessment type has such an influence on learning, an inappropriately designed assessment can negate its usefulness as a teaching and learning strategy. Richard et al. (1999) described teachers' perceptions of students in grades 5 to 8 who had peer assessed each other using a team sport performance assessment procedure. The results show that teachers appreciated peer assessment, despite it being time consuming and requiring frequent prompting. Peer assessments require more teacher management skills but planning and experience can allow an instructor to provide a positive and enjoyable experience for all students (Chng & Lund, 2018; Johnson, 2004).

Mosston and Ashworth (2008) suggest that peer assessment reinforces cognitive understanding and helps to develop skills and dispositions within the affective domain, such as criticizing others, accepting criticism, and being more tolerant and accepting of others, in addition to addressing psychomotor objectives. Peer assessments, like the reciprocal teaching style, can enhance the learning experience (Mosston & Ashworth, 2008) by:

(1) Increasing student involvement in learning as students take on teaching responsibilities
(2) Increasing social interactions and trust in others
(3) Allowing performers to receive individual feedback
(4) Focusing student attention on the process, rather than the product, and,
(5) Providing possibilities for enrichment for skilled or quick learners.

2.3.4.2 Formative and Summative Peer Assessments

When used as a formative assessment, peer assessment enhances the learning experience while also influencing achievement. Motor skill evaluations in PE can cause students to become confused, anxious, and dissatisfied (Mohnsen, 2012; Mosston & Ashworth, 2008). This is due in part to the traditional practice of a "one-shot assessment," in which a skill test is introduced and evaluated at the end of a unit with no opportunity for improvement or re-testing (Siedentop & Tannehill, 2000). Furthermore, students' perceptions of the purpose, grading, criteria, or procedures used in a skill test may be inaccurate. When used as a formative assessment before the formal summative assessment, peer assessment may (Chng & Lund, 2018; Johnson, 2004):

(1) Provide additional practice opportunities with individual feedback for the performer;
(2) Alleviate the students’ misconceptions about how the skill test is graded;
(3) Allow test-anxious students to learn in a less stressful, non-evaluative assessment;
(4) Identify mistakes for students to correct before the skill test is administered; and,

(5) Motivate students to persevere, especially if they are given time to practice and improve before the graded skill test.

Summative assessments, such as skill tests in PE, are rarely functional learning experiences, especially when there is little feedback and no opportunity to improve or retest. When individual feedback and additional practice opportunities are offered, peer assessment can provide a more beneficial learning experience. Peer assessment, when used as a formative assessment, allows students and physical educators to have a more positive skill test experience (Chng & Lund, 2018; Johnson, 2004). First, students will be able to see how much they have learned and improved by the time they take the skill test. Second, students who suffer from test anxiety should feel less stressed during the skill test because they had the opportunity to practice before the assessment (Martin et al., 2002). Third, when students have experienced the same test conditions in peer assessment, the teacher should find it easier to deliver the skill test.

2.3.4.3 Advantages of Peer Assessments in Physical Education

Although peer assessments take more time to plan and create, they allow students to be more involved and responsible for their own learning and the learning of their peers (Chng & Lund, 2018; Johnson, 2004). Furthermore, although it takes time to train students to perform their roles responsibly and accurately in this complex instructional and assessment activity, the experience and training may result in the transfer of these behaviors to other settings, such as other classes or professional settings. Peer learning encourages lifelong learning and is linked to generic skills such as teamwork and interpersonal skills, which are highly valued by employers (Tan, 2021). Finally, frequent peer assessment experiences will require less orientation and training for students.

While peer assessments may result in less physical practice time than traditional teaching methods, the student who is assessing is cognitively engaged. The time a student spends as an observer may provide educational benefits beyond those experienced in teacher-directed activities. Even though the observer's ability to diagnose skill components and provide feedback may not be as advanced as the PE teacher's advanced skills, knowledge, and experience, this issue can be overcome with a well-designed recording sheet, an effective demonstration, accuracy training, and active supervision (Chng & Lund, 2018; Johnson, 2004).

2.3.4.4 Using Peer Assessments in Physical Education

To improve skill mechanics, peer assessments should be used early in the learning process. A peer assessment activity should be planned as soon as students can perform the skill, and before skill errors become habitual (Chng & Lund, 2018; Johnson, 2004). The peer assessments can be used to help inform the PE teacher's comments on report cards as well as to provide feedback on the more difficult components to facilitate additional training (Jenkinson et al., 2014). Physical educators who understand the benefits of this instructional format, the advantages of incorporating assessment into teaching (Melograno, 1997), and the value of formative assessments (Siedentop & Tannehill, 2000), can create appropriate peer assessment materials and activities to address the dilemmas in their classroom.

When refining and mastering a motor skill, a learner will benefit from meaningful and substantive feedback. However, even if a physical educator has excellent teaching skills, a student may receive insufficient feedback from the teacher for a variety of reasons (e.g., class size, space, equipment, time, organization, and safety considerations) (Evans, 2017). A peer assessment, when used as a learning activity, can provide a student with immediate and relevant feedback after a motor skill performance; immediate feedback has been shown to increase student motivation to try again and to seek mastery (Chatoupis, 2018; Chng & Lund, 2018; Johnson, 2004; Mosston & Ashworth, 2008; Ward & Lee, 2005).

Butler and Hodge (2001) used a case study to illustrate the use of peer assessment with high school students during a softball unit. Students indicated that peer assessment enhanced feedback and increased trust, and that providing feedback was an important activity for them to engage in during class. Sackstein (2017) adds to these findings by stating that students respond more positively to peer feedback on performance than to teacher comments. Providing this opportunity for peer feedback without teacher intervention may provide a "safe space" in which students feel comfortable sharing their ideas on critical movement components without fear of revealing their concerns in front of the teacher.

A final aspect to consider when providing feedback is how the information is presented and communicated. Wulf (2013) argues that focusing a young learner's attention on internal factors such as body movements (internal focus) is less effective than focusing attention on the movement effect (external focus).

2.3.5 Accuracy and Training of Peer Assessment

Students benefit from fast and accurate feedback on motor performance, since this maximizes learning and student achievement in the PE environment (Mohnsen, 2012). However, as class sizes increase, teachers face the challenge of providing direct and timely feedback to all students (Evans, 2017). One potential solution is the use of peer assessments, in which students observe and evaluate the quality of a partner’s performance of a motor skill (Alstot, 2018; Chng & Lund, 2018; Kniffin & Baert, 2015; Lund & Veal, 2013; Veal, 1995). This will enable all students in a class to receive feedback on their skill development, increasing the potential for motor learning (Alstot, 2018; Chng & Lund, 2018; Veal, 1995).

2.3.5.1 Accuracy of Peer Assessment

Veal (1995) suggested that students of any age can be taught to conduct peer assessments, and research has examined the accuracy of assessments completed by student peers (Alstot, 2018). Ward, Crouch, and colleagues conducted a series of studies (Crouch et al., 1997; Ward et al., 1998) on the effects of a peer-mediated accountability intervention in PE and its effects on student performance. One of the overlapping components of these studies involved students conducting a peer assessment, with the findings showing that students in grades four, five, and six can accurately assess the process and product of performing volleyball skills (Crouch et al., 1997) and basketball skills (Ward et al., 1998). Similarly, Kolovelonis and Goudas (2012) discovered that fifth and sixth grade students conducted moderately accurate peer assessments on the product and process of performing a basketball pass, and that students who received more accurate feedback outperformed those who received less accurate feedback. Finally, Nadeau et al. (2008) discovered that 14- to 17-year-old hockey players assessed peers during game play with a high degree of accuracy. These studies show that students of any age can provide accurate feedback to peers when assessing both the product and the process of performing a motor skill.

2.3.5.2 Reliability of Peer Assessment

The assumption that students can accurately observe and assess their peers is central to peer-assisted learning strategies. If students can provide accurate and reliable assessments, the PE teacher will have a more detailed understanding of the student’s performances and will be able to provide relevant feedback, increasing the chance that the students will master the skills. Several studies provide evidence to support this assumption, including Crouch et al. (1997), Hill and Miller (1997), Johnson and Ward (2001), Kolovelonis and Goudas (2012), and Ward et al. (1998). The similarity between student assessments of peers and researcher agreement was examined in these studies using correlational analysis of interobserver reliability testing. The findings from these studies indicate that well-trained students from third grade to high school can reliably assess their peers (70-96 percent agreement with researchers), and that the reported levels of accuracy were sufficient to demonstrate a significant improvement in student performance.

2.3.5.3 Recommendations for Teachers

Teachers must take the time to educate and train their students in peer-assisted learning for it to be successful. Since training takes time, it should be incorporated into the structure of multiple lessons, rather than being used sporadically across random lessons (Chng & Lund, 2018; Johnson 2004; Ward & Lee, 2005). Such training can focus on a variety of activities, such as social or assessment skills, but it must include adequate practice and feedback. Most studies using the P2P learning model included training that lasted one to three days (Ward & Lee, 2005).

2.4 Observational Learning

Observational learning is the ability to learn a motor skill through observation of another individual performing the skill (Cross et al., 2009), and is a successful strategy in PE to promote the learning of a wide variety of motor skills (Hodges et al., 2007; Lago-Rodríguez et al., 2014; Ste-Marie et al., 2012; Vogt & Thomaschke, 2007). Although some aspects of learning are exclusive to physical practice and cannot be experienced during observation, observational practice is viewed as an effective method to teach general skill characteristics, although it might not be as effective as physical practice (Granados & Wulf, 2007).

The finding that observational practice is not as effective as physical practice (Bandura, 1969; Carroll & Bandura, 1982; McCullagh et al., 1989; Scully & Newell, 1985) suggests that observers are not able to experience all the necessary processing that physical practice provides. However, observational practice offers unique opportunities for information processing that would otherwise be unavailable early in practice, when most of the cognitive resources are required to physically perform the task (Shea et al., 2000). As a result, the observer may be able to gain insight into specific aspects of the coordination pattern or assess the efficiency of various techniques, which might be difficult to do when trying a new skill (Granados & Wulf, 2007).

2.4.1 Experienced vs. Novice Models

Demonstrations are one of the most common instructional methods used to convey information to the learner in the context of motor skill acquisition (Williams et al., 1999). Research has shown that observing both experienced models (Al-Abood et al., 2001; Hodges et al., 2003; Lee et al., 1995) and novice models (Black & Wright, 2000; Buchanan & Dean, 2010; Hayes et al., 2010; McCullagh & Meyer, 1997) leads to significant learning. While observation of a skilled model might facilitate the development of a correct movement interpretation, observing another learner might help performers identify and correct errors (Granados & Wulf, 2007). Recent research suggests that these benefits are optimized if the observer is told beforehand about the quality of the performance they will observe (Andrieux & Proteau, 2014; Rohbanfard & Proteau, 2011). This insight could be useful in situations that benefit from having a video observation model, such as in PE (Andrieux & Proteau, 2016).

Observing another performer has been shown to be beneficial for learning (McCullagh & Weiss, 2001; Wulf & Shea, 2002), especially for more complex skills (Bird et al., 2005). Lee and White (1990) indicated that the model does not have to be an expert for observation to be beneficial, and that by observing a novice model the observer gains access to the cognitive processes associated with detecting and correcting errors, which are thought to be important in motor skill learning. The suggestion by Lee and White (1990) that both the performer and the observer engage in error detection and correction processes is consistent with the hypothesis that at least some of the cognitive processes associated with observational practice are like those used when a motor skill is physically practiced (Adams, 1986; Bandura, 1977; Bandura, 1986).

2.4.2 Physical and Observational Practice Offer Unique Learning Opportunities

Observational practice, especially when combined with physical practice, can offer significant and unique contributions to learning (Shea et al., 2000; Wulf et al., 2010). Shea et al. (1999), for example, demonstrated that learning in pairs was superior to learning alone since it allowed each person to observe a peer doing the skill. One explanation for the benefits of the combined physical and observational training approach is that being allowed to observe, as well as to physically execute, the skill provided two unique perspectives (i.e., first and third person) on the task being learned. Trials involving those two perspectives may allow participants to improve their understanding of the motor skill by focusing more on one dimension of the task while physically practicing the skill (e.g., proprioceptive feedback) and on another dimension (e.g., coordination patterns) while observing another learner perform the skill (Shea et al., 2000).

The role this additional processing can play in learning is best demonstrated when participants alternate between physical and observational practice in pairs (Wulf et al., 2010). Retention performance in experiments using this training strategy have demonstrated the unique contributions of observational practice (Shea et al., 2000). On these retention tests, participants who practice in pairs perform as well as participants who simply do physical practice, despite the fact that the dyad participants do only half the physical practice trials that the physical-practice-only group does (Shea et al., 2000; Shebilske et al., 1992). Granados and Wulf (2007) investigated mastery of cup stacking skills to determine if the benefits of learning in dyads came from observation or from dialogue between participants. Their findings showed that participants who could observe another learner learned more (i.e., produced faster movement times, made fewer errors) than participants who could converse with their peers.

The learning benefits of dyad practice may be due to increased motivation, resulting perhaps from competition with the partner, the setting of higher goals, or the loss of self-consciousness as individuals fulfil interdependent dyadic roles and find somebody in the same learning boat (Wulf et al., 2010). It is probably not coincidental that participants in collaborative or cooperative learning environments often subjectively report more enjoyment than when learning alone (Mueller et al., 2007). In summary, observational practice, especially when combined with physical practice, can contribute significantly to skill learning. Observational practice not only provides the learner with information about the goal movement or potential mistakes to avoid, but it can also influence goal setting by introducing a competitive element to the practice situation (Wulf et al., 2010).

2.5 Assessment

2.5.1 Assessment in Physical Education

Assessment can be defined as any deliberate procedure used to test or evaluate a student's achievement and to draw conclusions based on that evidence for a variety of purposes (Doolittle, 1996), and is an essential component of the teaching and learning processes in PE. Physical literacy and motor development are important aspects of both elementary and secondary school PE curricula, but the numerous developmental and learning goals within a PE curriculum can make assessment of students' progress challenging for PE teachers (Chng & Lund, 2018; Lund & Tannehill, 2014; Lund & Veal, 2013). Given the number and varying complexity of the specific activities involved, learning objectives in this part of the PE curriculum can be quite varied. As a result, PE teachers are interested in discovering and identifying effective assessment techniques (Nadeau et al., 2008; Payne & Issacs, 2020).

Using assessments within PE has a variety of benefits and formalizing the assessment process shows students that the information in PE has value (Kniffin & Baert, 2015; Lund & Veal, 2013; Reeves, 1986). Therefore, students will recognise the importance of the activities being evaluated which increases their motivation to participate in PE activities (Kniffin & Baert, 2015; San-Tan & Wright, 2004). Assessments allow students to receive feedback about their performance in the short and long term, which facilitates learning in a PE class (Chng & Lund, 2018; DeJong et al., 2002; Reeves, 1986; Veal, 1988; Veal, 1992, Veal, 1995). Teachers may use the assessment data to make decisions about their teaching effectiveness (Fencl, 2014). A teacher's main role is to facilitate learning, and the only way to know whether learning has occurred in a PE class is for assessments to be administered successfully (Kniffin & Baert, 2015). However, the main objective of using assessments in PE should be to help students learn (Chng & Lund, 2018; DeJong et al., 2002; Veal, 1992).

There are two types of assessment: formative assessment and summative assessment. Formative assessments occur concurrent with instruction, are ongoing and recurring, and emphasize the process of student learning (Chng & Lund, 2018; Veal, 1992). Summative assessments are usually done at the end of a unit and assess a student's progress by providing a formal evaluation (Chng & Lund, 2018; Veal, 1992). While both formative and summative assessments play important roles within the PE curriculum, it is important to remember that the main goal of incorporating peer assessment into PE should be to help students learn (Chng & Lund, 2018; Veal, 1992), which suggests that formative assessments should be used more frequently because of their effective use as tools to aid student learning (Chng & Lund, 2018).

Efficient and reliable assessment tools are needed to measure student achievement to help teachers deliver the best programs possible and to provide necessary feedback to the students. Veal (1995) proposed ways for teachers to incorporate the assessment process into PE to measure psychomotor skills, as described below:

(1) Assessments can be performed by the teachers, by the students independently, or by student peers.

(2) Assessments can measure the product of performing motor skills (i.e., the number of shots made in basketball).

(3) Assessment can evaluate the process of performing a skill (i.e., the technique the student uses to perform an overhand throw in softball).

Fencl (2014) suggested using multiple forms of assessment within a class. Since students have different learning styles, using a variety of assessments can assist educators in effectively targeting a variety of learning styles, therefore improving the potential for learning and achievement among PE students. To accomplish this, teachers can use previously published assessments (Fisette & Franck, 2012), or create their own (Lund & Veal, 2013).

2.5.2 Product-Oriented vs. Process-Oriented Assessment

Assessment is a vital aspect of teaching and learning for any skill. The two primary approaches used in evaluating movement are product-oriented and process-oriented. The product-oriented approach examines the outcome or end-product of the movement, while the process-oriented approach emphasizes the technique of the movement (Payne & Issacs, 2020). Product-oriented assessment does not allow for a description of movement skill performance; instead, it only provides the outcome of a performance with no reference to the way the performance result was obtained (Barnett et al., 2014; Malina et al., 2004). When using this method, the researcher is more interested in performance outcomes than in the technique, or process, used to perform the skill. The process-oriented approach, on the other hand, emphasizes the movement technique itself while paying little attention to the movement outcome, or product. This technique is based on the idea that motor assessments should involve a segmental or component approach, because development occurs at various times within different body components (Payne & Issacs, 2020).

Consider an assessment of a child's running ability. A product-oriented assessment uses running speed or distance as the outcome variable, providing little information about the skill's movement characteristics; whereas a process-oriented assessment provides an evaluation of the child's movement characteristics while running (e.g., Do the child's arms move in opposition to their legs and along the sagittal plane? Is there a definite flight phase?). The inability to compare investigations employing product-orientated assessments to those using process-orientated measurements is apparent and results in difficulties when attempting to establish a clear picture of FMS proficiency among children.

However, in certain situations, the movement product and process are the same; for example, in many gymnastics-related movements (e.g. a forward roll) the process involved is also the product. A process-oriented approach would focus on the technique used to perform the movement, whilst using a product-oriented approach the technique can also be considered the desired outcome, because in competitive situations such movements are judged based on a certain level of perfection (Payne & Issacs, 2020).

The process-oriented approach has increased in popularity over the last few decades. The advantage of using process-oriented assessments to evaluate an individual's progress is that they permit educators and coaches to identify specific skill components that may need improvement (Barnett et al., 2014). However, the product-oriented approach, which has been criticized for its disregard for the underlying movement processes, can be valuable in movement research that is designed to have educational implications. Children's success in movement outcomes is an important factor in keeping children interested and motivated in the activity. Product-oriented research may show that certain factors impact movement outcomes, thereby affecting a child's likelihood of engaging in physical activity in the future. So, although the process-oriented approach emerged from dissatisfaction with the product-oriented approach, both methods of evaluating movement have value in motor development research (Payne & Issacs, 2020).

2.5.3 iPads as an Assessment Tool in Physical Education

Traditional methods of performance assessment documentation are time consuming and cumbersome. The main barriers to effective assessment in PE are a lack of time to evaluate and provide explicit feedback to many students, complicated assessment procedures, overwhelming record-keeping, and assessing students without distracting them from their performance (Gallo et al., 2006). Research suggests that replacing traditional methods of written assessment and feedback will improve the assessment process (Barnett et al., 2002; Bennett, 2002; Buzzetto-More & Alade, 2006; Byers, 2001; Vendlinski & Stevens, 2002). Byers (2001) claims that interactive technology will facilitate dynamic feedback and enhance assessment in PE. However, little research has been done on how technology could improve the traditional method of assessing students' performance in PE (Franklin & Smith, 2015).

There is growing interest in extending the use of technology in education, with several researchers agreeing that novel instructional technologies can enhance the teaching of PE (Kretschmann, 2010; Mohnsen, 2012; Roblyer & Doering, 2013; Sinelnikov (2012). Specifically, McFarlane (2013) and Melhuish and Falloon (2010) claim that the important advantages of an iPad are its portability and its potential for real-time experiential learning. Furthermore, using video technology in PE is useful for both instruction and assessment; video provides visual feedback and allows for movement analysis, as well as documenting student improvement and motor performance (Mohnsen, 2012). In summary, mobile devices such as iPads can provide assessors with a tool to conveniently record grades and deliver written, audio, or even video feedback (VFB) by teachers using the iPad's dictation functionality. As a result, mobile devices such as iPads offer a potential solution to the challenges faced by PE teachers.

2.6 Technology in Physical Education

2.6.1 Digital Technologies in Physical Education

In PE, technology can be defined as something that helps students improve their physical performance, social interaction, or cognitive understanding of PE concepts (Mohnsen, 2012). One of the challenges in PE is to use technology in meaningful ways that enhance student's learning (Harris, 2009; Wyant & Baek, 2019). Traditionally, the use of technology in PE has been limited to administrative purposes such as tracking attendance, in addition to assessing, recording, and reporting the student's work (Thomas & Stratton, 2006). Within the past decade, PE teachers have started using digital technologies to enhance instruction, to assist with classroom management, and to facilitate assessments (Beseler & Plumb, 2019; Cummiskey, 2013; Koekoek & van Hilvoorde, 2018). With these technological advances, attempts have been made to create and use mobile applications (apps) for learning motor skills (Hall, 2012; Wulf et al., 2010). The introduction of movement analysis techniques can improve young people's understanding of their bodies and movements (Beseler & Plumb, 2019; Mohnsen, 2012; Spittle, 2021), and these improvements may positively influence their experiences outside of PE. Therefore, digital technologies could be seen to increase the relevance and authenticity of learning that occurs in PE (Casey et al., 2017; Koekoek & van Hilvoorde, 2018; Spittle, 2021).

According to education experts, mobile apps play several roles in enhancing the quality of PE. First, they are used as interactive communication tools, such as scoreboards, whiteboards, and display platforms. Second, they are used as classroom management tools, such as clocks, music screens, and microphones. Finally, they are valuable tools for self-assessment and feedback (Beseler & Plumb, 2019; Casey et al., 2017; Pyle & Esslinger, 2014). These features and functions are tools that can help students engage with motor skills learning while also managing and organizing their learning either individually, in pairs, or in small groups (Casey et al., 2017; Spittle, 2021).

It has been suggested that when it comes to using digital technology in PE, collaborative use, whether in pairs or small groups, is typically more effective than individual use (Dyson et al., 2010; Higgins et al., 2012; Koekoek & van Hilvoorde, 2018; Spittle, 2021). Digital technologies improve the learning process by allowing for real-time collaboration and P2P learning outside of the classroom or gymnasium (Beseler & Plumb, 2019; Sackstein, 2017). Lee & Gao (2020) have suggested that incorporating a video recording application into a P2P learning model can help children gain self-efficacy and social support from classmates because of the personalized and visualized feedback they receive from peers and video analysis. Furthermore, using a technological approach such as video analysis could improve the quality of P2P assessment. By using an approach that promotes both cognitive and motor learning, teachers create a high-quality learning environment in which students function productively and maintain a strong emphasis on learning. This student-centered approach is a positive development that can be linked to the increased use of technology in PE (Casey et al., 2017; Spittle, 2021).

Integrating digital technologies (i.e., laptops, tablets, mobile phones, and mobile apps) into the PE curriculum increases children's motivation (Pyle & Esslinger, 2014), cognitive understanding (Casey & Jones, 2011), and motor skill development (Beseler & Plumb, 2019; Liebermann et al., 2002; Oñate et al., 2005; Palao et al., 2015). However, we must be cautious when attributing these benefits to technology rather than a teacher's expertise in their use of digitally constructed pedagogies; digital technology alone does not accelerate student learning (Casey et al., 2017; Koekoek & van Hilvoorde, 2018; Spittle, 2021). Technology must be appropriately scaffolded into lessons such that students are correctly prompted into the desired actions, maintain interest, remain on task, and remain engaged in higher levels of thinking and performance while using it (Chatoupis, 2018; Mosston & Ashworth, 2008; Ward & Lee, 2005; Zhu & Dragon, 2016). As a result, teachers must align the digital technology use with their PE learning objectives. Furthermore, the selection of mobile apps and their integration into PE must be carefully considered to avoid compromising children's physical activity for the sake of efficient classroom management (Lee & Gao, 2020). While digital technology allows teachers to transform their lessons, it is how the devices are used that enables deep learning and helps teachers in changing their roles (Koekoek & van Hilvoorde, 2018; Spittle, 2021; Tay, 2016). It is important to note that the use of digital tools does not replace or diminish the role of the teacher, but rather shifts their role to that of a facilitator of physical movements.

Physical educators should use digital technology innovations and the creation of mobile apps to improve teaching and learning in PE and sport environments; however, research indicates that teacher experience and school background are important factors that can affect whether technology is implemented effectively (Pountney & Schimmel, 2015). Teachers are now facing a generation of digital natives and are expected to have a deep understanding of how to use educational technologies to facilitate student learning (Sun, 2015). Some studies conclude that one potential barrier to the use of digital technology in PE teaching is the teachers' lack of confidence in their own pedagogical-technology competency (Palao et al., 2015; Weir & Connor, 2009). However, other studies show that many PE teachers are confident in their abilities to use different types of technology (Angers & Machtmes, 2005; Woods et al., 2008).

Digital technology continues to flourish and expand, providing teachers with new educational opportunities (Beseler & Plumb, 2019; Koekoek & van Hilvoorde, 2018; Spittle, 2021; Wyant & Baek, 2019). With an increasing number of technologically advanced students and a drive to incorporate innovative and inspiring teaching methods, it is critical for physical educators to reinvent the teaching-and-learning process to meet the needs of these tech-savvy students (Koekoek et al., 2018). In a study conducted by Harris (2009), 12 school PE departments in New Brunswick, Canada, used Dartfish technology, a software technology that records and evaluates movement in a range of sporting circumstances. All the teachers agreed that by using features such as live capture and instant replay, video analysis software could improve student learning in PE (Harris, 2009). However, the author offered no empirical evidence to support their claims.

2.6.2 Utilizing Video Technology in Physical Education

PE teachers have used video technology to enhance their teaching (Ammah & Hodge, 2005; Calandra et al., 2008) and students' learning experiences (Finkenberg et al., 2005; Foster, 2004). Video allows students to better visualise and reflect on their performance so they can see themselves from the perspective of others and notice their interactions with the surrounding environment (Beseler & Plumb, 2019; Casey et al., 2017; Spittle, 2021; Zhang & Li, 2018). When students record videos of themselves, or of a peer, practicing a skill and compare their performance to a model demonstration, they gain a better understanding of the reasoning behind the correct technique (Beseler & Plumb, 2019; Spittle, 2021; Zhang & Li, 2018); model demonstrations can improve learning, particularly for visual learners (Mohnsen, 2012). The video not only replaces teacher demonstration, but also allows students to watch the skill many times at different speeds and to pause at various points to comment on the critical features (Beseler & Plumb, 2019; Spittle, 2021; Zhang & Li, 2018). Weir and Connor (2009) discovered that students found digital video most useful in the following areas: (a) identifying strengths and weaknesses; (b) technical skills and (c) identifying the key points of a skill. In conclusion, video technology allows students to have a better understanding of their body, their movements, and arguably, themselves (Beseler & Plumb, 2019; McNicol et al., 2014; Spittle, 2021; Zhang & Li, 2018), resulting in an objective assessment of their skill development and progress over time.

Students' use of video technology in PE is influenced by two factors: age and ability level (Leight et al., 2009). An inexperienced student may use video replay to view practise trials to establish a connection between a kinesthetic experience and a visual experience. Video replay is most successful with students who have at least an intermediate ability level, as students need prior knowledge of the skill as well as viable mental imagery to use the information given by the video (Beseler & Plumb, 2019; Mohnsen, 2012; Zhang & Li, 2018). Students with advanced skills find video useful for improving motor performance because it allows them to see their movements and analyze their mistakes (Leight et al., 2009; Spittle, 2021; Zhang & Li, 2018). There has recently been an increase in the number of mobile devices capable of capturing high-quality video, including advances in increased frame rates for slow motion capture and replay (Casey et al., 2017; Leight et al., 2009). Since advanced movement is often rapid and can be difficult to analyze at normal speed, slow motion replay and freeze frame capabilities are essential for skill improvement (Mohnsen, 2012; Spittle, 2021). In conclusion, video replay can be used across all abilities and ages to engage students and increase motor learning by providing opportunities to visualize and evaluate their performance (Beseler & Plumb, 2019; Casey et al., 2017; Leight et al., 2009; Mohnsen, 2012; Spittle, 2021; Zhang & Li, 2018).

Mobile devices, such as iPads, have become increasingly popular in PE and provide a multimedia platform with multiple applications and video recording capabilities (Casey et al., 2017; McManis & Gunnewig, 2012; Weir & Connor, 2009). Compared to desktop computers, iPads are portable which allows students to learn in a variety of settings (McFarlane, 2013; Melhuish & Falloon, 2010). Thus, teachers can plan lessons with student independence in mind because the digital content can be accessed from anywhere and attempted when students feel ready, ensuring that students progress at their own pace (Beseler & Plumb, 2019; Casey et al., 2017; Spittle, 2021). There is also some evidence that mobile devices give youth learning freedom by providing immediate guidance and feedback (Beseler & Plumb, 2019; Higgins et al., 2012). Intuitive touch gestures of iPads, such as pinching and spreading, allow the students to zoom in or out and explore the video of the model demonstration on the screen and motivate the students to reflect more deeply on their own performance (Hung et al., 2018). More importantly, iPads incorporate the use of video cameras and health-related mobile applications to reinforce observational learning through visual digital feedback and video analysis (Beseler & Plumb, 2019; Koekoek & van Hilvoorde, 2018; Weir & Connor, 2009; Zhang & Li, 2018). This feedback can increase learning, motivate students, and reinforce desired behaviours or movements (Beseler & Plumb, 2019; Leight et al., 2009; O'Loughlin et al., 2013; Palao et al., 2015; Spittle, 2021; Weir & Connor, 2009).

2.6.3 WISER Model to Enhance Learning in Physical Education

Hung et al. (2018) conducted a study to investigate how mobile technology, specifically the iPad, could be integrated to facilitate badminton skill learning and increase student motivation. The participants were divided into two groups: one that would use the iPad and another that would use traditional PE methods. According to the findings, students who used the iPads outperformed those who used traditional PE methods in terms of both badminton skill learning and motivation. The researchers concluded that using iPads enabled students to change their overall perspectives of badminton skill learning, achieve the desired skills at their own pace, and receive immediate visual feedback on motor skill learning.

In the traditional approach to teaching PE, the teacher demonstrates the correct skills and movements in front of the entire class; however, this method only allows students to view the teacher's demonstration in a fixed location at the same time. During the skill development process, students must practice and receive feedback from the teacher on a regular basis. (Wulf et al., 2010). However, with class sizes growing larger, educators are facing the challenge of giving complete feedback to all students in a timely fashion (Beseler & Plumb, 2019; Evans, 2017). Furthermore, the instructor may encounter difficulties when evaluating each student's individual movements and then providing appropriate and sufficient feedback based on different levels of physical ability. The WISER model developed by Hung et al. (2018) outlines how iPads could be incorporated into PE settings to enhance student learning. It consists of a five-step model (Hung et al., 2018), which are as follows:

Step 1: **Watching** model demonstrations presented in handheld technology,

Step 2: **Imitating** demonstrations and immediately recording via handheld technology,

Step 3: **Self-examining** the recorded videos for identification,

Step 4: **Enhancing** the motor skills via comparing the videos, and

Step 5: **Repeating** movements and seeking advice from the teacher.

Individuals prefer iPads over desktop computers due to their mobility and more intuitive operation methods (Beseler & Plumb, 2019; Reychav & Wu, 2015). Compared to the traditional approach, iPads provide a more versatile and personalized pace of learning with the WISER model, which could play an important role in: (1) learning motor skills from a holistic viewpoint, (2) achieving the desired skills at their own pace supplemented by digitized videos on iPads, and (3) offering precise and immediate feedback on motor skill learning. The WISER model has been identified as a successful pedagogical model for PE in terms of promoting student self-paced learning and reducing the instructor's teaching burden by creating a more flexible teaching environment (Hung et al., 2018). The use of iPads in combination with the innovative WISER model could be applied to a variety of learning environments and teaching styles, such as the reciprocal teaching style described by Mosston and Ashworth (2008).

2.6.4 Research Using iPads in Physical Education

Over the last decade, there has been an increase in studies investigating app-integrated PE using the iPad to facilitate the teaching and learning process. In one of the first studies that examined the roles of iPads in PE, Sinelnikov (2012) described the experiences of 12- and 13-year-old students in a 20-lesson volleyball unit. Throughout the study, a variety of mobile apps were used to accomplish a range of tasks, such as creating and editing images, developing spreadsheets and graphs for statistical analysis, and giving team presentations to the entire class. Although the focus was on iPads as tools in the learning process, Sinelnikov (2012) concluded that "the iPad seems to offer numerous innovative outlets for advancing and using technology in PE classes" (p.45). It has been suggested that mobile apps may provide more convenient and efficient content delivery while also serving as venues to motivate the iGeneration to participate in physical activity (Koekoek et al., 2018; Krause & Sanchez, 2014).

Lee and Gao (2020) investigated the short-term impact of mobile app-integrated PE classes on children's physical activity levels and psychosocial beliefs (i.e., self-efficacy, social support, and enjoyment). They hypothesized that incorporating mobile applications using an iPad could improve children's psychosocial beliefs and increase their physical activity levels in PE. The study included 157 fourth and fifth grade students; 77 were assigned to the app-integrated group, while the remaining 80 were assigned to the traditional PE group. In terms of physical activity levels, the findings revealed that children in both groups had lower levels of moderate-to-vigorous physical activity; however, the decrease in the app-integrated group was substantially greater than the decrease in the traditional PE group. With regards to psychosocial beliefs, the results showed that there were no significant differences between the two groups.

Clearly these findings contradict their hypotheses, and it is interesting to see that similar results had been obtained in an earlier study (Zhu & Dragon, 2016).

Zhu and Dragon (2016) investigated the effects of integrating mobile technology, specifically iPads and mobile apps, on student situational interest and physical activity levels in PE. Sixth-grade students were assigned at random to either an experimental group that used mobile technology or a comparison group that did not; both groups received five identical PE lessons. According to the results, students in the experimental group reported significantly lower levels of physical activity and situational interest than their peers in the comparison group. However, further statistical analysis showed that the experimental group's student steps/minute rate gradually increased throughout the classes, while the comparison group remained relatively stable. This increase may be partly attributed to students becoming more comfortable with the mobile technology, thus allowing them to spend more time participating in physical activities.

Overall, the results from these studies demonstrate that mobile technologies with no physical activity prompts have little effect on improving physical activity levels, situational interest, or psychosocial beliefs in the short term (Lee & Gao, 2020; Zhu & Dragon, 2016). As a result, it is important to consider classroom dynamics when evaluating the constraints and strengths that technology-integrated PE lessons can present in a traditional PE setting.

2.6.5 Video Feedback in Physical Education

Feedback is inextricably linked to the processes of learning and teaching (Bangert-Drowns et al. 1991) and its use during the teaching process has been the focus of many studies (Georges & Pansu, 2011). In PE, the learning process when acquiring motor skills focuses on exploring actions and movements, which requires a lot of practice and feedback (Schmidt, 1988).

According to motor learning research, feedback has been found to improve the acquisition of fine and gross motor skills (Schmidt & Wrisberg, 2008; Wrisberg, 2007) and is one of the most important instructional variables affecting skill learning. Feedback is the return of performance information that occurs within a behavioural regulation loop, where error detection and correction are essential to motor learning (Mulder & Hulstijn, 1985; Schmidt & Lee, 2005). It can be delivered in several ways, including visual, auditory, touch, and multimodal, and can be given at differing points in time (Casey et al., 2017). While the use of feedback enhances motor learning, the most effective method of providing feedback is still unknown (Sigrist et al., 2013); traditionally, coaches and teachers have used verbal feedback (Casey et al., 2017).

Technological progress has led coaches and physical educators to reconsider their methods for giving movement-related feedback and to experiment with innovative learning aids (Kretschmann, 2012; Kretschmann, 2015; Leight et al., 2009; Trout, 2013; Wilson, 2008; Wyant & Baek, 2019); specifically, providing PE teachers with an increased number of digital tablets has led them to create learning aids based on VFB (Beseler & Plumb, 2019; Gubacs-Collins & Juniu, 2009; Kretschmann, 2015). VFB is defined as the replay of a learner’s own static or dynamic image to themselves (Potdevin et al., 2018; Spittle, 2021). It is an extrinsic or augmented source of feedback, since it provides additional knowledge about an individual’s actions that would not be available without the use of an external aid (Schmidt & Lee, 2005; Spittle, 2021). It differs from intrinsic feedback, which is defined as information that can be detected without the use of external aids. VFB can be used to guide the actions of students who have difficulty interpreting intrinsic feedback or who have less stable movement patterns (Hodges et al., 2003; Swinnen 1996). In fact, VFB has been shown to improve an individual's ability to learn and perform motor skills (Boyce et al., 1996; Deakin & Proteau, 2000;

Finkenberg et al., 2001; Kretschmann, 2017), because digital video recordings allow students to better visualize and reflect on their errors, strengths, and weaknesses (Beseler & Plumb, 2019; Leight et al., 2009; Spittle, 2021; Weir & Connor, 2009; Zhang & Li, 2018).

