EHLERT, ALEX M., M.S. Evaluation of the Reliability of a Goalkeeper-Specific Adaptation To The Yo Yo Intermittent Recovery Test Level 1 (2017)
Directed by Dr. Allan Goldfarb. 74 pp.

The evaluation of physical fitness is important for top soccer clubs to guide fitness programs and match preparation. Field tests are commonly used by coaches and trainers as they are more convenient and require less equipment than laboratory testing. The most commonly utilized and studied field test for soccer fitness is the Yo-Yo Intermittent Recovery Test Level 1 (YYIR1), which has been consistently validated with match performance (Krustrup et al., 2005; Rampinini et al., 2007). YYIR1 test results have shown strong correlations with distance covered at high-intensity running speeds, an important measure of physical match performance (Krustrup et al., 2005). The YYIR1 has also been able to differentiate between skill levels and shown a gender difference similar to that seen in match play (Krustrup et al. 2003). The goalkeeper is an important position to match outcome but has been largely neglected in the study of physical match demands and fitness evaluation. Current literature has shown only a handful of attempts at tests using the actions of the goalkeeper position, with most of them focused more on cognitive, reactive, or technical aspects rather than evaluating the ability to meet the physical demands of match play (Knoop, Fernandez, and Ferruati, 2013). In addition, very few studies involving the physical demands of goalkeepers have included females. As the YYIR1 is the most popular and validated fitness test for soccer players, establishing a goalkeeper-specific adaptation to the test (YYIR1-GK)

could benefit coaches and trainers in terms of developing fitness programs and evaluating the ability to perform high-intensity actions common to the position.

The primary objectives of this study are 1) to evaluate the test-retest reliability of the YYIR1-GK in both male and female collegiate soccer goalkeepers, 2) to determine whether performance on the YYIR1-GK correlates with performance on a validated cycling test of repeated sprint ability, and 3) to analyze group differences in performance on both tests in terms of NCAA division level and gender. To accomplish these objectives, male and female NCAA goalkeepers of various division levels will be recruited for two identical visits during which they will complete both the YYIR1-GK and the 5x6 Second RSA Cycle Test to determine reliability and group differences on both tests.

Results from this study will be useful in determining if the GK-YYIR1 could be a convenient and reliable test of intermittent fitness using actions commonly seen in goalkeeper match play. It would also be among the first studies to include both males and females as well as various division level goalkeepers in terms of fitness testing and physical performance specific to the position.

EVALUATION OF THE RELIABILITY OF A GOALKEEPER-SPECIFIC

ADAPTATION TO THE YO YO INTERMITTENT RECOVERY

TEST LEVEL 1

by

Alex M. Ehlert

A Thesis Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2017

Approved by

____
Committee Chair

APPROVAL PAGE

This thesis written by Alex M. Ehlert has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair ______________________________

Committee Members ______________________________

____

____
Date of Acceptance by Committee

____
Date of Final Oral Examination

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Soccer is the world's most popular sport with estimates of around 200,000 professionals and 240 million recreational players (Junge & Dvorak, 2004). With this popularity, competitive soccer has become a major financial industry with an emphasis on finding and developing the best players possible (Reilly, Bangsbo, & Franks, 2000). The last couple decades have seen a massive surge in the acceptance of sports science as a method of optimizing training programs and match preparation, particularly in the field of exercise physiology (Carling, Bloomfield, Nelsen, & Reilly, 2008). It should be noted that performance in soccer depends on a variety of factors such as technical, tactical, psychological, and physiological (Stølen, Chamari, Castagna, & Wisloff, 2005), yet coaches recognize the importance of having a team that is physically capable of meeting the demands of both individual games, and entire seasons of competitive play at the highest level. Modern match analysis techniques have allowed sport scientists to evaluate the physical requirements of match play in order to determine the demands of various positions and style of play (Carling et al., 2008).

One major revelation from match analysis and sport science research with soccer is the intermittent nature of the sport. A large majority of the activity during a soccer match is considered primarily aerobic and field players typically cover around ten to twelve kilometers of total distance during a competitive match (Stølen et al., 2005). Yet despite a large amount of activity in an aerobic capacity, 1000-1400 high speed activity changes also occur, as well as a sprint activity performed every ninety seconds (Bangsbo, Norregaard, & Thorsoe, 1991; Reilly, 2007). In addition to the high number of intense actions, the highly anaerobic periods of play are typically some of the most important and decisive in the match (Stølen et al., 2005). When looking at physiological values, the intensity of match play is often found to average around 80-90% of HR_{max}, which is considered close to the anaerobic threshold (Stølen et al., 2005). But Stølen et al. (2005) mentioned that despite an average around the anaerobic threshold, the player is constantly balancing short, intense actions such as sprints or changes of direction with slightly longer periods of walking and jogging in order to recover and position for the next play.

The goalkeeper position is also considered intermittent, but with different action requirements than field players due to their need to perform explosive movements within a small area (Ziv & Lidor, 2011). Match analysis has shown that professional goalkeepers were required to make a high-intensity displacement of about four meters at an occurrence of about one per minute (Padulo, Haddad, Ardigo, Chamari, & Pizzolato, 2015). Further analysis of the 2002 World Cup also found an average of 23.4 defensive technical actions in response to opposing attacks (Sainz De Baranda et al., 2008). Between these high-intensity actions, the majority of the movement of goalkeepers is at a low intensity (Di Salvo, Benito, Calderon, Di Salve, & Pigozzi, 2008), indicating that goalkeepers also operate in an intermittent manner, alternating between high-intensity actions and periods of passive or active recovery.

Performance in soccer is difficult to predict due to the complexity of the sport and the variety of subjective factors that play a role (Reilly, 2003). Yet there is a tremendous advantage in determining the specific physical demands of competitive match play and evaluating the ability of players to meet them through the use of various fitness and physiological tests. There have been a variety of studies using laboratory tests to evaluate typical physiological measures such as maximal oxygen consumption, muscular power, and blood lactate concentrations of soccer players (Rampini et al., 2007; Stewart, 2002; Balsom, 1994; Svenson & Drust, 2005). Yet these procedures require expensive equipment, qualified professionals to run the tests, take up valuable training time, and may not have sufficient ecological validity to match performance (Balsom, 1994; Stølen et al., 2005). To combat these issues, a large variety of field tests have been developed. Field tests have the benefit of being inexpensive, easy to operate and less time-consuming for coaches and trainers. The validity of many of these field tests are based off of logical validity with very little direct evidence of improved performance except for the Yo-Yo Intermittent Recovery Tests that are among the most studied and utilized fitness tests in the world (Rampinini et al., 2007).

The Yo-Yo Intermittent Recovery Test Level 1 (YYIR1) was created to fill the need for a fitness test that could easily and accurately measure intermittent exercise fitness in team sports athletes (Bangsbo, Iaia, & Krustrup, 2008). The test had originally started as a continuous fitness but the two levels of the intermittent test were later added for use with intermittent sport athletes. (Bangsbo, 1994). It is still one of the most used and studied tests in the world due to findings of excellent reliability as well as validity in terms of predicting physical high-intensity performance (measured in distance covered at high-intensity) during match play and correlations to important physiological factors (Rampinini et al., 2007; Krustrup et al., 2003). For example, strong significant correlations between YYIR1 and maximal oxygen consumption have been reported (Krustrup et al., 2003; Thomas, Dawson, & Goodman, 2006), which is considered one of the most important physiological factors for aerobic and intermittent physical performance (Stølen et al., 2005). Additionally, YYIR1 results have shown strong correlations with distance at high-intensity running, sprinting, and total distance covered in match play in both males (Krustrup et al., 2003; Rampinini et al., 2007) and females (Krustrup, Mohr, Ellingsgaard, Helga, & Bangsbo, 2005). It is also considered a sensitive enough measure to detect seasonal and training-induced changes (Krustup et al., 2005). Due to the widespread use of the YYIR1 to evaluate the ability to perform intermittent exercise by many top soccer clubs, there could be a benefit to evaluating a goalkeeper-specific adaptation to the test. The goalkeeper actions are intermittent in nature, but do not resemble the activity of field players (Padulo et al., 2015). Goalkeepers only cover about half the total distance that a typical field player might (Di Salvo et al., 2008) but they also perform a large number of explosive actions while confined to a small space (Ziv & Lidor, 2011). When analyzing the high-intensity actions of goalkeepers, it has been found that they perform about ninety-two explosive displacements of about four meters per match (Padulo et al., 2015). These actions are typically moving laterally across the goal line or forward and back in order to obtain optimal position for oncoming attacks or to cut off the angle of an attacking player (Padulo et al., 2015). Rarely do goalkeepers perform all-out sprints of any appreciable distance, except to occasionally clear out a crossed pass across the goalkeeper area or to come out to challenge an attacking player in a one-on-one situation (Aziz, Mukherjee, Chia, & Teh, 2008). The only attempts at goalkeeper-specific tests have either been developed for technical aspects (Rebelo-Goncalves, Figuierdo, Coelho Silva, & Tessitore, 2016), or cognitive and perceptual movement tests (Knoop, Fernandez, & Ferrauti, 2013), with little focus specifically on the intermittent fitness requirements of the position. A goalkeeper-adaptation of the YYIR1 for intermittent fitness would therefore focus on side-to-side lateral movement as well as forward and backwards displacements of around four meters in the confined space that the position operates in (Padulo et al., 2015).