2.6.6 Video Feedback Research in Physical Education

Research in sport settings has validated the use of VFB in developing motor skills (Kretschmann, 2017; Liebermann et al., 2002; Oñate et al., 2005; Palao et al., 2015). Numerous studies have demonstrated that VFB is useful in the development of various sport skills over relatively short learning periods. Examples include the golf swing (Guadagnoli et al., 2002), flip turns in swimming (Hazen et al. 1990), swimming front crawl (Kretschmann, 2017), gymnastics (Potdevin et al., 2018; Winfrey & Weeks, 1993), soccer skills (Ziegler, 1994), diving (Thow et al., 2012), weightlifting (Rucci & Tomporowski, 2010), spike jump in volleyball (Parsons & Alexander, 2012) and hurdling (Palao et al., 2015). While the findings demonstrate that providing VFB on motor learning was effective, how it was used in the studies varied depending on the learning environment.

However, the use of technology in general, and videotaping in particular, appears to be less common in PE than in sport settings (Palao et al., 2015; Ste-Marie et al., 2012). Reasons for this gap in the scientific literature may include the fact that the number of students per group is larger than in sport training groups, or that athletes and coaches theoretically have greater levels of investments in specific skill improvements (Guadagnoli et al., 2002; Kretschmann, 2012; Smith & Loschner, 2002; Wyant & Baek, 2019). Meanwhile, PE teachers may emphasize different aims such as motor, cognitive, social, moral, spiritual, or cultural development (Sallis & McKenzie, 1991). Additional disincentives for PE teachers to evaluate the effectiveness of VFB could be related to time and financial constraints (Kretschmann, 2012; Norris et al., 2002; Weir

& Connor, 2009; Wyant & Baek, 2019). Indeed, in the study conducted by Weir and Conner (2009), the major negative aspect reported by both teachers and students was that incorporating technology was particularly time consuming and detrimental to the students' use of practice time.

Nonetheless, research has shown that using VFB in PE can improve the effectiveness of demonstrations for skill learning (Kretschmann, 2017; Lhuisset & Margnes, 2014; Potdevin et al., 2018), as it gives students a better understanding of the movements they have performed or are about to attempt (Beseler & Plumb, 2019; Casey et al., 2017; Leight et al., 2009; Mohnsen, 2012; Spittle, 2021; Weir & Connor, 2009; Zhang & Li, 2018). Research on the impact of VFB on motor skill acquisition in a PE setting across education levels has shown its effectiveness when combined with teacher feedback (Amara et al., 2015; Kretschmann, 2015; Potdevin et al., 2013; Uhl & Dillon, 2009). By providing performance analysis sheets listing the critical skill features, students can use the VFB to evaluate their own performance against a set of criteria (Beseler & Plumb, 2019; Hamlin, 2005; McMillan & Hearn, 2008; Mohnsen, 2012).

Most of the studies that have investigated the impact of VFB on learning experiences in PE have used qualitative methods to examine learning outcomes. Using a semi-structured interview methodology, Kretschmann (2015) discovered that 10-year-old students found VFB useful for learning swimming skills. The same technique was used by O'Loughlin et al. (2013) to determine the effect of VFB on student learning outcomes in PE. The study revealed that VFB influenced self-reported motivation, self-assessment, and engagement when learning basketball skills in students aged 9 to 10. This was demonstrated, for example, by an increase in children's engagement with practicing basketball skills both inside and outside of class time to "look good" on video (O'Loughlin et al., 2013). Similarly, Casey and Jones (2011) showed the usefulness of using VFB to improve engagement with disaffected Year 7 students who developed a greater depth of knowledge about throwing and catching skills. Other studies have confirmed that VFB improves motivation during PE learning (Potdevin et al., 2013; Weir & Connor, 2009).

In comparison, Palao et al. (2015) evaluated the effectiveness of VFB on student learning in PE using a quantitative approach. Their study examined skill and knowledge outcomes in hurdling using three groups: verbal feedback from teacher; VFB plus verbal feedback from teacher, and; VFB plus student feedback. Regarding skill technique, significant results were found between initial and final means between the VFB plus verbal feedback from teacher group (20.9%, $p = 0.01$) in addition to the VFB and student feedback group (27.2%, $p = 0.009$). For knowledge outcomes, the VFB plus verbal feedback from teacher group saw an increase of 7.4% over the only verbal feedback from teacher group. This study demonstrates that the presence of VFB can positively improve student learning and the quality of their motor performances.

Kretschmann (2017) conducted a study to determine the impact of using a tablet to provide video feedback on swimming performance in PE. Two 5th grade PE swimming classes were randomly assigned as an experimental group (n=16) and a control group (n=15). A trained PE teacher administered the program, which lasted for seven weeks. Experimental group students were exposed to a standardized video analysis and feedback program using a tablet computer and were given feedback using slow motion right after their individual front crawl performance. Meanwhile, the control group didn't integrate any media or technology and used traditional teaching methods such as verbal feedback only. Pre- and post-test study design was used to measure student's swimming performance in front crawl at baseline and after the 7-week class period. Experimental group students significantly ($p<0.05$) improved in front crawl racing-results from pre- to post-test. Semi-structured interviews with selected experimental group students revealed that they found the tabled-based video feedback to be helpful and motivating for their learning process with regards to improving their front crawl technique and eventually their race results. In conclusion, video feedback via tablet technology in PE swimming classes was an effective teaching method for improving swimming performance in 5th grade students. The tablet-based video feedback technology was found to be superior to traditional teaching methods, was feasible in an aquatic environment, and students saw it as a beneficial addition to PE (Kretschmann, 2017). This finding is similar to Tanaka et al. (2014), who used tablets for instant video feedback in Japanese PE students. Evaluation questionnaires completed in their study showed that most of the students found video necessary and useful during the learning process.

A review of the current literature reveals that much remains to be discovered about the effects of VFB and peer modelling on the overall performance of adolescents, especially in terms of FMS performance in middle school students. Much of the existing research concerning the use of VFB to enhance motor learning has been undertaken under strictly controlled experimental conditions. Few studies have sought to explore the impact of VFB on the skill learning experience of the students in a structured, school-based PE setting. Most of those studies have used only qualitative approaches to implicate the potential value of VFB to enhance skill acquisition, students' engagement, or self-assessment ability (Potdevin et al., 2018). Additionally, research on the effect that the order of performance between peers has on overall skill performance outcomes is currently lacking. As a result, the combination of a reciprocal teaching model with VFB and peer modelling has yet to be studied.

2.6.7 Move Improve® Application

The VFB technology used in this study is called Move Improve® (MI). This thesis will use the MI FMS module and a reciprocal teaching model to determine if the order of performance between peers has any effect on their overall skill performance. Appendix A contains more information about the MI application and how it was used during data collection.

2.7 Reliability

2.7.1 Defining Reliability

The reliability of an assessment tool is an important and necessary quantitative property and is defined as the extent to which a measurement is consistent and free from error; it can also be thought of as reproducibility or dependability (Portney, 2020). For a measure to produce valid scores, it must demonstrate an acceptable level of reliability (Burton & Miller, 1998; Lacy & Williams, 2018; Portney, 2020). The reliability coefficient is formally quantified using an intraclass correlation coefficient (ICC). The ICC ranges from 0.00 to 1.00, with values closer to 1.00 indicating greater reliability (Weir & Vincent, 2020). The ICC is an established statistic for assessing measurement reliability and is defined by using variance components (Lacy & Williams, 2018; Shrout & Fleiss, 1979). A high-quality measure should have acceptable reliability across time (i.e., test-retest), across individuals administering the assessment (i.e., inter-rater), and across repeated scoring attempts (i.e., intra-rater) (Lacy & Williams, 2018). Assessment tools with lower reliability are assumed to have more measurement error in a score (Portney, 2020; Weir & Vincent, 2020).

Once an assessment tool has been developed, there are two types of reliability that are commonly examined, intra-rater reliability and inter-rater reliability. Intra-rater reliability is the degree of consistency of scores on an assessment across at least two occasions (e.g., days, weeks, months) by a single rater (Portney, 2020; Weir & Vincent, 2020). It represents the consistency of a rater’s score on different occasions. According to Maeng et al. (2017), intra-rater reliability is dependent on the abilities of trained raters, and on good standardization of the task or item being assessed. Inter-rater reliability is the consistency of scores obtained from two or more raters independently scoring the same individuals (Portney, 2020; Weir & Vincent, 2020). Inter-rater reliability is best assessed when all raters can measure a performance during a single trial, where they can observe an individual simultaneously and independently (Portney, 2020). Previous research has shown that inter-rater reliability is an important part of assessing FMS competence (Barnett et al., 2014). Since the MI application is an observational assessment tool which uses video performance technology, both intra-rater and inter-rater reliability are important.

2.7.2 Reliability of Assessment Tools

A variety of assessments are used to evaluate movement or motor performance, with most skill assessments being developed and tested for reliability and validity in students (Goodway et al., 2014). Although there is currently no gold-standard assessment tool or rubric available to evaluate FMS performance in adolescence, many of the commonly used process-oriented assessments in PE follow the criteria and physiological prompts established by Roberton (1977). Due to the differences in the tools and techniques used in motor skill assessments, the assessment results may not be interchangeable. Assessors should be consistent in their use, selection, and interpretation of assessment tools to ensure proper evaluation of progression over time (Logan et al., 2012; Minick et al., 2010; Palmer & Brian, 2016).

The intra-rater and inter-rater reliability of an assessment tool are important factors to consider when determining its objectivity. In many studies, raters undergo a period of training so that techniques are standardized. This is especially important when measuring devices are new or unfamiliar, or when subjective observations are used (Weir & Vincent, 2020). However, the accuracy and reliability of these assessments can be influenced by several factors, including rater background and characteristics (Nadeau et al., 2008; Palmer & Brian, 2016), access to motor skill observation training (Haynes & Miller, 2015), ability or level of expertise (Borghouts et al., 2017), and the population being assessed (Bailey et al., 1995). As a result, when one rater makes two measurements, we must consider the possibility of rater bias.

Raters can be influenced by their memory of the first score. This is most relevant when observers use subjective criteria to rate performances (Weir & Vincent, 2020). The major precautions against rater bias are to develop objective grading criteria, to train raters in the use of the instrument, and to record reliability among raters. However, even with detailed operational definitions and equal skill levels, raters may not agree about the quality of the motor performance (Portney, 2020). Videotapes of individuals performing activities have proved useful for allowing multiple raters to observe the exact same performance (Carballo-Fazanes et al., 2021; Portney, 2020; Rintala et al., 2017; Slotte et al., 2015).

2.7.3 Reliability of Observational Assessment Tools

Ethical issues must be considered when conducting research with children. Observation as a research method is unobtrusive and thus appropriate for use with children. Unfortunately, the reliability of observational tools is in question. According to a recent systematic review conducted by Eddy et al. (2020) that examined the validity and reliability of a variety of observational assessment tools to measure FMS in school-aged children, there is currently insufficient evidence to justify the use of any observational FMS assessment tool for universal screening in schools. Earlier studies evaluating FMS performance have used either live assessments or video recordings. The TGMD-2 (Ulrich, 2000) was used in a study by Barnett et al. (2014), which recorded inter-rater reliability based on live observation in six object-control skills. The ICC for object-control skills was 0.93, ranging in individual skills from 0.71 (catch) to 0.94 (dribble). Another study, conducted by Slotte et al. (2015), examined children's motor skills using video recordings and reported intra-rater reliability for 24 children's motor skills. In their study, reliability as ICC was 0.978 for locomotor skills and 0.995 for object-control skills. Additional reliability studies will provide valuable information for test developers about the characteristics of such tests and will inform future test development.

2.7.4 Studies Examining the Reliability of Assessment Instruments

A study conducted by Lee (2016), assessed the inter- and intra-rater reliability of four process-related scoring protocols:

(1) The Sport Technology Research Lab (STRL) Overhand Throwing Rubric;

(2) TGMD-2

(3) PHE Canada; and,

(4) New Hampshire Association of Health, Physical Education, Recreation, and Dance (NHAHPERD) overhand throwing assessments.

Five videos of elementary students performing an overhand throw were analysed by ten raters of diverse expertise and physical education backgrounds. Results showed that despite the similarities in assessment batteries, they were not equally reliable (Lee, 2016). Others have found that due to the differences in the tools and techniques used in motor skill assessments, the assessment results may not be interchangeable (Logan et al., 2012; Minick et al., 2010; Palmer & Brian, 2016). Therefore, assessors need to be consistent in their use, selection, and interpretation of measurement tools to ensure proper evaluation of skill progression over time.

Palmer and Brian (2016) assessed the difference in scoring of elementary school children between adult expert and novice coders using the TGMD-2. Novice coders with no previous TGMD-2 coding experience were recruited and subsequently introduced to the protocol during a two-hour intensive training session with an expert. During the training session, novice coders independently rated an example video and then participated in an in-depth discussion regarding results and discrepancies. The results indicated that novice coders were unable to achieve significant agreement with expert coders (Palmer & Brian, 2016). Specifically, locomotor skills had a higher percent difference (59.9%) compared with object control skills (47.1%) (Palmer & Brian, 2016). These large differences in both subscales imply that additional training is required, but that perhaps more training should be dedicated to locomotor skills as opposed to object control skills.

Another study evaluating novice and expert coders, examined the inter-rater reliability of the Functional Movement Screen in university-aged students (Minick et al., 2010). Four raters (two experts and two novices) independently coded the participants. Novices were defined as having taken an introductory training course and having used the assessment tool for less than a year. Conversely, experts were defined as having over 10 years of experience using the Functional Movement Screen, in addition to contributing to the development of the tool. The results demonstrated excellent agreement between expert and novice coders for the Functional Movement Screen assessment tool when assessing university-aged students.

Nadeau et al. (2008) investigated the validity and reliability of the Team Sport Assessment Procedure for ice hockey by using student-observers and two physical educators' observers. Student-observers were recruited from the ice hockey teams and partook in a one-hour training session. The percentage of agreement between student-observers ranged from 59-95%, while physical educator raters ranged from 80-82%. These results suggest that there is greater variability in less experienced coders, also known as novice coders. Although the student-observers only had a brief one-hour training session with the assessment tool, it is important to note that the evaluators were familiar with ice hockey, and therefore did not require substantial training on observation techniques and sport-related technical skills.

The contradictory findings between these studies may be a result of the assessment tool itself, the population being evaluated, and/or the training sessions for the novice coders. For example, Palmer and Brian (2016) provided specific lessons with sample videos from the population of interest and allowed for a consensus discussion to ensue between coders. Meanwhile, Minick et al. (2010) reported that the novice coders attended a standardized introductory training course. Roberton and Halverson (1984), known for their use of process-oriented assessments, has pointed out some limitations of the component approach. Specifically, they conclude that a comprehensive understanding of motor development is required, in addition to a prolonged period of study and practice using the tool, if the results obtained from a process-oriented assessment tool are to be considered reliable and accurate (Weir & Vincent, 2020). Their reasoning explains why inexperienced coders exhibit greater variability in their assessments compared to experienced coders. For example, the study conducted by Nadeau et al. (2008) demonstrated a total of 36% variability in the percentage of agreements (range of 59-95%) amongst novice coders. In contrast, the expert coders only had 2% variability in their percentage of agreements (range of 80-82%).

Chapter Three: Manuscript 1

Peer to Peer Learning: The Impact of Order of Performance on Learning Fundamental Movement Skills Through Video Analysis with Middle School Children

3.1 Abstract

Purpose: Through video analysis, this paper explores the impact order of performance has on middle school students' performance of fundamental movement skills within a peer to peer (P2P) learning model. Order of performance refers to the order in which a student performed a skill while paired up with a peer.

Method: Using a mobile application, Move Improve®, 18 students (eight males, 10 females) completed a standing jump (SJ) and hollow body roll (HBR) in partners assigned to order of performance (evaluator/performer). An independent samples *t*-test was conducted to evaluate the differences in the mean scores between students that performed first and those who performed second for each skill.

Results: There was a significant difference in SJ scores ($p < 0.01$), where students who performed second had a higher average score than their peers who went first. Although not statistically significant ($p = 0.293$), results for HBR also showed a similar performance pattern for students who went second compared to those who performed first.

Conclusion: The order of performance within a P2P learning model may have a significant effect on performance scores for SJ but not for HBR. Reasons for the discrepancy may be due to a combination of skill familiarity, skill complexity, and training of observational learning.

Keywords: fundamental skill performance; observational learning; peer learning; technology

3.2 Introduction

For children, technology is an integral part of daily life (Cappella, 2000; Prensky, 2001) thereby influencing how they spend their spare time. Excessive sedentary screen time has contributed to an increase in childhood obesity, disease, and difficulties learning in school (Hancox & Poulton, 2005; Robinson, 2001; Vandewater et al., 2005, 2006). According to the 2018 ParticipACTION (2018) Report Card on Physical Activity for Children and Youth, only 35% of 5-17-year olds in Canada currently meet the physical activity recommendation of performing at least 60 minutes of daily moderate-to-vigorous physical activity. Because children are attracted to screen-based technologies (Mohnsen, 2012), by integrating them into PE, they can become physical activity tools which may be used to create more engaging lesson plans and reduce sedentarism in PE class (Casey et al., 2017; Tanaka et al., 2018).

FMS are defined as the building blocks that lead to specialized movement sequences required for participation in both organized and non-organized physical activity (Gallahue & Ozmun, 2006; Hardy et al., 2012). Commonly developed in childhood and refined into more complex or sport-specific movement patterns, they are classified into three separate categories: locomotor (e.g., running and skipping), manipulative or object-control (e.g., catching and throwing), and non-locomotor (e.g., balancing and twisting; Payne & Issacs, 2020). The majority of children are developmentally able to achieve most of the basic FMS by the age of 6, with the mastery of more complex FMS skills being achieved by age 10 or 11 (O'Brien et al., 2016). Therefore, it is reasonable to expect that adolescent youth (12-17 years) should demonstrate competency in FMS. The development and mastery of FMS are essential for the growth and development of a physically literate child, which increases the likelihood of continuous engagement in physical activity and sport throughout the lifespan (Barnett et al., 2009; Fisher et al., 2005; Hardy et al., 2012; Higgs, 2010; Stodden et al., 2009). If children are unable to grasp these constructs or skills, the likelihood of them withdrawing from participation in organized sports and play experiences increases (Hardy et al., 2012).

Traditional teaching methods in PE have mostly revolved around the ‘mimic/practice’ style where the teacher demonstrates the skill, and the students attempt to replicate the movement individually and practice privately while receiving minimal feedback from the instructor (Mosston & Ashworth, 2008). An alternate approach is the P2P learning model, where peers are directly involved in the teaching and evaluation process (Jenkinson et al., 2014; Kelly & Katz, 2016; Topping & Ehly, 1998; Ward & Lee, 2005). Multiple studies have demonstrated that P2P learning can be utilized in PE to achieve increased performance scores (Ernst & Byra, 1998; Goldberger et al., 1982; Houston-Wilson et al., 1997). This can be explained by the reciprocal nature of P2P learning where students are required to observe and evaluate peers, resulting in students performing at an equal or improved level with less practice time (Goldberger et al., 1982; Kolovelonis et al., 2011). When students are required to observe and evaluate their peers in a P2P model, instead of the passive mimic/practice teaching style, they are actively engaged through observational learning.

Observational learning is the ability to learn a motor skill through observation of another individual performing said skill (Cross et al., 2008). Observational practice is often overlooked when it comes to learning simple and complex motor skills. This idea is partially based on previous findings that observational practice is typically less effective than physical practice; however, it has consistently shown to be more effective in the learning of motor skills than no practice at all (McCullagh & Weiss, 2001). Futhermore, research has shown that observational practice can make important and unique contributions to learning, especially when combined with physical practice (Shea et al., 2000). Observational learning is a succesful strategy to implement in PE because observation has been shown to promote the learning of a wide variety of motor skills (Hodges et al., 2007; Lago-Rodríguez et al., 2014; McCullagh et al., 1989; Ste-Marie et al., 2012; Vogt & Thomaschke, 2007). Research has shown that observing both experienced (Al-Abood et al., 2001; Hodges et al., 2003; Lee et al., 1995) and novice (Black & Wright, 2000; Buchanan & Dean, 2010; Hayes et al., 2010; McCullagh & Meyer, 1997) models leads to significant learning. However, recent research suggests that those benefits are optimized if the observer is told beforehand the quality of the performance that they are about to observe (Andrieux & Proteau, 2014; Rohbanfard & Proteau, 2011). This insight could be useful in situations that benefit from having a video observation model in PE (Andrieux & Proteau, 2016).

With the advances in modern technology, increasing efforts in developing and using multimedia for the learning of motor skills has been observed (Katz, 2003; Leser et al., 2011; Wulf et al., 2010). The past decade has seen a rise in studies exploring app-integrated PE using the iPad to facilitate teaching and learning processes (Krause & Sanchez, 2014; Lee & Gao, 2020; Watterson, 2012; Zhu & Dragon, 2016). The iPad supports multiple representations of the information by using various media modalities, such as animations, audio, graphics, text, and video (Mas et al., 2003), which allows it to be compatible with the modern nature of PE (Casey et al., 2017). For example, iPads integrate the use of video cameras and health-related apps to reinforce observational learning through video analysis (Weir & Connor, 2009). Learning from multimedia gives the student control over their own learning, as they can easily and freely control their navigation with the learning content (Cairncross & Mannion, 2001; Mason & Rennie, 2006; Mayer, 1997).

Previous studies have shown that digital feedback can have a positive impact on primary children's learning in PE. For example, digital video is commonly used in PE to provide augmented visual feedback to learners, teachers, and coaches. This feedback can be defined as an external stimulus used to increase learning, motivate learners, and reinforce behaviours (Leight et al., 2009; O'Loughlin et al., 2013; Palao et al., 2015; Weir & Connor, 2009). Interviews with children revealed that the use of digital video as part of basketball skill instruction proved motivational in increasing classroom engagement and practice outside of class time so children could 'look good' on video (O'Loughlin et al., 2013). In a study conducted by Weir and Connor (2009), students stated that identifying strengths and weaknesses, technical skills, and key points related to certain skills, were improved using digital video. Furthermore, a quantitative video feedback examination of skill and knowledge outcomes in hurdling using three groups (verbal feedback from teacher, video feedback plus verbal feedback from teacher, and video feedback plus student feedback) found significant differences in skill technique between the video plus teacher feedback group (20.9%, $p = 0.01$) and the student plus video group (27.2%, $p = 0.009$; Palao et al., 2015). With regards to knowledge outcomes, the video plus student feedback group saw an increase of 7.4% over the verbal feedback only group (Palao et al., 2015). Recently, it has been suggested that incorporating an application with video recording capabilities into the P2P learning model may help children gain self-efficacy and perceived social support due to the tailored and visualized feedback they receive through peers and video analysis (Lee & Gao, 2020). The outcomes of these studies are promising as they demonstrate that the presence of digital video can positively augment student learning and the quality of their performances.

As the integration and prevalence of digital technologies in PE continues to rise, more research is needed to evaluate the effectiveness of multimedia-supported teaching on motor skill development. Research has shown that the use of digital technologies and hardware, such as the iPad, has the potential to improve motor skill performance (Hung et al., 2018; Weir & Connor, 2009). Furthermore, multiple studies have demonstrated that P2P learning can be utilized in PE to achieve increased performance scores (Ernst & Byra, 1998; Goldberger et al.,1982; Kolovelonis et al., 2011). The combination of video analysis and observational learning provides an ideal solution for the mastery of FMS, as it allows students to receive rapid feedback on their performance. However, the effectiveness of this combination facilitated through mobile applications in PE settings has not been well researched, especially as it relates to the mastery of FMS within a P2P learning model. Therefore, the purpose of this study is to determine the impact order of performance has on middle school students' performance of FMS within a P2P learning model using a mobile application called MI. It was hypothesized that students who performed second would have greater performance scores than those who performed first.

3.3 Methods

3.3.1 Research Design and Participants

The research activities using the iPads were integrated into the students' daily lesson plan, and the total duration of each PE class was 30 minutes in length. The data was collected during two of their regularly scheduled PE classes over the course of two school days in the same week. In the first class, students completed the training. Meanwhile, in the second class, students engaged in P2P learning of two FMS during which additional data for the study was collected.

Purposive sampling was used to acquire the 18 middle school students from a small private parochial school in Alberta, Canada. The participants included eight male students and 10 female students, all ranging from grades 5 to 8. See Table 3.1 for participant information. They were split into two, separate groups based on gender. Each group had two 30-minute classes. In accordance with the inclusion and exclusion criteria, all participants in the study were able bodied, without any physical impairments or disabilities. Parental consent and student assent for the collection of data was obtained prior to the study. Ethical approval was obtained through the Conjoint Health Research Ethics Board (CHREB) at the University of Calgary.

Table 3.1 Participant Information

Grade Level	# of participants	# of males	# of females
5	3	2	1
6	8	4	4
7	4	1	3
8	3	1	2

3.3.2 Move Improve® Mobile Application

MI is a mobile application which allows users to evaluate how well an individual performs a movement skill. First, MI works by having a pre-created list of available skills that users can select to test themselves or their peers on (see Figure 3.1). Once a skill is selected (see Figure 3.2), users are then provided with an instructional video on how the skill should properly be performed. Finally, when the video ends, a component list that breaks down sequential movements and key points for correctly completing the skill appears (see Figure 3.3). With peers, one peer becomes the evaluator and the other peer takes the role of performer. The evaluator takes a video of the performer who attempts to do the skill (see Figure 3.4). The video is viewed through the MI app. The evaluator works with the performer to discuss the performance and begins an assessment using the performer's video (see Figure 3.5). MI allows the evaluator to step through each of the checklist components one at a time, providing an image of the correct action and providing a rating choice for each component: 'Yes' (the component was performed correctly by the partner (matches image and instructional video), 'Partial' (the partner performed the component, but not entirely correctly), or 'No' (the component was not performed correctly; see Figure 3.6). Each of the three values is given a numerical score (Yes =

3, Partial = 2, No = 1). After going through all the components, a summary list of them is shown on the MI app with the corresponding score, along with a total score out of a given value depending upon the number of components. This provides evaluators and performers with tangible feedback on the performance, in addition to directing their attention to certain components of a skill on which to focus for improvement. Evaluators and performers are encouraged to do the assessment together and discuss the options. Then they reverse the evaluator and performer roles and begin a new assessment.

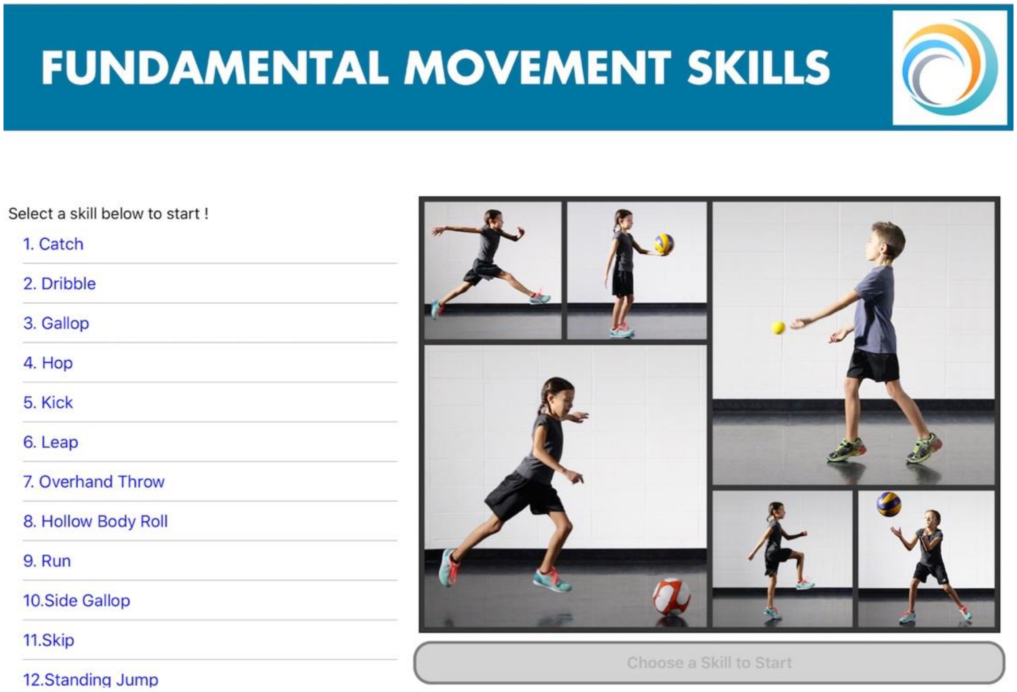


Figure 3.1 Move Improve® Fundamental Movement Skills List

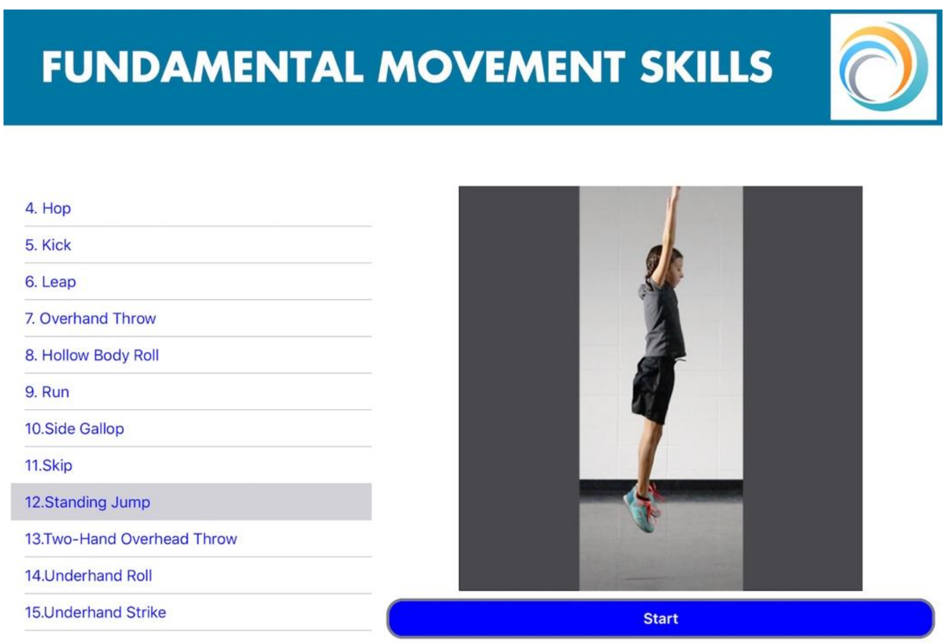


Figure 3.2 Choosing a Skill, Standing Jump, from Fundamental Movement Skills List

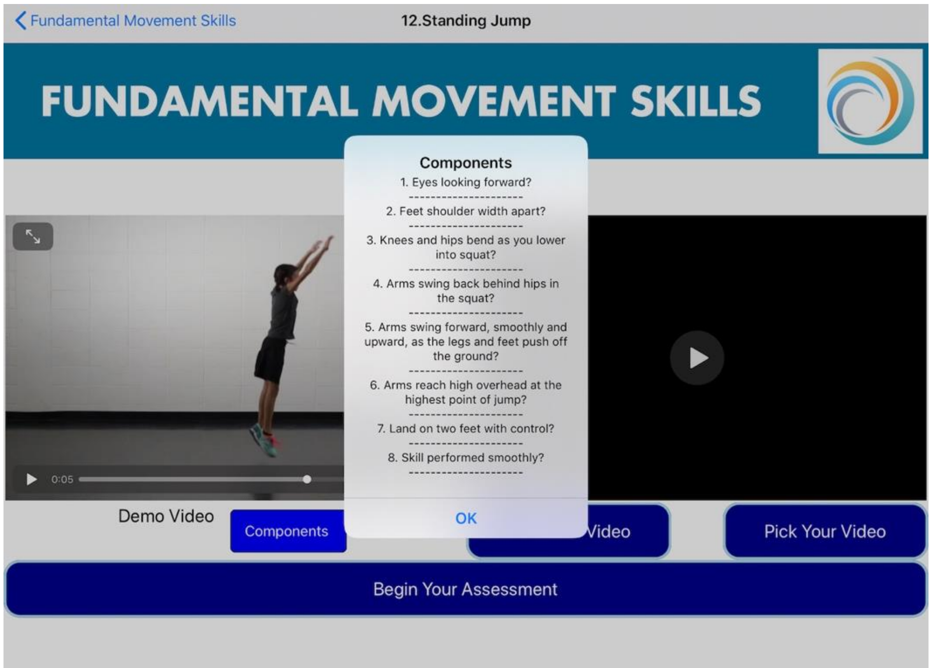


Figure 3.3 Fundamental Movement Skill Instructional Video and Component List for Standing Jump

Figure 3.4 Evaluator Filming the Performer During the Hollow Body Roll Performance

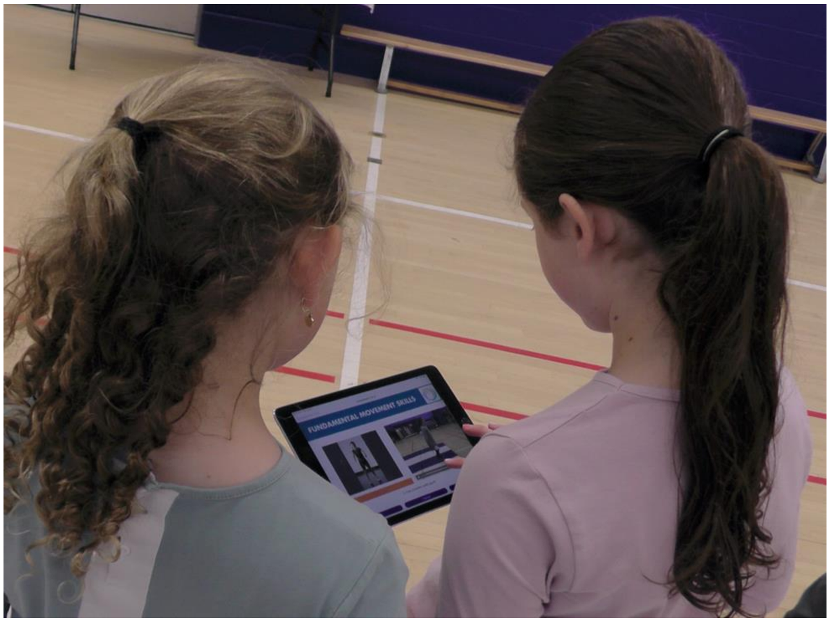


Figure 3.5 Evaluator and Performer using the Move Improve® Application to Discuss and Evaluate Performance

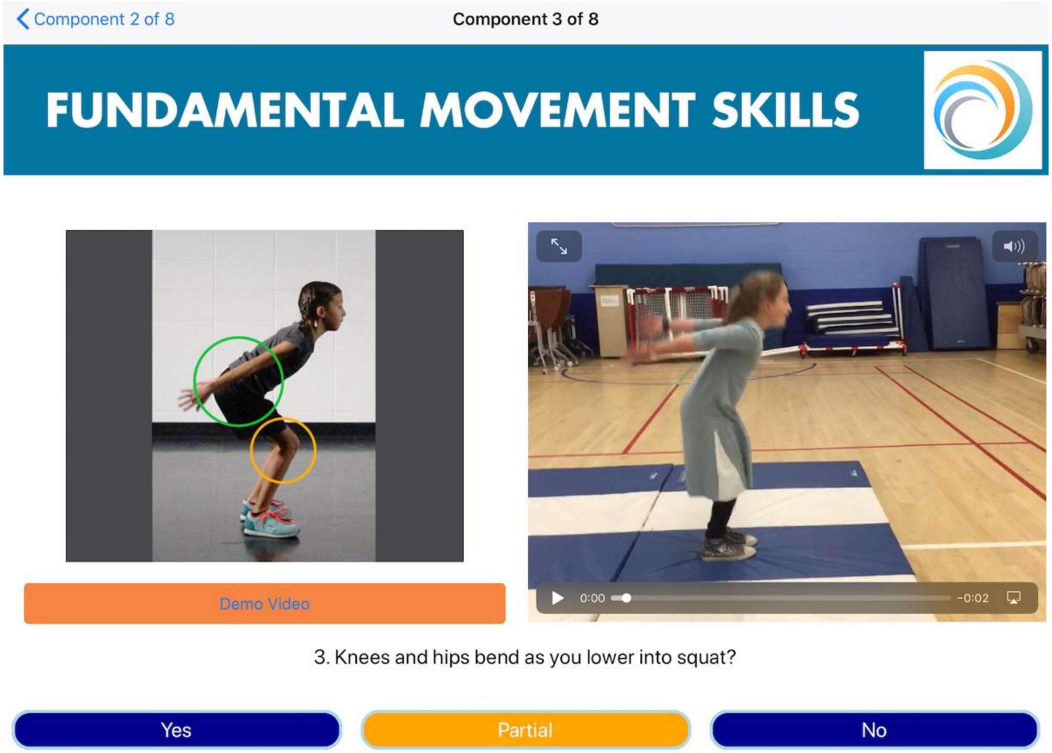


Figure 3.6 Evaluation Component Using Yes, Partially, and No Options

3.3.3 Consensus and Training Session

A 30-minute training session was held on the first day of data collection. The goals of this session were two-fold: primarily to teach the students how to effectively give feedback to their peers when evaluating, and learning how to use the MI software.

Students were presented with a slide presentation which outlined the basics of FMS, P2P learning, and the MI software they would be using. Students were then randomly paired up and allowed to practice taking videos of each other on the iPads using the built-in camera. They were asked to do two tasks: running and jumping. These two tasks allowed the students to practice capturing actions in both horizontal and vertical dimensions. Students were instructed to record the video in landscape form, as this layout is preferred for the MI software. Additionally, the investigators mentioned the importance of being able to completely see the entire student in the video recording, as it would make it easier to evaluate later when using MI.

Once the students gained familiarity with how to properly take video recordings of each other, they were called in to start the consensus session. Two student volunteers were chosen from the group, with one taking on the role of ‘evaluator’ and the other being the ‘performer’. After watching a demonstration video on MI of an individual lifting a box, the two students attempted to replicate what they had seen. Using the MI software, the evaluator filmed the performer lifting an empty box. This recording was then projected onto a television screen to allow the entire class to clearly see it.