One other measure of intermittent physical performance that is gaining popularity with team sport athletes is repeated sprint ability (RSA). RSA revolves around the ability to generate maximal or near-maximal actions repeatedly with brief periods of recovery (Bishop, Girard, & Mendez-Villanueva, 2011). Soccer players must utilize actions of an intermittent nature, including a variety of explosive actions and sprints (Stølen et al., 2005), so there is possibly some utility in testing RSA in the soccer athlete. There have been a few positive results when correlating RSA test performance to physical match action in elite soccer players (Rampinini et al., 2007) as well as findings of usefulness in other intermittent sports (Spencer et al., 2004). But many of the running sprint-oriented tests of RSA would not translate well when adapting to goalkeeper action. Therefore any tests of RSA with goalkeepers should involve short, explosive bursts of actions such as the reliable and validated 5x6 Second RSA cycle test (Bishop, Spencer, Duffield, & Lawrence, 2008; Fitzsimons, Dawson, Ward, & Wilkinson, 1993), with the goal of measuring RSA as a general physiological quality to compare to match performance or commonly used tests such as the YYIR1.

Current gaps in the literature call for a need to evaluate a fitness test specifically designed for goalkeeper activity. The commonly-used fitness tests with elite soccer players do not correlate to the explosive displacements and technical actions of the goalkeeper position. Prediction of goalkeeper performance is nearly impossible due to the complexity of the position, but an objective test of fitness related to the actions required in match play should be useful for coaches in guiding training programs and observing physical match preparation. There is also a nearly complete disregard for the physical demands of female goalkeepers, with all of the studies on the actions and match analysis of goalkeepers recruiting males. Therefore there is a need for studies including the female goalkeeper in terms of physical fitness and demands of this position.

The purpose of this study was to (1) evaluate the reliability of a goalkeeper-specific adaptation to the Yo-Yo Intermittent Recovery Test Level 1 (YYIR1-GK) using NCAA collegiate goalkeepers, (2) to correlate results on the YYIR1-GK to performance on the validated 5x6 Second cycle RSA test, and (3) to determine performance differences on both tests between various levels of NCAA goalkeepers and gender. Based on the cited literature, the following hypotheses were constructed:

Hypothesis 1: The Goalkeeper-specific adaptation of the YYIR1 will prove to have strong-to-excellent test-retest reliability with the NCAA goalkeeper population.

Hypothesis 2: There will be a negative correlation between the distance covered on the YYIR1-GK and the rate of fatigue development on the 5x6 Second cycle RSA test, indicating a relationship between the two.

Hypothesis 3: The Division 1 goalkeepers will complete more stages on the YYIR1-GK than lower Division players of the same gender, indicating a superior intermittent fitness level in actions normally performed during match play.

Hypothesis 4: The male NCAA goalkeepers will complete more stages on the YYIR1-GK than the female goalkeepers of the same division and on average as a whole, showing similar gender gaps as those seen in the field players.

CHAPTER II

REVIEW OF THE LITERATURE

Overview

This literature review will briefly discuss the use of sport science in competitive soccer to determine the physical and physiological requirements of match play. Next, the review will cover what the research has shown in terms of physical demands and fitness levels of competitive soccer players through match analysis and how this has led to the development and evaluation of various fitness tests. A specific focus will be on one test known as the Yo-Yo Intermittent Recovery Test Level 1 which is one of the most studied and widely utilized fitness tests in the world. Additionally, this literature review will discuss the lack of scientific study on the specific physical demands of the soccer goalkeeper, despite their massive importance in the outcome of matches. Finally, this review will highlight the need for a goalkeeper-specific fitness test to evaluate their physical ability in order to guide training programs and match preparation for an often neglected soccer position in sport science research.

Soccer and Sport Science Research

Soccer is the most popular sport in the world with estimates of about 200,000 professional and 240 million amateur players (Junge & Dvorak, 2004). As a result, competitive soccer has become a massive financial industry with an emphasis on finding and developing the best possible players (Reilly et al., 2000). In an attempt to maximize performance, soccer coaches and trainers have been very accepting of the use of scientific research and analysis in order to guide their individual player and team preparations (Carling et al., 2008). In particular, the soccer industry has embraced the field of exercise physiology and the information it can provide coaches on the specific demands of match play and the best methods to optimize fitness development (Williams & Hodges, 2005; Carling et al., 2008). Soccer performance is complex and relies on the combination of a variety of factors beyond just physical fitness; including tactical understanding, technical skill, and psychological factors (Stølen et al., 2005). Due to the complexity of the sport, talent evaluation relies on various subjective as well as objective aspects (Reilly, 2003). A player is considered to be prepared for a match when they can meet the demands required in each aspect of game play, with technical, tactical, and behavioral factors often better evaluated in a subjective manner (Reilly, 2007). Despite the importance of subjective evaluation, the fitness capacity of a soccer player or team can play a major role in the outcome of a match (Bangsbo et al., 2008) and it is much easier to objectively evaluate physiological and fitness measures than behavioral, tactical, or technical aspects (Williams & Hodges, 2005). For this reason, research in the field of exercise physiology has commonly been used by soccer coaches and trainers to determine the specific physical demands and guide the training for the players (Carling et al., 2008).

Most of the studies of the physical demands required for soccer have focused on match analysis. Early studies involved taking basic physiological measures such as heart rate and lactate pre-game, at halftime, and post-game to estimate match demands, but more recent studies have observed demands throughout the match with a keen interest in the most demanding and decisive periods of play (Bangsbo et al., 2008). Current motion systems can analyze a variety of players and provide a pool of data on the speed of movement, distance covered, and the various actions performed that can be used to determine individual work profiles (Carling, Williams, & Reilly, 2005). This has shifted the focus from general demands of the sport to a more specific analysis based on the player's training status, position, and tactical role, resulting in a more individualized training approach (Bangsbo et al., 2008).

Match Analysis

A common method of match analysis is to observe the distance covered at various intensities categorized by the speed of the actions. The average distance covered by an elite male field player is usually in the range of ten to twelve kilometers, with midfielders typically covering the most total distance (Stølen et al., 2005). Of that total distance, it has been found that 58.2-69.4% is covered by walking or jogging, and another 13.4-16.3% at a low speed run, while high-speed runs and sprints constituted 3.9-6.1% and 2.1-3.7% respectively (Di Salvo et al., 2007). There is also typically a decrease in high-intensity action as the second half progresses, indicating the presence of fatigue in field players (Mohr, Krustrup, & Bangsbo, 2003). Observing the distance performed at high-intensity is considered an important measure of physical performance for soccer players (Sporis, Jukic, Ostojic, & Milanovic, 2009).

Studies with elite female soccer players have yielded similar results in terms of total distance, with an average of ten to twelve kilometers covered and a similar number of activity changes per match. But they covered only about two-thirds of the distance at high-intensity compared to what is normally reported in elite men (Krustrup et al., 2005). In fact, the difference in distance at high-intensity is seen as the biggest observable gender difference since males and females seem to tax the aerobic and anaerobic systems to a similar degree (Helgerud, Hoff, &Wisloff, 2002). The gaps between elite male and female soccer players in studies evaluating physiological and physical performance are similar to the gap seen in their sedentary counterparts (Stølen et al., 2005), indicating that the female soccer athlete has likely developed their fitness to a similar extent. Yet the total number of studies performed with females is low compared to the number performed with men.

Despite the abundance of motion analysis data available on field players, very little has been performed on the work profiles of goalkeepers. One study found that goalkeepers covered about 4,000 meters per match (Reilly & Bowen, 1984). A more recent motion capture study using the Prozone analysis system observed sixty-two English Premier League goalkeepers over 109 matches and found an average distance covered of 5,611 meters per match while spending only 2% of the match running at high-intensity and 73% at a walk (Di Salvo et al., 2008). One issue related to the use of the Prozone system is that it requires a velocity to be maintained for at least 0.5 seconds to classify a movement into a category of running intensity (Di Salvo et al., 2008). Similar to the Prozone system, many of the motion analysis systems used previously in the study of soccer match demands are relatively accurate when the speed of an action is constant and in a straight-line but some issues arise with sudden changes of speed or direction (Witte & Wilson, 2004). This suggests that the Prozone system is likely unreliable in creating an accurate work profile for goalkeepers as many of their actions occur in a small space and consist of quick, explosive actions that are rarely linear (Padulo et al., 2015).

A more recent study decided to utilize a different motion analysis system in an attempt to accurately discern the actions of high level goalkeepers. The study used a variety of high frequency digital cameras over ten official matches with ten professional, Italian goalkeepers and analyzed the video with pro-motion software to calculate high-intensity goalkeeper displacement (Padulo et al., 2015). Padulo et al. (2015) found an average of 92 high intensity actions (52 forward and 40 lateral) during a 90 minute match, or an average of about one per minute. The mean distance covered for each one of the displacements was about four meters and the higher category goalkeepers performed more total forward displacements, with each being at a higher velocity than the lower category goalkeepers (Padulo et al., 2015). The authors speculated this increased velocity to be due to the best goalkeepers being more effective at "closing out the space" to cut off the angle of an attacking opponent (Padulo et al., 2015).

One other study analyzed match play of thirty-four goalkeepers during fifty two matches of the 2002 World Cup with a focus on the actions performed protecting the goal and attempting to regain ball possession from the opponent (Sainz De Baranda et al., 2008). They found an average of 26.3 opposing attacks requiring an average of 23.4 technical actions that could include making a save, foot control of the ball, clearing out a crossed ball, and deflections among others (Sainz De Baranda et al., 2008). Outside of these few studies, very little match analysis has been performed with goalkeepers, and none of the major studies used females.

Soccer as an Intermittent Sport

Traditionally the physical activity of soccer was viewed differently than it is today since aerobic metabolism is the primary energy pathway for about 90% of a soccer match (Stølen et al., 2005). This is further supported by findings that top players typically cover about ten to twelve kilometers per match and often have above average $VO2_{max}$ values; with male field players having a typical range of 50-75 ml/kg/min and females around 39-57 ml/kg/min (Stølen et al., 2005). In contrast, goalkeepers have not been studied as often as field players. What has been reported is that goalkeepers only cover around five kilometers a match and have $VO2_{max}$ in the range of 50-55 ml/kg/min (Stølen et al., 2005). There has been very little reported data on work profiles or characteristics for female goalkeepers. Yet despite a majority of the activity being aerobic, it is typically the actions dictated by anaerobic metabolism that serve as the most important actions in the match (Stølen et al., 2005). It was reported that a two to four second sprint was performed on average every 90 seconds, with an additional 1000-1400 high speed activity changes throughout a match, correlating to about one every four to six seconds (Bangsbo et al., 1991; Reilly, 2007). Soccer match play can be said to have an ever-changing intensity due to the constant need to change speed and perform high-intensity actions such as runs, jumps, and kicks, while also requiring periods of low intensity to recover and position for the next high-intensity action or play (Bangsbo & Michalski, 2002). Given this information, it has become evident that soccer is an intermittent sport that relies on the ability to repeatedly perform high-intensity actions to succeed (Krustrup et al., 2003).

Various physiological measures have been used to calculate game intensity. Heart rate values during match play are on average around 80-90% of HR_{max} which is considered close to the anaerobic threshold, yet due to the intermittent nature of the sport, the general demands of the game typically rely on bouts above the anaerobic threshold followed by periods well below it to recover from the high-intensity action (Stølen et al., 2005). Values of oxygen uptake have been unreliable at best and are largely considered underestimated due to the assumption that equipment inhibited optimal performance.

Goalkeeper action is considered intermittent as well, requiring high levels of agility while operating in a small space with rapid movements to react to the play (Ziv & Lidor, 2011). As noted before, ten matches in Italian professional league play required goalkeepers to make high intensity actions at an average of about one per minute (Padulo et al., 2015) as well as another study using the 2002 World Cup finding an average of 23.4 defensive technical actions in response to opposing attacks (Sainz De Baranda et al., 2008). However, movement analysis of English Premier League goalkeepers found that 73% of their movement was performed at a walk (Di Salvo et al., 2008). This information would indicate that the high level goalkeeper must perform rapid, explosive actions while also having brief periods of recovery when they are able to walk. In comparison, field players are performing more prolonged, but less intense actions with slightly shorter periods of recovery between each action. Both goalkeeper and field player activity would be classified as intermittent.

The Use of Field Tests for Intermittent Sports

The coaching process in soccer is aided by the ability to evaluate performance in different areas in order to implement the best training strategies (Carling et al., 2005). For that reason, the study of the physical demands through match analysis has led to the development and utilization of a variety of lab and field tests to attempt to predict performance and guide training practices (Krustrup et al., 2005). Lab tests involved evaluating elite soccer players in traditional physiological factors such as maximal oxygen consumption, heart rate responses, and blood lactate. For example the $VO2_{max}$ values for elite male field players typically reside in the range of 50-75 ml/kg/min, while females have been found to be 38.6-57.6 ml/kg/min (Stølen et al., 2005). Other variables have been commonly measured with soccer players such as muscular power and repeated sprint ability (Rampinini et al., 2007; Stewart, 2002; Balsom, 1994). Though lab tests are the ideal method for obtaining reliable scientific data due to the ability to control many extraneous factors, they can often require expensive equipment, training to operate equipment properly, and are often inconvenient for repeated use of entire teams. Therefore field tests have been used to evaluate physical fitness with soccer teams for years because they are inexpensive, convenient, and usually easy to perform by coaches and trainers (Krustrup et al., 2005). Many field tests originally used continuous exercise, such as the twelve minute run and twenty meter shuttle run tests, but this type of test has been questioned due to the intermittent nature of soccer match play (Bangsbo et al., 2008).

The need for a field test that could evaluate the intermittent physical fitness of team sports athletes led to the development of the Yo-Yo Intermittent Recovery Tests, which are now among the most studied and used fitness tests in the world (Bangsbo et al., 2008). The tests involve stages of two 20 meter shuttle runs, with ten second active recovery and the time to complete each stage progressively decreasing (Bangsbo et al., 2008). The goal of the two versions of the Yo-Yo Intermittent Recovery Tests (YYIR) is to evaluate an athlete's ability to repeatedly perform intense exercise with brief time for recovery (Krustrup et al., 2003). There has been concern about how accurately field tests predict on-field performance, yet the YYIR tests, in particular level 1 has exhibited excellent reliability and multiple sources of validity (Rampinini et al., 2007). For example, the YYIR level 1 (YYIR1) has been found to have good test-retest reliability, with an average intra-individual difference of 4.9% (Krustrup et al., 2003).

In addition to reliability, there has been a variety of correlations between factors that are thought to improve match performance. The YYIR1 was found to have a significant correlation with maximal oxygen uptake ($r=0.71$, $P<0.05$), an important physiological factor for soccer match performance and general aerobic capacity (Krustrup et al., 2003). Another study found excellent test-retest reliability with the YYIR1 ($ICC=0.86$-0.95) while also finding a strong correlation with maximal oxygen consumption ($r=0.87$, $P<0.01$) in recreational athletes (Thomas et al., 2006). Maximal oxygen consumption is considered one of the most important physiological factors for aerobic performance and possibly the recovery from intense activity in soccer (Stølen et al., 2005), offering evidence that the YYIR1 is valuable in terms of the estimation of aerobic capacity for the soccer athlete. YYIR1 performance was also found to have significant correlations with the amount of high-intensity running ($r=0.71$, $P<0.05$), sprinting ($r=0.58$, $P<0.05$), and total distance covered ($r=0.53$, $P<0.05$) during actual match play while sensitive enough to detect changes in fitness during the competitive season (Krustrup et al., 2003). Similar results have also been found with elite female soccer players, with correlations between YYIR1 performance and high-intensity running ($r=0.76$, $P<0.05$) as well as total distance covered ($r=0.53$, $P<0.05$) during soccer matches (Krustrup et al., 2005).