As a group, the researchers and students completed the evaluation process of the student volunteer lifting a box, making sure to discuss each component individually. To ensure that students were actively engaged in the discussion process, students were each given an iPad and were asked to interact with a student response application called Kahoot, which was loaded on the iPad tablets. Individually, students chose whether they believed each component deserved a score of ‘Yes’, ‘Partial’, or ‘No’. Kahoot compiles the student responses and then organizes them into box charts. The box charts for each component were projected onto the screen one at a time and discussed as a group. For initial responses, there was no consensus for any of the components. The facilitator used this as an opportunity to engage students in discussion about each component and the reasons for differing opinions. For example, why does one partner think Component 2 should be scored as ‘Partial’, while the other thinks it should be scored as ‘Yes’? These discussions are extremely important to the P2P learning process, as it is during these conversations that learning can occur (Mosston & Ashworth, 2008). Once a consensus had been reached within the group, the researchers moved onto the next component, until a final performance score had been determined. Finally, students were instructed on how to take a screenshot of the final screen with the iPads, as this final score screen contained the crucial data needed for the study. The facilitator encouraged students to work together to discuss the performance when they start, but that ultimately, the evaluator would make the final decision.

3.3.4 Data Collection

On the second day, the students were permitted to start using the MI application to collect their data. Students in each class were randomly assigned partners and were instructed to take turns performing the standing jump (SJ) and the hollow body roll (HBR) skills in the MI program. They received no specific training on how to perform the skills beforehand. Instead, they reviewed the demonstration video along with the skill component list prior to performing each skill. Students were randomly given either the evaluator or the performer role for SJ and were instructed to reverse orders for the HBR. They had the option of using mats or the gym floor to perform the skills and were advised to take screenshots of the summary score page that was received at the end of each skill attempt. These screenshots and the corresponding video for each participant were then exported to external hard drives. Afterwards, the consensus team, comprising of three of the study's researchers and all of whom holding a Bachelor of Science in Kinesiology, used MI to provide 'expert' scores for each performance. Once completed, all videos of participants were then deleted from the iPads.

3.3.5 Data Analysis

In the data analysis for order of performance, 'Order 1' represents the participants who performed first, while 'Order 2' represents the participants who performed the skill second. Using the rubric from MI, the consensus team discussed and evaluated each component together to determine the performance scores. Data was analyzed using IBM SPSS for Windows Version 25.0 (SPSS Inc., 2017).

The data was collected for each skill and participant via MI software. An independent samples *t*-test was conducted to evaluate the differences in the mean scores between students that performed the skill first and those who performed the skill second for each skill (SJ and HBR). A Cohen's *d* (Cohen, 1988) effect size, which measures the strength of the relationship between two variables, was also calculated for each skill between the students that performed first and the students that performed second. Based on benchmarks suggested by Cohen (1988), effect sizes are commonly interpreted as small ($d = 0.2$), moderate ($d = 0.5$), and large ($d = 0.8$).

3.4 Results

Table 3.2 shows the descriptive statistics for the total performance scores of all the students who completed both the SJ and HBR skills. Scores are expressed both in absolute terms and as percentages. These results demonstrate that students had lower performance scores and larger standard deviations when performing the HBR (12.8 ± 2.99) compared to the SJ (20.5 ± 2.60). Additionally, HBR scores had a wider range than those from SJ. This increased variability in student scores suggest that students may have found the HBR more challenging to perform than the SJ.

Table 3.2 Student Performance Scores in Standing Jump and Hollow Body Roll

Skill		**Minimum**	**Maximum**	**Mean**	**Standard Deviation (*s*)**
Standing Jump	**Score (out of 24 pts)**	14	23	20.5	2.60
	Percentage (%)	58.3	95.8	85.4	10.81
Hollow Body Roll	**Score (out of 18 pts)**	8	18	12.8	2.99
	Percentage (%)	44.4	100	71.2	16.63

Figure 3.7 provides a box plot of the performance scores for the participants by order of performance for SJ. There is more variability in order one despite the two outliers in order two.

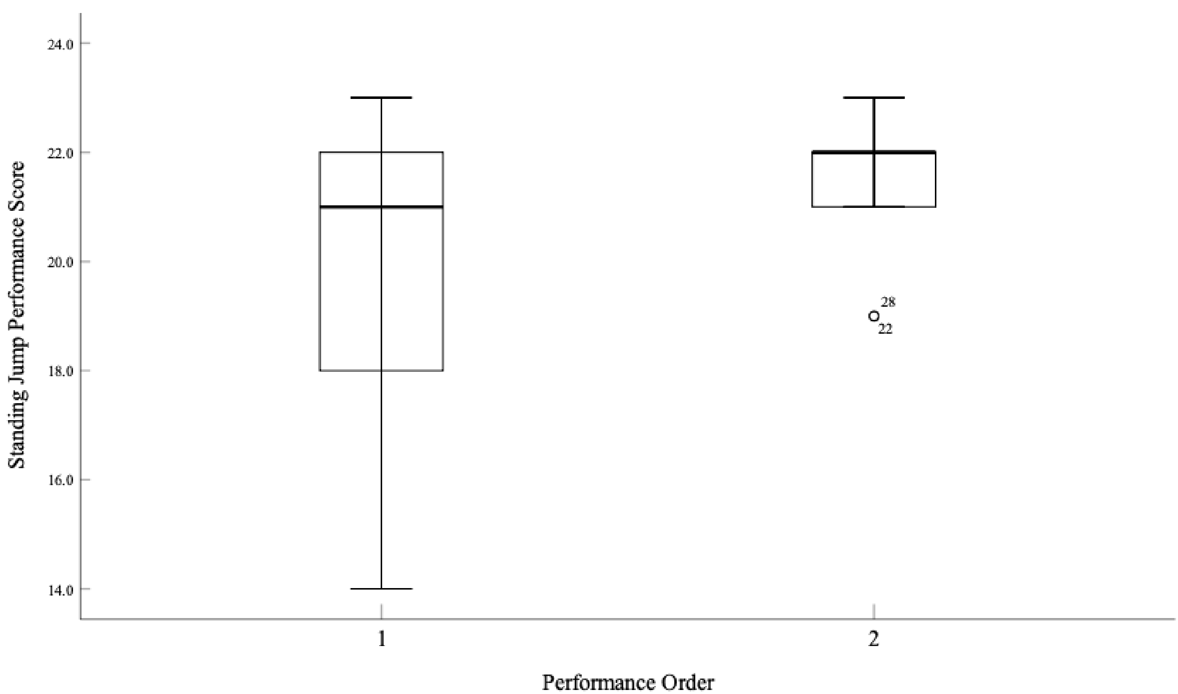


Figure 3.7 Standing Jump Performance Scores for Group 1 and Group 2

Figure 3.8 shows a box plot of the performance scores for the participants by order of performance for HBR. There is more variability in order one.

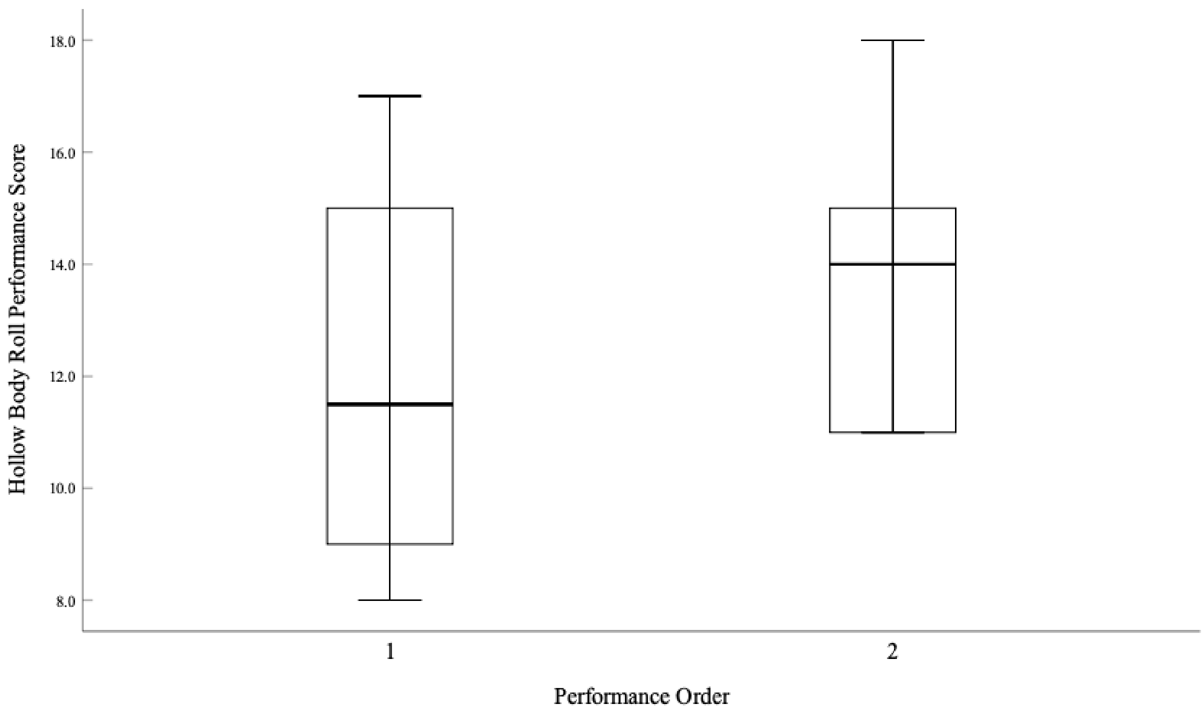


Figure 3.8 Hollow Body Roll Performance Scores for Group 1 and Group 2

Table 3.3 shows the independent *t*-test results comparing order for both skills, SJ and HBR. There was a significant difference in the order when students performed SJ of $t(16) = 1.4$, $p < 0.01$, and a Cohen's $d = 0.660$, indicating a moderate effect size. The mean performance scores show that students who performed second (M = 21.3 ± 1.41) scored higher than their peers who performed the skill first (M = 19.7 ± 3.28). While the direction was the same, there was no significant difference between the students who performed first and students who performed second on the HBR.

Table 3.3 Independent t-test Results Comparing Order for Standing Jump and Hollow Body Roll

Skill	***n***	**Order**	**Mean**	**Standard Deviation**	***t***	**Significance (*p*)**	**Cohen's D (*d*)**
Standing Jump	9	1	19.7	3.28	1.400	0.005**	0.660
	9	2	21.3	1.41			
Hollow Body Roll	8	1	12.0	3.38	1.093	0.293	0.546
	8	2	13.6	2.50			

** $p < 0.01$

3.5 Discussion

The purpose of this study was to examine whether there was a significant impact due to the order in which a student performed a SJ and HBR within a P2P learning model using the MI application. The hypothesis, that students who performed second (Order 2) would have greater performance scores than those who performed first (Order 1), was retained for SJ but not for HBR. The results from this study suggest that there was a significant difference in SJ scores ($p < 0.01$), where participants who performed second had greater scores than those who performed first. While the HBR mean scores were greater for the students that went second than those who went first, the difference was not significant. The overall findings seem to support the concept of observational learning having an impact on performance. However, order having a significant effect on SJ and not on HBR did raise some questions as to why there was such a discernible difference between the two skills.

3.5.1 Impact of the Quality of Video on Students' Learning

In retrospect, the most likely cause of the observed difference between SJ and HBR results could be due to the quality of the HBR skill in the MI app. The video provided for the HBR is taken from the front of the feet, meaning that it is much harder for the participants to notice the proper arm and feet positioning. As well, the lighting in the video caused the feet of the demonstrator to blend with the floor, potentially causing more confusion to participants as to whether they needed to lift their feet off the ground or not. Another issue was that the component descriptions were a bit too wordy for the students to properly understand, specifically components 3 and 5, listed below:

Component 3: Stomach muscles are strong and tight when on the back

Component 5: Back muscles are strong and tight when on the stomach

3.5.2 Observational Learning Playing a Role in Order Effect of Learning

Another variable that might account for the difference in results between SJ and HBR could be participant familiarity with the two skills. Students may be more familiar with jumping than trying to perform a HBR. Certainly, percentage scores were higher on SJ than on HBR.

Moreover, students may have had an easier time understanding the cues required to transform their current jump form to match the form depicted in MI, which may be attributed to the phenomenon of observational learning. By observing another individual perform a desired skill, it provides a mental blueprint on how the skill should look, along with segmental timing and sequential movements of various joints to perform the skill (Blandin & Proteau, 2000; Cross et al., 2009). Research has also found that subjects who observe a task they are familiar with, or have had experience attempting, will experience some benefit through observational learning (watching peer perform skill, or watching MI instructional video of skill; Shea et al., 2000), as this learning pathway shares neural paths with physical learning networks (similar or same skill done during active play; Rizzolatti & Fogassi, 2014). Yet, research has also found the opposite, in which subjects that were exposed to motor skills they were unfamiliar with did not experience a significant learning benefit through observation (Cross et al., 2009; Jennings et al., 2013). This could explain the non-significant outcome of the HBR scores: since HBR was an unfamiliar skill to the students, having them watch the MI instructional video and their peer's performance would have caused minimal to no learning (mirror neurons were silent, as there was no physical learning established beforehand; Jennings et al., 2013; Rizzolatti & Fogassi, 2014).

While it is incorrect to state that the participants in this study have never performed a side-to-side roll in any fashion during their growth, a HBR is much different in terms of execution, compared to the comparison between a strict SJ and a general jump during active play. The HBR is generally used as a coaching drill for gymnastic athletes to develop the ability to transition rapidly from a hollow-body position to an arched back position (Živčić-Marković et al., 2015). The HBR is a roll that requires an individual to keep their arms and legs off the ground during the entire movement, done so by tightening the core when on the back and then arching the lower back when rolled over to the stomach. In contrast, general active play between children does not necessarily require them to hold either a hollow-body position or an arched back position, let alone, rapidly switch between the two. Rather, children in active play would roll with a looser body, segmenting the roll between the torso and lower body, which was evident in the videos of nearly all participants in the study. This would suggest that unfamiliar skills can reduce the benefit from order of performance in P2P learning.

3.5.3 Benefits of Using Peer to Peer Learning and Technology in Physical Education

Traditional teaching methods in PE often have students standing around and not actively engaged (Mosston & Ashworth, 2008). Technology-based learning, like MI, hinges upon students taking an active role in their learning, so that they are not waiting around for the teacher’s individual feedback. Scholars such as Casey et al. (2017) have argued that with appropriate pedagogical considerations, digital technology can offer potential benefits to learning and performance. While digital video has been shown to be beneficial in the PE classroom (Hung et al., 2018), providing a structured assessment can further enhance learning. Technology must be appropriately incorporated into lessons in a manner such that students are correctly prompted into the desired actions, maintain interest, remain on task, and remain engaged in higher levels of thinking and performance while using it (Lee & Gao, 2020; Mosston & Ashworth, 2008; Ward & Lee, 2005; Zhu & Dragon, 2016).

With the appropriate pedagogical framework and effective use of technology, tablet computers may offer a way to provide scaffolded content knowledge in PE classrooms (O'Loughlin et al., 2013; Sinelnikov, 2012; Zhu & Dragon, 2016). As a result, teachers are recommended to align the use of digital technology with their PE learning objectives. We believe the results found in this study will encourage teachers to use MI through P2P learning in their PE lessons, as it permits students to be engaged in structured evaluation of their FMS performance.

3.5.4 Limitations of the Study

The study was conducted with a small, heterogeneous (grade 5-8) sample of students. They were separated by gender which is different from standard classrooms in North America. The participation of mixed grade levels separated by gender may have subtly affected the results of the study.

Due to the nature of P2P learning, partner cohesion and cooperation were things to consider. When students are expected to rate their peers in any sort of activity, those ratings are often related more to subjective social dynamics than objective academic criteria (McDermid et al., 2014). Students may have disliked their assigned peer, which may have led to instances where participants did not engage in discussions during the evaluation process. For example, one student may make a claim about their peer's performance, and their peer could either automatically agree with the statement or appear uninterested in the outcome. Finally, it is not uncommon for middle school students to dislike being filmed and critiqued on their performance, possibly leading to disagreements and confrontations between peers.

It would also be useful to control for prior physical ability differences between students by conducting an initial FMS ability test prior to data collection. By conducting a pre-test, it would be possible to control the possible confounding factor, that is, the influence of students' previous skill levels. The results from this pre-test would inform student pairings, as students with similar skill levels could be paired together. This would eliminate the possibility that students with higher skill abilities went first, therefore making it harder to see if there truly is an order effect.

3.5.5 Future Directions

This appears to be the first evaluation of a structured P2P learning tool to assess the teaching of fundamental movement skills to middle school students. Based on the findings, it would be useful to more closely evaluate the quality of each skill to ensure that they are clearly performed, and that the terminology of the components are easily understood.

More time should be given to teach participants how to take clear videos of their peers and how to properly discuss and critique a movement skill. The time constraint of 30 minutes for a gym class may not provide enough time to effectively teach the fundamentals of critiquing and analysing a movement performance. Encouraging schools to utilize tools like MI using a cross curricular approach could improve student learning. For example, consensus learning could be taught in a class on communications. Future studies will benefit from having multiple structured practice sessions before the participants attempt any performance skills and data collection.

Measuring and evaluating the quality and accuracy of student feedback is essential. Ensuring student feedback is consistent with the teacher's standard of good performance would reinforce desired knowledge content outcomes and improve the accuracy of peer assessment. Therefore, additional research should be conducted to evaluate MI's reliability and validity.

Further, future research should account for how friendship levels between peers may affect performance scores.

3.5.6 Conclusion

The results of this study indicate that order of performance may have had a significant effect on performance scores when performing a standing jump. The observed effect of order on performance appears to be impacted by observational learning and complexity of skill. However, skill complexity may have been confounded by the quality of video demonstration. Future studies should incorporate more homogeneous and larger sample sizes, using skills with different levels of complexity.

Using tablet-based structured video analysis for P2P learning has the potential to improve student learning and performance. Instead of the traditional PE teaching style of mimic/practice, through P2P learning and the use of digital technologies such as a mobile application, students are actively participating in their learning and evaluation, which may lead to more engaging lesson plans, minimize classroom sedentarism, and ultimately, develop mastery of FMS.

References

Al-Abood, S. A., Davids, K., & Bennett, S. J. (2001). Specificity of task constraints and effects of visual demonstrations and verbal instructions in directing learner' search during skill acquisition. *Journal of Motor Behaviour*, *33*(3), 295-305.

Andrieux, M., & Proteau, L. (2014). Mixed observation favors motor learning through better estimation of the model's performance. *Experimental Brain Research*, *232*(10), 3121-3132.

Andrieux, M., & Proteau, L. (2016). Observational learning: Tell beginners what they are about to watch and they will learn better. *Frontiers in Psychology*, *7*(51).

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.

Black, C. B., & Wright, D. L. (2000). Can observational practice facilitate error recognition and movement production? *Research Quarterly for Exercise and Sport*, *71*(4), 331-339.

Blandin, Y. & Proteau, L. (2000). On the cognitive basis of observational learning: Development of mechanisms for the detection and correction of errors. *The Quarterly Journal of Experimental Psychology*, *53A*(3), 846-867.

Buchanan, J. J., & Dean, N. J. (2010). Specificity in practice benefits learning in novice models and variability in demonstration benefits observational practice. *Psychological Research*, *74*, 313-326.

Cairncross, S., & Mannion, M. (2001). Interactive multimedia and learning: Realizing the benefits. *Innovations in Education & teaching International*, *38*(2), 156-164.

Cappella, E. (2000). What children think about computers. *The Future of Children, 10*(2), 189-191.

Casey, A., Goodyear, V. A., & Armour, K. M. (2017). *Digital Technologies and Learning in Physical Education: Pedagogical Cases.* Routledge.

Cohen J. W. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Hillsdale: Lawrence Erlbaum Associates.

Cross, E. S., Kraemer, D. J. M., Hamilton, A. F. de C., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex*, *19*(1), 315-326.

Ernst, M. & Byra, M. (1998). Pairing learners in the reciprocal style of teaching: Influence on student skill, knowledge, and socialization. *Physical Educator, 55*(1) 24-37.

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.

Gallahue, D., Ozmun, J. & Goodway, J. (2012). *Understanding Motor Development: Infants, Children, Adolescents, Adults*. 7th ed. New York: McGraw-Hill.

Goldberger, M., Gerney, P., & Chamberlain, J. (1982). The effects of three styles of teaching on the psychomotor performance and social skill development of fifth grade children. *Research Quarterly for Exercise and Sport, 53*(2), 116-124.

Hancox, R. J., Milne, B. J.., & Poulton, R. (2005). Association of television viewing during childhood with poor educational achievement. *Archives of Pediatric and Adolescent Medicine, 159*(1), 614-618.

Hardy, L. L., Reinten-Reynolds, T., Espinel, P., Zask, A., & Okely, A. D. (2012). Prevalence and correlates of low fundamental movement skill competency in children. *Pediatrics*, *130*(2), 390-398.

Hayes, S. J., Elliott, D., & Bennett, S. J. (2010). General motor representations are developed during action-observation. *Experimental Brain Research, 204*(2), 199-206.

Higgs, C., (2010). Physical literacy – two approaches, one concept. *Physical & Health Education Journal*, *76*(1), 6–7.

Hodges, N. J., Chua, R., & Franks, I. M. (2003). The role of video in facilitating perception and action of a novel coordination movement. *Journal of Motor Behaviour*, *35*(3), 247-260.

Hodges, N. J., Williams, A. M., Hayes, S. J., & Breslin, G. (2007). What is modelled during observational learning? *Journal of Sports Sciences*, *25*(5), 531-545.

Houston-Wilson, C., Dunn, J.M., van der Mars, H., & McCubbin, J. (1997). The effect of peer tutors on motor performance in integrated physical education classes. *Adapted Physical Activity Quarterly*, *14*(4), 298-313.

Hung, H. C., Young, S. S. C., & Lin, K. C. (2018). Exploring the effects of integrating the iPad to improve students' motivation and badminton skills: A WISER model for physical education. *Technology, Pedagogy, and Education, 27*(3) 265-278.

Jenkinson, K.A., Naughton, G., & Benson, A.C. (2014). Peer-assisted learning in school physical education, sport and physical activity programmes: A systematic review. *Physical Education and Sport Pedagogy*, *19*(3), 253-277.

Jennings, C. T., Reaburn, P., & Rynne, S. B. (2013). The effect of a self-modelling video intervention on motor skill acquisition and retention of a novice track cyclist's standing start performance. *International Journal of Sports Science and Coaching*, *8*(3), 467-480.

Johnson, R. (2004). Peer assessments in physical education. *Journal of Physical Education, Recreation, and Dance, 75*(8), 33-40.

Katz, L. (2003). Multimedia and the internet for sport sciences: Applications and innovations. *International Journal of Computer Science in Sport*, *2*(1), 4-18.

Kelly, P., & Katz, L. (2016). Comparing peer-to-peer and individual learning: Teaching basic computer skills to disadvantaged adults. *International Journal of Adult Vocational Education and Technology (IJAVET)*, *7*(4), 1-15.

Koekoek, J., & Knoppers, A. (2015). The role of perceptions of friendships and peers in learning skills in physical education. *Physical Education and Sport Pedagogy, 20*(3), 231-249.

Kolovelonis, A., Goudas, M., & Gerodimos V. (2011). The effects of the reciprocal and the selfcheck styles on pupils' performance in primary physical education. *European Physical Education Review, 17*(1) 35-50.

Krause, J., & Sanchez, Y. (2014). Meeting the national standards: There's an app for that! *Strategies*, *27*(4), 3-12.

Lago-Rodriguez, A., Cheeran, B., Koch, G., Hortobagyi, T., & Fernandez-del-Olmo, M. (2014). The role of mirror neurons in observational learning: An integrative review. *European Journal of Human Movement*, *32*, 82-103.

Lee, J. E., & Gao, Z. (2020). Effects of the iPad and mobile application-integrated physical education on children's physical activity and psychosocial beliefs. *Physical Education and Sport Pedagogy*, *25*(6), 567-584.

Lee, T. D., Swinnen, S. P., & Verschueren, S. (1995). Relative phase alterations during bimanual skill acquisition. *Journal of Motor Behaviour*, *27*(3), 263-274.

Leight, J., Banville, D., & Polifko, M.F. (2009). Using digital video recorders in physical education. *Journal of Physical Education, Recreation & Dance, 80*(1), 17-21.

Leser, R., Baca, A., & Uhlig, J. (2011). Effectiveness of multimedia-supported education in practical sports courses. *Journal of Sports Science and Medicine*, *10*(1), 184-192.

Mas, F. G., Plass, J., Kane, W. M., & Papenfuss, R. L. (2003). Health education and multimedia learning: Connecting theory and practice (part 2). *Health Promotion Practice, 4*(4), 464-469.

Mason, R., & Rennie, F. (2006). *E-learning: The Key Concepts.* Routledge.

Mayer, R. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.

McCullagh, P., & Meyer, K. N. (1997). Learning versus correct models: Influence of model type on the learning of a free-weight squat lift. *Research Quarterly for Exercise and Sport*, *68*(1), 56-61.

McCullagh, P., & Weiss, M. (2001). Modelling: Considerations for motor skill performance and psychological responses. *Handbook of Sport Psychology*, 205-238.

McCullagh, P., Weiss, M. R., & Ross, D. (1989). Modeling considerations in motor skill acquisition and performance: An integrated approach. *Exercise and Sport Sciences Reviews, 17,* 475-513.

McDermid, F., Peters, K., Jackson, D., & Daly, J. (2014). Conducting qualitative research in the context of pre-existing peer and collegial relationships. *Nurse Researcher*, *21*(5), 28-33.

Metzler, M. (2017). *Instructional models in physical education* (3rd ed.). Routledge.

Mohnsen, B. S. (2012). *Using Technology in Physical Education*. Bonnie's Fitware, Inc.

Mosston, M. & Ashworth, S. (2008). *Teaching Physical Education*. First Online Edition. https://spectrumofteachingstyles.org/assets/files/book/Teaching_Physical_Edu_1st_Online.pdf

O'Brien, W., Belton, S., & Issartel, J. (2016). Fundamental movement skill proficiency amongst adolescent youth. *Physical Education and Sport Pedagogy, 21*(6), 557-571.

O'Loughlin, J., Chróinín, D. N., & O'Grady, D. (2013). Digital video: The impact on children's learning experiences in primary physical education. *European Physical Education Review*, *19*(2), 165–182.

Palao, J. M., Hastie, P. A., Cruz, P. G., & Ortega, E. (2015). The impact of video technology on student performance in physical education. *Technology, Pedagogy and Education*, *24*(1), 51–63.

ParticipACTION. (2018). The brain + body equation: Canadian kids need active bodies to build their best brains. *The 2018 ParticipACTION Report Card on Physical Activity for Children and Youth.*

Payne, G., & Isaacs, L. D. (2020). *Human Motor Development: A Lifespan Approach*. Routledge.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon, 9*(5).

Rizzolatti, G., & Fogassi, L. (2014). The mirror mechanism: Recent findings and perspectives. *Philosophical Transactions of the Royal Society B, 369*(1644).

Robinson, T. N. (2001). Television viewing and childhood obesity. *Pediatric Clinics of North America*, *48*(4), 1017-1025.

Rohbanford, H., & Proteau, L. (2011). Learning through observation: A combination of expert and novice models favors learning. *Experimental Brain Research*, *215*(3-4), 183-197.

Sinelnikov, O. A. (2012). Using the iPad in a sport education season. *Journal of Physical Education, Recreation & Dance*, *83*(1), 39-45.

Shea, C. H., Wright, D. L., Wulf, G., & Whitacre, C. (2000). Physical and observational practice afford unique learning opportunities. *Journal of Motor Behaviour*, *32*(1), 27-36.

Ste-Marie, D. M., Law, B., Rymal, A. M., Jenny, O., Hall, C., & McCullagh, P. (2012). Observation interventions for motor skill learning and performance: An applied model for the use of observation. *International Review of Sport and Exercise Psychology*, *5*(2), 145-176.

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.

Tanaka, C., Tanaka, M., & Tanaka, S. (2018). Objectively evaluated physical activity and sedentary time in primary school children by gender, grade and types of physical education lessons. *BMC Public Health, 18*(1), 948-958.

Topping, K., & Ehly, S. (1998). *Peer-Assisted Learning*. Taylor & Francis. https://books.google.ca/books?id=i76aRj-xC2cC

Vandewater, E. A., Bickham, D. S., & Lee, J. H. (2006). Time well spent? Relating television use to children's free-time activities. *Pediatrics*, *117*(2), e181-e191.

Vandewater, E. A., Bickham, D. S., Lee, J. H., Cummings, H. M., Wartella, E. A., & Rideout, V. J. (2005). When the television is always on: Heavy television exposure and young children's development. *American Behavioural Scientist, 48*(5), 562-577.

Vogt, S., & Thomaschke, R. (2007). From visuo-motor interactions to imitation learning: Behavioural and brain imaging studies. *Journal of Sports Sciences*, *25*(5), 497-517.

Ward, P., & Lee, M. A. (2005). Peer-assisted learning in physical education: A review of theory and research. *Journal of Teaching in Physical Education*, *24*(3), 205–225.

Watterson, T. (2012). *Changes in attitudes and behaviours toward physical activity, nutrition, and social support for middle school students using AFIT app as a supplement to instruction in a physical education class* [Unpublished doctoral dissertation]. University of South Florida.

Weir, T. & Connor, S. (2009). The use of digital video in physical education. *Technology, Pedagogy and Education*, *18*(2), 155-171.

Yoncalik, O., Yetim, A. A., & Senel, Ö. (2010). Effects of teaching with Mosston's command, practice, and reciprocal styles on affective reactions of sixth-grade students toward physical education lessons. *International Journal of Educational Reform, 18*(4), 311-326.

Zhu, X., & Dragon, L. A. (2016). Physical activity and situational interest in mobile technology integrated physical education: A preliminary study. *Acta Gymnica, 46*(2). 59-67.

Živčić-Marković, K., Krističević, T., & Aleksić-Veljković, A. (2015). A suggested model of handstand teaching method. *Physical Culture, 69*(2), 138-149.

Chapter Four: Manuscript 2

Intra- and Inter-rater Reliability of a Peer-to-Peer Video Analysis Tool to Measure Fundamental Movement Skills with Middle School Children

4.1 Abstract

This study examined the intra- and inter-rater reliability of the Move Improve® video performance analysis tool in middle school children when performing two fundamental movement skills, a standing jump and hollow body roll. Eighteen students participated in the study. Four distinct groups of raters (students, two independent raters, and a three-person consensus panel) evaluated a total of 34 videos. The intra-rater reliability of the consensus panel was found to be good for both skills. For standing jump, ICC (2,5) values showed good inter-rater reliability amongst all raters ($\alpha > 0.75$). Hollow body roll showed poor inter-rater reliability for the student raters compared to the individual raters and the consensus panel ($\alpha < 0.5$). Further analysis into the individual skill components revealed some inconsistencies in the scores given by each coder for the hollow body roll. Modifying the content of that skill and improving training protocols for users may improve reliability.

Keywords: fundamental movement skills; assessment; reliability

4.2 Introduction

FMS are movement patterns that involve various body parts and provide the basis of physical literacy (Gallahue et al., 2012). FMS are important for development as they are the foundation for building specialized movement skills which are essential for proper growth (Barnett et al., 2009; Fisher et al., 2005; Gallahue et al., 2012; Higgs, 2010). Gallahue et al. (2012) divide motor skills into three distinct categories: locomotor (e.g., running and hopping), manipulative or object control (e.g., ball-handling skills), and stability or balance (e.g., rolling and balancing on one foot). If deficiencies are not detected at an early age, children may experience lifelong problems engaging in physical activity and sports (Hardy et al., 2012; Ulrich, 2000).

Children's motor competence becomes visible through their FMS performances, and mastering these basic skills has been shown to be positively associated with greater physical activity levels (Stodden et al., 2009). It is thus imperative to track the development and level of children's motor competence using reliable observational tools. There are two primary types of motor skill assessment: process-oriented (skill development) and product-oriented (outcomes) (Foulkes et al., 2015; Payne & Isaacs, 2020). Although there is currently no gold-standard assessment tool or rubric available to evaluate FMS performance in adolescence, many of the commonly used process-oriented assessments in PE follow the criteria established by Roberton (1977). Previous research has shown that rater reliability is an important part of assessing FMS competence (Barnett et al., 2014). However, the accuracy and reliability of process-oriented assessments can be influenced by a variety of factors including: the background and characteristics of raters (Nadeau et al., 2008; Palmer & Brian, 2016); access to training for motor skill observation (Haynes & Miller, 2015); ability or level of expertise (Borghouts et al., 2017); and the population being assessed (Bailey et al., 1995).

Several reliability studies have examined the differences between expert and novice coders regarding the evaluation of a wide range of movements. Nadeau et al. (2008) examined the reliability of the Team Sport Assessment Procedure for ice hockey by student observers, ranging from 14 to 17 years of age, and physical educators' observers. Despite being familiar with ice hockey, student observers were considered novices when it came to using the measurement tool, so they participated in a one-hour training session. Results showed that student observers' percent agreement ranged from 59-95%, whereas the physical educators ranged from 80-82%, suggesting that inexperienced coders have greater variability. Minick et al. (2010) evaluated the inter-rater reliability of the Functional Movement Screen with university students. Four raters, two experts and two novices, independently coded the participants. Because they had less than a year of experience using the Functional Movement Screen, novice coders participated in an introductory training course. The results from this study are promising, as it showed excellent agreement between expert and novice coders. The final study, conducted by Palmer and Brian (2016), assessed the difference in scoring of elementary school children between adult expert and novice coders using the TGMD-2. Novice coders did not have previous experience using the TGMD-2, so they participated in an intensive two-hour long training session with an expert, where they individually rated an example video, which was then followed by a thorough discussion. The results from this study indicated that novice coders were unable to reach significant agreement with the expert coders. The differences in results between these three studies may be due to the assessment tool itself, the population being observed, and/or the effectiveness of the training sessions for the novice coders.

PE teachers play a key role in monitoring the physical development of children (Ng, 2002); therefore, it is important that they structure and implement appropriate activities and assessments with sufficient feedback (Claxton et al., 2006; Logan et al., 2012). Students benefit greatly from receiving fast and accurate feedback on motor performance, as it is this feedback that maximizes learning and student achievement in the PE environment (Mohnsen, 2012). However, as class sizes increase, teachers are facing the challenge of providing complete feedback to all students directly in a timely manner (Evans, 2017). One potential solution to this issue is the use of peer assessments (Kniffin & Baert, 2015; Lund & Veal, 2013; Veal, 1995).

Peer assessments, where students observe and assess a partner's performance of a motor skill, ensure students receive immediate feedback, thus maximizing the opportunity for motor learning in PE (Alstot, 2018). If students can provide accurate and reliable assessments, then the PE teacher will have more detailed knowledge about the student's performance. Furthermore, these peer evaluations can be used to help inform the PE teacher's comments on report cards and provide feedback on the more difficult components to facilitate further training (Jenkinson et al., 2014). It is beneficial to use process-oriented batteries to teach and assess an individual's skill progression, because it allows educators and coaches to identify specific components that may need improvement (Barnett et al., 2014). Factors such as age and experience are presumed to influence the reliability of peer assessment (Fry, 1990). If students are not trained in the application of rubrics, then the reliability and validity of an assessment cannot be improved (Chang et al., 2011; Rezaei & Lovorn, 2010). However, providing sufficient training and a clear marking rubric may improve peer reliability (Fry, 1990).

This paper examines the intra-rater and inter-rater reliability of the mobile application and novel VFB tool, MI, when assessing the performance of a SJ and a HBR by students in grades 5 through 8. Since the MI application is an observational assessment tool which uses video performance technology, both intra- and inter-rater reliability are particularly important.

4.3 Methods

Conducted within the STRL, data for this research were collected as part of a more extensive study concerning the effect of observational learning on students' execution of an FMS when using the MI application. The data were screened for completeness, and any missing or incomplete data were excluded from the analysis. Parental consent for the collection of data was obtained prior to the investigation, and the only data that were collected and analyzed came from students whose parents had granted permission. Assent was not required as students were required to use the MI application as part of their regularly scheduled PE class. Ethical approval was obtained through the Conjoint Health Research Ethics Board at the University of Calgary.

4.3.1 Move Improve® Mobile Application

MI is a mobile application which contains a video performance analysis tool with multiple modules (i.e., FMS, Workplace Safety, etc.) designed to evaluate an individual's ability to perform a variety of movement skills. The FMS module contains 16 different skills that users can select from to test themselves or their peers. Users are first provided with an instructional video which contains a visual demonstration of the skill, followed by a component list that outlines the sequential movements required to properly complete the skill (Figure 4.1). After viewing the demonstration and a breakdown of the components, users film themselves performing the skill in a video which is later examined in the MI application. MI allows users to evaluate one component of a skill at a time, providing both an image of the correct action and a rating choice of "Yes", "Partial", or "No" (Figure 4.2). "Yes" indicates the component was performed correctly, "Partial" infers that the component was performed somewhat correctly, while “No” denotes that the component was not performed accurately. Each of these rating choices are assigned a corresponding numerical score (Yes = 3, Partial = 2, No = 1) – termed a “component score”. Once all the components are given a rating, the evaluation is complete and a summary list of all the components and their corresponding scores is displayed on the screen. Finally, the component scores are then totalled at the end to give a “performance score” for the entire skill.