More evidence of the YYIR1's ability to predict performance is seen when comparing genders and high-intensity running activity. The distance a player runs at a high intensity is considered more important than the total distance covered during a match (Sporis et al., 2009). Elite female soccer players were found to cover similar total distances as their male counterparts but covered significantly less at a high-intensity (Krustrup et al., 2005). Comparing male and female performance on the YYIR1 showed that the test was able to differentiate between male and female due to the gender gap being similar on YYIR1 performance and in physical match performance in terms of high-intensity running (Mujika, Santisteban, Impellizzer, & Castagna, 2009). The results also found significant differences between senior and junior level players in both males and females, indicating that the YYIR1 was able to differentiate between skill levels as well (Mujika et al., 2009). Various sources of evidence highlighting the ability to predict distance covered at high-intensity and differentiate between typical gender and skill level differences in physical match performance indicates that the YYIR1 is a reliable and valid measure of intermittent fitness for both male and female soccer players (Krustrup et al., 2005). It also has a high enough sensitivity to detect seasonal changes and the effects of training programs, making it a useful test to track progress at various points of time (Krustrup et al., 2003).

Repeated Sprint Ability

One measure of physical performance that is becoming popular to test with intermittent sport athletes has been labeled repeated sprint ability (RSA). Team sports such as soccer require the athlete to repeatedly produce maximal or near-maximal actions followed by brief periods of recovery throughout the time of a match (Girard, Mendez-Villanueva, & Bishop, 2011). Current research has spent significant effort evaluating intermittent sport athletes and their ability to repeat high intensity actions with various recovery durations in an attempt to evaluate the intermittent physical requirements of their sport (Bishop et al., 2011). The literature seems to acknowledge that there is currently no “gold standard” for RSA (Green & Dawson, 1993), but there have been some promising results with various tests using intermittent sport athletes (Aziz et al., 2008).

One study found a moderate correlation between the results of a shuttle run RSA test and high-intensity distance covered in elite soccer match play (Rampinini et al., 2007), and a 6x4 second repeated sprint test was considered a useful method for assessing field hockey fitness, which has been noted to have similar intermittent demands as elite soccer (Spencer et al., 2004). Various tests of RSA are regularly incorporated fitness measures with soccer teams (Rampinini et al., 2007; Stewart, 2002; Stølen et al., 2005) but there have been concerns about which RSA tests are best suited for specific athletes. For example, a study comparing soccer players by age, position, and competitive level using a 20-meter sprint RSA test found that the highest level professional players performed best while the lower level amateur players performed worse, with goalkeepers having poorer results than players in the field (Aziz et al., 2008). First, this suggested that the running RSA test was able to discriminate between players of different levels. However, it is likely not a valid measure of goalkeeper intermittent activity due to the rarity of sprint performance by that position (Aziz et al., 2008).

There has not been an established RSA test that would mimic goalkeeper activity but the best method would be one with actions that are of short duration, with the goal of measuring general RSA rather than sport-specific. For example, the most commonly used anaerobic test is the 30 second Wingate Cycling Test (Granier, Mercier, Mercier, Anselme, & Prefaut, 1995). Yet, there was found to be an aerobic component to the standard Wingate cycling test, and goalkeeper actions are very short, explosive motions, indicating that a shorter protocol would be better (Granier et al., 1995). One example of this could be the 5x6 second RSA cycling test that has been found to be both reliable and valid with intermittent activity (Bishop et al., 2008; Fitzsimons et al., 1993). One discussion has been that RSA could be a general quality in team sports athletes (Bishop, Spencer, Duffield, & Lawrence, 2001), indicating that an athlete fit for their intermittent sport would perform well on most RSA tests. If this is true, there could be an advantage in comparing RSA test results to more specific intermittent fitness tests such as the YYIR1 with intermittent athletes.

Physical Demands of the Goalkeeper

Despite the massive amount of research performed with the sport of soccer and the importance of the goalkeeper to match outcome, a majority of studies on the physical demands and fitness tests have been tailored towards field players. For example, a large majority of studies that included goalkeepers were comparing the results on field test measures such as aerobic fitness, repeated sprint ability, and countermeasure jump by position, which do not translate well to goalkeeper-specific actions and physical demands (Al-Hazzaa et al., 2001; Knoop, Fernandez, & Ferrauti, 2013). Additional studies targeting goalkeepers did not specifically target their physical demands, but rather focused on various other isolated parameters such as injuries (Luthje & Nurmi, 2002; Mihalik, Myers, Sell, & Anish, 2005; Narayana, Josty, & Dickson, 2000), situational cognitive factors (Wood & Wilson, 2010; Kerwin & Bray, 2006) or discussions of tactical strategies and training (Mulqueen & Woitalla, 2011; McMorris & Hauxwell, 1997). There are also a variety of handbooks and guides on technical skill development of the goalkeeper (Rebelo-Goncalves et al., 2016) and studies on the biomechanics of the more technically demanding movements of the position (Sorensen, Ingvaldsen, & Whiting, 2001; Matsukura & Asai, 2013).

The area that has been all but ignored is the study of goalkeeper-specific fitness requirements as well as the development of goalkeeper-specific fitness tests to evaluate movements that are common to the position. This is surprising due to the large amount of research devoted to analyzing the physical demands and fitness tests for field players (Rampinini et al., 2007; Krustrup et al., 2005). There have been a couple of goalkeeper-specific tests developed in recent years, with the most recent targeting the evaluation of technical skills (Rebelo-Goncalves et al., 2016). The others attempted to produce a test that evaluated the perceptual and movement responses of goalkeeper action, known as the Reaction and Action Speed (RAS) test (Knoop et al., 2013). This test was successful at discriminating between different age-matched skill levels and identifying a variety of weaknesses and strengths of the various goalkeepers but it was complex and would likely run into issues of convenience as a regularly-used field test. The authors even discussed that despite their best efforts, it is nearly impossible to create a test that truly evaluates the skill of a goalkeeper, and making the test any more complex than it was already would result in poor reliability (Knoop et al., 2013).

Prediction of performance in soccer is a difficult task compared to some other sports and activities due to the complexity of match play and the variety of factors that can affect the results, some of which can only be reliably evaluated subjectively (Reilly, 2003). Goalkeeper performance is determined by many factors that are extremely difficult to test in a reliable fashion such as tactical understanding, perception, situational awareness, and anticipation (Knoop et al., 2013). The advantage of field tests is that they are typically inexpensive, simple to operate and can be performed conveniently throughout the season (Krustrup et al., 2005). The most effective fitness tests are designed to evaluate one specific measure or a general foundation of fitness, such as the YYIR1 that evaluates the ability to perform intermittent exercise at high-intensity (Krustrup et al., 2003). The YYIR1 is a simple fitness test that targets the general activity patterns of team sports athletes, especially field soccer players rather than attempting to encompass all of the various aspects of match play. In order to properly evaluate goalkeeper-specific fitness, it is important to understand the physical actions goalkeepers actually perform, as most current tests do not address the physical demands of the position.

The first motion analysis study with goalkeepers found an average total distance covered of around 4,000 meters (Reilly & Bowen, 1984). A more recent study using the popular Prozone match analysis system found an average of 5,611 meters covered with only about 2% being at high intensity, and 73% performed at a walk (Di Salvo et al., 2008). Prozone analysis software requires velocity to be maintained for at least 0.5 seconds to categorize a movement (Di Salvo et al., 2008), and does not include energy-demanding actions like rapid accelerations, jumping, and change of direction into the analysis (Bangsbo, Mohr, & Krustrup, 2006). Goalkeepers typically perform highly explosive movements such as jumping and diving along with rapid displacements within a small space and only occasionally engage in all-out sprints of any significant distance (Aziz et al., 2008). This would indicate the Prozone system is inadequate for providing a true measure of the various high-intensity actions a goalkeeper is expected to perform.

Therefore, a more recent study focused on in-depth analysis of goalkeeper-specific actions and found that professional goalkeepers performed an average of 92 high-intensity actions, each having a mean distance of about four meters (Padulo et al., 2015). This corresponds to about one high intensity displacement (about four meters either forward, back, or laterally) per minute. Elite goalkeepers during the 2002 World Cup were also found to perform an average of 23.4 defensive technical actions per match, which could include saves, deflections, ball control, dives, and a variety of other activities (Sainz De Baranda et al., 2008). This information suggests that the goalkeeper is performing high-intensity displacements at an average rate of about one per minute while also performing a variety of defensive technical actions interspersed throughout the match. In-between these explosive bouts of action are typically periods of low-intensity or passive recovery that classify the position as highly reliant on intermittent exercise.

Gaps in the Literature

The sport of soccer has been heavily studied in terms of sport and exercise science yet a vast majority of the studies focused on field players. Most studies involving the evaluation of physical demands of match play or fitness testing have either excluded the goalkeeper or compared them to field player performance in tests that do not translate well to goalkeeper-specific action. There have only been a few authors who have attempted to quantify the physical demands of goalkeepers during match play, with only the most recent using analysis techniques that can accurately depict the rapid, explosive actions of the goalkeeper in the small space that they typically perform their most important plays. The goalkeeper plays a major role in the most decisive defensive plays of a match, yet there has been very little quality research on the physical aspects of the position. The literature also has very few examples of fitness tests that include the specific actions that a goalkeeper is required to perform. Many of the fitness tests that have been studied and used the most by soccer coaches such as various sprint tests have little to no carryover to the position resulting in very little practical significance to goalkeepers. Additionally, none of the match analysis studies have been performed with female goalkeepers of any level.

The tests that have been developed to evaluate the goalkeeper position have often revolved around technical skills or observing cognitive reactions of the goalkeeper in complex fashions. The ability to reliably test and predict goalkeeper performance is challenging due to the many factors that play a role in goalkeeping skill such as anticipation, situational awareness, positioning, and technical skills. With a position as complex as the goalkeeper, it is necessary to understand that subjective analysis by coaches and trainers will play a large role in talent identification due to the difficulty of objectively testing such factors.

The advantage to field tests is the ease of use, the lack of expensive equipment required or the trained professionals to operate it, and the ability to frequently test players to determine any seasonal changes or effects of training programs. Previous studies have repeatedly evaluated various fitness tests for field players, with the Yo-Yo Intermittent Recovery Test level 1 being viewed as one of the most effective due to its strong correlations with important factors of match play such as high-intensity distance covered and maximal oxygen consumption. But no studies have evaluated a field test designed to assess the goalkeeper's fitness using actions specific to the position in the small space that they operate. A reliable test of goalkeeper-specific fitness will not predict performance in match play but will give coaches and trainers valuable information about the training status of the goalkeeper using actions that are commonly seen during match play, in order to best guide training and match preparation.

CHAPTER III

OUTLINE OF PROCEDURES

Participants

Thirteen total NCAA collegiate goalkeepers, male (n=7) and female (n=6) were recruited for this study. All goalkeepers met the inclusion criteria, including currently being between 18-32 years of age with four years of goalie experience, currently on a NCAA Division 1, 2, or 3 soccer roster, no injuries that would prohibit intense exercise participation, and no history of cardiovascular events or disorders. Prior to their inclusion as subjects they were given the consent form and were asked questions about the study. Upon agreement to participate, the consent form was signed. The consent form was approved by the Institutional Review Board (IRB) at UNCG. They also filled out the AHA-ACSM health history screening questionnaire form and activity form to ensure they had no known injuries that might impact their participation or any health-related issues.

Instrumentation

Equipment required for baseline measures included a sphygmomanometer and stethoscopes to obtain baseline blood pressures, a Seca weight scale to measure weight (to the nearest 0.1 kg), a stadiometer to determine height (within 0.1 cm). Cosmed Bod Pod body composition system (Rome, Italy) using the Siri equation for body composition (Siri, 1961). Barometric pressure (to the nearest 0.01 mmHg), and room temperature (to the nearest 0.1 °C) was measured using a Davis indoor climate monitor (Hayward, California). For the performance of the YYIR1-GK fitness test, equipment used included an audio file containing the YYIR1 test, an audio device to broadcast the file, four cones per participant, a chart of rating of perceived exertion (RPE) using the Borg 6-20 scale (Borg & Kaijser, 2006) and a Polar heart rate monitor (Kempele, Finland) placed snugly around the chest to continuously monitor and store heart rates which were downloaded for further analysis. For the 5x6 Second RSA cycle test, an electronically braked Lode cycle ergometer (Groningen, Netherlands), a Borg scale of RPE, and a continuously monitoring heart rate monitor as noted above were used.

Procedure

Each subject participated in two testing sessions three to seven days apart. Before the first visit, each participant was told to take the three days leading up to the visit as though they are preparing for a match in terms of hydration and nutrition. They were also asked to refrain from intense physical activity in the 24 hours prior to each visit. Food logs were provided prior to each visit and they were instructed on how to fill out the forms to record nutritional information on the three days leading up to the visit. They were given a copy of the nutrition logs and asked to replicate the diet on the next visit. The subjects were asked to wear the same shoes for both sessions. Upon arrival for the first visit, participants were screened and signed the consent forms before proceeding on with the testing session. The two testing sessions were identical in structure with the exception of Bod Pod measures which were taken only during visit one. Additionally, the internal environment of the testing locations was monitored to ensure consistency in terms of barometric pressure, temperature, and humidity.

Testing Protocol

Baseline measures were taken at the beginning of each testing session. This included measurements of height, weight, resting blood pressure, resting heart rate and a Bod Pod body composition analysis. The participant was then fitted with the heart rate monitor with a chest strap to allow for continuous monitoring, with data being downloaded onto a computer for further analysis at a later time. There was a five minute period for collecting resting heart rate using the R-R method prior to any exercise test. The subject remained seated and relaxed and was asked not to move or talk while breathing normally. They were then put through a standardized soccer warm up that involved three minutes of low intensity jogging, side shuffling, and back pedaling followed by about nine minutes of dynamic flexibility movements and a standardized familiarization warm up to the testing protocol. When the warm up was complete, the YYIR1-GK test began.

YYIR1- GK.

The YYIR1-GK test was designed to be similar to a graded exercise test with a short period of recovery between stages and the speed of actions gradually increasing as the test progresses. The test was guided through the use of a digitized audio file of the YYIR1. The standard YYIR1 involves field players performing runs of 20 meters before an audible beep, and then returning the 20 meters back to the starting line before a second audible beep. The YYIR1-GK used the same audio file and format with a few differences in protocol. There were two cones set up parallel to each other, creating a "starting line" on which the goalkeeper began each stage. Two additional cones were set up four meters away, forming a second line to which actions were performed. Each stage of the test required the goalkeeper to perform four different actions, each requiring four meters of movement; (1) side-shuffle (left or right), (2) side-shuffle back to the starting line; (3) backwards run; and (4) forwards run back to the starting line. Each stage therefore covered 16 meters of total distance between the two sets of cones. During an action, the goalkeeper must touch the line with their foot before proceeding on to the next action. The four actions were completed successively before the first audible beep from the audio file. They then stood for passive recovery through the next audio beep (during which field players would be performing their second 20 meter run), and then had an additional ten second of passive recovery before the next stage commenced.

Field players typically perform a 20 meter run in this time that is designed for the GK to perform the 16 meter action sequence. In addition, the field players would then have to perform a second 20 meter run back to the starting line before the second beep. Goalkeepers rested through this second period of activity and then joined the field players in the normal ten second recovery period, leading to greater periods of rest than the field players. The time for passive recovery gradually decreases as the time given to complete each stage decreases. There were no field players performing this test simultaneously in this study so the goalkeepers were continuously prompted on when to rest and when to start their next 16 meter action sequence.

The test progressed in this manner until two failures were recorded or until volitional exhaustion was reached by the participant. A failure was recorded when the goalkeeper was unable to perform the 16 meter sequence in the required time. They were warned of their first failure and the second failure resulted in termination of the test and the last stage completed successively recorded as their score. Throughout the test, heart rate was monitored continuously and rating of perceived exertion (RPE) was obtained by asking each participant their rating upon completion of each stage and immediately after termination of the test.

5x6 Second RSA Cycle Test

Upon completion of the YYIR1-GK test, the participant was allowed a 30 minute cool down recovery period during which they were encouraged to rehydrate before the commencement of the 5x6 Second RSA cycle test with the volume consumed recorded. During this period, recovery heart rate values were recorded for further analysis, including a standardized walking recovery immediately after termination of the YYIR1-GK for two minutes. The participant was also fitted to the Lode cycle ergometer, adjusting aspects such as seat and handle bar height, distance from the handlebar, and seat tilt to ensure the participant was as comfortable as possible. These measurements were recorded on the first visit and replicated for visit two. They then performed five minutes of warm-up on the Lode ergometer at a self-selected pace at a workload of 50-75 watts. Once the warm up was completed, the participant performed an all-out six second sprint during which peak power output in watts was measured. This served as their baseline power output. They were then allowed a five minute recovery time before the commencement of the 5x6 second RSA cycle test.

After the five minute recovery, the participant was prompted to gradually ramp up their revolutions per minute (RPM) to around 100 during a ten second transition period with no resistance. During this ten second period, the participant received a countdown to the commencement of the first six second sprint.