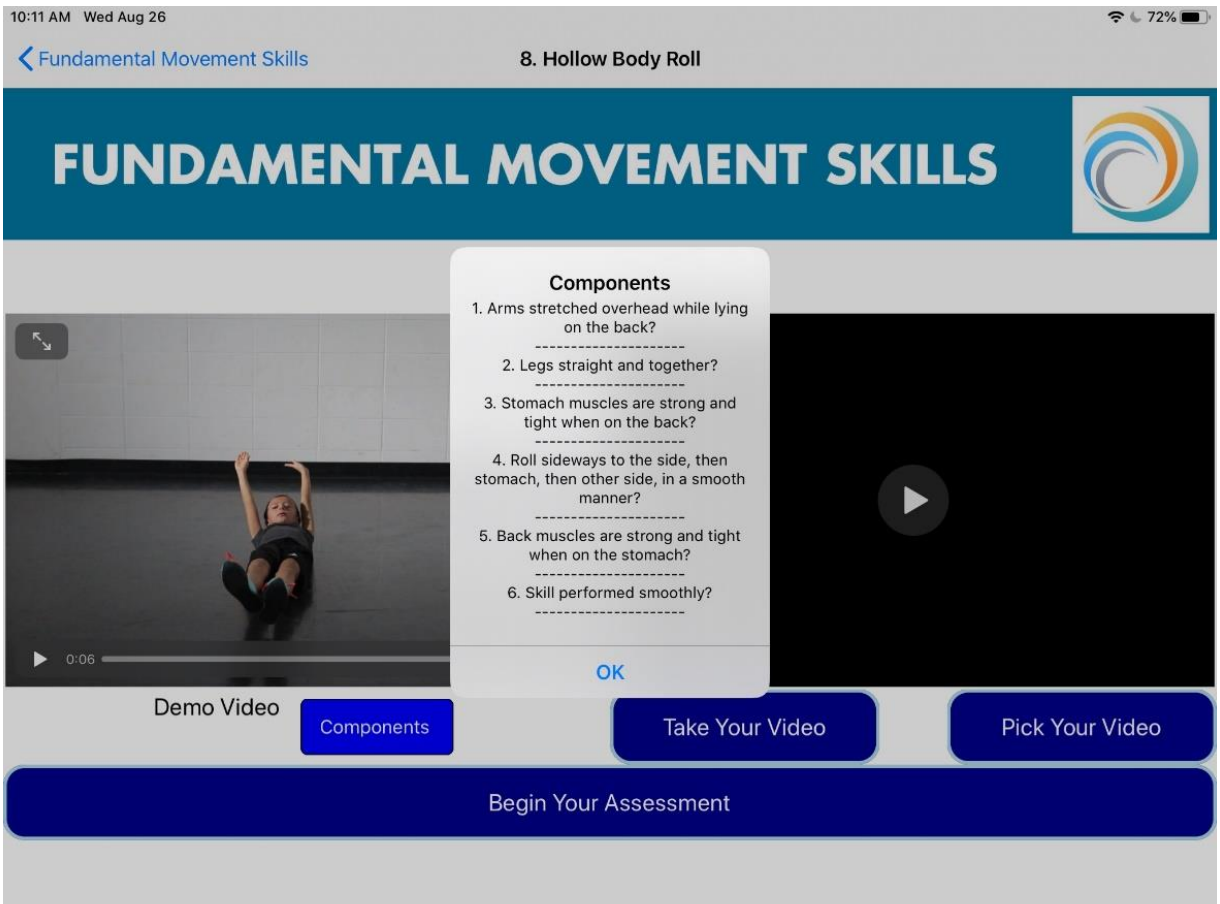


Figure 4.1 Fundamental Movement Skill Instructional Video and Component List for Hollow Body Roll

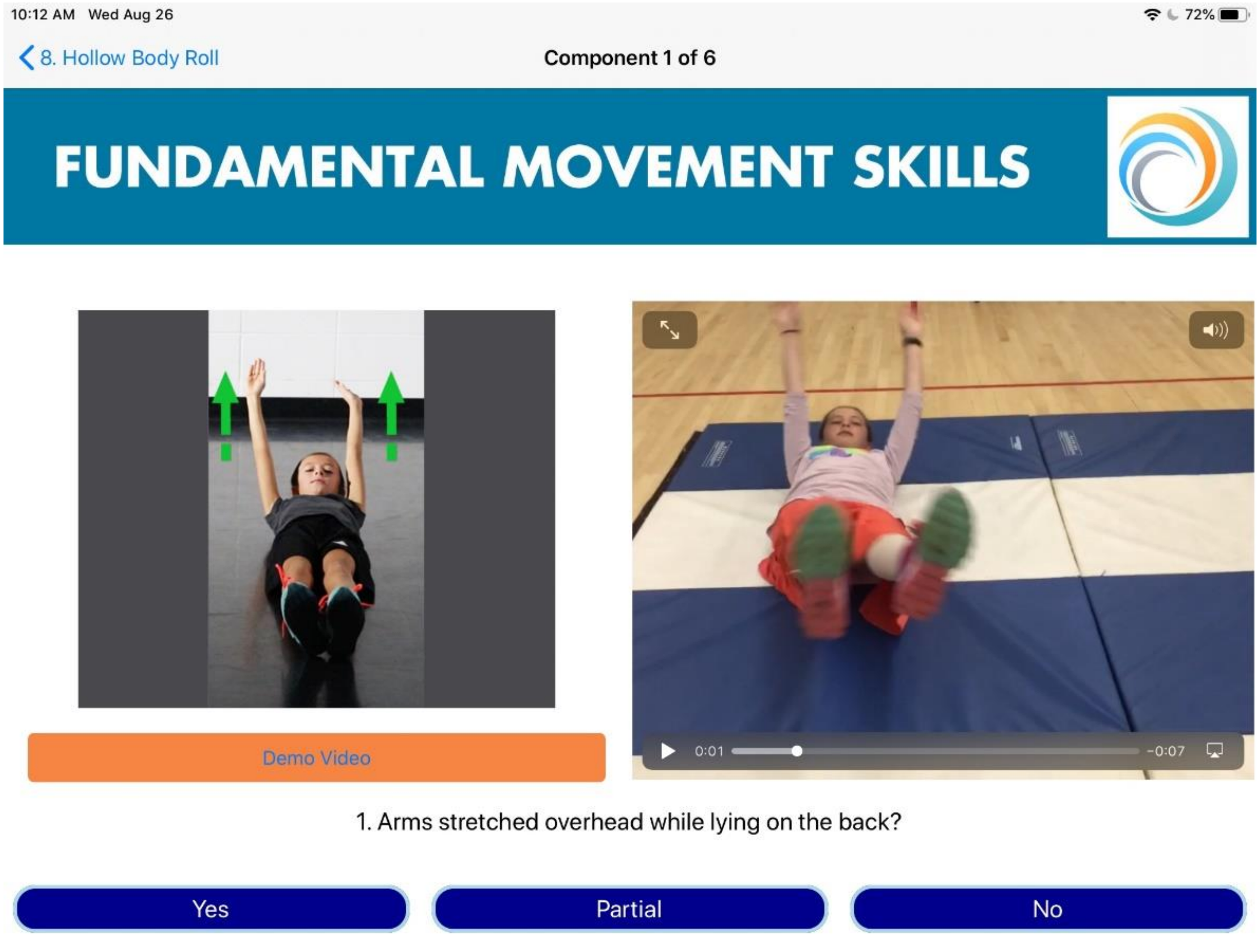


Figure 4.2 Evaluation Component Using Yes, Partially, and No Options

4.3.2 Dataset and Participants

Eighteen participants (44% male, 56% female) ranging from grades 5 to 8 were included in this study. Two fundamental movement skills were chosen from the MI software for the students to perform: the SJ and the HBR. All 18 participants (8 male, 10 female) completed the SJ, while only 16 participants (8 male, 8 female) completed the HBR due to two female students feeling uncomfortable filming themselves performing the HBR while wearing skirts. All together, there were 18 SJ videos and 16 HBR videos, for a total of 34 videos in this dataset.

4.3.3 Standing Jump and Hollow Body Roll

A SJ was the first FMS selected for the students to perform. This skill has been broken down into 8 different components by the MI software. These components, in sequential order, are: (1) Eyes looking forward; (2) Feet shoulder width apart; (3) Knees and hips bend as you lower into squat; (4) Arms swing back behind hips in the squat; (5) Arms swing forward smoothly and upward as the legs and feet push off the ground; (6) Arms reach high overhead at the highest point of jump; (7) Land on two feet with control; and (8) Skill performed smoothly? The maximum performance score for this skill is 24 points.

The HBR was the second FMS the students were asked to perform. This skill is broken down into 6 different components by the MI software. These components, in sequential order, are: (1) Arms stretched overhead while lying on the back; (2) Legs straight and together; (3) Stomach muscles are strong and tight when on the back; (4) Roll sideways to their side, then stomach, then side, in a smooth manner; (5) Back muscles are strong and tight when on the stomach; and (6) Skill performed smoothly. The maximum performance score for this skill is 18 points.

4.3.4 Consensus and Training Session

Prior to data collection, all students participated in a 30-minute training session in which they learned how to effectively provide feedback to their peers when evaluating, as well as how to use the MI application. The students were called in to start the consensus session once they were comfortable taking video recordings of each other. After watching a demonstration video on MI of an individual lifting a box, two student volunteers were chosen from the group and attempted to replicate what they had seen. The researchers and students completed the evaluation of the student volunteer lifting a box as a group, making sure to discuss each component individually. Students decided as a class whether each component deserved a ‘Yes,’ ‘Partial,’ or ‘No’ score; there was no consensus on any of the components in the initial responses. The researchers used this as an opportunity to engage students in discussion about each component and the reasons for their differing opinions. Once a group consensus was reached, the researchers moved on to the next component until a final performance score was determined.

4.3.5 Raters

Four groups of coders were included in the rating of each video using the MI application evaluation feature: students, two independent raters, and a three-person consensus panel.

Using a P2P learning model, students were randomly assigned partners and instructed to perform and evaluate the SJ and HBR skills using the MI application. Students were encouraged to do the assessment together and to engage in discussions regarding their performance. Screenshots of the summary page containing both the component and performance scores were taken by each student once they had completed their assessment. Subsequently, all participants’ videos were re-evaluated by two independent raters and an interdisciplinary consensus panel to provide “expert” scores for each student’s performance.

Two independent raters, the lead researcher (Rater 1) and a biomechanics expert (Rater 2), both of whom were also members of the interdisciplinary consensus panel, viewed and rated the 34 videos independently. Rater 1 was a kinesiology graduate student and Rater 2 had completed a Bachelor of Science in Kinesiology, specializing in Biomechanics.

In addition to Raters 1 and 2, the consensus panel also included a kinesiologist. All members of the interdisciplinary consensus panel were knowledgeable about physical literacy and FMS, with each holding a Bachelor of Science in Kinesiology. Due to previous involvement in research and pilot studies, the lead researcher was the only member who had previous experience evaluating FMS videos using the MI application.

4.3.6 Data Collection Protocol

Data were uploaded to a secure server, labelled using participant ID numbers, and exported to external hard drives. Each coder was given access to all 34 videos. This digital format provides multiple benefits, as it allows raters to pause, slow down, and re-watch the video as many times as required (Haynes & Miller, 2015).

Prior to data collection, the lead researcher conducted a 1-hour training session, where members of the consensus panel practiced using the MI application and spent time discussing how to evaluate the FMS using the software. Topics covered during this training sessions included instructions on how to log in and watch the videos, how to adjust video playback speed, and troubleshooting video playback issues.

4.3.7 Scoring

The consensus panel rated each video twice; however, the scores were recorded 6 months apart. The first round of consensus panel scoring was termed “Consensus 1”, while the second round of consensus panel scoring was called “Consensus 2”. During this time frame the raters were not permitted to view the videos to limit recall. Additionally, raters were not allowed to reference the first scoring round results during the second rating session. The consensus panel viewed and scored all videos together by using the built-in evaluation feature within the MI application. After each assessment, any discrepancies were discussed until all members came to an agreement. There were no limits to how many times a video could be viewed.

In comparison, individual coders rated each video once, independently, using the built-in evaluation feature within the MI application as well. It is important to note that coders were not granted access to view each others' scores throughout this process. There were no limits to how many times a video could be viewed.

4.3.8 Data and Statistical Analysis

All statistics were conducted in SPSS version 26, and alpha levels were set to 0.05 *a priori.*

The inter-rater reliability analysis was conducted by addressing the ICC (2,5) values of the overall skill scores for the raters by pairs (Rater 1, Rater 2, Consensus 1, Consensus 2, and Student). ICC reliability values are characterized as poor (<0.50), moderate (0.51-0.75), or good (>0.75; Portney 2020; Sheehan et al., 2011). An independent t-test was conducted to determine a statistically significant difference between the overall scores for each skill provided by each rater. Intra-rater analysis was conducted by examining the ICC (2,2) values of the total scores and component scores for Consensus 1 and Consensus 2. Additionally, a dependent t-test was conducted to determine a statistically significant difference between the Consensus 1 and Consensus 2 scores for each skill. Finally, percent agreement was conducted to further analyse the interrater and intra-rater reliability of the MI application by looking at the agreement between raters for each component of the skill.

4.4 Results

4.4.1 Inter-rater Reliability

Table 4.1 summarizes the inter-rater ICC (2,5) values calculated for SJ performance scores for all the inter-rater combinations. The findings showed an $\alpha > 0.75$ between all pairs of raters, with α values ranging from $\alpha = 0.827\text{-}0.936$. According to these ICC (2,5) reliability values, the SJ performance score inter-rater reliability results can be categorized as good for all rater pairings.

Table 4.1 Inter-rater Intraclass Correlation (2,5) for Overall Performance Scores Between Rater 1, Rater 2, Students, Consensus 1 and Consensus 2 for Standing Jump

	Rater 1	**Rater 2**	**Students**
Rater 1	-	-	-
Rater 2	0.936	-	-
Students	0.886	0.841	-
Consensus 1	0.866	0.827	0.926
Consensus 2	0.916	0.872	0.882

Table 4.2 summarizes the inter-rater ICC (2,5) values calculated for HBR performance scores for all the inter-rater combinations. Both individual raters were found to have good inter-rater reliability between them, as demonstrated by the ICC (2,5) value of $\alpha = 0.802$. Additionally, Rater 2 had good inter-rater reliability with both Consensus 1 ($\alpha = 0.913$) and Consensus 2 ($\alpha = 0.909$); meanwhile, Rater 1 only had good inter-rater reliability with Consensus 1 ($\alpha = 0.861$), and moderate inter-rater reliability with Consensus 2 ($\alpha = 0.679$). Overall, students showed poor inter-rater reliability with both individual raters and both rounds of consensus scoring, as demonstrated by ICC (2,5) values of $\alpha < 0.5$.

Table 4.2 Inter-rater Intraclass Correlation (2,5) for Overall Performance Scores Between Rater 1, Rater 2, Students, Consensus 1 and Consensus 2 for Hollow Body Roll

	Rater 1	**Rater 2**	**Students**
Rater 1	-	-	-
Rater 2	0.802	-	-
Students	0.337	0.346	-
Consensus 1	0.861	0.913	0.278
Consensus 2	0.679	0.909	0.145

Table 4.3 contains the inter-rater ICC values calculated for each of the SJ and HBR components. The ICC (2,5) values were calculated between all raters (e.g., Rater 1, Rater 2, Consensus 1, Consensus 2, and Students), whereas the ICC (2,4) values were calculated without

student raters (e.g., Rater 1, Rater 2, Consensus 1, and Consensus 2). With regards to the SJ results, the ICC (2,5) values ranged from α = 0.585-0.989, while the ICC (2,4) values ranged from α = 0.741-0.990. In both cases, the α value was >0.75 for all the components, except for component 8, which had an ICC (2,5) of α = 0.585 and an ICC (2,4) of α = 0.741. According to these findings, component 8 can be categorized as having moderate inter-rater reliability, while the remaining seven SJ components can be categorized as having good inter-rater reliability. With regards to the HBR results, the ICC (2,5) values ranged from α = 0.391-0.875, while the ICC (2,4) values ranged from α = 0.375-0.906. When student raters were included in the statistical analysis, components 1 and 3 had poor inter-rater reliability, component 5 had moderate inter-rater reliability, and components 2, 4, and 6 all had good inter-rater reliability. Meanwhile, when student raters were excluded from the statistical analysis, component 3 was the only component to demonstrate poor inter-rater reliability, and component 1 now demonstrated moderate inter-rater reliability. Finally, the remaining four HBR components all had good inter-rater reliability. Generally, both the SJ and HBR ICC (2,4) results demonstrated increased reliability compared to the ICC (2,5) results.

Table 4.3 Inter-rater Intraclass Correlation Coefficients for Standing Jump and Hollow Body Roll Individual Components With and Without Student Raters

Skill	**Component**	**Intraclass Correlation Coefficient (α)**	
		ICC (2,5) – With Student Raters (*n*=14)	**ICC (2,4) – Without Student Raters (*n*=18)**
Standing Jump	1	0.812	0.853
	2	0.891	0.878
	3	0.907	0.895
	4	0.919	0.882
	5	0.892	0.905
	6	0.989	0.990
	7	0.893	0.930
	8	0.585	0.741
		ICC (2,5) – With Student Raters (*n*=15)	**ICC (2,4) – Without Student Raters (*n*=16)**
Hollow Body Roll	1	0.467	0.693
	2	0.875	0.858
	3	0.391	0.375
	4	0.808	0.906
	5	0.680	0.804
	6	0.871	0.902

Table 4.4 shows the results from the independent t-test for SJ performance scores between all raters, including Rater 1, Rater 2, Consensus 1, Consensus 2, and Students. According to the data, there were no significant differences between performance scores given by each pair of raters.

Table 4.4 Descriptive Statistics and Independent Sample t-Test for Standing Jump Performance Scores Between Raters

Rater	***n***	**Mean (x̄)**	**Standard Deviation (*s*)**	***t***	**Significance (*p*)**	**Cohen's D (*d*)**
1	18	20.2	2.29	1.022	0.314	0.327
2	18	19.4	2.59			
1	18	20.2	2.29	0.524	0.604	0.180
S	14	19.7	3.20			
1	18	20.2	2.29	-0.062	0.951	0.037
C #1	18	20.3	3.03			
1	18	20.2	2.29	-0.341	0.736	0.122
C #2	18	20.5	2.60			
2	18	19.4	2.59	-0.318	0.753	0.103
S	14	19.7	3.20			
2	18	19.4	2.59	-0.946	0.351	0.319
C #1	18	20.3	3.03			
2	18	19.4	2.59	-1.285	0.207	0.424
C #2	18	20.5	2.60			
S	14	19.7	3.20	-0.510	0.614	0.193
C#1	18	20.3	3.03			
S	14	19.7	3.20	-0.768	0.449	0.274
C #2	18	20.5	2.60			

Note. *$p < 0.05$; ** $p < 0.01$; C = consensus; S = student

Table 4.5 shows the results from the independent t-test for HBR performance scores between all raters, including Rater 1, Rater 2, Consensus 1, Consensus 2, and Students. The pairings of Rater 1 and Rater 2 ($p = 0.019$), Rater 1 and Consensus 1 ($p = 0.015$), Rater 2 and Students ($p = 0.011$) and Students and Consensus 1 ($p = 0.009$) showed significant differences between scores. The t-test revealed no significant differences between scores in the remaining pairs.

Table 4.5 Descriptive Statistics and Independent Sample t-Test for Hollow Body Roll Performance Scores Between Raters

Rater	*n*	Mean (x̄)	Standard Deviation (*s*)	*t*	Significance (*p*)	Cohen's D (*d*)
1	16	14.4	2.68	4.2.474	0.019*	0.871
2	16	12.1	2.60			
1	16	14.4	2.68	-0.106	0.916	0.040
S	15	14.5	2.33			
1	16	14.4	2.68	2.575	0.015*	0.889
C #1	16	11.9	2.94			
1	16	14.4	2.68	1.617	0.116	0.564
C #2	16	12.8	2.99			
2	16	12.1	2.60	-2.709	0.011*	0.972
S	15	14.5	2.33			
2	16	12.1	2.60	0.255	0.801	0.072
C #1	16	11.9	2.94			
2	16	12.1	2.60	-0.693	0.494	0.250
C #2	16	12.8	2.99			
S	15	14.5	2.33	2.779	0.009**	0.980
C #1	16	11.9	2.94			
S	15	14.5	2.33	1.779	0.086	0.632
C #2	16	12.8	2.99			

Note. *$p < 0.05$; ** $p < 0.01$; C = consensus; S = student

To analyze the measure of agreement between groups, percent agreement was determined for each component. The SJ skill consisted of 8 components, while the HBR skill consisted of 6 components.

As indicated in Table 4.6, raters had an average percent agreement of 64% for SJ, compared to a lower percent agreement of 39% for HBR. In addition, the mean percent agreement varied between the individual components, ranging from 33%-89% for SJ and 0%-56% for HBR. When comparing the percent agreement between components for SJ, it became apparent that raters disagreed most on component 2 (percent agreement = 39%) and component 8 (percent agreement = 33%). The SJ component with the highest percent agreement was component 6, with a percent agreement of 89%. When comparing the percent agreement between components for HBR, results showed that raters disagreed most on component 3 and component 5, with a percent agreement of 0% and 19%, respectively. Finally, the mean percent agreement for each HBR component was <60%.

Table 4.6 Mean Inter-rater Percent Agreement Based on Individual Components for Standing Jump and Hollow Body Roll (Rater 1, Rater 2, Consensus 1, Consensus 2, and Students)

Skill	**Component**	**Mean Percent Agreement (%)**
Standing Jump	1	72
	2	39
	3	72
	4	67
	5	72
	6	89
	7	67
	8	33
	Average Score	64
Hollow Body Roll	1	56
	2	50
	3	0
	4	56
	5	19
	6	50
	Average Score	39

Table 4.7 summarizes the inter-rater percent agreement of individual components for SJ. When analyzing the percent agreement between pairings, the pattern observed was that the lowest percent agreements were found in components 2 and 8. Furthermore, the pairings that demonstrated the lowest average inter-rater percent agreement (<55%) were those that were compared with the student evaluations. An example is the percent agreement between Rater 2 and Students, where components 2 and 8 have a percent agreement of 28% and 39%,

respectively. In this specific example, there was also <70% percent agreement for all the other components.

Table 4.7 Inter-rater Percent Agreement of Individual Components for Standing Jump

Percent Agreement (%)	**Components**								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Average**
Rater 1 & Rater 2	72	44	72	72	78	89	89	56	**72**
Rater 1 & Consensus 1	83	78	83	89	83	100	72	61	**81**
Rater 2 & Consensus 1	83	44	89	72	78	89	72	44	**72**
Rater 1 & Consensus 2	83	83	83	89	83	100	72	61	**82**
Rater 2 & Consensus 2	83	39	89	72	78	89	72	44	**71**
Rater 1 & Students	50	44	61	61	56	72	56	22	**53**
Rater 2 & Students	67	28	56	56	56	61	56	39	**52**
Students & Consensus 1	56	44	56	61	56	72	61	22	**53**
Students & Consensus 2	56	50	56	61	56	72	61	22	**54**

Table 4.8 summarizes the inter-rater percent agreement of individual components for HBR. When analyzing the percent agreement between pairings, components 3 and 5 were found to have the lowest percent agreement. However, there was <70% percent agreement for most pairings and their corresponding component ratings. Once again, the pairings that demonstrated the lowest average inter-rater percent agreement (<50%) were those that were compared with the student evaluations. An example is the percent agreement between Rater 2 and Students, where components 3 and 5 have a percent agreement of 13% and 19%, respectively. In this specific example, there was also <70% percent agreement for all the other components.

Table 4.8 Inter-rater Percent Agreement of Individual Components for Hollow Body Roll

Percent Agreement (%)	**Components**						
	1	**2**	**3**	**4**	**5**	**6**	**Average**
Rater 1 & Rater 2	75	63	6	63	38	75	**53**
Rater 1 & Consensus 1	69	69	6	69	38	63	**52**
Rater 2 & Consensus 1	63	63	75	69	63	56	**65**
Rater 1 & Consensus 2	69	69	6	69	38	63	**52**
Rater 2 & Consensus 2	63	63	75	69	63	56	**65**
Rater 1 & Students	50	50	56	44	44	38	**47**
Rater 2 & Students	31	56	13	50	19	50	**36**
Students & Consensus 1	38	38	25	38	19	38	**32**
Students & Consensus 2	38	38	25	38	19	38	**32**

4.4.2 Intra-rater Reliability

Table 4.9 summarizes the intra-rater ICC (2,2) values calculated for both SJ and HBR performance scores between Consensus 1 and Consensus 2. The findings showed an $\alpha > 0.75$ for both skills, with SJ and HBR receiving α values of 0.896 and 0.920, respectively. According to these ICC (2,2) reliability values, the SJ and HBR intra-rater reliability performance score results can be categorized as good for both skills.

Table 4.9 Intra-rater Intraclass Correlation (2,2) for Overall Performance Scores Consensus 1 and Consensus 2 for Standing Jump and Hollow Body Roll

Skill	***n***	**Consensus**	**α**
Standing Jump	18	1	0.896
	18	2	
Hollow Body Roll	16	1	0.920
	16	2	

According to the data in Table 4.10, there was no significant difference between the first consensus scores and the second consensus scores for SJ (t = -0.544, p = 0.594). However, there was a significant difference between the first consensus scores and the second consensus scores for the HBR (t= -2.270, p = 0.038), with the second consensus scores being greater than the first consensus scores (12.8 $\pm$ 2.88 > 11.9 $\pm$ 2.94).

Table 4.10 Descriptive Statistics and Paired Samples t-Test for Standing Jump and Hollow Body Roll Performance Scores

Skill	***n***	**Consensus**	**Score (x)**	**Standard Deviation (*s*)**	***t***	**Significance (*p*)**	**Cohen's D (*d*)**
Standing Jump	18	1	20.3	3.03	-0.544	0.594	0.0788
	18	2	20.5	2.60			
Hollow Body Roll	16	1	11.9	2.94	-2.328	0.034*	0.322
	16	2	12.8	2.99			

Note. *p < 0.05

Table 4.11 contains the intra-rater ICC (2,2) values calculated for each of the eight SJ components. The findings showed α values ranging from α = 0.517-0.983. The α value was <0.75 for components 4, 7 and 8, with Consensus 2 giving higher scores for components 4 and 8, and a lower score for component 7; the five remaining components all had α >0.75. According to these results, components 4, 7, and 8 can be categorized as having moderate intra-rater reliability, while the five remaining components can be categorized as having good intra-rater reliability.

Table 4.11 Intra-rater Intraclass Correlation (2,2) Between Consensus 1 and Consensus 2 by Individual Components for Standing Jump

Component	Consensus	Score (x)	Standard Deviation (*s*)	α
1	1	2.7	0.767	0.876
	2	2.8	0.647	
2	1	2.1	0.832	0.914
	2	2.0	0.840	
3	1	2.8	0.515	0.940
	2	2.9	0.471	
4	1	2.6	0.784	0.746
	2	2.7	0.575	
5	1	2.7	0.686	0.933
	2	2.7	0.686	
6	1	2.4	0.916	0.983
	2	2.3	0.908	
7	1	2.8	0.384	0.598
	2	2.4	0.705	
8	1	2.2	0.878	0.517

Table 4.12 contains the intra-rater ICC (2,2) values calculated for each of the six HBR components. The findings showed α values ranging from α = 0.064-0.906. The α value was >0.75 for all the components except component 1, which had an α=0.064. According to these results, component 1 can be categorized as having poor intra-rater reliability, while the remaining seven components can be categorized as having good intra-rater reliability.

Table 4.12 Intra-rater Intraclass Correlation (2,2) Between Consensus 1 and Consensus 2 by Individual Components for Hollow Body Roll

Component	Consensus	Score (x)	Standard Deviation (*s*)	α
1	1	2.4	0.629	0.064
	2	2.7	0.602	
2	1	2.1	0.885	0.757
	2	2.7	0.479	
3	1	1.4	0.727	0.897
	2	1.4	0.719	
4	1	2.1	0.998	0.906
	2	2.3	0.856	
5	1	1.9	0.885	0.783
	2	1.5	0.817	
6	1	1.9	0.772	0.899
	2	2.3	0.873	

Table 4.13 summarizes the intra-rater percent agreement of the individual components for SJ between Consensus 1 and Consensus 2. The average percent agreement across components was 80%, but it ranged from 44%-94%. When comparing the percent agreement between components, it became apparent that the consensus panel disagreed most on component 7 (percent agreement = 67%) and component 8 (percent agreement = 44%). Meanwhile, components 1, 3, and 6 were tied for highest agreement, with a percent agreement of 94%.

Table 4.13 Intra-rater Percent Agreement of Individual Components for Standing Jump

Percent Agreement (%)	Components								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Average**
Consensus 1 & Consensus 2	94	78	94	78	89	94	67	44	80

Table 4.14 summarizes the intra-rater percent agreement of the individual components for HBR between Consensus 1 and Consensus 2. The average percent agreement across components was 66%, but it ranged from 44%-81%. When comparing the percent agreement between components, it became apparent that the consensus panel disagreed most on component 1 (percent agreement = 44%) and component 2 (percent agreement = 50%), while component 3 had the highest agreement, with a percent agreement of 81%.

Table 4.14 Intra-rater Percent Agreement of Individual Components for Hollow Body Roll

Percent Agreement (%)	**Components**						
	1	**2**	**3**	**4**	**5**	**6**	**Average**
Consensus 1 & Consensus 2	44	50	81	75	75	69	66

4.5 Discussion

The primary purpose of this study was to examine the inter-rater and intra-rater reliability of the MI application when assessing the performance of a SJ and HBR by grade 5-8 students. Additionally, this paper examined the differences between student's peer assessments, individual raters, and a consensus panel. Using video and performance analysis technology, the MI software allows users to evaluate a variety of movement-based skills. To our knowledge, this is the first study to assess the inter-rater and intra-rater reliability of the MI application in middle school children. It was hypothesized that the MI application would be reliable when assessing healthy middle school students' SJ and HBR technique. However, the results of this study suggest that although the overall scores were generally reliable, some of the individual components have low reliabilities.

4.5.1 Inter-rater Reliability

According to the ICC (2,5) data, the SJ performance scores demonstrated good inter-rater reliability between all rater pairings, with results ranging from $\alpha = 0.827$-0.936, but the inter-rater reliability results for HBR performance scores showed more variability. The results from the HBR performance scores had poor, moderate, and good inter-rater reliability, depending on the rater pairing in question, and ranged from $\alpha = 0.145$-0.913. Both Rater 1 and Rater 2 demonstrated moderate or good inter-rater reliability with Consensus 1 and Consensus 2 HBR performance scores. However, the lowest inter-rater reliability scores for HBR performance scores came from the pairings involving students; they showed poor inter-rater reliability with both individual raters, and both rounds of consensus scoring, as demonstrated by ICC (2,5) values of $\alpha < 0.5$.

To expand on the inter-rater reliability results, the individual components were compared between all raters (ICC (2,5)) and without student raters (ICC (2,4)). In both cases, SJ components 1 through 7 demonstrated good inter-rater reliability, whereas component 8 had moderate inter-rater reliability (Table 4.3). These findings support the data in Table 4.6, which showed that component 8 of SJ had the lowest mean inter-rater percent agreement. When looking at the HBR results, the statistical analysis that included the student raters revealed that components 1 and 3 had poor inter-rater reliability, but when students were excluded from the analysis, only component 3 demonstrated poor inter-rater reliability. These findings support the data in Table 4.6, which showed that component 3 of HBR had the lowest mean inter-rater percent agreement. In general, when student raters were excluded from the statistical analysis, both the SJ and HBR inter-rater reliability results for the individual components showed increased reliability (Table 4.3). This finding is supported by the data in Tables 4.7 and 4.8, which show that that the rater pairings with the lowest average inter-rater percent agreement were those that were compared with the student evaluations.

To further explore the reasonings for the ICC (2,5) results, the data were analyzed using percentage agreements. Results showed that raters had a higher mean inter-rater percent agreement for SJ in comparison to HBR. While the ICC (2,5) results show that most of the rater pairings were found to have moderate or good inter-rater reliability, the component scores between raters had lower levels of agreement, as demonstrated by the mean inter-rater percent agreement data. This indicates that the MI application allows raters to achieve a good level of inter-rater reliability for overall performance scores, despite the variability in mean inter-rater percent agreements between individual raters for the component scores.

With regards to the reliability of peer assessment, discrepancies were found between student and professional assessments. For both skills, the pairings that demonstrated the lowest inter-rater percent agreement were those that were compared with the student evaluations. Results from Tables 4.6 and 4.7 indicate that the student evaluations had low agreement when compared with the assessments completed by the professionals (Rater 1, Rater 2, Consensus 1, and Consensus 2). This observation is similar to that found by Nadeau et al. (2008), who determined that students had lower percent agreements when compared to physical educators, suggesting that inexperienced coders have greater variability. Similarly, a study conducted by Johnson and Ward (2001) found that the range of elementary school students’ accuracy in determining each other’s performance in a striking task was quite large, ranging from 46-100%. It has been suggested that a lack of training in discriminating performance may result in recording inaccuracy (Kolovelonis & Goudas, 2012). Thus, increasing students’ knowledge regarding the critical elements of a sport skill may increase their recording accuracy.

Similarly skilled raters (e.g., Rater 1 and Rater 2) should achieve high consistency and agreements during an evaluation. Indeed, both individual raters were found to have good inter-rater reliability for both skills, as demonstrated by the ICC (2,5) values of α = 0.936 and α = 0.802 for SJ and HBR, respectively. Although the ICC (2,5) data suggests the individual raters (i.e., Rater 1 and Rater 2) are of similar skill levels, the *t*-test results show significant differences for HBR (Table 4.4). When analyzing the percentage agreement data, individual raters were found to have higher percent agreement values when compared to the mean percent agreement values.

Although individual raters showed a trend of inflated scores compared to the consensus panel for HBR, there was still moderate to good reliability overall in the ICC scores. Upon further investigation of the raw data and percentage agreements, it appears the consensus group was more discriminating of performance, while the individual raters were more generous with giving a “yes” score. The data reflects trends like those described by Palmer and Brian (2016), suggesting that the consensus panel acted similarly to “expert raters”, while the individual raters followed similar patterns to “novice raters”.

4.5.2 Intra-rater Reliability

According to the data, there was no significant difference between the Consensus 1 and Consensus 2 scores for SJ. However, a significant difference was detected between the Consensus 1 and Consensus 2 scores for HBR which revealed that the consensus panel gave higher scores during the second round of evaluations. Interestingly, the ICC (2,2) values for SJ and HBR are α = 0.896 and α = 0.912, respectively, indicating a good level of intra-rater reliability. However, it is important to note that a high ICC value does not imply that the scores are unidimensional (Cohen, 1988).

Comparisons between the alpha levels for each component were performed to determine the reason for the statistical difference between the first and second consensus score for HBR. Component 1 of the HBR reported an α level of 0.064, which is very low (Portney, 2020). This component assesses if the students' hands are overhead during the initiation of the HBR. The consensus panel may have given greater scores the second time around because the raters were more experienced at evaluating HBR videos and determining the execution of the skill itself. Illing (1965) and Michaelsen et al. (1989) argue that group-based decision making allows for groups to combine participant's knowledge through discussion. Through discussions in the first evaluation, it became clear to the consensus panel that the students were simply attempting to copy what they had seen in the HBR demonstration video, where component 1 was filmed at a poor angle. This new knowledge was taken into consideration during the second round of evaluations which likely led to higher Consensus 2 scores.

4.5.3 Limitations

As the consensus panel had more combined assessment experience with the MI application compared to the individual raters, there is a possibility that the group dynamic, in particular the chance to discuss student performance with other "experts" affected scores, as opposed to the individual raters who gave scores in isolation. Furthermore, as the first author was both an individual rater and part of the consensus panel, there is the potential for exposure misclassification bias. However, this bias was mitigated by including the first author as part of a consensus panel whereby the scoring occurred six months after the data was collected, and the final score was recorded after a consensus discussion. The use of consensus panels in inter-rater reliability studies appears to be a new phenomenon and warrants further investigation.

Students may be able to replicate components 3 and 5 of the HBR because they themselves were able to feel if their core and back muscles were engaged throughout the manoeuvre. However, this is much more challenging to assess through video recordings, since it is more difficult to determine if the appropriate muscles are engaged, especially if the student is wearing loose fitting clothing. Therefore, there is a possibility that students were more accurate than the researchers in their evaluations for these specific components. This issue could be slightly alleviated by asking participants to tuck in their shirts when performing the skill, or by wearing leotards, which is commonly done in gymnastics.

The use of percent agreements may be an inadequate measure of reliability, despite being one of the most commonly used methods to assess agreement (Hallgren, 2012). This study rated all subjects and utilized a fully crossed design to limit systematic bias, but nonetheless there remains a potential for bias. As such, additional statistical methods were applied in the analysis. Finally, only two individual raters and one consensus group were used for comparison; additional raters would provide more clarity in assessing the reliability of the assessment tool.

4.5.4 Conclusion

The purpose of this study was to examine the reliability of the MI application when assessing the performance of two fundamental movement skills, SJ and HBR, by middle school students. In this study, inter-rater and intra-rater reliability were examined using ICC and percent agreement. The consensus panel results from this study are promising because both the SJ and HBR overall performance scores showed good intra-rater reliability. SJ assessments also showed good inter-rater reliability between all groups of coders, while the student peer assessments for the HBR were inconsistent when compared to the individual coders and consensus panel. Further analysis into the individual skill components revealed some inconsistencies in the scores given by each coder, particularly component 1 of the HBR, which must be addressed before MI can be considered a reliable assessment tool for that skill. Improving the training protocols for students and professionals may help to improve the reliability of the instrument as well. Future research should examine inter-rater and intra-rater differences between consensus panels, individual raters, and students. Finally, additional research should be conducted to re-evaluate MI's reliability and validity in different populations and with different skill-level coders.

References

Alstot, A. E. (2018). Accuracy of a peer process assessment performed by elementary physical education students. *The Physical Educator, 75*(1), 739-755.