The 5x6 Second RSA cycle test involved five bouts of six second sprints that were performed every thirty second interval. After each sprint, the participant had a fourteen second window for passive recovery. In the ten seconds prior to each sprint, the participant was asked to gradually increase cadence to 100 RPM against no resistance and a count-down was given prior to the onset of the sprint. Subjects were encouraged to give full effort while remaining seated. Immediately after the completion of each sprint, RPE and HR were recorded. Peak power output and RPM during each sprint were recorded by the Lode, while HR was monitored continuously using the heart rate monitor and downloaded to a computer for further analysis. Upon completion of the test, the participant performed a brief cool down at 50 watts of resistance until their HR dropped below 120 beats per minute.

Data Collection and Analysis

Heart rate information was collected continuously by the Polar heart rate monitor and downloaded onto a computer for further analysis. RPE during both tests was recorded by hand and inserted into an excel spreadsheet. Results from the YYIR1-GK as recorded by hand in stages completed successfully, correlating to a total distance covered. Power output information during the repeated sprint cycle test was recorded automatically onto the Lode program and exported from the attached computer to the nearest 0.1 Watts. All data was stored on Microsoft Excel and all statistical tests were run using SPSS (IBM, Armonk, NY). The alpha level as set at priori to 0.05.

To address hypothesis 1, coefficient of variance (CV), Pearson-Product Moment correlation, and Intraclass Correlation (ICC) were calculated as measurements of test-retest reliability. For hypothesis 2, Pearson's Product-Moment Correlation was calculated to identify correlations between the results from the two tests. Finally, Hypotheses 3 and 4 were addressed through independent T tests to compare results of each test between groups.

CHAPTER IV

MANUSCRIPT

Introduction

The use of field-based fitness tests is considered important for soccer teams as it allows coaches to evaluate a player in terms of physical preparedness and to guide training programs. The most commonly used fitness test for soccer players is the Yo Yo Intermittent Recovery Test Level 1 (Bangsbo et al., 2008). This test is among the most utilized and studied fitness tests in the world (Rampinini et al., 2007) and has been reported to have excellent reliability and validity on numerous occasions (Krustrup et al., 2003; Krustrup et al., 2005; Thomas et al., 2006). Because of the significant evidence supporting the use of the YYIR1, it has had widespread use as a means to evaluate intermittent fitness levels for soccer players. Goalkeepers are often included in this testing despite the vast differences in actions performed by this position when compared to field players. Goalkeeper actions are also intermittent in nature but do not resemble those of field players (Padulo et al., 2015). Goalkeepers only cover about one half of the total distance that a typical field player has been reported to travel (Di Salvo et al., 2008) but their actions are often explosive in nature while confined to a small space (Ziv & Lidor, 2011). Analysis of the actions of high level goalkeepers has recently found that they perform ninety-two high intensity displacements of about four meters per match such as moving laterally across the goal line or moving forward and backwards off the line to position themselves for oncoming attacks or to cut off the angle of an attacking player (Padulo et al., 2015). Rarely do goalkeepers perform sprints or high-intensity runs of any appreciable distance (Aziz et al., 2008). For this reason, goalkeepers should be evaluated for intermittent fitness using actions that are more appropriate to their position. Yet a large majority of the goalkeeper-specific tests in the literature have focused on cognitive, reactive, or technical factors rather than the ability to meet or exceed the physical demands of match play (Knoop, Fernandez, & Ferruati, 2013).

Tests of repeated sprint ability (RSA) have also recently gained some attention as a means to evaluate the intermittent fitness of team sport athletes. RSA involves testing the athlete's ability to generate maximal or near-maximal actions repeatedly with brief recovery periods (Bishop et al., 2011). Soccer players are required to perform intermittent bouts of high-intensity activities throughout a match (Stølen et al., 2005) so there could be some utility in testing RSA in these athletes. There have been a couple of reports that have shown positive results with various intermittent sport athletes (Spencer et al., 2004; Rampinini et al., 2007; Aziz et al., 2008) but there has not been an established RSA test that would mimic the actions of soccer goalkeepers. Therefore, the best option for evaluating RSA in goalkeepers would be a test involving short, explosive actions. The 5x6 Second RSA

Cycling test has been found to be both reliable and valid in evaluating intermittent fitness in team sport athletes (Bishop et al., 2008; Fitzgerald et al., 1993). Therefore this test could potentially be useful in evaluating general RSA in soccer goalkeepers, despite a different modality of physical activity.

Although it is important to evaluate the physical fitness of soccer players for optimal match preparation, there are very few fitness tests that target goalkeeper actions. A majority of the studies involving goalkeepers have been more technical or reactive in nature (Knoop, Fernandez, & Ferrauti, 2013) or have used fitness tests that are more appropriate for the activity of field players than goalkeepers (Al-Hazzaa et al., 2001; Knoop, Fernandez, & Ferrauti, 2013). Additionally, the YYIR1 is among the most heavily utilized and studied fitness tests in the world and has been proven to be a valid test of intermittent fitness for soccer field players (Krustrup et al., 2003). Establishing a goalkeeper-specific adaptation using actions more suitable to the position could provide valuable information to coaches about the physical preparedness of players that are crucial to match outcome and team success. Tests of RSA have also gained popularity in the literature as a means to evaluate the intermittent fitness levels of team sport athletes. As the goalkeeper position involves repeated bouts of explosive activity, tests of RSA could be of use with this population. Thus, the purpose of this study was to 1) evaluate the test-retest reliability of a goalkeeper-specific adaptation to the YYIR1 (YYIR1-GK), 2) to determine if there is a correlation between performance on the YYIR1-GK and performance on a validated test of RSA, the 5x6 Second RSA Cycle test, and 3) to determine if there is a gender gap in performance on the YYIR1-GK that is similar to that seen with the standard YYIR1. Three hypotheses were tested: 1) the YYIR1-GK and the 5x6 Second RSA Cycle test would exhibit strong test-retest reliability, with results on the YYIR1-GK being similar to reports on the standard test, 2) There would be strong correlations between performance on the YYIR1-GK and performance outcomes on the 5x6 Second RSA Cycle Test, and 3) the male goalkeepers would outperform the female goalkeepers on the YYIR1-GK to a similar extent as seen with the standard test.

Methods

Participants

A total of thirteen NCAA collegiate soccer goalkeepers participated in the study (see Table 1 for demographics). The subjects consisted of a combination of male (n=7) and female (n=6) goalkeepers on teams of either Division level 1 (n=9) or Division level 3 (n=4) in the NCAA. All participants were recruited from schools in the Triad area and had competed on a NCAA soccer roster at the goalkeeper position this past season, had no history of cardiovascular disease or events, were currently healthy and fit to play with no musculoskeletal injuries, normotensive as measured upon their first visit, and not currently pregnant. Each participant provided an informed consent and were screened for basic health history and physical activity habits through questionnaires. All procedures for the study were approved by the University of North Carolina at Greensboro Institutional Review Board.

Instrumentation

Equipment required for baseline measures included a sphygmomanometer and stethoscopes to obtain baseline blood pressures, a Seca weight scale to measure weight to the nearest 0.1 kg, a stadiometer to determine height (within 0.1 cm) and a Cosmed Bod Pod body composition system (Rome, Italy) using the Siri equation for body composition (Siri, 1961). Barometric pressure (to the nearest 0.01 mmHg), humidity, and room temperature (to the nearest 0.1 °C) was measured using a Davis indoor climate monitor (Hayward, California). For the performance of the YYIR1-GK fitness test, equipment included an audio file containing the YYIR1 test, an audio device to broadcast the file, four cones per participant, a chart of rating of perceived exertion (RPE) using the Borg 6-20 scale (Borg & Kaijser, 2006) and a Polar heart rate monitor (Kempele, Finland) placed around the chest to continuously monitor and store heart rates which were downloaded for further analysis. For the 5x6 Second RSA cycle test, an electronically braked Lode cycle ergometer (Groningen, Netherlands), a Borg scale of RPE, and a continuously monitoring heart rate monitor as noted above were used.

Experimental Design

Participants signed a consent form and provided basic health and physical activity history. Blood pressure and height were then collected before assessing body composition using the Bod Pod system. The heart rate monitor was then fit to the individual and was worn for the entire duration of the testing sessions. The subject remained seated and relaxed for five minutes of resting heart rate data. The subject was familiarized with the RPE scale before walking to the gym to begin the standardized warm up. The warm up consisted of three phases and lasted approximately fifteen minutes. The first three minutes was an active warming phase during which the participant performed moderate intensity jogging, side shuffles, and backwards jogging. This phase was followed by a series of seventeen dynamic flexibility movements common to soccer training sessions. These dynamic movements progressed from low intensity to high intensity, from single joint to multiple joint, and from simple to complex movements. The participant then performed a series of familiarization runs through the YYIR1-GK action sequence in progressing intensity, including two at maximal effort. The participant was prompted on the YYIR1-GK protocol and the testing began shortly after.

The YYIR1-GK involved bouts of high-intensity actions prompted by a digitized audio file. Two sets of cones were set up exactly four meters apart. The participant performed a series of four actions back and forth in this area in the following order; 1) side shuffle from one set of cones to the next, 2) side shuffle back to the starting cones, 3) Backwards run from the starting cones to the second set, and 4) forwards run back to the starting line. Each action was four meters in distance, and were performed in succession during each bout of activity for a total of sixteen meters. This sixteen meter action sequence was performed in standardized stages as prompted by the audio file. There was a countdown to begin each bout of activity followed by a beep. The goalkeeper then had to successfully complete the four actions before the next audible beep. Upon completion of the bout, they would passively rest on the starting line until the next countdown. The time to complete these actions decreased as stages progressed and the participant would continue until failing to meet the required pace on two different occasions or until they reached volitional exhaustion. Heart rate and RPE was measured at the end of each stage and the same tester evaluated each of the participants for consistency.

Upon completion of the YYIR1-GK, the participant was immediately taken through a standardized two minute walking recovery period during which the pace was set by the tester. The participant was then given a period of recovery during which they sat in the lab and were encouraged to hydrate before beginning the cycle warm up. The exact time for recovery was recorded and kept identical for each visit. Just before the end of the recovery period, the cycle ergometer was set up for the participant's optimal comfort. Seat height, handlebar height, and distance from the handlebar were measured and replicated on the second visit. The cycle warm up consisted of five minutes of cycling at a self-selected intensity. The workload and chosen revolutions per minute (RPM) were recorded and repeated during the second visit.

Upon completion of the five minute warm up, the 5x6 Second RSA Cycle test began. The test consisted of five sprints of six seconds each at intervals of thirty seconds. Prior to each sprint, there was a ramp up period of ten seconds during which the participant gradually increased their RPM to 100 against no resistance.

There was an upper limit set on RPM during this period to keep revolutions at the desired speed. The participant was then given a countdown into each sprint. Sprints were maximal effort against a weight-dependent resistance while remaining seated at all times. Upon completion of each sprint, there was a fourteen second period of passive recovery during which the subject sat on the bike until the next ramp up period. Immediately after each sprint, heart rate was recorded and RPE was asked from the participant. Upon completion of the fifth sprint, the subject was given a cool-down period to allow the heart rate to drop to a safe level before getting off the ergometer.

The participant returned at the same time a week later to perform the same two tests, with an effort to replicate conditions as best as possible. Prior to each visit, the participant was asked to eat and hydrate as though they would for a match and to replicate these diet habits as much as possible prior to the second visit. They were also asked to wear the same shoes and similar clothing to each visit.

Data Collection and Analysis

Heart rate information was collected continuously by the Polar heart rate monitor and exported onto a computer for further analysis. RPE during both tests was recorded by hand and inserted into an excel spreadsheet. Results from the YYIR1-GK was recorded by hand in stages completed successfully and correlated to a total distance covered in meters. Measurements of power and work during the repeated sprint cycle test was recorded automatically onto the Lode program and exported from the attached computer. All data was stored on Microsoft Excel and statistical tests were run using SPSS (IBM, Armonk, NY).

Statistical Approach

To address hypothesis 1, coefficient of variance (CV), Pearson-Product Moment correlation coefficient, and Intraclass Correlation Coefficient (ICC) was calculated as measurements of test-retest reliability. For hypothesis 2, Pearson's Product-Moment Correlation was calculated to identify correlations between the results from the two tests. Finally, Hypothesis 3 was addressed through independent T tests to compare results of each test between groups. Additionally, paired sample t tests were used to determine that testing conditions were not significantly different between visits. Alpha level was set a priori to 0.05.

Results

Table 1. Participant Demographics by Group. Demographics [mean ± (SD)] for participants including age, height, mass, body fat percentage (BF), resting heart rate (RHR), years of goalkeeper experience (GKExp), resting systolic (RBP-S) and diastolic (RBP-D) blood pressure.

Group	**Age (yrs)**	**Height (cm)**	**Mass (kg)**	**BF (%)**	**RHR (bpm)**	**GKExp (yrs)**	**RBP-S (mmHg)**	**RBP-D (mmHg)**
Male (n=7)	19.86 (1.68)	184.31 (2.86)	81.77 (9.24)	14.56 (6.70)	71.30 (4.08)	7.33 (1.75)	118.00 (2.58)	77.71 (2.93)
Female (n=6)	19.67 (1.21)	170.03 (3.25)	76.92 (7.00)	27.42 (7.08)	70.94 (9.65)	11.50 (3.87)	118.33 (5.85)	74.67 (8.36)

Table 2. Summary of Results. Results include total work, total power, peak power, work decrement (WorkDec), and power decrement (Power Dec) for the 5x6 Second RSA Cycle test and distance covered, peak RPE, and peak heart rate for the YYIR1-GK divided by gender. Results are divided into visit 1 and visit 2 and significance value from independent t tests (Sig) is included for each variable. Shading indicates the results were insignificant at an alpha of $p < 0.05$.

		Group	Visit 1	Sig	Visit 2	Sig
5x6 Second RSA	*Total Work (KJ)*	Men	21.26(2.48)	p = 0.008*	22.37(2.28)	p = 0.002*
		Women	17.53(2.09)		17.54(2.07)	
	Total Power (W)	Men	4383.5(614.9)	p =0.031*	4587.1(492.6)	p = 0.004*
		Women	3635.2(445.6)		3689.4(398.8)	
	Peak Power (W)	Men	1024.2(165.9)	p = 0.034*	1055.2(117.2)	p = 0.002*
		Women	836.1(98.8)		829.1(82.7)	
	WorkDec (%)	Men	11.42(7.80)	p = 0.739	14.18(7.44)	p = 0.963
		Women	12.97(8.58)		13.98(7.94)	
	Power Dec (%)	Men	15.75(8.68)	p = 0.449	12.85(6.23)	p = 0.544
		Women	12.31(6.78)		10.96(4.25)	
YYIR1-GK	*Distance (m)*	Men	377.14(193.5)	p =0.022*	416.00(193.5)	p = 0.010*
		Women	154.67(15.8)		165.33(38.8)	
	Peak RPE	Men	16.57(2.07)	p = 0.251	16.00(1.79)	p = 0.271
		Women	15.00(2.61)		14.33(3.01)	
	Peak HR (bpm)	Men	189.14(7.94)	p = 0.911	188.67(8.31)	p = 0.703
		Women	188.67(6.86)		186.67(9.33)	

Table 3. Confounding Factors by Visit. Mean and SD for each factor along with mean difference (MD) and significance result from pair t test for body mass, volume of water consumed, temperature of the gym (GymTemp), humidity of the gym (GymHum), barometric pressure of the gym (GymBaro), temperature of the lab (LabTemp), humidity of the lab (LabHum) and barometric pressure of the lab (LabBaro). Significance (Sig) from paired t tests at $p < 0.05$ is indicated by an asterisk.

	Visit 1	Visit 2	MD	Sig
Mass (kg)	82.48 (8.75)	82.53 (8.69)	-0.51	p = 0.852
Water (ml)	403.90 (305.07)	358.41 (211.44)	45.49	p = 0.220
GymTemp (°C)	20.92 (0.95)	21.46 (0.97)	-0.54	p = 0.068
GymHum (%)	28.15 (5.70)	36.15 (9.66)	-8.0%	p = 0.011*
GymBaro (mmHg)	740.48 (4.23)	738.53 (4.13)	1.94	p = 0.239
LabTemp (°C)	20.80 (0.45)	21.50 (0.55)	-0.80	p = 0.099
LabHum (%)	33.40 (7.99)	34.67 (9.73)	-1.80%	p = 0.650
LabBaro (mmHg)	737.70 (4.23)	739.62 (3.40)	-1.12	p = 0.673

Table 4. Summary of Reliability Measures. Results are included for the YYIR1-GK and 5x6 Second RSA Cycle test, including Pearson-Product Moment Correlation Coefficient (Pearson), Intraclass Correlation Coefficient (ICC) and Coefficient of Variation (CV) for distance covered on the YYIR1-GK and Total Work, Peak Power, and total power on the 5x6 Second RSA Cycle test. Significance at an alpha of p <0.05 on the Pearson correlation is indicated by an asterisk.