Bailey, D. B. J., Buysse, V., Simeonsson, R. J., Smith, T., & Keyes, L. (1995). Individual and team consensus ratings of child functioning. *Developmental Medicine & Child Neurology*, *37*(3), 246–259.

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.

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.

Borghouts, L. B., Slingerland, M., & Haerens, L. (2017). Assessment quality and practices in secondary PE in the Netherlands. *Physical Education and Sport Pedagogy*, *22*(5), 473–489.

Chang, C,C., Tseng, K,H., Chou, P,N., & Chen, Y,H. (2011). Reliability and validity of web-based portfolio peer assessment: A case study for a senior high school's students taking computer course. *Computers & Education, 57*(1), 1306–1316.

Claxton, D. B., Troy, M., & Dupree, S. (2006). A question of balance. *Journal of Physical Education, Recreation and Dance, 77*(1), 32–37.

Cohen, J. W. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale: Lawrence Erlbaum Associates.

Evans, G. L. (2017). Handheld Technology Use with EAL Learners in Lincolnshire Primary Schools (Doctoral dissertation, Bishop Grosseteste University).

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.

Fry, S.A. (1990). Implementation and evaluation of peer marking in higher education. *Assessment & Evaluation in Higher Education*, *15*(3), 177–189.

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. *Perceptual and Motor Skills*, *121*(1), 260–283.

Gallahue, D., Ozmun, J. & Goodway, J. (2012). *Understanding motor development: Infants, children, adolescents, adults*. 7th ed. New York: McGraw-Hill.

Goodway, J., Brian, A., Chang, S., & Park, S. (2014). Methodological approaches to evaluate motor skill development, physical activity, health-related fitness, and perceived motor competence in young children. *Handbook of Research Methods in Early Childhood Education*, *2*(1), 349-391.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34.

Hardy, L. L., Reinten-Reynolds, T., Espinel, P., Zask, A., & Okely, A. D. (2012). Prevalence and correlates of low fundamental movement skill competency in children. *Pediatrics*, *130*(2), 390-398.

Haynes, J., & Miller, J. (2015). Preparing pre-service primary school teachers to assess fundamental motor skills: Two skills and two approaches. *Physical Education and Sport Pedagogy*, *20*(4), 397–408.

Higgs, C., (2010). Physical literacy – two approaches, one concept. *Physical & Health Education Journal*, *76*(1), 6–7.

Illing, H. A. (1965). " A Social Psychology of Group Processes for Decision-Making," by Barry E. Collins and Harold Guetzkow (Book Review). *International Journal of Group Psychotherapy*, *15*(1), 405-406.

Jenkinson, K.A, Naughton, G., & Benson, A.C. (2014). Peer-assisted learning in school physical education, sport and physical activity programmes: A systematic review. *Physical Education and Sport Pedagogy*, *19*(3), 253-277.

Johnson, M., & Ward, P. (2001). Effects of classwide peer tutoring on correct performance of striking skills in 3rd grade physical education. *Journal of Teaching in Physical Education*, *20*(3), 247-263.

Kniffin, K. M., & baert, H. (2015). Maximizing learning through assessment in middle and high school physical education. *Journal of Physical Education, Recreation, and Dance, 86*(4), 7-16.

Kolovelonis, A., & Goudas, M. (2012). Students' recording accuracy in the reciprocal and the self-check teaching styles in physical education. *Educational Research and Evaluation, 18*(8), 733-747.

Logan, S., Robinson, L., Wilson, A., & Lucas, W. (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.

Lund, J., & Veal, M. L. (2013). Assessment-driven instruction in physical education: A standards-based approach to promoting and documenting learning. Champaign, IL: Human Kinetics.

Magin, D.J. (2001). A novel technique for comparing the reliability of multiple peer assessments with that of single teacher assessments of group process work. *Assessment & Evaluation in Higher Education, 26*(2), 139–152.

Michaelsen, L. K., Watson, W. E., & Black, R. H. (1989). A realistic test of individual versus group consensus decision making. *Journal of Applied Psychology, 74*(5), 834–839.

Minick, K. I., Kiesel, K. B., Burton, L., Taylor, A., Plisky, P., & Butler, R. J. (2010). Interrater reliability of the functional movement screen. *Journal of Strength and Conditioning Research*, *24*(2), 479–486.

Mohnsen, B. S. (2012). *Using technology in physical education*. Bonnie's Fitware, Inc.

Nadeau, L., Richard, J.F., & Godbout, P. (2008). The validity and reliability of a performance assessment procedure in ice hockey. *Physical Education & Sport Pedagogy, 13*(1), 65–83.

Ng, S. K. (2002). A discussion on role conflict between Hong Kong physical education teachers and coaches. *Journal of Physical Education and Recreation (Hong Kong), 8*(1), 59-63.

Palmer, K. K., & Brian, A. (2016). Test of Gross Motor Development-2 scores differ between expert and novice coders. *Journal of Motor Learning and Development*, *4*(2), 142–151.

Payne, G., & Isaacs, L. D. (2020). *Human motor development: A lifespan approach*. Routledge.

Portney, L. G. (2020). *Foundations of Clinical Research: Applications to Evidence-Based Practice*. (4th ed.). Philadelphia: F. A. Davis.

Rezaei, A. R. & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, *15*(1), 18–39.

Roberton, M. A. (1977). Stability of stage categorizations across trials: implications for the "stage theory" of overarm throw development. *Journal of Human Movement Studies, 3*(1), 49–59.

Sheehan, D. P., Lafave, M. R., & Katz, L. (2011). Intra-rater and inter-rater reliability of the balance error scoring system in pre-adolescent school children. *Measurement in Physical Education and Exercise Science*, *15*(3), 234–243.

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.

Ulrich, D.A. (2000). *Test of gross motor development*: 2nd Edition. Austin, Texas: Pro-ed.

Veal, M. L. (1995). Assessment as an instructional tool. *A Journal for Physical and Sport Educators*, *8*(6), 10-15.

Chapter Five: Discussion and Conclusion

The purpose of this thesis was twofold: (1) to determine the impact that order of performance within a reciprocal teaching style has on the performance outcomes of FMS in middle school students, and (2) to evaluate the inter-rater reliability and intra-rater reliability of the MI software when assessing the performance outcomes of middle school students' FMS. Using video and performance analysis technology, the MI software permits users to evaluate a variety of movement-based skills.

5.1 Research Aim #1 – Impact of order of performance within a reciprocal teaching style on the performance outcomes of fundamental movement skills in middle school students

The main findings of the first research question are as follows: This study explored the impact that order of performance has on middle school students' performance of FMS within a P2P learning model. Eighteen students (eight male, ten females) ranging from grades 5-8 from a private parochial school participated in the study. Using the MI iOS application, students completed a SJ and HBR in partners randomly assigned to order of performance (evaluator/performer). SJ was completed first and the order of performance was reversed for the HBR. Overall performance scores were calculated using the component criteria located within the MI software and recorded for further analysis. The results from this study suggest that there was a significant difference in SJ scores ($p < 0.01$), where students who performed second had higher scores than their peers who performed first. Although not significant, the HBR results also showed higher mean performance scores for students who went second compared to those who performed first. The overall findings seem to support the concept of "order of performance" having a positive impact on learning and performance outcomes in middle school students.

However, order of performance within a P2P learning model having a significant effect on SJ and not HBR does raise some questions as to why there was such a discernible difference between the two skills.

5.2 Research Aim #2 – Inter-rater reliability and intra-rater reliability of the Move Improve® application in middle school students

The main findings of the second research question are as follows: This study aimed to examine the differences between student's peer assessments, individual raters, and consensus panel coding. Using the MI iOS application, four distinct groups of raters (students, two independent raters, and a three-person consensus panel) evaluated a total of 34 videos, which included 18 SJ videos and 16 HBR videos. Students evaluated their individual videos once during their PE class using peer assessment. Individual raters assessed all videos once, while the consensus panel completed two rounds of evaluation with 6 months between the sessions. ICC and percent agreement were conducted to assess inter-rater and intra-rater reliability. The intra-rater reliability of the consensus panel was found to be good for both SJ and HBR, with values of $\alpha = 0.896$ and $\alpha = 0.920$, respectively. With regards to inter-rater reliability, SJ demonstrated good inter-rater reliability between all raters ($\alpha > 0.75$), whereas HBR had poor inter-rater reliability for the student raters compared to the individual raters and the consensus panel ($\alpha < 0.5$). As such, it seems that higher reliability was evident when raters with similar skill levels evaluated performances. Additionally, the high results in terms of intra-rater reliability but lower in the case of inter-rater reliability, suggest that some of the skill's component criteria can be interpreted differently. Further analysis into the individual task components demonstrated major inconsistencies depending on the rater, limiting the acceptability of the instrument. More specifically, the mean inter-rater percent agreement for SJ ranged from 33-89%, with an average of 64%, while HBR had a mean inter-rater percent agreement of 39% (range 0-56%).

To our knowledge, this is the first study to examine the intra-rater and inter-rater reliability differences between student, individual, and consensus raters when using the MI iOS application to evaluate FMS performance in middle school children. As such, it is unreasonable to generalize these results to suggest the trends between the different groups of raters will be consistent from various data sources (e.g., developmental groups, skill level of raters, motor skill assessment). Researchers have found that due to the differences in motor skill assessments, the assessment results are not interchangeable (Logan et al., 2012; Minick et al., 2010; Palmer & Brian, 2016). Therefore, assessors should be consistent in their use, selection, and interpretation of measurement tools to ensure proper evaluation of progression over time.

5.3 General Discussion

Technological advancements continue to offer new opportunities for PE. Scholars such as Casey et al. (2017) have argued that with appropriate pedagogical considerations, digital technology can offer potential benefits to learning and performance. Educators and researchers should develop new tools to facilitate structured video learning with P2P methods. With the appropriate pedagogical framework, effective use of technology, specifically through tablet computers, may offer a way to provide scaffolded content knowledge in PE classrooms (Gubacs-Collins & Juniu, 2009; Leight et al., 2009; O'Loughlin et al., 2013; Palao et al., 2015; Sinelnikov, 2012; Sun, 2012; Zhu & Dragon, 2016).

Recent studies have demonstrated young people's general acceptance of health applications in learning settings (Goodyear & Armour, 2018). Current literature suggests that the use of tablets can result in observational learning and increased engagement especially with more frequent use of the technology (Diemer et al., 2012). However, it remains critical that the software on the tablets be adaptable to the developmental level of the students, provide adequate support for implementation, and be sufficiently flexible to smoothly integrate into existing curricula in a variety of classroom settings (McManis & Gunnewig, 2012). Furthermore, it is critical that technology use is appropriately scaffolded into lesson plans to keep students on-task and engaged in higher levels of thinking and performance when using it (Chatoupis, 2018; Chng & Lund, 2018; Koehler et al., 2013; Metzler, 2001; Mosston & Ashworth, 2008; Ward & Lee, 2005; Yelland & Masters, 2007).

The nature of PE involves motor performances that can be effectively captured by digital video. While digital video has been shown to be beneficial in the PE classroom (Hung et al., 2018), providing structure can further enhance learning. Designing digital video tools which incorporate structure and feedback for self and peer evaluation can increase participant engagement and performance outcomes. Using educational tools, like the MI iOS application, the learning experience can be enhanced through scaffolding, criteria evaluation, video, and an assessment tool. Instructional scaffolding offers critical learning supports for students to develop mastery through temporary and adjustable assistance. As students master assigned tasks, these supports can be gradually removed; however, it is important to offer a variable level of scaffolding in an educational tool since scaffolding is most useful for teaching new tasks with multiple steps or components (Chai & Koh, 2017).

Finally, this tool also provides students with the opportunity to easily repeat their attempts to gain mastery. Once an individual has mastery of skills, they will have the confidence to perform the skill with competence, therefore increasing ones’ self-efficacy and self-esteem associated with being physically active (Mandigo et al., 2009). With higher levels of self-efficacy and self-esteem, individuals are more motivated to remain physically active throughout their lifespan (Higgs, 2010).

P2P learning is already being adopted in many PE classrooms due to increased engagement and improved performance (Ernst & Byra, 1998; Goldberger et al., 1982; Jenkinson et al., 2014; Koekoek & Knoppers, 2015; Metzler, 2017; Yoncalik et al., 2010). However, in practice, there appears to be a lack of strategies for providing students with the skills to provide accurate and reliable peer evaluations (Chng & Lund, 2018; Johnson, 2004; Jenkinson et al., 2014). Therefore, the development and testing of effective approaches for teaching students to be accurate and reliable in their evaluation and assessment for both self-assessment and P2P learning is an important first step. One example is the use of consensus learning, such as was used in this study. In this scenario, two students are video recorded while engaging in the assessment process. Then the class discusses the process and the evaluator's decision-making to come to a consensus on the issues. This allows the students to accurately evaluate each other and themselves. Additionally, the quality of interaction between students must be made a focus of the training, as for P2P learning to be more effective, it needs to provide students with training on how to deliver quality feedback with their partners. Measuring the quality of interaction between students offers a pathway to more effective teaching and learning in a P2P environment, which could also improve performance and assessment (Jenkinson et al., 2014; Johnson & Ward, 2001). If we teach students to take responsibility for their own learning by using reliable and accurate self and peer evaluation, then student engagement, communication skills, enthusiasm, and participation may all improve. Eddy et al. (2020) conducted a systematic review of the validity and reliability of observational assessment tools available to measure FMS in school-age children. Although they concluded that there is insufficient evidence to justify the use of any observational FMS assessment tool for universal screening in schools (Eddy et al., 2020), this thesis found that similar skilled coders had good inter-rater rater reliability when using the MI observational assessment tool. Although the students were found to have poor inter-rater reliability, undergoing training could increase their overall reliability. Nevertheless, the process of performing an assessment in peers seemed to enhance the overall learning experience. If students provide accurate performance assessments, the PE teacher will have additional detailed knowledge about the students' performance, enabling the instructor to provide enhanced feedback.

5.4 Future Directions

More time should be given to teach participants how to take clear videos of their peers and how to properly discuss and critique a movement skill. The time constraint of 30 minutes for a gym class may not provide enough time to effectively teach the fundamentals of critiquing and analysing a movement performance. Future studies will benefit from having multiple structured practice sessions before the participants attempt any performance skills and data collection.

It would also be useful to control for prior physical ability differences between students by conducting an initial FMS ability test prior to data collection. By conducting a pre-test, it would be possible to control the possible confounding factor, that is, the influence of students' previous skill levels. The results from this pre-test would inform student pairings, as students with similar skill levels could be paired together. This would eliminate the possibility that students with higher skill abilities went first, therefore making it harder to see if there truly is an order effect. Unfortunately, due to time constraints this was not done in this study.

In future studies, a larger sample size would be beneficial. Additionally, selecting participants from a single grade level would remove any possible age effect in their performance. The participation of mixed grade levels may have subtly affected the results of the study.

Future research should look at developing and testing training protocols to improve peer assessment techniques. Frequent changes in attitude, as well as in peer relationships, may contribute to students rating each other inconsistently on assessment tools or rating others based on popularity within peer groups (Lindblom-ylänne et al., 2006). When students are expected to rate their peers in any sort of activity, those ratings are often related more to subjective social dynamics than objective academic criteria (McDermid et al., 2014). Finally, peer interactions are volatile; on a daily or even hourly basis, separate interactions between peers or groups can influence the way adolescents perceive or relate to one another (Brown & Larson, 2009).

More research is needed to examine learners' perceptions of peer evaluation. Evaluating learner satisfaction with an assessment based on an understanding of the evaluation process itself could improve P2P evaluation further. Satisfaction with an assessment would not mean a learner was pleased because they received a high grade, but rather that the score was an accurate reflection of their performance. Therefore, increasing clarity and transparency of the principles guiding the evaluation will allow learners to better evaluate their peers and could possibly improve performance. Additionally, future research should account for friendship levels between peers.

Future studies should be undertaken to measure and evaluate not only the quality, but also the accuracy of student feedback. Ensuring student feedback correlates to a teacher-provided standard of good performance would reinforce desired knowledge content outcomes and improve the accuracy of peer assessment. Finally, additional research should be conducted to evaluate MI's reliability and validity in different populations.

The emerging technologies of augmented reality (AR) and virtual reality (VR) are gaining traction in the educational field (Calabuig-Moreno et al., 2020). Researchers have highlighted the suitability of these tools to PE (Díaz et al., 2018; Pasco, 2013), both to teach key concepts and to increase students' physical activity. In particular, VR can enrich the quality of PE instruction by allowing students to gain immersive feelings and experiences, which gives students a better understanding of sports knowledge and skills and promotes their physical and mental development (Zhang & Liu. 2016). As costs and accessibility become more reasonable (Zhang & Liu, 2016), these tools have potential to enhance the P2P learning environment and should be considered for future iterations of the structured learning tool described in this thesis.

5.5 Conclusion

The MI iOS application is a video performance analysis tool that can be used to facilitate structured video peer evaluation among student pairs. Using tablet-based structured video analysis for P2P learning has the potential to improve student learning and performance. This appears to be the first study examining the effect that order of performance has on middle school students' performance in PE. To our knowledge, this is also the first study to assess the inter-rater and intra-rater reliability of the MI application in middle school children.

The findings suggest that the order of performance within a P2P learning model may have a significant effect on performance scores for some, but not all FMS. The consensus panel results from this study are promising because both fundamental movement skills exhibited good intra-rater reliability. However, further analysis into the individual skill components revealed some inconsistencies in the scores given by each rater. Based on this data, it is difficult to decide if the MI assessment tool is reliable.

Based on the results, it would be beneficial to conduct a more thorough evaluation of the video quality of each skill in the MI application to ensure that they are clearly performed, and that the terminology of the components is easily understood. Improving training protocols for students and professionals may improve the reliability of the instrument.

References

Adams, J. (1986). Use of the model’s knowledge of results to increase the observer’s performance. *Journal of Human Movement Studies, 12*, 89-98.

Al-Abood, S. A., Davids, K., & Bennett, S. J. (2001). Specificity of task constraints and effects of visual demonstrations and verbal instructions in directing learner’ search during skill acquisition. *Journal of Motor Behaviour*, *33*(3), 295-305.

Alstot, A. E. (2018). Accuracy of a peer process assessment performed by elementary physical education students. *The Physical Educator, 75*(1), 739-755.

Amara, S., Mkaouer, B., Chaabene, H., & Nassib, S. H. (2015). Effect of Video Modeling Process on Teaching/Learning Hurdle Clearance Situations on Physical Education Students. *Advances in Physical Education, 5*(5), 225-233.

Ammah, J. O., & Hodge, S. R. (2005). Secondary physical education teachers' beliefs and practices in teaching students with severe disabilities: A descriptive analysis. *The High School Journal*, *89*(2), 40-54.

Anderson, A., & Goode, R. (1997). Assessment informs instruction. *Journal of Physical Education, Recreation & Dance*, *68*(3), 42-49.

Andrieux, M., & Proteau, L. (2014). Mixed observation favors motor learning through better estimation of the model’s performance. *Experimental Brain Research*, *232*(10), 3121-3132.

Andrieux, M., & Proteau, L. (2016). Observational learning: Tell beginners what they are about to watch and they will learn better. *Frontiers in Psychology*, *7*(51).

Angers, J., & Machtmes, K. (2005). An ethnographic-case study of beliefs, context factors, and practices of teachers integrating technology. *The Qualitative Report*, *10*(4), 771-794.

Ayvazo, S., & Ward, P. (2009). Effects of classwide peer tutoring on the performance of sixth grade students during a volleyball unit. *Physical Educator*, *66*(1), 12-22.

Bai, P., Thornton, A., Lester, L., Schipperijn, J., Trapp, G., Boruff, B., . . . Christian, H. (2020). Nature Play and Fundamental Movement Skills Training Programs Improve Childcare Educator Supportive Physical Activity Behavior. *International Journal of Environmental Research and Public Health*, *17*(1), 223.

Bailey, D. B. J., Buysse, V., Simeonsson, R. J., Smith, T., & Keyes, L. (1995). Individual and team consensus ratings of child functioning. *Developmental Medicine & Child Neurology*, *37*(3), 246–259.

Bandura, A. (1969). Social-learning theory of identificatory processes. *Handbook of Socialization Theory and Research*, *213*, 262.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, *61*(2), 213-238.

Barker, D., Nyberg, G., & Larsson, H. (2020). Joy, fear and resignation: investigating emotions in physical education using a symbolic interactionist approach. *Sport, Education and Society*, *25*(8), 872-888.

Barnett, M., Harwood, W., Keating, T., & Saam, J. (2002). Using emerging technologies to help bridge the gap between university theory and classroom practice: Challenges and successes. *School Science and Mathematics*, *102*(6), 299-313.

Barnett, L., van Beurden, E., Morgan, P. J., Brooks, L. O., & Beard, J. R. (2008). Does childhood motor skill proficiency predict adolescent fitness? Medicine and Science in Sports and Exercise, 40(12), 2137-2144.

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.

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.

Beighle, A. E., Erwin, H., Morgan, C. F., & Alderman, B. (2012). Children's In-School and Outof-School Physical Activity During Two Seasons. *Research Quarterly for Exercise and Sport*, *83*(1), 103–107.

Bennett, R. E. (2002). Inexorable and Inevitable: The Continuing Story of Technology and Assessment, *The Journal of Technology, Learning and Assessment*, *1*(1), 1-22.

Beseler, B., & Plumb, M. S. (2019). 10 Tips to Using Video Analysis More Effectively in Physical Education. *Journal of Physical Education, Recreation & Dance*, *90*(1), 52-56.

Bird, G., Osman, M., Saggerson, A., & Heyes, C. (2005). Sequence learning by action, observation and action observation. *British Journal of Psychology*, *96*(3), 371-388.

Black, C. B., & Wright, D. L. (2000). Can observational practice facilitate error recognition and movement production? *Research Quarterly for Exercise and Sport*, *71*(4), 331-339.

Blandin, Y. & Proteau, L. (2000). On the cognitive basis of observational learning: Development of mechanisms for the detection and correction of errors. *The Quarterly Journal of Experimental Psychology*, *53A*(3), 846-867.

Booth, M. L., Okely, T., McLellan, L., Phongsavan, P., Macaskill, P., Patternson, J., Wright, J., & Holland, B. (1999). Mastery of Fundamental Movement Skills among New South Wales School Students: Prevalence and Sociodemographic Distribution. *Journal of Science and Medicine in Sport*, *2*(2), 93-105.

Borghouts, L. B., Slingerland, M., & Haerens, L. (2017). Assessment quality and practices in secondary PE in the Netherlands. *Physical Education and Sport Pedagogy*, *22*(5), 473–489.

Boud, D. (2001). Making the move to peer learning. *Peer learning in higher education: Learning from and with each other*, *1*, 20.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment and Evaluation in Higher Education*, *24*(4), 413-426.

Boyce, B. A., Markos, N. J., Jenkins, D. W., & Loftus, J. R. (1996). How should feedback be delivered?. *Journal of Physical Education, Recreation & Dance*, *67*(1), 18-22.

Brown, B. B., & Larson, J. (2009). Peer Relationships in Adolescence. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of Adolescent Psychology: Contextual Influences on Adolescent Development* (pp. 74-103). John Wiley & Sons, Inc.

Buchanan, J. J., & Dean, N. J. (2010). Specificity in practice benefits learning in novice models and variability in demonstration benefits observational practice. *Psychological Research*, *74*, 313-326.

Burnham, J. M. (1998). Exercise is medicine: health benefits of regular physical activity. *The Journal of the Louisiana State Medical Society: Official Organ of the Louisiana State Medical Society*, *150*(7), 319-323.

Burton, A. W., & Miller, D. E. (1998). *Movement Skill Assessment*. Human Kinetics.

Butler, S. A., & Hodge, S. R. (2001). Enhancing student trust through peer assessment in physical education. *Physical Educator, 58*(1), 30-42.

Buzzetto-More, N. A., & Alade, A. J. (2006). Best practices in e-assessment. *Journal of Information Technology Education: Research*, *5*(1), 251-269.

Byers, C. (2001). Interactive assessment: an approach to enhance teaching and learning. *Journal of Interactive Learning Research*, *12*(4), 359-374.

Byra, M. (2006). Teaching styles and inclusive pedagogies. In D. Kirk, D. MacDonald, & M. O'Sullivan (Eds.), *The Handbook of Physical Education* (pp.449-466). London: Sage Publications.

Byra M., & Marks M.C. (1993). The effect of two pairing techniques on specific feedback and comfort levels of learners in the reciprocal style of teaching. *Journal of Teaching in Physical Education*, *12*(3), 286-300.

Cairncross, S., & Mannion, M. (2001). Interactive multimedia and learning: Realizing the benefits. *Innovations in Education & teaching International, 38*(2), 156-164.

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 (Auckland, N.Z.)*, *49*(3), 371–383.

Calabuig-Moreno, F., González-Serrano, M. H., Fombona, J., & García-Tascón, M. (2020). The emergence of technology in physical education: A general bibliometric analysis with a focus on virtual and augmented reality. *Sustainability*, *12*(7), 2728.

Calandra, B., Gurvitch, R., & Lund, J. (2008). An Exploratory Study of Digital Video Editing as a Tool for Teacher Preparation, *Journal of Technology and Teacher Education, 16*(2), 137-153.

Cappella, E. (2000). What children think about computers. *The Future of Children, 10*(2), 189-191.

Carballo-Fazanes, A., Rey, E., Valentini, N. C., Rodríguez-Fernández, J. E., Varela-Casal, C., Rico-Díaz, J., ... & Abelairas-Gómez, C. (2021). Intra-Rater (Live vs. Video Assessment) and Inter-Rater (Expert vs. Novice) Reliability of the Test of Gross Motor Development—Third Edition. *International Journal of Environmental Research and Public Health*, *18*(4), 1652.

Carroll, W. R., & Bandura, A. (1982). The role of visual monitoring in observational learning of action patterns: Making the unobservable observable. *Journal of Motor Behavior*, *14*(2), 153-167.

Casey, A., Goodyear, V. A., & Armour, K. M. (2017). *Digital Technologies and Learning in Physical Education: Pedagogical Cases.* Routledge.

Casey, A., & Jones, B. (2011). Using digital technology to enhance student engagement in physical education. *Asia-Pacific Journal of Health, Sport and Physical Education*, *2*(2), 51-66.

Castelli, D. M., Centeio, E. E., Beighle, A. E., Carson, R. L., & Nicksic, H. M. (2014). Physical literacy and Comprehensive School Physical Activity Programs. *Preventive Medicine, 66*, 95–100.

Cekin, R. (2015). Psychological Benefits of Regular Physical Activity: Evidence from Emerging Adults. *Universal Journal of Educational Research*, *3*(10), 710-717.

Chai, C. S., & Koh, J. H. L. (2017). Changing teachers' TPACK and design beliefs through the Scaffolded TPACK Lesson Design Model (STLDM). *Learning: Research and Practice*, *3*(2), 114-129.

Chang, C,C., Tseng, K,H., Chou, P,N., & Chen, Y,H. (2011). Reliability and validity of web-based portfolio peer assessment: A case study for a senior high school's students taking computer course. *Computers & Education, 57*(1), 1306–1316.

Chatoupis, C. (2015). Pairing learners by companionship: effects on motor skill performance and comfort levels in the reciprocal style of teaching. *The Physical Educator*, *72*(5).

Chatoupis, C. C. (2018). Physical education teachers' use of Mosston and Ashworth's teaching styles: A literature review. *Physical Educator*, *75*(5), 880-900.

Chng, L. S., & Lund, J. (2018). Assessment for learning in physical education: The what, why and how. *Journal of Physical Education, Recreation & Dance*, *89*(8), 29-34.

Clark, J. E. (2007). On the problem of motor skill development. Journal of Physical Education, Recreation & Dance, 78(5), 39-44.

Clark, J.E., & Metcalfe, J.S. (2002). The Mountain of Motor Development: A Metaphor. In J.E. Clark & J.H. Humphrey (Eds.), *Motor Development: Research and Reviews* (Vol. 2, pp. 163–190). Reston, VA: National Association of Sport and Physical Education.

Claxton, D. B., Troy, M., & Dupree, S. (2006). A question of balance. *Journal of Physical Education, Recreation and Dance*, *77*(1), 32–37.

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.

Cohen, J. W. (1986). Theoretical considerations of peer tutoring. *Psychology in the Schools, 23*(2), 175-186.

Cohen J. W. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Hillsdale: Lawrence Erlbaum Associates.

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), 1-9.

Collier, D. (2011). Increasing the value of physical education: The role of assessment. *Journal of Physical Education, Recreation & Dance*, *82*(7), 38-41.

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.

Corlett, J., & Mandigo, J. (2013). A day in the life: Teaching physical literacy. *Physical and Health Education Journal*, *78*(4), 18.

Cross, E. S., Kraemer, D. J. M., Hamilton, A. F. de C., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral Cortex*, *19*(1), 315-326.

Crouch, D. W., Ward, P., & Patrick, C. A. (1997). The Effects of Peer-Mediated Accountability on Task Accomplishment during Volleyball Drills in Elementary Physical Education. *Journal of Teaching in Physical Education, 17*(1), 26-39.

Cummiskey, M. (2011). There's an app for that smartphone use in health and physical education. *Journal of Physical Education, Recreation & Dance, 82*(8), 24-30.

D'Arripe-Longueville, F., Gernigon, C., Huet, M. L., Winnykamen, F., & Cadopi, M. (2002). Peer-Assisted Learning in the Physical Activity Domain: Dyad Type and Gender Differences. *Journal of Sport and Exercise Psychology, 24*(3), 219-238.

Deakin, J. M., & Proteau, L. (2000). The role of scheduling in learning through observation. *Journal of Motor Behavior*, *32*(3), 268-276.

DeJong, G., Kokinakis, C. L., & Kuntzleman, C. (2002). The role of assessment in meeting the NASPE physical education content standards. *Journal of Physical Education, Recreation & Dance*, *73*(7), 22-25.

Díaz, I. A., Rodríguez, J. M. R., & García, A. M. R. (2018). La tecnología móvil de Realidad Virtual en educación: una revisión del estado de la literatura científica en España. *Edmetic*, *7*(1), 256-274.

Diemer, T. T., Fernandez, E., & Streepey, J. W. (2012). Student perceptions of classroom engagement and learning using iPads. *Journal of Teaching and Learning with Technology, 1*(2), 13-25.

Doolittle, S. (1996). Practical assessment for physical education teachers. *Journal of Physical Education, Recreation and Dance*, *67*(8), 35-37.

Duran, D., & Monereo, C. (2005). Styles and sequences of cooperative interaction in fixed and reciprocal peer tutoring. *Learning and Instruction*, *15*(3), 179-199.

Dyson, B., Griffin, L. L., & Hastie, P. (2004). Sport education, tactical games, and cooperative learning: Theoretical and pedagogical considerations. *Quest*, *56*(2), 226-240.

Eddy, L. H., Bingham, D. D., Crossley, K. L., Shahid, N. F., Ellingham-Khan, M., Otteslev, A., ... & Hill, L. J. (2020). The validity and reliability of observational assessment tools available to measure fundamental movement skills in school-age children: A systematic review. *PloS one*, *15*(8), e0237919.

Ensergueix, P. J., & Lucile, L. (2010). Reciprocal peer tutoring in a physical education setting: Influence of peer tutor training and gender on motor performance and self-efficacy outcome. *European Journal of Psychology of Education*, *25*(2), 222-242.

Ellis, L. A., Marsh, H. W., & Craven, R G. (2009). Addressing the Challenges Faced by Early Adolescents: A Mixed-Method Evaluation of the Benefits of Peer Support, *American Journal of Community Psychology*, *44*(1), 54-75.

Ernst, M. & Byra, M. (1998). Pairing learners in the reciprocal style of teaching: Influence on student skill, knowledge, and socialization. *Physical Educator*, *55*(1) 24-37.

Erwin, H. E., & Castelli, D. M. (2008). National physical education standards: A summary of student performance and its correlates. *Research Quarterly for Exercise and Sport*, *79*(4), 495–505.

Evans, G. L. (2017). Handheld Technology Use with EAL Learners in Lincolnshire Primary Schools (Doctoral dissertation, Bishop Grosseteste University).

Fencl, M. J. (2014). Fun and creative unit assessment ideas for all students in physical education. *Journal of Physical Education, Recreation & Dance*, *85*(1), 16-21.

Finkenberg, M. E., Fiorentino, L. H., & Castelli, D. (2005). Creating a virtual gymnasium. *Journal of Physical Education, Recreation & Dance*, *76*(4), 16-18.

Finkenberg, M. E., Ryan, S., Marzilli, S., & Martindale, T. (2001). Using digital cameras to assess motor learning. *Journal of Physical Education, Recreation & Dance*, *72*(8), 13-16.

Fisette, J. L., & Franck, M. D. (2012). How teachers can use PE metrics for formative assessment. *Journal of Physical Education, Recreation & Dance*, *83*(5), 23-34.

Foster, B. (2004). Video analysis of muscle motion. *Strategies*, *17*(4), 11-12.

Franklin, R., & Smith, J. (2015). Practical assessment on the run–iPads as an effective mobile and paperless tool in physical education and teaching. *Research in Learning Technology*, *23*(1), 1-19.

Fry, S.A. (1990). Implementation and evaluation of peer marking in higher education. *Assessment & Evaluation in Higher Education*, *15*(3), 177–189.

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. *Perceptual and Motor Skills*, *121*(1), 260–283.

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.

Gabbard, C. (2021). *Lifelong Motor Development*. (7th ed.). Lippincott Williams & Wilkins.

Gallahue, D., & Cleland-Donnelly, F. (2007). *Developmental Physical Education for All Children* (4th ed.). Champaign, IL: Human Kinetics.

Gallahue, D., Ozmun, J. & Goodway, J. (2012). *Understanding Motor Development: Infants, Children, Adolescents, Adults*. 7th ed. New York: McGraw-Hill.

Gallo, A. M., Sheehy, D. A., Patton, K., & Griffin, L. (2006). Assessment benefits and barriers: What are you committed to?. *Journal of Physical Education, Recreation & Dance*, *77*(8), 46-50.

Georges, F. & Pansu, P. (2011). Feedback at School: A Guarantee for Regulating School Behavior. *Revue française de pédagogie*, *3*(3), 101-124.

Giblin, S., Collins, D., & Button, C. (2014). Physical literacy: importance, assessment, and future directions. *Sports Medicine*, *44*(9) 1177-1184.

Goldberger, M., & Gerney, P. (1986) The Effects of Direct Teaching Styles on Motor Skill Acquisition of Fifth Grade Children. *Research Quarterly for Exercise and Sport*, *57*(3), 215-219.

Goldberger, M., & SueSee, B. (2020). Effects of the Reciprocal Teaching Style on skill acquisition, verbal interaction and ability to analyse in fifth grade children in physical education. In M. Goldberger and B. SueSee (eds.). *The Spectrum of Teaching Styles in Physical Education* (pp. 116-127). Routledge.

Goldberger, M., Gerney, P., & Chamberlain, J. (1982). The effects of three styles of teaching on the psychomotor performance and social skill development of fifth grade children. *Research Quarterly for Exercise and Sport, 53*(2), 116-124.

Goodway, J. & Branta, C. (2003). Influence of a motor skill intervention on fundamental movement skill development of disadvantaged preschool children. *Research Quarterly for Exercise and Sport, 74*(1), 36-46.

Goodway, J., Brian, A., Chang, S., & Park, S. (2014). Methodological approaches to evaluate motor skill development, physical activity, health-related fitness, and perceived motor competence in young children. *Handbook of Research Methods in Early Childhood Education*, *2*(1), 349-391.

Goodway, J. D., Crowe, H., & Ward, P. (2003). Effects of motor skill instruction on fundamental motor skill development. *Adapted Physical Activity Quarterly*, *20*(3), 298-314.

Goodyear, V. A., & Armour, K. M. (2018). Young people's perspectives on and experiences of health-related social media, apps, and wearable health devices. *Social Sciences*, *7*(8), 137.

Granados, C., & Wulf, G. (2007). Enhancing motor learning through dyad practice: contributions of observation and dialogue. *Research Quarterly for Exercise and Sport*, *78*(3), 197-203.

Green, N. R. (2021). Assessment in Physical Education: Charting progress through Authentic Core Tasks. *International Journal of Physical Education, Health & Sports Sciences*, *10*(1), 32-41.

Greenwood C, Terry B, Delquadri J, et al. (1995). ClassWide peer tutoring (CWPT): Effective teaching and research review. *Kansas City, KS: Juniper Gardens Children's Project*.

Guadagnoli, M., Holcomb, W., & Davis, M. (2002). The efficacy of video feedback for learning the golf swing. *Journal of Sports Sciences*, *20*(8), 615-622.

Gubacs-Collins, K., & Juniu, S. (2009). The mobile gymnasium: Using tablet PCs in physical education. *Journal of Physical Education, Recreation & Dance, 80*(2), 24-31.

Hall, T. (2012). Emplotment, embodiment, engagement: Narrative technology in support of physical education, sport and physical activity. *Quest*, *64*(2), 105-115.

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 23–34.

Hamlin, B. (2005). Motor competency and video analysis. *Teaching Elementary Physical Education*, *16*(5), 8-13.

Hancox, R. J., Milne, B. J.., & Poulton, R. (2005). Association of television viewing during childhood with poor educational achievement. *Archives of Pediatric and Adolescent Medicine*, *159*(1), 614-618.

Hardy, L. L., Reinten-Reynolds, T., Espinel, P., Zask, A., & Okely, A. D. (2012). Prevalence and correlates of low fundamental movement skill competency in children. *Pediatrics*, *130*(2), 390-398.

Harris, F. (2009). Visual technology in physical education. *Physical & Health Education Journal*, *74*(4), 24.

Hayes, S. J., Elliott, D., & Bennett, S. J. (2010). General motor representations are developed during action-observation. *Experimental Brain Research*, *204*(2), 199-206.

Haynes, J., & Miller, J. (2015). Preparing pre-service primary school teachers to assess fundamental motor skills: Two skills and two approaches. *Physical Education and Sport Pedagogy*, *20*(4), 397–408.

Haywood, K., & Getchell, N. (2019). *Lifespan Motor Development* (7th ed.). Human Kinetics.

Hazen, A., Johnstone, C., Martin, G. L., & Srikameswaran, S. (1990). A videotaping feedback package for improving skills of youth competitive swimmers. *The Sport Psychologist*, *4*(3), 213-227.

Hennings, J., Wallhead, T., & Byra, M. (2010). A Didactic Analysis of Student Content Learning during the Reciprocal Style of Teaching. *Journal of Teaching in Physical Education*, *29*(3), 227-244.

Hensley, L. D. (1997). Alternative assessment for physical education. *Journal of Physical Education, Recreation & Dance*, *68*(7), 19-24.

Higgins, S., Xiao, Z., & Katsipataki, M. (2012). The Impact of Digital Technology on Learning: A Summary for the Education Endowment Foundation. Full Report. *Education Endowment Foundation*.

Higgs, C., (2010). Physical literacy – two approaches, one concept. *Physical & Health Education Journal*, *76*(1), 6–7.

Higgs, C., Balyi, I., Way, R., Cardinal, C., Norris, S., & Bluechardt, M. (2008). Development physical literacy: A guide for parents of children ages 0 to 12. Retrieved from Vancouver, B.C.

Hill, G., & Miller, T. (1997). A comparison of peer and teacher assessment of students' physical fitness performance. *Physical Educator*, *54*(1), 40-46.

Hodges, N. J., Chua, R., & Franks, I. M. (2003). The role of video in facilitating perception and action of a novel coordination movement. *Journal of Motor Behaviour*, *35*(3), 247-260.

Hodges, N. J., Williams, A. M., Hayes, S. J., & Breslin, G. (2007). What is modelled during observational learning? *Journal of Sports Sciences*, *25*(5), 531-545.

Horak, F. B. (2006). Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls?. *Age and Ageing*, *35*(Suppl 2), ii7-ii11.

Houston-Wilson, C., Dunn, J.M., van der Mars, H., & McCubbin, J. (1997). The effect of peer tutors on motor performance in integrated physical education classes. *Adapted Physical Activity Quarterly*, *14*(4), 298-313.

Hung, H. C., Young, S. S. C., & Lin, K. C. (2018). Exploring the effects of integrating the iPad to improve students' motivation and badminton skills: A WISER model for physical education. *Technology, Pedagogy, and Education*, *27*(3) 265-278.

Illing, H. A. (1965). " A Social Psychology of Group Processes for Decision-Making," by Barry E. Collins and Harold Guetzkow (Book Review). *International Journal of Group Psychotherapy*, *15*(1), 405-406.

International Physical Literacy Association: Canada's Physical Literacy Consensus Statement (2015, June).

Jenkinson, K. A., & Benson, A. C. (2010). Barriers to Providing Physical Education and Physical Activity in Victoria State Secondary Schools. *Australian Journal of Teacher Education, 35*(8), 1-17.

Jenkinson, K.A., Naughton, G., & Benson, A.C. (2014). Peer-assisted learning in school physical education, sport and physical activity programmes: A systematic review. *Physical Education and Sport Pedagogy*, *19*(3), 253-277.

Jennings, C. T., Reaburn, P., & Rynne, S. B. (2013). The effect of a self-modelling video intervention on motor skill acquisition and retention of a novice track cyclist's standing start performance. *International Journal of Sports Science and Coaching*, *8*(3), 467-480.

Johnson, R. (2004). Peer assessments in physical education. *Journal of Physical Education, Recreation, and Dance*, *75*(8), 33-40.

Johnson, M., & Ward, P. (2001). Effects of classwide peer tutoring on correct performance of striking skills in 3rd grade physical education. *Journal of Teaching in Physical Education*, *20*(3), 247-263.

Katz, L. (2003). Multimedia and the internet for sport sciences: Applications and innovations. *International Journal of Computer Science in Sport*, *2*(1), 4-18.

Kelly, P., & Katz, L. (2016). Comparing peer-to-peer and individual learning: Teaching basic computer skills to disadvantaged adults. *International Journal of Adult Vocational Education and Technology (IJAVET)*, *7*(4), 1-15.

Kieth, D. (1980). *Soccer Handbook*. Vancouver, Canada: Hancock House.

Kirk, D. (2013). Educational value and models-based practice in physical education. *Educational Philosophy and Theory*, *45*, 973-986.

Kniffin, K. M., & baert, H. (2015). Maximizing learning through assessment in middle and high school physical education. *Journal of Physical Education, Recreation, and Dance*, *86*(4), 7-16.

Koehler, M. J., Mishra, P., & Cain, W. (2013). What is technological pedagogical content knowledge (TPACK)?. *Journal of Education*, *193*(3), 13-19.

Koekoek, J., & Knoppers, A. (2015). The role of perceptions of friendships and peers in learning skills in physical education. *Physical Education and Sport Pedagogy*, *20*(3), 231-249.

Koekoek, J., Knoppers, A., & Stegeman, H. (2009). How do Children Think They Learn Skills in Physical Education?. *Journal of Teaching in Physical Education, 28*(3), 310-332.

Koekoek, J., & Van Hilvoorde, I. (2018). *Digital technology in physical education: Global perspectives*. (1st ed.). Routledge.

Koekoek, J., Van Der Mars, H., van der Kamp, J., Walinga, W., & van Hilvoorde, I. (2018). Aligning digital video technology with game pedagogy in physical education. *Journal of Physical Education, Recreation & Dance*, *89*(1), 12-22.

Kolovelonis, A., & Goudas, M. (2012). Students' recording accuracy in the reciprocal and the self-check teaching styles in physical education. *Educational Research and Evaluation*, *18*(8), 733-747.

Kolovelonis, A., Goudas, M., & Gerodimos V. (2011). The effects of the reciprocal and the selfcheck styles on pupils' performance in primary physical education. *European Physical Education Review*, *17*(1) 35-50.

Krause, J., & Sanchez, Y. (2014). Meeting the national standards: There's an app for that! *Strategies*, *27*(4), 3-12.

Kretschmann, R. (2010). Physical Education 2.0. In M. Ebner & M. Schiefner (Eds.), *Looking Toward the Future of Technology-Enhanced Education: Ubiquitous Learning and the Digital Native* (pp.432-454). Hershey, PA: IGI Publishing.

Kretschmann, R. (2012), What do Physical Education Teachers Think about Integrating Technology in Physical Education? European Journal of Social Sciences, 27(3), 444-448.

Kretschmann, R. (2015). Physical Education Teachers' Subjective Theories about Integrating Information and Communication Technology (ICT) into Physical Education. *Turkish Online Journal of Educational Technology-TOJET*, *14*(1), 68-96.

Kretschmann, R. (2017). Employing tablet technology for video feedback in physical education swimming class. *Journal of e-Learning and Knowledge Society*, *13*(2), 103-115.

Lacy, A. C., & Williams, S. M. (2018). *Measurement and Evaluation in Physical Education and Exercise science*. (8th ed.). Routledge.

Lago-Rodriguez, A., Cheeran, B., Koch, G., Hortobagyi, T., & Fernandez-del-Olmo, M. (2014). The role of mirror neurons in observational learning: An integrative review. *European Journal of Human Movement*, *32*, 82-103.

Lee, H. S. (2016). *A Comparison of an Eastern Philosophical Approach to Human Movement with a Western Approach through Teaching Overhand Throwing Skills with Grade Two Students (Unpublished Doctoral Thesis)* [University of Calgary].

Lee, J. E., & Gao, Z. (2020). Effects of the iPad and mobile application-integrated physical education on children's physical activity and psychosocial beliefs. *Physical Education and Sport Pedagogy*, *25*(6), 567-584.

Lee, T. D., Swinnen, S. P., & Verschueren, S. (1995). Relative phase alterations during bimanual skill acquisition. *Journal of Motor Behaviour*, *27*(3), 263-274.

Lee, T. D., & White, M. A. (1990). Influence of an unskilled model's practice schedule on observational motor learning. *Human Movement Science*, *9*(3-5), 349-367.

Leight, J., Banville, D., & Polifko, M.F. (2009). Using digital video recorders in physical education. *Journal of Physical Education, Recreation & Dance, 80*(1), 17-21.

Leser, R., Baca, A., & Uhlig, J. (2011). Effectiveness of multimedia-supported education in practical sports courses. *Journal of Sports Science and Medicine, 10*(1), 184-192.

Lewis, S. F., & Hennekens, C. H. (2016). Regular physical activity: forgotten benefits. *The American journal of medicine*, *129*(2), 137-138.

Lhuisset, L., & Margnes, E. (2015). The influence of live-vs. video-model presentation on the early acquisition of a new complex coordination. *Physical Education and Sport Pedagogy, 20*(5), 490-502.

Lieberman, L.J., Dunn, J.M., van der Mars, H., & McCubbin, J. (2000). Peer tutors' effects on activity levels of deaf students in inclusive elementary physical education. *Adapted Physical Activity Quarterly,17*(1), 20-39.

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

Lindblom-ylänne, S., Pihlajamäki, H., & Kotkas, T. (2006). Self-, peer-and teacher-assessment of student essays. *Active Learning in Higher Education*, *7*(1), 51-62.

Logan, S., Robinson, L., Wilson, A., & Lucas, W. (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.

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.

Lund, J. (1997). Authentic assessment: Its development & applications. *Journal of Physical Education, Recreation & Dance, 68*(7), 25-28.

Lund, J., & Tannehill, D. (2014). *Standards-Based Physical Education Curriculum Development*. (3rd ed.). Jones & Bartlett Publishers.

Lund, J., & Veal, M. L. (2013). Assessment-driven instruction in physical education: A standards-based approach to promoting and documenting learning. Champaign, IL: Human Kinetics.

Macphail, A., Kirk, D., & Griffin, L. (2008). Throwing and Catching as Relational Skills in Game Play: Situated Learning in a Modified Game Unit. *Journal of Teaching in Physical Education*, *27*(1), 100-115.

Madou, T., & Iserbyt, P. (2018). Effect of Pairing by Ability on Performance, Physical Activity, and Time-on-Task During Reciprocal Peer Teaching in Swimming. *Physical Educator*, *75*(5), 756-773.

Maeng, H., Webster, E.K., Pitchford, E., & Ulrich, D. (2017). Inter- and Intrarater Reliabilities of the Test of Gross Motor Development—Third Edition among Experienced TGMD-2 Raters. *Adapted Physical Activity Quarterly, 34*(4), 442–455.

Magill, R., & Anderson, D. (2014). *Motor Learning and Control: Concepts and Applications* (10th ed.). McGraw-Hill.

Magin, D.J. (2001). A novel technique for comparing the reliability of multiple peer assessments with that of single teacher assessments of group process work. *Assessment & Evaluation in Higher Education, 26*(2), 139–152.

Malina, R. M., Bouchard, C., & Bar-Or,O. (2004). *Growth, Maturation, and Physical Activity* (2nd ed.). Champaign, IL: Human Kinetics.

Mandigo, J., Francis, N., Lodewyk, K., & Lopez, R. (2009). Physical Literacy for Educators. *Physical & Health Education Journal*, *75*(3), 27–30.

Martens, R. (1996). Turning kids on to physical activity for a lifetime. *Quest*, *48*(3), 303-310.

Martin, J. J., Kulinna, P. H., & Cothran, D. (2002). Motivating students through assessment. *Journal of Physical Education, Recreation & Dance*, *73*(8), 18-19.

Martin, E. H., Rudisill, M. E., & Hastie, P. A. (2009). Motivational climate and fundamental movement skill performance in a naturalistic physical education setting. *Physical Education and Sport Pedagogy, 14*(3), 227-240.

Mas, F. G., Plass, J., Kane, W. M., & Papenfuss, R. L. (2003). Health education and multimedia learning: Connecting theory and practice (part 2). *Health Promotion Practice, 4*(4), 464-469.

Mason, R., & Rennie, F. (2006). *E-learning: The Key Concepts.* Routledge.

Mayer, R. (1997). Multimedia learning: Are we asking the right questions? *Educational Psychologist, 32*(1), 1-19.

McCullagh, P., & Meyer, K. N. (1997). Learning versus correct models: Influence of model type on the learning of a free-weight squat lift. *Research Quarterly for Exercise and Sport*, *68*(1), 56-61.

McCullagh, P., & Weiss, M. (2001). Modelling: Considerations for motor skill performance and psychological responses. *Handbook of Sport Psychology*, 205-238.

McCullagh, P., Weiss, M. R., & Ross, D. (1989). Modeling considerations in motor skill acquisition and performance: An integrated approach. *Exercise and Sport Sciences Reviews, 17*, 475-513.

McDermid, F., Peters, K., Jackson, D., & Daly, J. (2014). Conducting qualitative research in the context of pre-existing peer and collegial relationships. *Nurse Researcher*, *21*(5), 28-33.

McFarlane, C. (2013). iPads and their potential to revolutionize learning. In *EdMedia+ Innovate Learning* (pp. 1690-1695). Association for the Advancement of Computing in Education (AACE).

McKenzie, T. L. (2010). Seeing is believing: Observing physical activity and its contexts. *Research Quarterly for Exercise and Sport*, *81*(2), 113-122.

McManis, L. D., & Gunnewig, S. B. (2012). Finding the education in educational technology with early learners. *Young Children*, *67*(3), 14-24.

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational Horizons*, *87*(1), 40-49.

McNicol, S., Lewin, C., Keune, A., & Toikkanen, T (2014). *Facilitating Student Reflection Through Digital Technologies in the iTEC Project: Pedagogically Led Change in the Classroom.* First International Conference, LCT 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part II.

Melhuish, K., & Falloon, G. (2010). Looking into the future: M-Learning with the iPad, *Computers in New Zealand Schools: Learning, Leading, Technology*, *22*(3), 1-16.

Melograno, V. J. (1997). Integrating assessment into physical education teaching. *Journal of Physical Education, Recreation & Dance*, *68*(7), 34-37.

Metzler, M. (2017). *Instructional models in physical education* (3rd ed.). Routledge.

Michaelsen, L. K., Watson, W. E., & Black, R. H. (1989). A realistic test of individual versus group consensus decision making. *Journal of Applied Psychology*, *74*(5), 834–839.

Minick, K. I., Kiesel, K. B., Burton, L., Taylor, A., Plisky, P., & Butler, R. J. (2010). Interrater reliability of the functional movement screen. *Journal of Strength and Conditioning Research*, *24*(2), 479–486.

Mohnsen, B. S. (2012). *Using Technology in Physical Education*. Bonnie's Fitware, Inc.

Mohr, D. J., & Townsend, J. S. (2002). Research to Practice: Using Comprehensive Teaching Models To Enhance Pedagogical Content Knowledge. *Teaching Elementary Physical Education*, *13*(4), 37-36.

Morey, R.S., & Karp, G. G. (1998). Why do some students who are good at physical education. *Physical Educator*, *55*(2), 89.

Morgan, K., Kingston, K., & Sproule, J. (2005). Effects of different teaching styles on the teacher behaviours that influence motivational climate and pupils' motivation in physical education. *European Physical Education Review*, *11*(3), 257-285.

Mosston, M. & Ashworth, S. (2008). *Teaching Physical Education*. First Online Edition. https://spectrumofteachingstyles.org/assets/files/book/Teaching_Physical_Edu_1st_Online.pdf

Mueller, D., Georges, A., & Vaslow, D. (2007). Cooperative learning as applied to resident instruction in radiology reporting. *Academic Radiology*, *14*(12), 1577-1583.

Mulder, T., & Hulstijn, W. (1985). Sensory feedback in the learning of a novel motor task. *Journal of Motor Behavior*, *17*(1), 110-128.

Mulvihill, C., Rivers, K., & Aggleton, P. (2000). A qualitative study investigating the views of primary-age children and parents on physical activity. *Health Education Journal*, *59*(2), 166-179.

Nadeau, L., Richard, J.F., & Godbout, P. (2008). The validity and reliability of a performance assessment procedure in ice hockey. *Physical Education & Sport Pedagogy*, *13*(1), 65–83.

Ng, S. K. (2002). A discussion on role conflict between Hong Kong physical education teachers and coaches. *Journal of Physical Education and Recreation (Hong Kong)*, *8*(1), 59-63.

Norris, C., Soloway, E., & Sullivan, T. (2002). Examining 25 years of technology in US education. *Communications of the ACM*, *45*(8), 15-18.

Nupponen, H. (1997). Motor development of 9–16-year-olds. *Research Reports on Sport and Health, 106.*

Nupponen, H., & Telama, R. (1998). Physical activity and motor performance as part of the lifestyle of 11-16-year-old European young people. *Research Reports on Sport and Health, 1.*

Nupponen, H., Soini, H., & Telama, R. (1999). Measurement of school students' physical fitness and motor skills. *Research Reports on Sport and Health, 118.*

O'Brien, W., Belton, S., & Issartel, J. (2016). Fundamental movement skill proficiency amongst adolescent youth. *Physical Education and Sport Pedagogy*, *21*(6), 557-571.

Okely, A. D., & Booth, M. L. (2004). Mastery of fundamental movement skills among children in New South Wales: Prevalence and sociodemographic distribution. *Journal of Science and Medicine in Sport*, 7, 358–372.

Okely, A. D., Booth, M. L., & Patterson, J. W. (2001) Relationship of physical activity to fundamental movement skill among adolescents. *Medicine and Science in Sports and Exercise, 33*(11), 1899-1904.

O'Loughlin, J., Chróinín, D. N., & O'Grady, D. (2013). Digital video: The impact on children's learning experiences in primary physical education. *European Physical Education Review*, *19*(2), 165–182.

Oñate, J. A., Guskiewicz, K. M., Marshall, S. W., Giuliani, C., Yu, B., & Garrett, W. E. (2005). Instruction of jump-landing technique using videotape feedback: altering lower extremity motion patterns. *The American Journal of Sports Medicine*, *33*(6), 831-842.

Palao, J. M., Hastie, P. A., Cruz, P. G., & Ortega, E. (2015). The impact of video technology on student performance in physical education. *Technology, Pedagogy and Education*, *24*(1), 51–63.

Palmer, K. K., & Brian, A. (2016). Test of Gross Motor Development-2 scores differ between expert and novice coders. *Journal of Motor Learning and Development*, *4*(2), 142–151.

Panicucci, J., & Hunt, L. F. (2002). *Adventure Curriculum for Physical Education: Middle School*. Project Adventure.

Papastergiou, M. (2009). Exploring the potential of computer and video games for health and physical education: A literature review. *Computers and Education,* 53(3), 603–622.

Parsons, J. L., & Alexander, M. J. (2012). Modifying spike jump landing biomechanics in female adolescent volleyball athletes using video and verbal feedback. *The Journal of Strength & Conditioning Research*, *26*(4), 1076-1084.

ParticipACTION. (2018). The brain + body equation: Canadian kids need active bodies to build their best brains. *The 2018 ParticipACTION Report Card on Physical Activity for Children and Youth.*

Pasco, D. (2013). The potential of using virtual reality technology in physical activity settings. *Quest*, *65*(4), 429-441.

Payne, G., & Isaacs, L. D. (2020). *Human Motor Development: A Lifespan Approach*. Routledge.

Physical and Health Education (PHE) Canada. (2013). Physical literacy resources. Retrieved from http://www.phecanada.ca/programs/physical-literacy

Ploetzner, R., Dillenbourg, P., Preier, M., & Traum, D. (1999). Learning by Explaining to Oneself and to Others. In: Dillenbourg P (ed) *Collaborative learning: Cognitive and Computational Approaches,*103-121.

Portney, L. G. (2020). *Foundations of Clinical Research: Applications to Evidence-Based Practice*. (4th ed.). Philadelphia: F. A. Davis.

Potdevin, F., Bernaert, F., Huchez, A., & Vors, O. (2013). Le feedback vidéo en EPS: une double stratégie de progrès et de motivations. Le cas de l'Appui Tendu Renversé en classe de 6ème. *eJRIEPS. Ejournal de la recherche sur l'intervention en éducation physique et sport*, (30).

Potdevin, F., Vors, O., Huchez, A., Lamour, M., Davids, K., & Schnitzler, C. (2018). How can video feedback be used in physical education to support novice learning in gymnastics? Effects on motor learning, self-assessment and motivation. *Physical Education and Sport Pedagogy*, *23*(6), 559-574.

Pountney, R., & Schimmel, H. (2015). Developing professional knowledge and expertise in educational technology: legacy, change and investment. *The Journal of Technology Enhanced Learning, Innovation & Change, 1*(1).

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon, 9*(5).

Pritchard, T., Hawkins, A., Wiegand, R., & N. Metzler, J. (2008). Effects of Two Instructional Approaches on Skill Development, Knowledge, and Game Performance. *Measurement in Physical Education and Exercise Science*, 12(4), 219-236.

Pyle, B., & Esslinger, K. (2014). Utilizing technology in physical education: Addressing the obstacles of integration. *Delta Kappa Gamma Bulletin*, *80*(2), 35.

Reeves, B. (1986). Assessment in physical education. *Bulletin of Physical Education, 22*(1), 29-31.

Reychav, I., & Wu, D. (2015). Mobile collaborative learning: The role of individual learning in groups through text and video content delivery in tablets. *Computers in Human Behavior*, *50*(1), 520-534.

Rezaei, A. R. & Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing, 15*(1), 18–39.

Richard, J. F., Godbout, P., Tousignant, M., & Gréhaigne, J. F. (1999). The try-out of a team sport performance assessment procedure in elementary and junior high school physical education classes. *Journal of Teaching in Physical Education*, *18*(3), 336-356.

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.

Rizzolatti, G., & Fogassi, L. (2014). The mirror mechanism: Recent findings and perspectives. *Philosophical Transactions of the Royal Society B, 369*(1644).

Roberton, M. A. (1977). Stability of stage categorizations across trials: implications for the “stage theory” of overarm throw development. *Journal of Human Movement Studies, 3*(1), 49–59.

Roberton, M. A., & Halverson, L. E. (1984). *Developing Children—Their Changing Movement.* Philadelphia: Lea & Febiger.

Robinson, T. N. (2001). Television viewing and childhood obesity. *Pediatric Clinics of North America*, *48*(4), 1017-1025.

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.

Roblyer, M. D., & Doering, A. H. (2013). *Integrating Educational Technology into Teaching* (6th ed.). Pearson.

Roetert, E. P., & Jefferies, S. C. (2014). Embracing physical literacy. *Journal of Physical Education, Recreation and Dance, 85*(8), 38-40.

Rohbanford, H., & Proteau, L. (2011). Learning through observation: A combination of expert and novice models favors learning. *Experimental Brain Research*, *215*(3-4), 183-197.

Rohrbeck C.A., Ginsburg-Block M.D., Fantuzzo J.W., & Miller, T.R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95*(2), 240-257.

Roychowdhury, D. (2020). Using physical activity to enhance health outcomes across the life span. *Journal of Functional Morphology and Kinesiology*, *5*(1), 2.

Rucci, J. A., & Tomporowski, P. D. (2010). Three types of kinematic feedback and the execution of the hang power clean. *The Journal of Strength & Conditioning Research*, *24*(3), 771-778.

Sackstein, S. (2017). *Peer Feedback in the Classroom: Empowering Students to Be the Experts*. ASCD. Retrieved from https://books.google.ca/books?id=DrT6MAAACAAJ

Sallis, J. F., & McKenzie, T. L. (1991). Physical education's role in public health. *Research Quarterly for Exercise and Sport*, *62*(2), 124-137.

Sallis, J.F. & Patrick, K. (1994). Physical Activity Guidelines for Adolescents: Consensus Statement. *Pediatric Exercise Science, 6*(4), 302-314.

San-Tan, S. K., & Wright, S. (2004). Authentic Assessment to Enhance Student Learning in Primary School Physical Education. *Asian Journal of Physical Education & Recreation*, *10*(2), 36-44.

Schmidt, S. R. (1988). Test expectancy and individual-item versus relational processing. *The American Journal of Psychology, 101*(1), 59-71.

Schmidt, R. A., & Lee, T. D. (2005). *Motor Control and Learning: A Behavioral Emphasis* (4th ed.). Human Kinetics.

Schmidt, R. A., & Wrisberg, C. A. (2008). *Motor Learning and Performance: A Situation-Based Learning Approach*. Human Kinetics.

Scully, D. M., & Newell, K. M. (1985). Observational-learning and the acquisition of motor-skills-toward a visual-perception perspective. *Journal of Human Movement Studies*, *11*(4), 169-186.

Siedentop, D., & Tannehill, D. (2000). *Developing Teaching Skills in Physical Education* (4th ed.). New York, NY: McGraw-Hill.

Sigrist, R., Rauter, G., Riener, R., & Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic Bulletin & Review*, *20*(1), 21-53.

Sinelnikov, O. A. (2012). Using the iPad in a sport education season. *Journal of Physical Education, Recreation & Dance*, *83*(1), 39-45.

Shea, C. H., Wright, D. L., Wulf, G., & Whitacre, C. (2000). Physical and observational practice afford unique learning opportunities. *Journal of Motor Behaviour*, *32*(1), 27-36.

Shea, C. H., Wulf, G., & Whltacre, C. (1999). Enhancing training efficiency and effectiveness through the use of dyad training. *Journal of Motor Behavior*, *31*(2), 119-125.

Shebilske, W. L., Regian, J. W., Arthur, W., & Jordan, J. A. (1992). A dyadic protocol for training complex skills. *Human Factors, 34*(3), 369–374.

Sheehan, D. P., Lafave, M. R., & Katz, L. (2011). Intra-rater and inter-rater reliability of the balance error scoring system in pre-adolescent school children. *Measurement in Physical Education and Exercise Science*, *15*(3), 234–243.

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

Slotte, S., Sääkslahti, A., Metsämuuronen, J., & Rintala, P. (2015). Fundamental movement skill proficiency and body composition measured by dual energy X-ray absorptiometry in eight-year-old children. *Early Child Development and Care*, *185*(3), 475-485.

Smith, T. K. (1997). Authentic assessment: Using a portfolio card in physical education. *Journal of Physical Education, Recreation & Dance*, *68*(4), 46-52.

Smith, R. M., & Loschner, C. (2002). Biomechanics feedback for rowing. *Journal of Sports Sciences*, *20*(10), 783-791.

Spittle, M. (2021). *Motor learning and skill acquisition: Applications for physical education and sport*. (2nd ed.). Palgrave Macmillan Australia.

Sport for Life Society. (2021). Sport for life. Retrieved from http://www.canadiansportforlife.ca

Stahl G. (2006). Group cognition: Computer support for building collaborative knowledge, Cambridge, MA: MIT Press.

Stanton, R., Happell, B., & Reaburn, P. (2014). The mental health benefits of regular physical activity, and its role in preventing future depressive illness. *Nursing: Research and Reviews*, *4*, 45-53.

Ste-Marie, D. M., Law, B., Rymal, A. M., Jenny, O., Hall, C., & McCullagh, P. (2012). Observation interventions for motor skill learning and performance: An applied model for the use of observation. *International Review of Sport and Exercise Psychology*, *5*(2), 145-176.

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.

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.

Sun, H. (2012). Exergaming impact on physical activity and interest in elementary school children. *Research Quarterly for Exercise and Sport*, *83*(2), 212-220.

Sun, H. (2015). Operationalizing physical literacy: The potential of active video games. *Journal of Sport and Health Science*, *4*(2), 145-149.

Swinnen, S. P. (1996). Information Feedback for Motor Skill Learning: A Review. In H. N. Zelaznik (Ed.), *Advances in Motor Learning and Control* (pp. 37-66). Champaign, IL: Human Kinetics.

Sykes, H. & McPhail, D. (2008). Unbearable lessons: Contesting fat phobia in physical education. *Sociology of Sport Journal*, *25*(1), 66-96.

Tan, O. S. (2021). *Problem-Based Learning Innovation: Using Problems to Power Learning in the 21st Century*. Gale Cengage Learning.

Tanaka, Y., Murakami, S., Kakoi, C., Wada, T. & Takahashi, H. (2014). Practice of Physical Education Classes Utilizing Video. In L. Gómez Chova, A. López Martínez & I. Candel Torres (Eds.), INTED 2014 Proceedings (pp. 725-729). IATED Academy.

Tanaka, C., Tanaka, M., & Tanaka, S. (2018). Objectively evaluated physical activity and sedentary time in primary school children by gender, grade and types of physical education lessons. *BMC Public Health*, *18*(1), 948-958.

Tannehill, D., McPhail, 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.

Tay, H. Y. (2016). Longitudinal study on impact of iPad use on teaching and learning. *Cogent Education*, *3*(1).

Taylor, I. M., Ntoumanis, N., & Smith, B. (2009). The social context as a determinant of teacher motivational strategies in physical education. *Psychology of Sport and Exercise*, *10*(2), 235-243.

Thomas, A., & Stratton, G. (2006). What we are really doing with ICT in physical education: a national audit of equipment, use, teacher attitudes, support, and training. *British Journal of Educational Technology*, *37*(4), 617-632.

Thow, J. L., Naemi, R., & Sanders, R. H. (2012). Comparison of modes of feedback on glide performance in swimming. *Journal of Sports Sciences*, *30*(1), 43-52.

Tinto V. (2003). Learning better together: The impact of learning communities on student success. *Higher Education Monograph Series*, *1*(8), 1-8.

Topping, K. J. (2005). Trends in peer learning. *Educational psychology*, *25*(6), 631-645.

Topping, K., & Ehly, S. (1998). *Peer-Assisted Learning*. Taylor & Francis. https://books.google.ca/books?id=i76aRj-xC2cC

Tremblay, M., & Lloyd, M. (2010). Physical literacy measurement: The missing piece. *Physical and Health Education Journal*, *76*(1), 26-30.

Trout, J. (2013). Digital movement analysis in physical education. *Journal of Physical Education, Recreation & Dance*, *84*(7), 47-50.

Uhl, B., & Dillon, S. (2009, March). Dartfish video analysis in secondary physical education: A pilot study. In *Poster Session at the American Alliance for Health, Physical Education, Recreation, and Dance Annual Convention, Tampa* (Vol. 31).

Ulrich, D.A. (2000). *Test of Gross Motor Development*: 2nd Edition. Austin, Texas: Pro-ed.

Ulrich, D.A., & Sanford, C. B. (2000) *Test of Gross Motor Development Examiner's Manual Austin.* TX: Pro-ed.

Van den Bossche, P., Gijselaers, W. H., Segers, M., & Kirschner, P. A. (2006). Social and cognitive factors driving teamwork in collaborative learning environments: Team learning beliefs and behaviors. *Small Group Research*, *37*(5), 490-521.

Vandewater, E. A., Bickham, D. S., & Lee, J. H. (2006). Time well spent? Relating television use to children's free-time activities. *Pediatrics*, *117*(2), e181-e191.

Vandewater, E. A., Bickham, D. S., Lee, J. H., Cummings, H. M., Wartella, E. A., & Rideout, V. J. (2005). When the television is always on: Heavy television exposure and young children's development. *American Behavioural Scientist*, *48*(5), 562-577.

Veal, M. L. (1988). Pupil assessment issues: A teacher educator's perspective. *Quest*, *40*(2), 151-161.

Veal, M. L. (1992). The role of assessment in secondary physical education—A pedagogical view. *Journal of Physical Education, Recreation & Dance*, *63*(7), 88-92.

Veal, M. L. (1995). Assessment as an instructional tool. *A Journal for Physical and Sport Educators*, *8*(6), 10-15.

Vendlinksi, T., & Stevens, R. (2002). Assessing student problem-solving skills with complex computer-based tasks. *The Journal of Technology, Learning and Assessment*, *1*(3). Available from http://www.jtla.org.

Vogt, S., & Thomaschke, R. (2007). From visuo-motor interactions to imitation learning: Behavioural and brain imaging studies. *Journal of Sports Sciences*, *25*(5), 497-517.

Warburton, D. E., Nicol, C. W., & Bredin, S. S. (2006). Health benefits of physical activity: the evidence. *Cmaj*, *174*(6), 801-809.

Ward, P., & Lee, M. A. (2005). Peer-assisted learning in physical education: A review of theory and research. *Journal of Teaching in Physical Education*, *24*(3), 205–225.

Ward, P., Smith, S. L., Makasci, K., & Crouch, D. W. (1998). Differential effects of peer-mediated accountability on task accomplishment in elementary physical education. *Journal of Teaching in Physical Education*, *17*(4), 442-452.

Watterson, T. (2012). *Changes in attitudes and behaviours toward physical activity, nutrition, and social support for middle school students using AFIT app as a supplement to instruction in a physical education class* [Unpublished doctoral dissertation]. University of South Florida.

Weir, T. & Connor, S. (2009). The use of digital video in physical education. *Technology, Pedagogy and Education*, *18*(2), 155-171.

Weir, J. P., & Vincent, W. J. (2020). *Statistics in Kinesiology*. (5th ed.). Human Kinetics Publishers.

Weiss, M. R., & Stuntz, C. P. (2004). A Little Friendly Competition: Peer Relationships and Psychosocial Development in Youth Sport and Physical Activity Contexts. In M. R. Weiss (Ed.), *Developmental Sport and Exercise Psychology: A Lifespan Perspective* (pp.165-196). Fitness Information Technology.

Westcott, S. L., Lowes, L. P., & Richardson, P. K. (1997). Evaluation of postural stability in children: Current theories and assessment tools. *Physical Therapy, 77*(6), 629-645.

Whitehead, M. (Ed.) (2010). *Physical Literacy: Throughout the Lifecourse*. New York, NY, Routeledge.

Williams, A. M., Davids, K., & Williams, J. G. P. (1999). *Visual Perception and Action in Sport*. Taylor & Francis.

Wilson, B. D. (2008). Development in video technology for coaching. *Sports Technology*, *1*(1), 34-40.

Winfrey, M. L., & Weeks, D. L. (1993). Effects of self-modeling on self-efficacy and balance beam performance. *Perceptual and Motor Skills*, *77*(3), 907-913.

Woods, M., Karp, G., Miao, H., & Perlman, D. (2008). Physical educators' technology competencies and usage. *Physical Educator, 65*(2), 82-99.

Wright, M. T., & van der Mars, H. (2004). Blending assessment into instruction: Practical applications and meaningful results. *Journal of Physical Education, Recreation & Dance*, *75*(9), 29-34.

Wrisberg, C. A. (2007). *Sport Skill Instruction for Coaches*. Human Kinetics.

Wulf, G. (2013). Attentional focus and motor learning: a review of 15 years. *International Review of Sport and Exercise Psychology*, *6*(1), 77-104.

Wulf, G., & Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, *9*(2), 185-211.

Wulf, G., Shea, C., & Lewthwaite, R. (2010). Motor skill learning and performance: a review of influential factors. *Medical Education*, *44*(1), 75–84.

Wyant, J., & Baek, J. H. (2019). Re-thinking technology adoption in physical education. *Curriculum Studies in Health and Physical Education*, *10*(1), 3-17.

Yelland, N., & Masters, J. (2007). Rethinking scaffolding in the information age. *Computers & Education*, *48*(3), 362-382.

Yoncalik, O., Yetim, A. A., & Senel, Ö. (2010). Effects of teaching with Mosston's command, practice, and reciprocal styles on affective reactions of sixth-grade students toward physical education lessons. *International Journal of Educational Reform, 18*(4), 311-326.

Zhang, T., & Li, H. (2018). Digital video and self-modeling in the PE classroom. In T. Zhang and H. Li (eds). *Digital Technology in Physical Education* (pp. 19-31). Routledge.

Zhang, K., & Liu, S. J. (2016, July). The application of virtual reality technology in physical education teaching and training. In *2016 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI)* (pp. 245-248). IEEE.

Zhu, X., & Dragon, L. A. (2016). Physical activity and situational interest in mobile technology integrated physical education: A preliminary study. *Acta Gymnica, 46*(2). 59-67.

Ziegler, S. G. (1994). The effects of attentional shift training on the execution of soccer skills: A preliminary investigation. *Journal of Applied Behavior Analysis, 27*(3), 545-552.

Živčić-Marković, K., Krističević, T., & Aleksić-Veljković, A. (2015). A suggested model of handstand teaching method. *Physical Culture, 69*(2), 138-149.

Appendix A: Extended Methodology

MI is a video performance analysis tool designed to evaluate an individual's ability to perform a variety of movement skills, by breaking down physical skills into easily comprehended components. The MI application is an assessment tool that works on smartphones and tablets. The primary purpose of this investigation was to explore the impact, if any, that order of performance within a P2P learning model has on the acquisition of FMS in middle school students. A secondary objective involved examining the inter-rater and intra-rater reliability of this assessment tool with regards to the FMS performance outcomes of middle school students.

Participants

Sampling Technique and Recruitment

The study utilized a purposive sampling method to acquire the middle school students from a small private parochial school in Calgary, Alberta, Canada.

Characteristics of Sample

A total of 18 middle school students were recruited into the study, ranging from Grades 5, 6, 7, and 8. There were slightly more females (n = 10) than males (n = 8), with most students being registered in Grade 6 (n = 8). In accordance with the inclusion and exclusion criteria, all participants in the study were able bodied, without any physical impairments or disabilities. The full breakdown for the number of participants in participants in each grade level, as well as their gender is presented in Table A.1.

Table A.1 Participant Breakdown by Grade Level and Gender

Grade Level	# of participants	# of males	# of females
5	3	2	1
6	8	4	4
7	4	1	3
8	3	1	2
Total	18	8	10

Inclusion Criteria

Children were included in the study if they were able-bodied and currently a student in Grades 5, 6, 7, or 8. Inclusion for participation was also determined by the obtainment of parental consent. Since the students used the tablets and MI program as part of their physical education class, assent was not required. Although all the participants willingly participated in the study, data was only collected and analyzed from those students from which parental consent had been granted.

Exclusion Criteria

Children were excluded from the study if they were not in Grades 5, 6, 7 or 8. Additionally, students who had injuries or chronic health issues that prevented then from completing the assigned task without pain or discomfort were excluded from the investigation. Due to the nature of the study being a pilot, there was not the ability to readily adapt the skills to meet the needs required by individuals who may have a physical disability. Lastly, if parents did not agree to sign the consent form (Appendix B), the child's information and corresponding data was excluded from the study.

Investigators

The investigators of the study included Anna Thacker, a graduate student at the University of Calgary, and her supervisor, Dr. Larry Katz.

Procedure

Ethics

Ethics was obtained and approved through the University of Calgary CHREB (Appendix C). The lead researcher in this study completed the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2) tutorial (Appendix D). The parents and/or guardians of the study participants signed an informed consent form prior to the start of the intervention. Assent was not required as students were required to use the MI program as part of their regularly scheduled PE class. Finally, the safety of participants was ensured by continuous supervision by trained and qualified members of the research team.

Location

All assessments and study participation took place at Akiva Academy in Calgary, Alberta on June 4th, and June 6th, 2019. The first day was primarily a training day for the students where the researchers conducted a consensus session and allotted time to allow students to get comfortable filming one another. The second day was used to collect the data analyzed in this thesis. The research took place in the school gymnasium during their regularly scheduled PE class.

Previous Pilot Study

A pilot study conducted by the STRL in June 2018 aimed to see if there was an order effect in learning with grade 3 students. This pilot study allowed us to pinpoint the flaws in our research methodology and make the appropriate amendments to ensure success in this study.

Most importantly, it was discovered that a training session with the students prior to data collection would alleviate a multitude of problems and inconsistencies. Firstly, students did not receive guidance when it came to using the iPads to film their partner performing the skill, which resulted in some videos being un-useable. This could be solved by allowing students to practice capturing both horizontal and vertical movements prior to data collection. Secondly, there was no clearly defined consensus on how to evaluate your partner. When the researchers reviewed the data, differences were found between the students' scores and those given by the investigators. This realization demonstrated that students needed to be taught how to discuss and give feedback on the skill being performed. To alleviate this issue, a consensus session was conducted prior to data collection, in which students could practice evaluating each other on a skill that was unrelated to the study (i.e., not a FMS).

Consensus and Training Session

A 30-minute training session was held on the first day of data collection. The goals of this session were two-fold – primarily to teach the students how to effectively give feedback to their peers when evaluating, in addition to learning how to use the MI software.

Students were presented with a slide presentation which outlined the basics of FMS, P2P learning, and the MI software they would be using (Appendix E). Students were then randomly paired up and allowed to practice taking videos of each other on the iPads using the built-in camera. They were asked to do two tasks – running and jumping. These two tasks allowed the students to practice capturing actions in both horizontal and vertical dimensions. Students were instructed to record the video in landscape form, as this layout is preferred for the MI software. Additionally, the investigators mentioned the importance of being able to completely see the entire student in the video recording, as it would make it easier to evaluate later when using MI.

Once the students had become familiar on how to properly take video recordings of each other, they were called in to start the consensus session. Two student volunteers were chosen from the group, with one taking on the role of "evaluator" and the other being the "performer". After watching a demonstration video on MI of an individual lifting a box, the two students attempted to replicate what they had seen (Appendix F). Using the MI software, the "evaluator" filmed the "performer" lifting an empty box. This recording was then projected onto a television screen to allow the entire class to clearly see it.

As a group, the researchers and students completed the evaluation process of the student volunteer lifting a box, making sure to discuss each component individually. To ensure that students were actively engaged in the discussion process, students were each given an iPad and were asked to interact with a student response application called Kahoot, which was loaded on the iPad tablets. Individually, students chose whether they believed each component deserved a score of "Yes", "Partial", or "No". Kahoot compiles student responses and puts them into box charts. The box charts for each component were projected onto the screen one at a time and discussed as a group. For initial responses, there was no consensus for any of the components (see Table A.2). The facilitator used this as an opportunity to engage students in discussion about each component and the reasons for different opinions. For example, why does one partner think

Component 2 should be scored as "Partial", while the other thinks it should be scored as "Yes"? These discussions are extremely important to the peer-to-peer learning process, as it is during these conversations that learning can occur (Mosston & Ashworth, 2008). Once a consensus had been reached within the group, the researchers moved onto the next component, until a final performance score had been reached. Finally, students were instructed on how to take a screenshot of the final screen with the iPads, as this final summary screen contained the crucial data needed for the study.

Table A.2 Frequency of Student Responses for the Individual Components of the 'Workplace Safety – Lifting a Box' Example Used During the Consensus and Training Session

	Males			Females		
Component	**Yes**	**Partial**	**No**	**Yes**	**Partial**	**No**
Feet shoulder width apart?	5	2	1	8	1	1
One foot slightly ahead of the other?	1	3	4	2	3	5
Squat down, bending at the hips and knees only?	5	2	1	8	2	0
Chest out and your shoulders back?	2	3	3	4	1	5
Lift by straightening your hips and knees (not your back)?	2	4	2	8	2	0
Hold the box as close to your body as possible?	2	5	1	4	5	1
Skill performed smoothly?	4	0	4	4	4	2

Tools and Equipment

Move Improve® Application

The VFB technology utilized in this study, MI, was invented by Dr. Larry Katz. MI is a new application created to facilitate structured video peer evaluation among student pairs, or dyads. It capitalizes on P2P evaluation and structured video analysis to provide specific, meaningful, and actionable feedback, therefore improving a student's competence in movement skills. Providing meaningful feedback to learners is critically important to the learning progress, including the learning and teaching of physical skills (Sigrist et al., 2013). MI is a process-oriented assessment tool that breaks down physical skills into easily comprehended components accompanied with an instructional video of the skill being properly executed. As previously mentioned, process-oriented batteries allow educators to identify specific skill components that may need improvement (Barnett et al., 2014). Therefore, by breaking down each skill into its individual components using video and then comparing the actions with pre-recorded evaluation components, students may be able to better understand their mistakes and strengths.

The Chosen Fundamental Movement Skills

MI is a video performance analysis tool with multiple modules (i.e., Fundamental Movement Skills, Workplace Safety, etc.) designed to evaluate an individual's ability to perform a variety of movement skills. The FMS module contains 16 different skills that users can select from to test themselves or their peers (see Figure A.1). Appendix G contains a complete and detailed breakdown of the 16 skills and their individual components. In this study, students were required to complete the following two skills from the FMS module – a SJ and a HBR.

The SJ was the first FMS selected for the students to perform. This skill is broken down into 8 different components within the MI software. These components, in sequential order, are: (1) Eyes looking forward; (2) Feet shoulder width apart; (3) Knees and hips bend as you lower into squat; (4) Arms swing back behind hips in the squat; (5) Arms swing forward smoothly and upward as the legs and feet push off the ground; (6) Arms reach high overhead at the highest point of jump; (7) Land on two feet with control; and (8) Skill performed smoothly? The maximum performance score for this skill is 24 points.

The HBR was the second FMS the students were asked to perform. This skill is broken down into 6 different components by the MI software. These components, in sequential order, are: (1) Arms stretched overhead while lying on the back; (2) Legs straight and together; (3) Stomach muscles are strong and tight when on the back; (4) Roll sideways to their side, then stomach, then side, in a smooth manner; (5) Back muscles are strong and tight when on the stomach; and (6) Skill performed smoothly. The maximum performance score for this skill is 18 points.

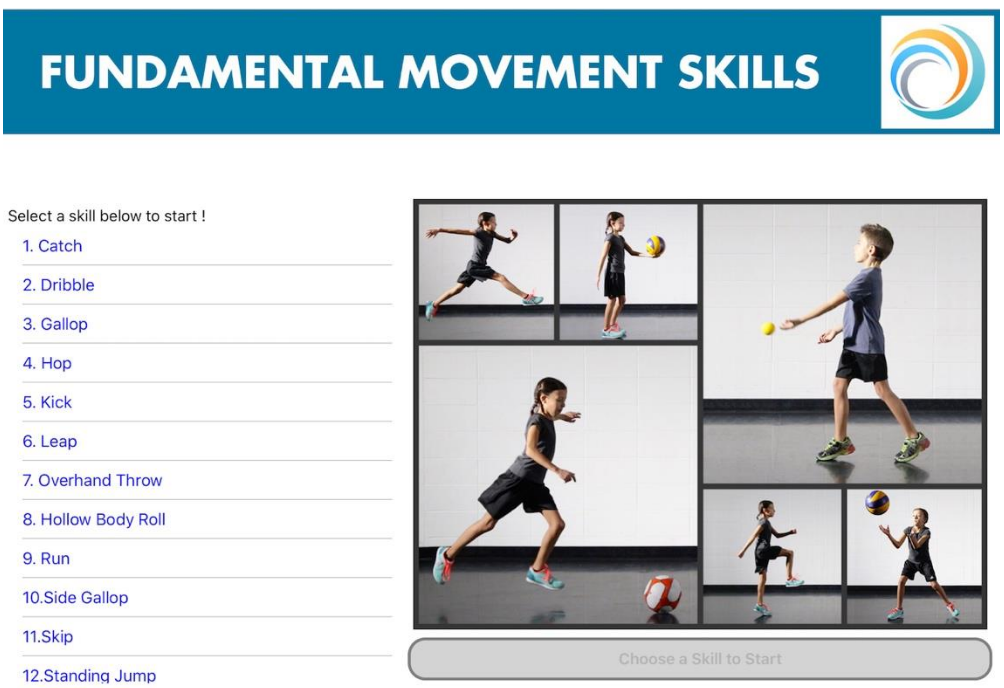


Figure A.1 Move Improve Fundamental Movement Skills List

Data Collection

Dataset and Participants

Eighteen participants (44% male, 56% female) ranging from grades 5, 6, 7, and 8 were included in this study. Two FMS were chosen from MI software for the students to perform – the SJ and the HBR. All 18 participants (eight male, ten female) completed the SJ, while only 16 participants (eight male, eight female) completed the HBR. This was because two female students felt uncomfortable filming themselves performing the HBR while wearing skirts. All together, there were 18 SJ videos and 16 HBR videos, for a total of 34 videos in this dataset.

Research Aim #1

<u>Data Collection Protocol</u>

The school agreed to participate in the study, and the research activities using the iPads were integrated into the students' daily lesson plan. The data was collected during two of their regularly scheduled PE classes, and the total duration of each class was 30 minutes in length. In the first class, students completed the consensus and training session, while the second class focused on data collection. Using the MI VFB software, students employed the reciprocal style of teaching and completed peer assessments for the two chosen skills.

On the second day of on-site data collection, the students were permitted to start using the MI software to collect their data. Using a random pair generator, students were randomly assigned partners and were instructed to take turns performing the SJ and the HBR skills in the MI application. Randomization was also used to determine which student within the dyad would perform the skill first. Students were also told to reverse the order of evaluation for the second skill. For example, student #1 would evaluate student #2 performing the SJ skill, and then, student #2 would evaluate student #1's SJ performance. Next, student #2 would evaluate student #1's HBR performance, and then, student #1 would evaluate student #2 performing the HBR skill. This ensured that within each dyad, each student got a chance to be an evaluator prior to performing the skill themselves. Finally, students were reminded to take screenshots of the summary score page after each skill attempt.

To commence, students launched the application on their iPad, selected peer evaluation mode, and selected either of the two chosen skills from within the FMS module (see Figure A.2). Once a skill is selected, users are provided with an instructional video demonstrating proper performance along with a detailed component list that breaks down the sequential movements and key points required to complete the skill correctly (see Figures A.3 and A.4). Within their dyad, one peer becomes the evaluator, while the other peer assumes the role of performer. The evaluator takes a video of the performer who attempts to do the skill (see Figure A.5). Once students had filmed themselves performing using the iPads, they viewed the video through the MI application while simultaneously assessing their performance with a peer.

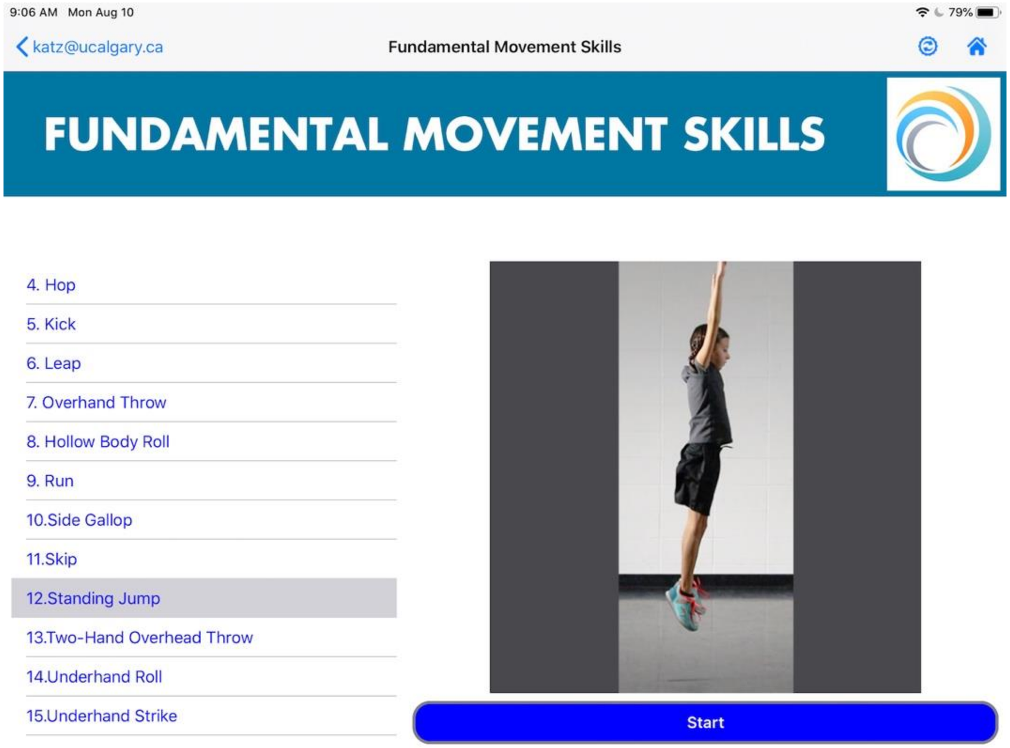


Figure A.2 Choosing a Skill, Standing Jump, from Fundamental Movement Skills Module List

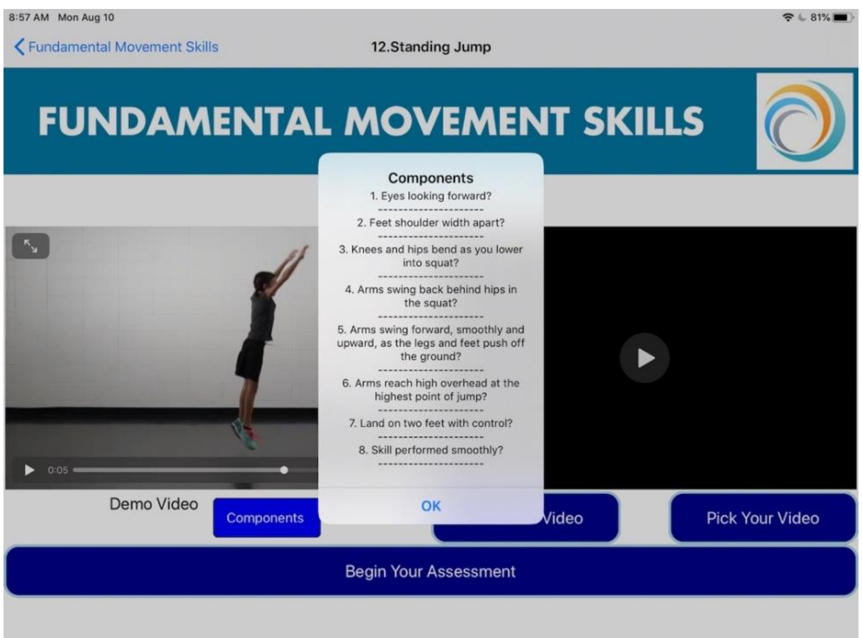


Figure A.3 Fundamental Movement Skill Instructional Video and Component List for Standing Jump

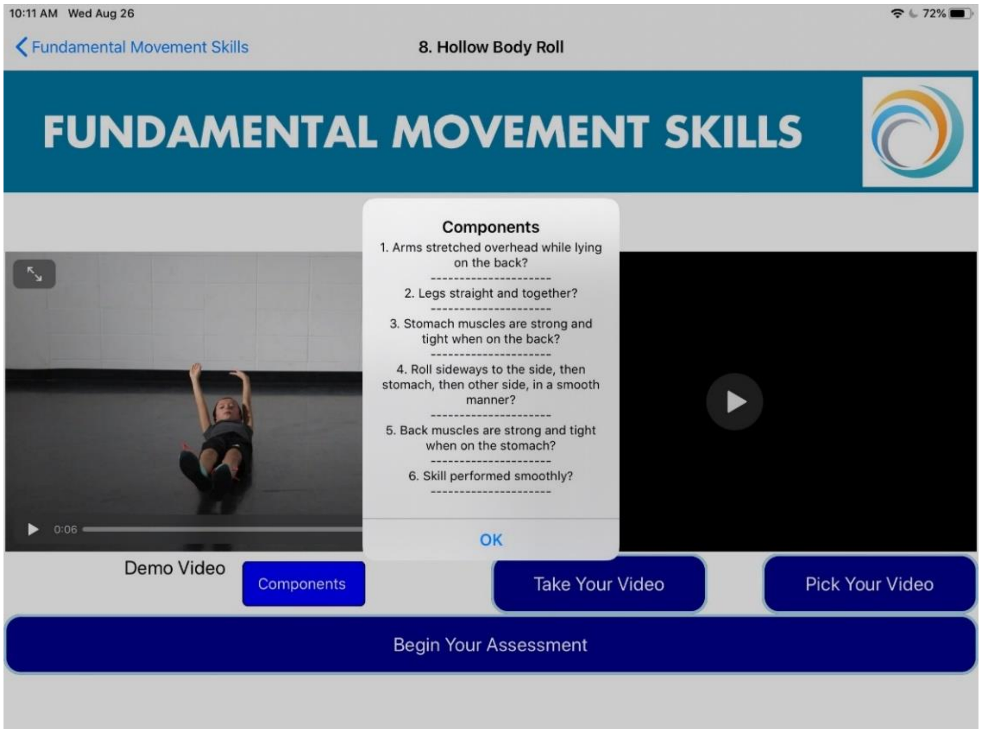


Figure A.4 Fundamental Movement Skill Instructional Video and Component List for Hollow Body Roll

Figure A.5 Evaluator Filming the Performer During the Hollow Body Roll Performance

During the evaluation or assessment phase, the evaluator and performer worked together to assess the performance based on what they saw during video replay (see Figure A.6). Students were encouraged to work together when discussing each other's performances, but were reminded that ultimately, the evaluator makes the final decision regarding scoring. Using the structured questions from the MI software, the dyad evaluated each of the checklist components using the following three-point rating scale – "Yes" (the component was performed correctly – i.e., matches image or instructional video), "Partial" (the component was performed, but not entirely correctly), or "No" (the component was either not performed at all or performed incorrectly) (see Figure A.7). Furthermore, each of these rating options are assigned a numerical score (Yes = 3, Partial = 2, No = 1) – termed a "component score". Evaluators and performers were encouraged to do the assessment together and discuss their opinions.

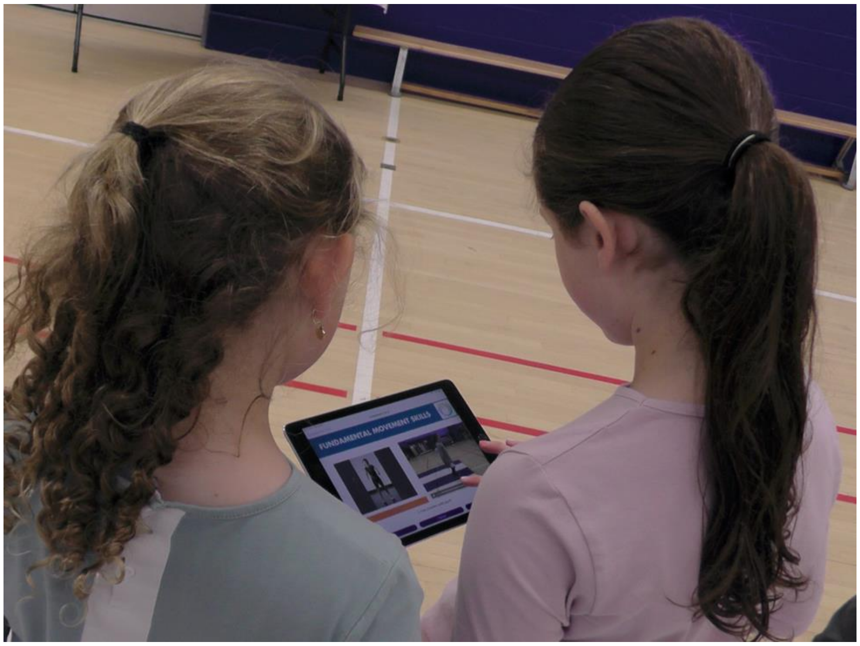


Figure A.6 Evaluator and Performer using the Move Improve® Application to Discuss and Evaluate Performance

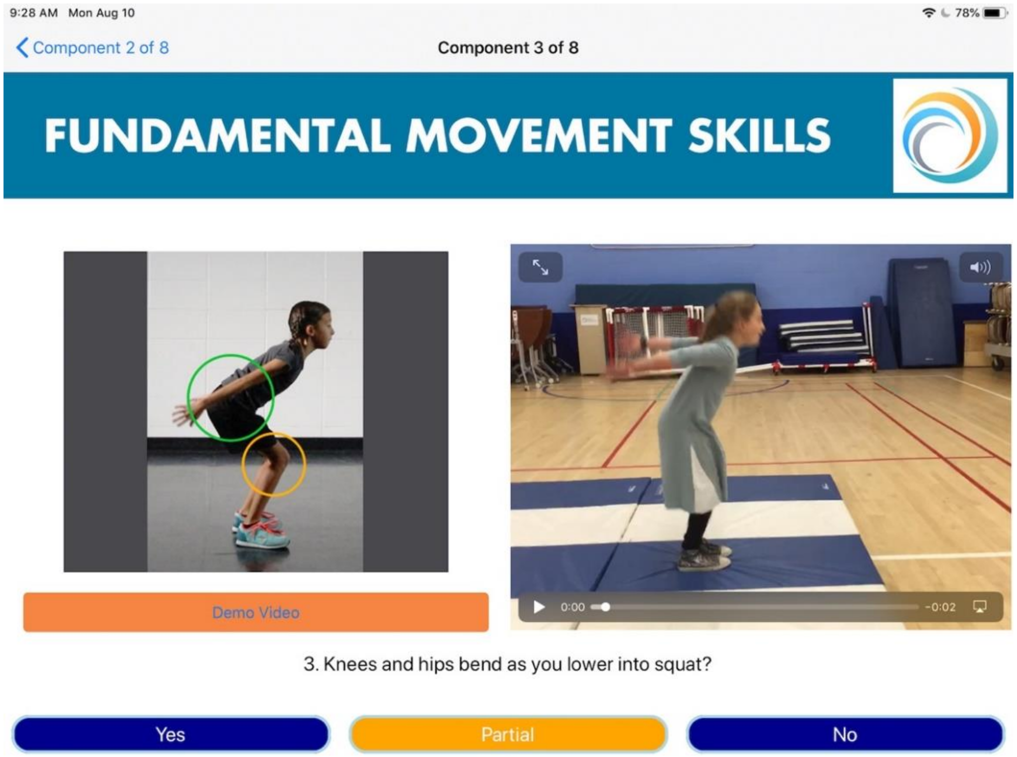


Figure A.7 Evaluation Component Using Yes, Partially, and No Options

When the assessment was complete, a summary list of all the components was shown on the iPad with their corresponding component score, which is totaled at the end to give the entire skill a "performance score". As mentioned in the consensus session, the students took a screenshot of their summary screen, which contained the component scores and performance score. The performance score was also recorded on the "Data Collection Worksheet (Appendix H) prior to the students switching roles. This knowledge offers objective feedback on the results to evaluators and performers, in addition to focusing their attention to certain aspects of the skill which require improvement. Then the students reversed evaluator and performer roles and began a new assessment. This process was repeated until both students had completed the SJ and HBR skills.

Following data collection, the screenshots and the corresponding video for each participant was exported to external hard drives. Subsequently, all participants' videos were re-evaluated by two independent raters and an interdisciplinary consensus panel to provide "expert" scores for each student's performance. This process is further expanded on later in this chapter. Once complete, all the participants' videos were deleted from the iPads.

Research Aim #2

Data was uploaded to a secure server, labelled using participant ID numbers, and exported to external hard drives. Each coder was given access to all 34 videos. This digital format provides multiple benefits, as it allows raters to pause, slow down, and re-watch the video as many times as required (Haynes & Miller, 2015).

Prior to reliability data collection, the graduate student conducted a 1-hour training session, where members of the consensus panel practiced using the MI application and spent time discussing how to evaluate the FMS using the software. Topics covered during this training sessions included instructions on how to log in and watch the videos, how to adjust video playback speed, and troubleshooting video playback issues.

Data Collection Protocol

Raters

Four groups of raters were included in the rating of each video using the MI application evaluation feature: students, two independent raters, and a three-person consensus panel.

Using the reciprocal style of teaching, students were randomly assigned partners and instructed to perform the SJ and HBR skills from the MI application. Prior to data collection, all students participated in a 30-minute training session where they were taught how to effectively give feedback to their peers when evaluating, in addition to learning how to use the MI software. Students were encouraged to do the assessment together and engage in discussions regarding their performance. Once students had completed their peer assessment, they took screenshots of the summary page which contained the component and performance scores for the skill, in addition to recording the performance score on the "Data Collection Worksheet" (see Appendix H).

Two independent raters, the graduate student (Rater 1) and a biomechanics expert (Rater 2), both of whom were also members of the interdisciplinary consensus panel, viewed and rated the 34 videos independently. Rater 1 was a kinesiology graduate student and Rater 2 had recently completed a Bachelor of Science in Kinesiology, specializing in Biomechanics.

The consensus panel consisted of a biomechanics expert, a kinesiologist, and the graduate student. All members of the interdisciplinary consensus panel were knowledgeable about physical literacy and FMS, having all completed a Bachelor of Science in Kinesiology.

Furthermore, all members were also familiar with the MI software as a VFB tool, however, the graduate student was the only one who had previous experience evaluating FMS videos using the MI software. This was due to previous involvement in research and pilot studies.

Scoring Protocol

Using peer assessment, students rated each video once within their dyad immediately after performing the assigned skills. There were no limits to how many times a video could be replayed, but only one performance score was recorded per skill. Screenshots of the summary page containing both the component and performance scores were taken by students once they completed their peer assessment, in addition to recording the performance score on the "Data Collection Worksheet" (see Appendix H).

The consensus panel rated each video twice, however, the scores were recorded 6 months apart. During this time frame raters were not permitted to view the videos to limit recall. Additionally, raters were not allowed to reference the first scoring round results during the second rating session. All videos were viewed and scored together by the consensus panel using the structured questions and built-in evaluation feature within the MI application. After each trial, any discrepancies were discussed until all members came to an agreement. Finally, there were no limits to how many times a video could be watched by the consensus panel.

Individual coders rated each video once, independently, using the built-in evaluation feature within the MI application. It is important to note that coders were not granted access to view each others' scores throughout this process. There were no limits to how many times a video could be viewed.

Scoring Criteria

Appendix I contains a detailed breakdown of the scoring criteria used in this study.

Statistical Analysis

Statistical Analysis Tools

The data was screened for completeness, and any missing or incomplete data was excluded from the analysis. Following the checking of all data for any errors, and cleaning the excel spreadsheet, all data was entered into SPSS statistical analysis software to be cross checked for any missing or incorrect values, as well as for analysis. All statistical analysis was conducted in SPSS Version 26, and alpha levels were set to 0.05 *a priori*. The statistical methods used on the data set varied depending on the research question.

Statistical Analysis for Research Aim #1

The first outcome of this investigation examined the impact that order of performance within a reciprocal teaching style has on the performance outcomes of two FMS, SJ and HBR, in middle school students. An independent samples *t*-test was conducted to evaluate the differences in the mean scores between students that performed the skill first and those who performed the skill second for each skill (SJ and HBR). A Cohen's *d* effect size, which measures the strength of the relationship between two variables (Cohen, 1988), was also calculated for each skill between the students that performed first and the students that performed second.

Statistical Analysis for Research Aim #2

The second aim of this investigation was to determine the inter-rater and intra-rater reliability of the MI application when assessing the performance outcomes of two FMS, SJ and

HBR, in middle school students. The inter-rater reliability analysis was conducted by calculating the ICC (2,k) values of the performance scores for the raters by pairs (Rater 1, Rater 2, Consensus 1, Consensus 2, and student). ICC (2,k) reliability values are characterized as poor (<0.50), moderate (0.51-0.75), or good (>0.75) (Portney, 2020; Sheehan et al., 2011). An independent t-test was conducted to determine a statistically significant difference between the performance scores for each skill provided by each rater. A dependent t-test was conducted to determine a statistically significant difference between the consensus 1 and consensus 2 scores for each skill. Percent agreement was calculated to further analyze the inter-rater reliability of each pair of raters by looking at the agreement between raters for each component of the skill. An intra-rater reliability analysis was conducted by calculating the ICC (2,k) values of the total performance scores and component scores for consensus 1 and consensus 2. Finally, percent agreement was used to further determine the intra-rater reliability between consensus 1 and consensus 2.

Appendix B: Parental Consent Form

UNIVERSITY OF CALGARY

TITLE: Using Tablets to Learn Physical Skills

SPONSOR: Sport Technology Research Lab

INVESTIGATORS: Dr. Larry Katz and Ms. Anna Thacker

This consent form is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your child's participation will involve. If you would like more detail about something mentioned here, or information not included here, please ask. Take the time to read this carefully and to understand any accompanying information. You will receive a copy of this form for your records.

BACKGROUND

Large class sizes make it difficult for educators to teach physical skills to 20+ students in the 30-minute duration allotted for Physical Education (PE) class. Computer technology, such as tablets (e.g., iPads) may be a valuable resource for teachers to use. A number of applications have been developed specifically for physical education that use these tablets. As these applications are starting to be used in the gym, it is important to evaluate the effectiveness of their use. One such application is Move Improve. This software will be used to evaluate your child's performance regarding fundamental movement skills. This software generates scores based on the child's performance of the skills.

WHAT IS THE PURPOSE OF THE STUDY?

This study will examine the impact of using Move Improve® software as an aid for students to learn and evaluate various physical skills performed during gym class. Performance will be assessed by examining results student performance obtained from the tablet and videos of their performance.

WHAT WOULD MY CHILD HAVE TO DO?

This study will take place during your child's regularly scheduled PE class. Their PE teacher will be teaching the class using iPads with specially designed programs.

Before PE class, your child will be asked to fill out a short questionnaire. Your child will be taught how to use the iPad program which will take about 5 minutes. The rest of the class will follow the curriculum set out by the teacher using the iPads as part of the lesson. Once the lesson is complete, the students will be asked to fill out another questionnaire about the experience.

When the students use the iPad, students will record each other's performances. The video of your child will be evaluated by their classmate using the pre-programmed cues in the Move Improve® software. The video will be evaluated using the predetermined criteria and both a performance and component score will be generated. The videos are stored on the iPads.

This consent form is asking your permission for the following:

- To have your child fill out the pre and post questionnaires,
- To use the video stored on the iPads, and
- To use the performance data collected on the iPads.

Finally, if your child is selected, they may participate in a small group discussion if they agree.

WHAT ARE THE RISKS?

As this is a regularly scheduled PE class, there are no risks beyond what is normally expected in these active situations.

ARE THERE ANY BENEFITS FOR MY CHILD?

This study may be beneficial as your child will have the opportunity to improve their physical abilities and learn how to use the tablets for physical activity.

DOES MY CHILD HAVE TO PARTICIPATE?

This is a regularly scheduled PE class, so your child will have to participate in the physical activities. They are not required to fill in the questionnaires, and you do not have to give permission for the researchers to use the data. Although video data will be collected on all of the children present in class, if you choose not to participate in this study, your child's video will not be accessed by the research team. You or your child may also withdraw permission for the researchers to use the data at any time up until the start of data analysis.

WHAT ELSE DOES MY CHILD'S PARTICIPATION INVOLVE?

In this research personal information will be collected including: name, gender and age. However, names will not be connected to the data for analysis and presentation of results.

WILL WE BE PAID FOR PARTICIPATING, OR DO WE HAVE TO PAY FOR ANYTHING?

There is no cost associated with participation in this study, there is also no remuneration.

WILL MY CHILD'S RECORDS BE KEPT PRIVATE?

Data collected during your child's time in this research study will be de-identified and will be held in a database for future use by other researchers. The name of each participant will be linked to an identification number and locked in a separate secure cabinet from the information collected during testing. All field notes, journals or observations shall remain in the investigator's possession or securely locked in a filing cabinet inside the Sports Technology Research Lab. Video files or photographic images will be kept for a very short period of time and used to verify testing results. Any future use of this research data is required to undergo review by a Research Ethics Board.

IF MY CHILD SUFFERS A RESEARCH-RELATED INJURY, WILL WE BE COMPENSATED?

In the event that your child suffers injury as a result of participating in this research, no compensation will be provided to you by *Sport Technology Research Lab*, the University of Calgary, Alberta Health Services or the Researchers. You still have all your legal rights. Nothing said in this consent form alters your right to seek damages.

☐ I consent to my child participating in this research project

☐ In addition to the data, your child will be videoed for the performance. We want permission to use the video for research purposes. Saying no to the video does not exclude your child from participating. I consent to my child's video being used for research purposes

SIGNATURES

Your signature on this form indicates that you have understood to your satisfaction the information regarding your child's participation in the research project and agree to their participation as a subject. In no way does this waive your legal rights nor release the investigators or involved institutions from their legal and professional responsibilities. If you have further questions concerning matters related to this research, please contact:

Dr. Larry Katz (403) 220-3418

Or

Ms. Anna Thacker (403) 542-3628

If you have any questions concerning your rights as a possible participant in this research, please contact the Chair of the Conjoint Health Research Ethics Board, University of Calgary at 403-220-7990.

Parent/Guardian's Name	**Signature and Date**
Child's Name	**Signature and Date**
Investigator/Delegate's Name	**Signature and Date**
Witness' Name	**Signature and Date**

The investigator or a member of the research team will, as appropriate, explain to your child the research and his or her involvement. They will seek your child's ongoing cooperation throughout the study.

The University of Calgary Conjoint Health Research Ethics Board has approved this research study.

A signed copy of this consent form has been given to you to keep for your records and reference.

Appendix C: Certification of Institutional Ethics Approval

UNIVERSITY OF CALGARY

Conjoint Health Research Ethics Board
Research Services Office
2500 University Drive, NW
Calgary AB T2N 1N4
Telephone: (403) 220-2297
chreb@ucalgary.ca

CERTIFICATION OF INSTITUTIONAL ETHICS APPROVAL

The Conjoint Health Research Ethics Board (CHREB), University of Calgary has reviewed and approved the following research protocol:

Ethics ID:	REB18-0301
Principal Investigator:	Larry Katz
Co-Investigator(s):	There are no items to display
Student Co-Investigator(s):	Ramandeep Jaswal Anna Thacker Alanna Weisberg
Study Title:	Peer to Peer Learning: The Impact of Order of Instruction on Learning
Sponsor:	

Effective: Tuesday, June 12, 2018 **Expires:** Wednesday, June 12, 2019

The following documents have been approved for use:

- Letter to parents from principal, 1, June 11, 2018
- Consent Form Clean, 3, June 11, 2018
- Mature Minor Consent Form, 1, June 11, 2018
- Assent Form Clean, 2, June 11, 2018
- PostQVer2, 2, June 4, 2018
- Focus Group Questions, 1, May 26, 2018
- PreQ Version 2 , 2, June 4, 2018
- Mi Proposal Clean, 2, June 11, 2018
- Data Collection Worksheet, 1, May 29, 2018

The CHREB is constituted and operates in accordance with the current version of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS); International Conference on Harmonization E6: Good Clinical Practice Guidelines (ICH-GCP); Part C, Division 5 of the Food and Drug regulations, Part 4 of the Natural Health Product Regulations and the Medical Device Regulations of Health Canada; Alberta's Health Information Act, RSA 2000 cH-5; and US Federal Regulations 45 CFR part 46, 21 CFR part 50 and 56.

You and your co-investigators are not members of the CHREB and did not participate in review or voting on this study.

Restrictions:

This Certification is subject to the following conditions:

1. Approval is granted only for the research and purposes described in the application.
2. Any modification to the approved research must be submitted to the CHREB for approval.
3. An annual application for renewal of ethics certification must be submitted and approved by the above expiry date.
4. A closure request must be sent to the CHREB when the research is complete or terminated.

Approval by the REB does not necessarily constitute authorization to initiate the conduct of this research. The Principal Investigator is responsible for ensuring required approvals from other involved organizations (e.g., Alberta Health Services, community organizations, school boards) are obtained.

Approved By:	**Date:**
Stacey A. Page, PhD, Chair , CHREB	Tuesday, June 12, 2018

Note: This correspondence includes an electronic signature (validation and approval via an online system).

Appendix D: TCPS 2: Core Certificate

PANEL ON
RESEARCH ETHICS
Navigating the ethics of human research

TCPS 2: CORE

Certificate of Completion

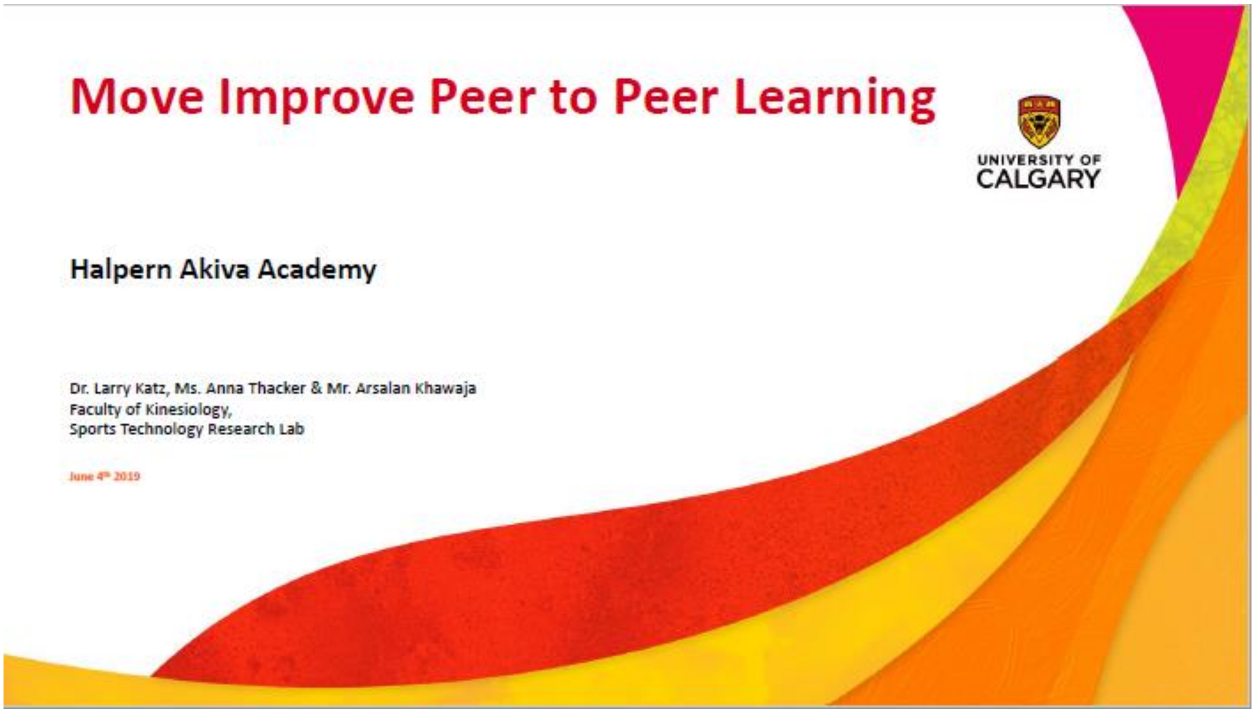
This document certifies that

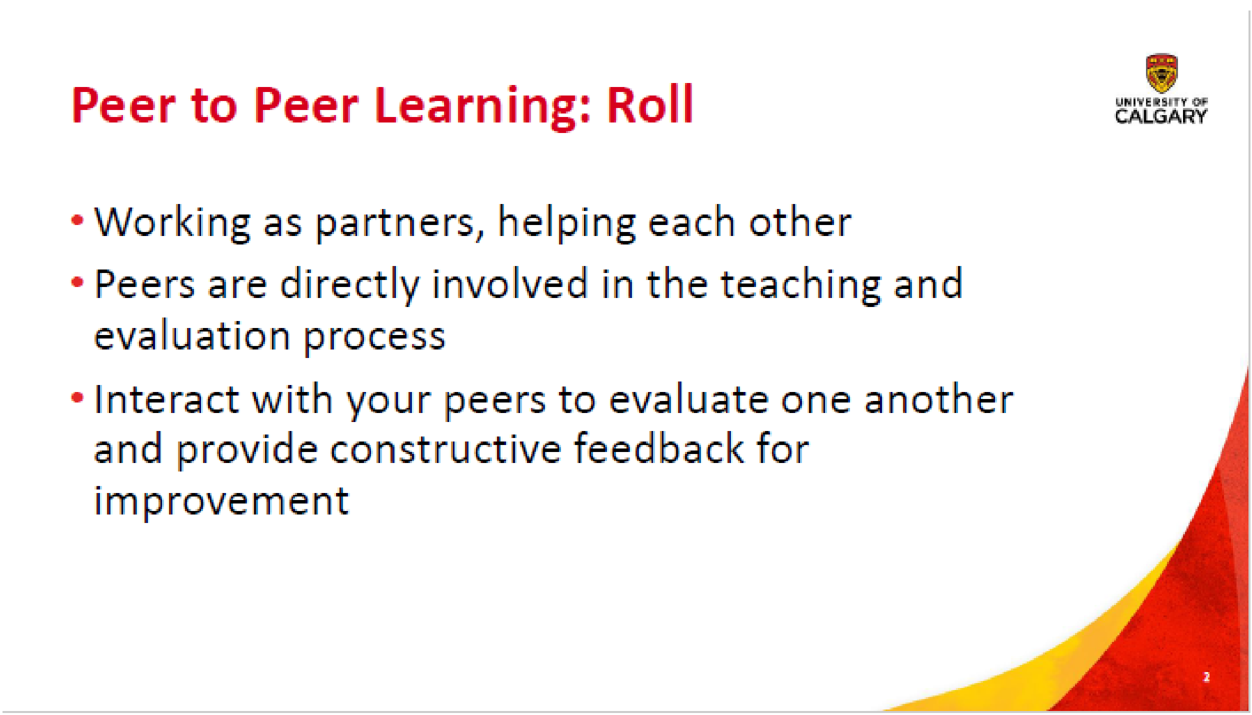
Anna Thacker

has completed the Tri-Council Policy Statement:
Ethical Conduct for Research Involving Humans
Course on Research Ethics (TCPS 2: CORE)

Date of Issue: **2 May, 2017**

Appendix E: Consensus Training PowerPoint Presentation





Move Improve Roles

UNIVERSITY OF CALGARY

- Evaluator
  - Filming student using iPad
- Performer
  - Person doing the task
- Discuss the performance

What is Move Improve?

UNIVERSITY OF CALGARY

Rating	Description
Yes	Student correctly demonstrates skills
Partial	Student partially demonstrates skills
No	Student does not demonstrate skills

How to Properly Evaluate

- After performing task, you will evaluate the video **TOGETHER**
- Communicate effectively with your partner
  - Provide constructive feedback
  - If you have a difference of opinion, discuss it with your partner!
    - Why do you think it's a "Partial" and why do they think it's a "Yes/No"?
- Remember, you are a **TEAM!**

How to use the iPad

- Taking videos
  - **Evaluator** will track the performer with the iPad
    - Camera in top left hand corner
    - iPad in TV – Landscape mode
    - Steady hands
    - Press the record button and then tell the performance to start
    - Keep the performer in the centre of the screen
  - **Performer tries their best**
  - **Discuss video – look at quality**
- Screenshot of performance results

UNIVERSITY OF CALGARY

Let's practice!

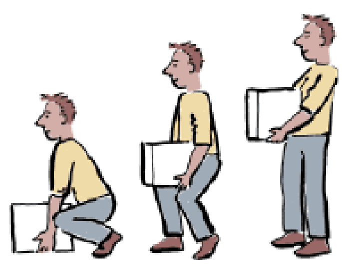
- Practice taking good video
- Choose who will be an evaluator vs. performer
- Take two separate videos
  - Running –
    - Together evaluate quality of video
    - Switch rolls and repeat
  - Jumping
    - Together evaluate quality of video
    - Switch rolls and repeat

- When reviewing the video together
  - Make sure it's clear (not blurry)
  - Performer is visible in the frame

UNIVERSITY OF CALGARY

Consensus Session

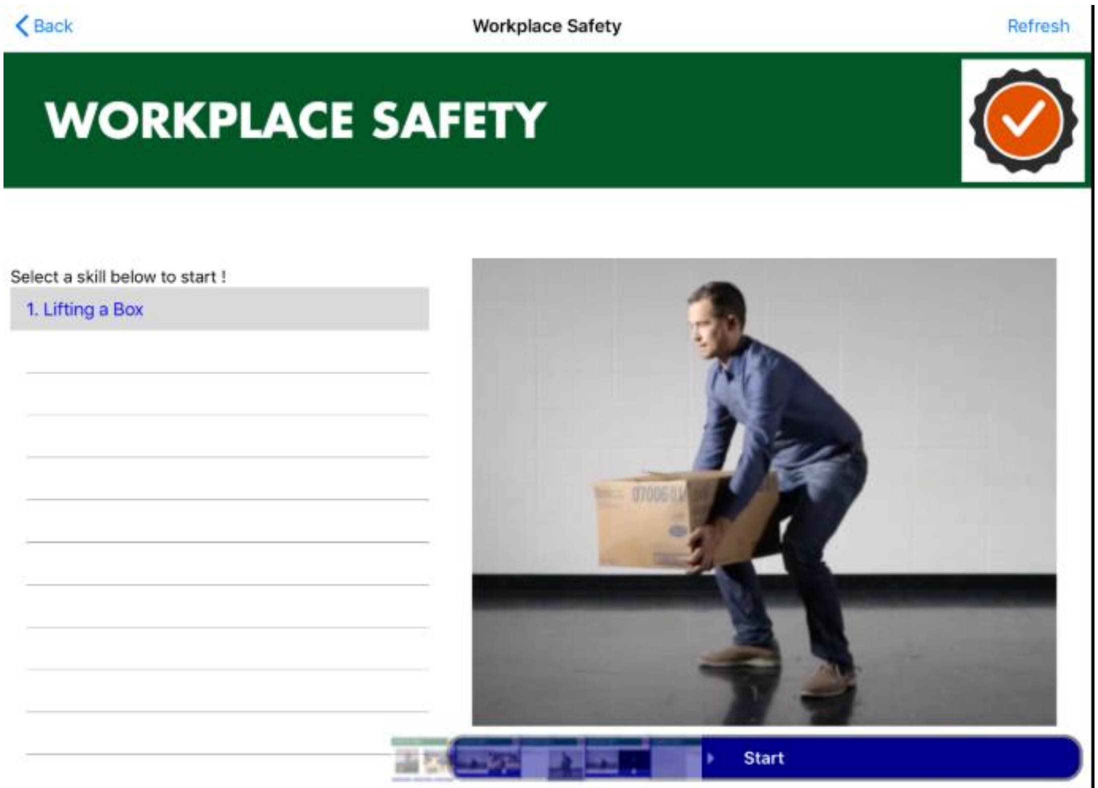
- Two volunteers
  - One "Evaluator" and one "Performer"
- Practice Lifting a Box

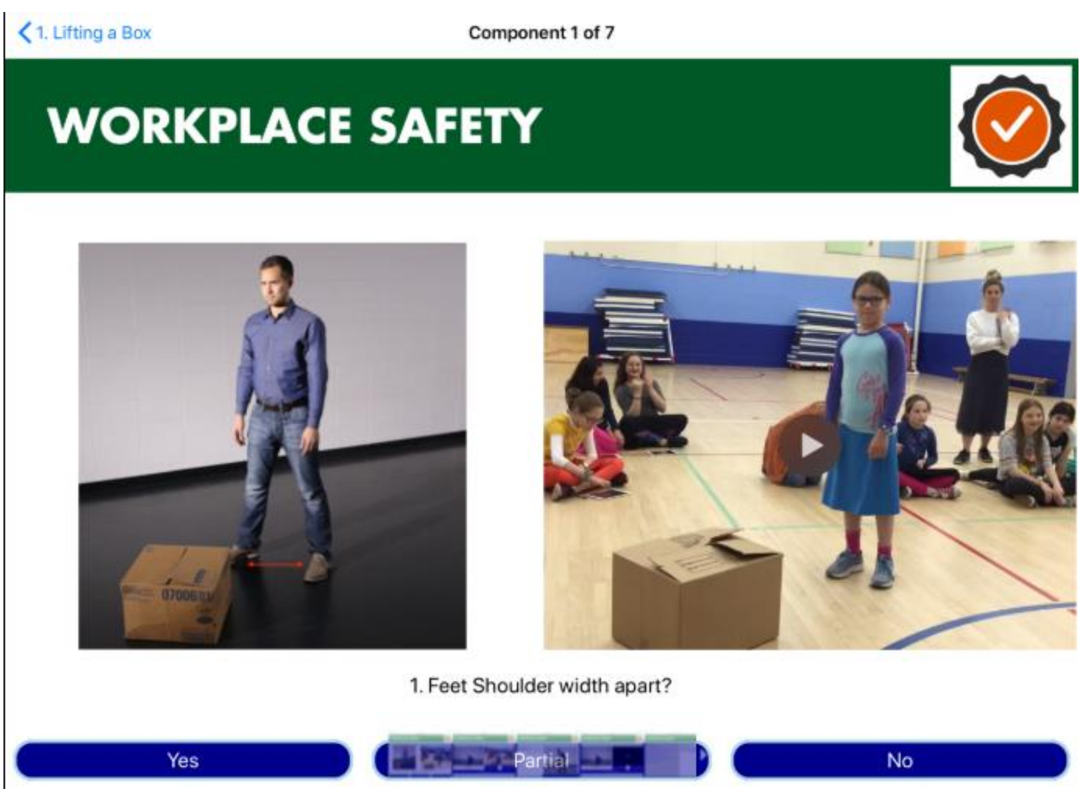


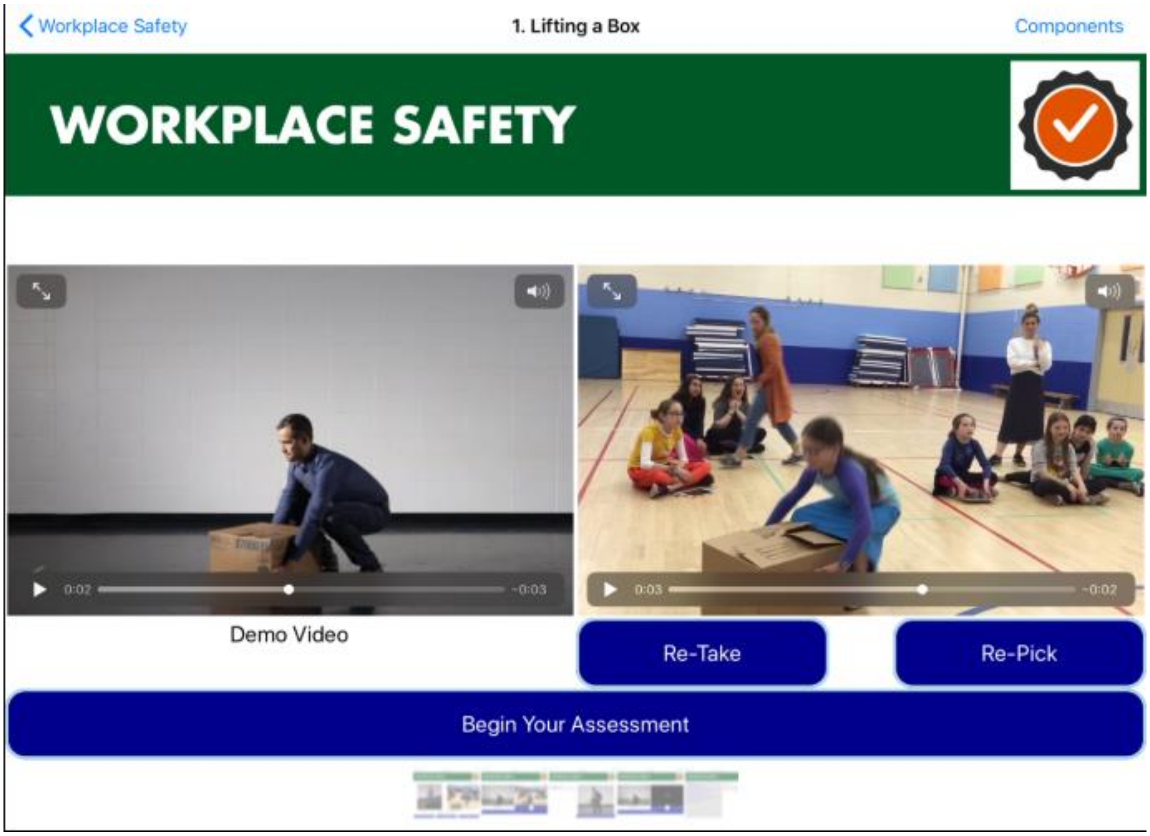
- Kahoot!
  - Enter PIN to join

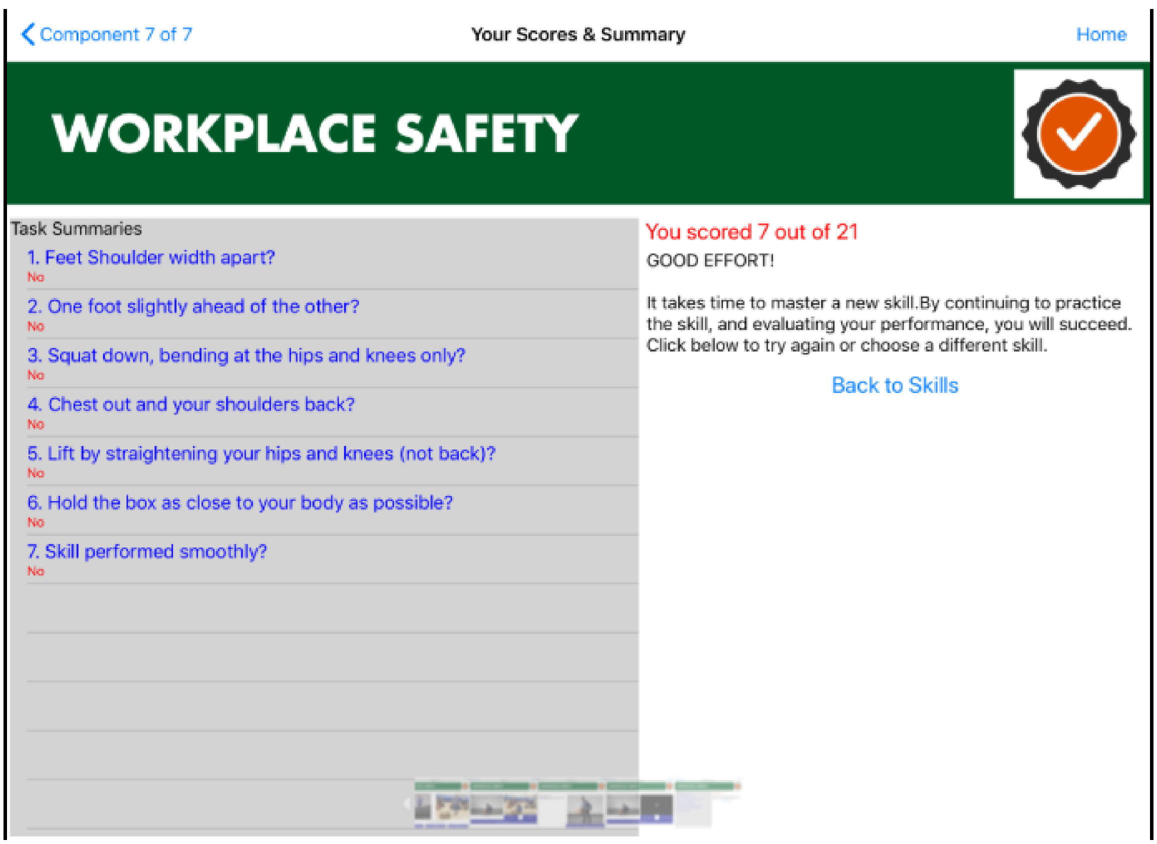
Kahoot!

Appendix F: Move Improve® Box Lifting Example






Workplace Safety
1. Lifting a Box
Components
WORKPLACE SAFETY
Demo Video
Re-Take
Re-Pick
Begin Your Assessment


Component 7 of 7
Your Scores & Summary
Home
WORKPLACE SAFETY
Task Summaries
1. Feet Shoulder width apart?
No
2. One foot slightly ahead of the other?
No
3. Squat down, bending at the hips and knees only?
No
4. Chest out and your shoulders back?
No
5. Lift by straightening your hips and knees (not back)?
No
6. Hold the box as close to your body as possible?
No
7. Skill performed smoothly?
No
You scored 7 out of 21
GOOD EFFORT!
It takes time to master a new skill.By continuing to practice the skill, and evaluating your performance, you will succeed. Click below to try again or choose a different skill.
Back to Skills

Appendix G: Move Improve® Fundamental Movement Skills Module – Skill Components

1. Catch – Total Number of Points = 21 points

1. Ready position with eyes looking at target?
2. Eyes track the object all the way to the hands?
3. Moves toward the travelling object?
4. Hands extend toward the travelling object?
5. Thumbs together if high or pinkies together if low?
6. Elbows and shoulders give way to softly cradle the object?
7. Skill performed smoothly?

2. Dribble – Total Number of Points = 18 points

1. Ready position with eyes looking in the direction of the travel?
2. Soft taps used with the inside and outside of the feet?
3. Both feet used?
4. Ball kept within 50cm of the feet?
5. Controlled movement of the ball in various pathways?
6. Skill performed smoothly?

3. Gallop – Total Number of Points = 18 points

1. Eyes looking forward?
2. Front foot steps forward?
3. Back foot quickly glides to meet the heel of the front foot in flight phase?
4. Step – glide pattern has rhythm?
5. Arms rock smoothly with each step – glide pattern?
6. Skill performed smoothly?

4. Hop – Total Number of Points = 27 points

1. Eyes looking forward?
2. Balance on one foot as arms swing back behind hips?
3. Knee bends slightly?
4. Opposite knee swings upwards?
5. Arms and shoulders help by lifting upward?
6. Soft landing on the same foot?
7. Quick controlled rhythm in each hop?
8. Arms help to stay balanced?
9. Skill performed smoothly?

5. Kick – Total Number of Points = 33 points

1. Ready position with eyes looking in the direction of the target?
2. Ball is placed approximately 1 to 1 ½ step(s) in front of the student?
3. Non-kicking foot used to step forward towards the ball like the toe will “kiss the ball on the cheek”?

4. Balance shown briefly on the non-kicking leg as the kicking leg prepares to kick?
5. Kicking leg swings through toward the ball with a slight knee bend?
6. The toes of the kicking leg foot are pointed down so as to kick with the laces of the shoe?
7. The 'chin' of the ball is contacted with the laces of the kicking foot?
8. Kicking leg follows through towards the target while stepping forward onto it?
9. Increased power is shown by leaping with the non-kicking foot from a greater distance from the ball?
10. Skill performed smoothly?
11. Performed with both feet?

6. Leap – Total Number of Points = 24 points

1. Eyes looking forward?
2. Back foot is used to push off the ground?
3. Front knee/foot drives up and forward?
4. Both feet are off the ground in flight?
5. Legs are spread in a large scissor like shape?
6. Front foot used for soft landing?
7. Arms help to stay balanced?
8. Skill performed smoothly?

7. Overhand Throw – Total Number of Points = 30 points

1. Ready position with eyes looking at target?
2. Throwing hand makes a muscle as the body turns sideways towards the target?
3. Sideways step used toward the target with the opposite foot?
4. Toe of the opposite foot pointing in the direction of the target?
5. Elbow of the throwing hand travel toward the target with the step of the opposite foot?
6. Shoulders rotate to square up with the target as the elbow travels forward past the ear?
7. Throwing arm extended toward the target as the object is released?
8. Throwing hand follows through across to the opposite knee?
9. Skill performed smoothly?
10. Performed with both arms?

8. Hollow Body Roll – Total Number of Points = 18 points

1. Arms stretched overhead while lying on the back?
2. Legs straight and together?
3. Stomach muscles are strong and tight when on the back?
4. Roll sideways to their side, then stomach, then side, in a smooth manner?
5. Back muscles are strong and tight when on the stomach?
6. Skill performed smoothly?

9. Run – Total Number of Points = 27 points

1. Eyes looking forward?
2. Body leaning forward?

3. Middle of the foot strikes ground first?
4. Back foot snaps toward bum after push off of the ground in flight phase?
5. Back leg swings through bent to become new front leg?
6. Front knee is bent as it drives forward?
7. Arms are bent?
8. Arm swings forward with opposite foot?
9. Skill performed smoothly?

10. Side Gallop – Total Number of Points = 24 points

1. Stand sideways to the direction of travel?
2. Head turned over the shoulder to look in the direction of travel?
3. Front foot points in the direction of travel?
4. Front foot steps forward?
5. Back foot quickly glides to meet the heel of the front foot in flight phase?
6. Step – glide pattern has rhythm while travelling sideways?
7. Arms rock smoothly with each step – glide pattern?
8. Skill performed smoothly?

11. Skip – Total Number of Points = 27 points

1. Eyes looking forward?
2. Lead foot steps forward?
3. A quick smooth hop is performed on the lead foot with the opposite knee driving up and forward?
4. Step forward with the opposite foot?
5. Step – hop pattern has rhythm?
6. Front arm is bent?
7. An arm moves forward with the opposite leg?
8. Arms help to stay balanced?
9. Skill performed smoothly?

12. Standing Jump – Total Number of Points = 24 points

1. Eyes looking forward?
2. Feet shoulder width apart?
3. Knees and hips bend as you lower into squat?
4. Arms swing back behind hips in the squat?
5. Arms swing forward smoothly and upward as the legs and feet push off the ground?
6. Arms reach high overhead at the highest point of jump?
7. Land on two feet with control?
8. Skill performed smoothly?

13. Two-Hand Overhead Throw – Total Number of Points = 21 points

1. Ready position with eyes looking at target?
2. Place hands on the ‘ears’ of the ball and hold it at waist level?

3. Ball lifted over and behind the head?
4. Step forward with the dominant foot as the elbows begin to travel forward past the student's ears?
5. Both elbows extend, releasing the ball towards the target?
6. Both hands follow through toward the target with the hands turning so the palms face outward?
7. Skill performed smoothly?

14. Underhand Roll – Total Number of Points = 27 points

1. Ready position with eyes looking at target?
2. Rolling arm drawn backwards behind the hip as a forward step is started?
3. Step forward with the opposite foot to the rolling arm?
4. Rolling arm swung forward in a straight smooth line?
5. Front knee bends to absorb and lower the body toward the ground?
6. Object released at a low level to roll along the ground?
7. Rolling hand follows through toward the target?
8. Skill performed smoothly?
9. Performed with both arms?

15. Underhand Strike – Total Number of Points = 24 points

1. Ready position with eyes looking in the direction of the target?
2. Object is balanced in the palm of the non-striking hand at waist level?
3. Striking arm drawn backwards behind the hip while controlling the balanced object in front?
4. Step forward with the opposite foot as the striking arm swings straight forward?
5. Ball is contacted with a straight arm and open palm of the striking hand at waist level?
6. Striking hand follows through towards the target?
7. Skill performed smoothly?
8. Performed with both arms?

16. Underhand Throw – Total Number of Points = 27 points

1. Ready position with eyes looking at target?
2. Throwing arm drawn backwards behind the hip?
3. Step forward with the opposite foot as the throwing arm swings forward?
4. Throwing arm swung forward in a straight smooth line?
5. Front knee bends slightly to absorb the step?
6. Object released at waist level to travel in a curved pathway in the air?
7. Throwing hand follows through toward the target?
8. Skill performed smoothly?
9. Performed with both arms?

Appendix H: Data Collection Worksheet

UNIVERSITY OF CALGARY

Sport Technology Research Laboratory

PARTNER'S NAMES: ___________________ & ____________________

DATE: ____________________

GRADE: (SK) (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (UNI)

Trial	Skill	Evaluator	Performer	Performance Score	Screen Capture ☑
1					☐
2					☐
1					☐
2					☐

Appendix I: Detailed Scoring Criteria for Standing Jump and Hollow Body Roll

Hollow Body Roll

Components:

1. Arms stretched overhead while lying on the back?
2. Legs straight and together?
3. Stomach muscles are strong and tight when on the back?
4. Roll sideways to their side, then stomach, then side, in a smooth manner?
5. Back muscles are strong and tight when on the stomach?
6. Skill performed smoothly?

Scoring Criteria:

1. Arms stretched overhead while lying on the back?
   - Yes – Arms stretched overhead and parallel to the floor
   - Partial – Arms stretched, but perpendicular to the floor
   - No – Arms remain down by their sides and not stretched overhead
2. Legs straight and together?
   - Yes – Legs are straight and together throughout the entire hollow body roll
   - Partial – Legs are straight and together at either the start or end of the hollow body roll, but are bent and separated during the middle of the movement
   - No – Legs are either bent and/or separated throughout the entire hollow body roll
3. Stomach muscles are strong and tight when on the back?
   - Yes – Core stomach muscles are fully engaged when lying on their back
   - Partial – Core stomach muscles are partially engaged when lying on their back
   - No – Core stomach muscles are unengaged when lying on their back
4. Roll sideways to their side, then stomach, then side, in a smooth manner?
   - Movement criteria:
     - (1) Roll sideways to their side
     - (2) Roll from their side to their stomach
     - (3) Roll from their stomach to their opposite side
   - Yes – All three of the movement criteria are completed in a smooth manner
   - Partial – At least 1/3 of the movement criteria are completed in a smooth manner
   - No – None of the movement criteria are completed in a smooth manner
5. Back muscles are strong and tight when on the stomach?
   - Yes – Back muscles are fully engaged when on their stomach
   - Partial – Back muscles are partially engaged when on their stomach
   - No – Back muscles are unengaged when on their stomach
6. Skill performed smoothly?
   - Yes – Performed smoothly throughout the entirety of the movement
   - Partial – Movement started smoothly, but transitioned to ending not smoothly, OR, movement started not smoothly, but transitioned to end smoothly
   - No – Not performed smoothly throughout the entire movement

<u>Standing Jump</u>

Components:

1. Eyes looking forward?
2. Feet shoulder width apart?
3. Knees and hips bend as you lower into squat?
4. Arms swing back behind hips in the squat?
5. Arms swing forward smoothly and upward as the legs and feet push off the ground?
6. Arms reach high overhead at the highest point of jump?
7. Land on two feet with control?
8. Skill performed smoothly?

Scoring Criteria:

1. Eyes looking forward?
   - Yes – Eyes looking forward throughout the entire standing jump
   - Partial – Eyes looking forward for part of the skill
   - No – Eyes looking anywhere else except forward
2. Feet shoulder width apart?
   - Yes – Feet shoulder width apart
   - Partial – N/A
   - No – Feet not shoulder width apart (either wider than or less than shoulder width)
3. Knees and hips bend as you lower into squat?
   - Movement criteria:
     - (1) Knees bend
     - (2) Hips bend
     - (3) Lower into a squat
   - Yes – All three of the movement criteria are met
   - Partial – At least 1/3 of the movement criteria are met
   - No – None of the movement criteria are met
4. Arms swing back behind hips in the squat?
   - Yes – Arms swing back behind hips
   - Partial – Arms swing back, but don't go behind hips (arms remain anterior to hips)
   - No – Arms remain stationary
5. Arms swing forward smoothly and upward as the legs and feet push off the ground?
   - Movement criteria:
     - (1) Arms swing forward smoothly
     - (2) Arms swing upwards
     - (3) Legs and feet push off the ground
   - Yes – All three movement criteria are met
   - Partial – At least 1/3 of the movement criteria are met
   - No – None of the movement criteria are met
6. Arms reach high overhead at the highest point of jump?
   - Yes – Arms are fully extended and reach above their head at the highest point of jump

- Partial – Arms are bent at highest point of jump, but are still above the head
- No – Arms are bent at highest point of jump and remain below shoulder height

7. Land on two feet with control?
   - Yes – Maintain full control of both feet upon landing
   - Partial – Take a single step forward or back to maintain balance upon landing
   - No – Take more than one step forward or back to maintain balance upon landing, or, lose balance and fall down
8. Skill performed smoothly?
   - Yes – Performed smoothly throughout the entirety of the movement
   - Partial – Movement started smoothly, but transitioned to ending not smoothly, or, movement started not smoothly, but transitioned to end smoothly
   - No – Not performed smoothly throughout the entire movement

Appendix J: Data Collection Plan

Pre-Session

- Parental consent forms sent home
    - Signed and returned
- Students assigned into partner pairings
- Students assigned a study ID number

Session 1 (30 minutes)

- Consensus Training PowerPoint Presentation ~10 minutes
- Students paired up by researchers to practice taking proper videos of each other performing two movements – running and jumping ~10 minutes
    - Purpose of practicing both movements – allowed students to learn how to capture horizonal and vertical movement on video
    - Students were instructed to capture the video in landscape form, as this layout is preferred in the Move Improve® software
    - Reminded students of the importance of being able to completely see the entire student throughout the duration of the video recording
        - This would make it easier to evaluate when using the Move Improve® software
- "Lifting a Box" consensus session with student volunteers ~10 minutes
    - Two student volunteers were chosen from the group – one "evaluator" and one "performer"
    - After watching a demonstration video of an individual lifting a box, the "performer" attempted to replicate what they had seen, while the "evaluator" filmed them using the Move Improve® software
    - Recording was projected onto a television screen to allow the entire class to clearly see it
    - Using the Move Improve® software, the researchers led the students through the evaluation process of the student volunteer lifting a box
        - Made sure to discuss each component individually
        - Once a consensus had been reached, the researchers moved onto the next component, until a final performance score had been reached
    - To ensure that all students were actively engaged in the discussion process, they were required to interact with the Kahoot application on their tablets

Session 2 (30 minutes)

- To ensure consistency, students worked in their previously assigned pairings and used the same iPad they had used during Session 1
- Partners were instructed to fill in the "Data Collection Worksheet" with their total performance scores for each of the skills

  - In addition to this, all students were instructed to capture screenshots of the final performance score – this allowed the researchers to see the individual component scores, something that the "Data Collection Worksheet" lacked
- Students were instructed to complete the following fundamental movement skills by the end of the 30-minute class:
  - Standing Jump
  - Hollow Body Roll

Appendix K: Observational Notes from Data Collection

Day	Class + Time	Observations

Tuesday June 4th, 2019	Grade 5-8 Boys (8:45-9:15am)	- Had trouble concentrating on the PowerPoint presentation
		- No students were absent today
		- Difficulty following instructions
		    o Not all the pairings practiced filming the two movements (running and jumping), and instead chose a skill at random
		        ▪ Had to be re-directed and re-briefed on what the task was – focus on learning how to capture horizontal and vertical movements on video
		    o Several partner pairings had to be reminded that we were not using the Move Improve® software until next class
		    o To ensure that all students knew how to capture a screenshot of the final performance score, the researchers individually checked-in with each student pairing

	Grade 5-8 Girls (9:15-9:45am)	- Were more attentive throughout the PowerPoint presentation than the boy's class
		- One of the girls was feeling sick today, so she did not physically participate
		    o Sat on a bench on the sidelines
		    o Was given a tablet so she could participate in the Kahoot portion of the training session
		- Had less difficulty following instructions regarding using the iPads to video their partner running and jumping
		- Most pairings completed screenshots without difficulty
		    o To make sure that all students were comfortable completing screenshots, the researchers still checked-in with each partner pairing
		        ▪ This was important as it ensured consistency in training between both gym classes (boys and girls)

Thursday June 6th, 2019	Grade 5-8 Boys (8:45-9:15am)	- Students worked with the same partner and used their assigned iPad
		- Students were instructed to complete the “Data Collection Sheet” and capture screenshots of their final performance scores for the following Fundamental Movement Skills:
		  o Standing Jump
		  o Hollow Body Roll
		  o Leap
		  o Skip
		- Conflicts within partner pairings:
		  o Grade 5 boy got upset with his partner (Grade 7) for giving him an evaluation he did not agree with
		    ▪ Grade 5 boy = “performer” and Grade 7 boy = “evaluator”
		  o Ended up fighting and not wanting to be partners anymore
		  o I was the first to intervene but was initially unsuccessful in convincing them to stay partners, so I went and recruited my supervisor, Dr. Larry Katz, for additional assistance
		  o Dr. Katz managed to convince them to remain partners, even though the Grade 5 boy was still mad at his Grade 7 partner
		  o At this point they switched roles and the “performer” did a Hollow Body Roll
		    ▪ Grade 7 boy = “performer” and Grade 5 boy = “evaluator”
		  o Immediately after filming his partner, the Grade 5 “evaluator” ran to the bench on the side of the gym to complete the evaluation individually, even though they were supposed to be engaging in discussions throughout the peer to peer assessment
		    ▪ Out of spite, he gave his Grade 7 partner all “No” for each of the six components for Hollow Body Roll
		    ▪ Additionally, he completed the evaluation without looking at the video which meant that his evaluation was not based on the performance of his partner, but

		rather was fueled by anger and emotions. - Due to time constraints a complete data set for all of the skills was not attained for all partner pairings o Decided to use the "Standing Jump" and "Hollow Body Roll" data, since they were the only two skills that had complete data sets

	Grade 5-8 Girls (9:15-9:45am)	- Students worked with the same partner and used their assigned iPad
		- Students were instructed to complete the "Data Collection Sheet" and capture screenshots of their final performance scores for the following Fundamental Movement Skills:
		  o Standing Jump
		  o Hollow Body Roll
		  o Leap
		  o Skip
		- Saw that some of the girls felt uncomfortable being filmed while performing an assigned skill
		  o As the class progressed, most of the girls became more comfortable filming and evaluating their skills
		- Additionally, some girls did not like being evaluated by someone they did not consider a friend
		  o These pairings had less discussion throughout their peer assessment
		- It was also apparent that some of the partners were friends, as there was more discussion between them while they completed the evaluation
		- Due to time constraints a complete data set for all the skills was not attained for all partner pairings
		  o Decided to use the "Standing Jump" and "Hollow Body Roll" data, since they were the only two skills that had complete data sets