Test	Pearson	ICC	CV
YYIR1-GK	r = 0.920*	0.91	19.49%
5x6 Total Work	r = 0.751*	0.749	8.52%
5x6 Peak Power	r = 0.837*	0.835	6.95%
5x6 Total Power	r = 0.607*	0.607	10.12%

Hypothesis 1: The Goalkeeper-specific adaptation to the YYIR1 and the 5x6 Second RSA Cycle Test would both prove to have strong test-retest reliability with NCAA goalkeepers.

The YYIR-GK was shown to have strong test-retest reliability as measured by Pearson Product-Moment Correlation Coefficient (r=0.920, n=13, p<0.001), Intraclass Correlation Coefficient ($ICC_{3,1}$ =0.91) and a Coefficient of Variation of 19.49%. The 5x6 Second RSA Cycle test exhibited moderate to strong test-retest reliability in a variety of outcome measures such as Total work performed (r=0.751, n=13, p=0.003, $ICC_{3,1}$ = 1.098, CV = 8.52%), Peak Power (r= 0.837, n=13, p<0.001, $ICC_{3,1}$ = 1.076, CV = 6.95%) and Total Power (r=0.607, n=13, p=0.028). Paired T tests for repeated measures also indicated that the YYIR1-GK performance from visit 1 and visit 2 were not significantly different [t(12) = -0.851, p=0.411, two-tailed]. Paired T tests for the outcomes from the 5x6 Second RSA Cycle test also indicated that there were not significant differences in total work [t(12) = -0.851, p=0.411, two-tailed], peak power [t(12) = -0.530, p=0.606, two-tailed], or total power [t(12) = -0.851, p=0.411, two-tailed].

Figure 1. The Relationship Between Visits for YYIR1-GK and Total Work.

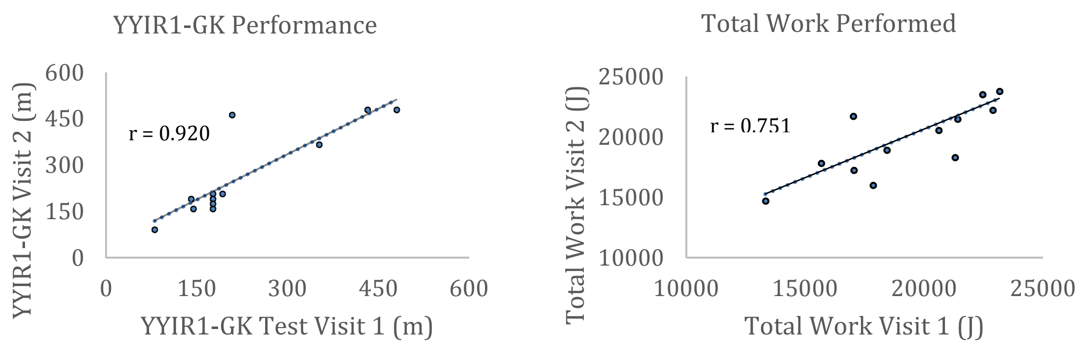
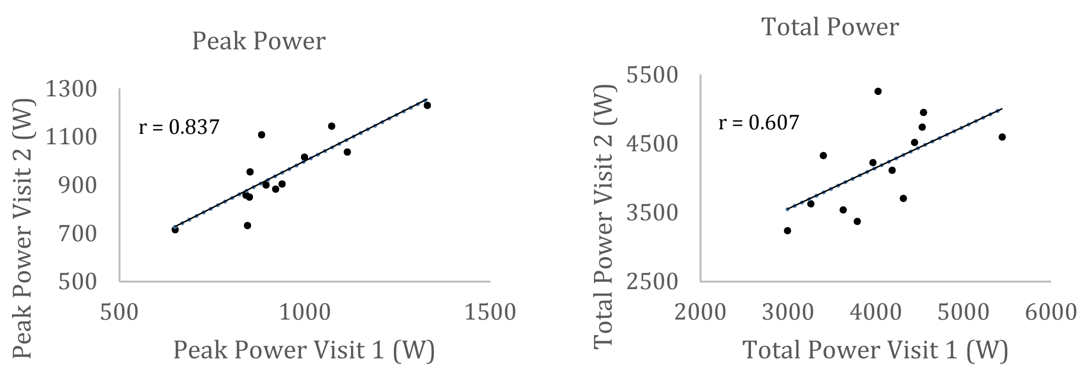


Figure 2. The Relationship Between Visits for Peak Power and Total Power.



Hypothesis 2: There will be a significant correlation between performance on the YYIR1-GK and performance on the 5x6 Second RSA Cycle Test.

There was not a significant correlation between performance on the YYIR1-GK and any outcome measure from the 5x6 Second RSA Cycle Test as measured by

Pearson Product-Moment Correlation Coefficient. There were moderate correlations between distance covered on the YYIR1-GK and total work performed on the cycle test for visit 1 (r=0.520, n=13, p=0.069) and visit 2 (r=0.449, n=13, p=0.124) respectively, but neither correlation was significant, though it seemed to be trending towards significance. Additionally, there were weak correlations between YYIR1-GK performance and peak power for visit 1 (r = 0.356, n=13, p=.256) and visit 2 (r=0.390, n=13, p=.210) respectively, with neither being significant. There were very weak, insignificant correlations between YYIR1-GK performance and work decrement for visit 1 (r= 0.186, n=13, p= 0.544) and visit 2 (r= 0.133, n=13, p=0.644).

Figure 3. Correlation of Total Work and Total Power with YYIR1-GK.

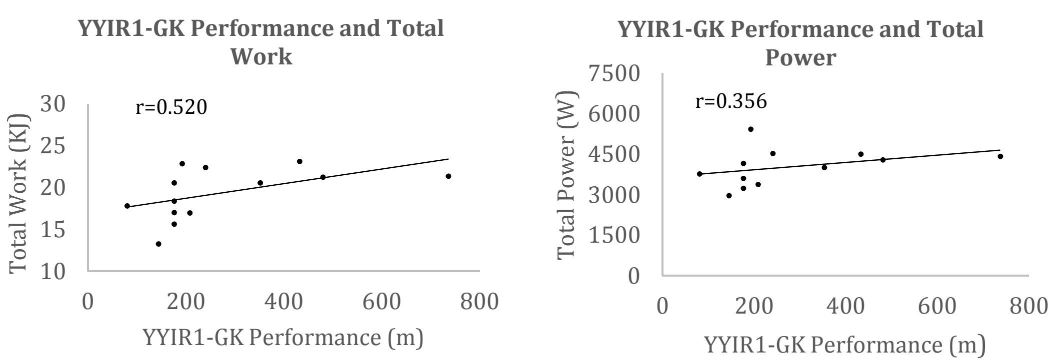
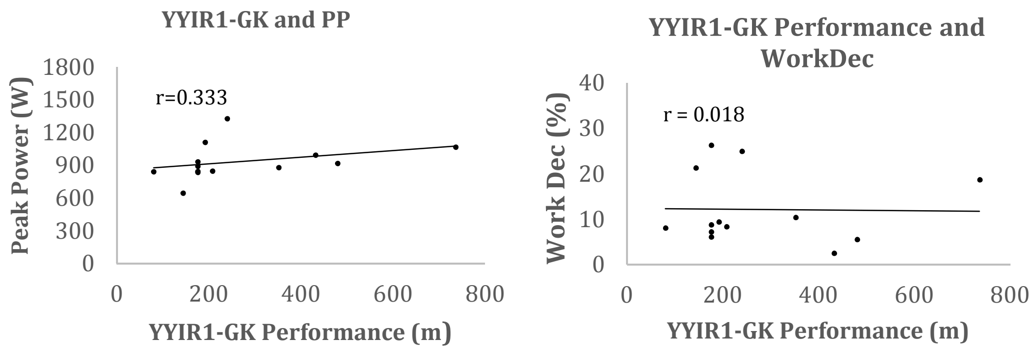


Figure 4. Correlation of Peak Power and Work Dec with YYIR1-GK.



Hypothesis 3: The male goalkeepers would outperform the female goalkeepers on the YYIR1-GK similar to the gender gaps as those seen in the standard YYIR1. Additionally, the male goalkeepers would outperform the female goalkeepers on the 5x6 Second RSA Cycle test.

The male goalkeepers covered significantly more distance on the YYIR1-GK than their female counterparts during visit 1 [t(12) = -2.973, p=0.022, two-tailed] and visit 2 [t(12) = -3.552, p=0.10, two-tailed). The gap between men and women was nearly 2.5 fold, even greater than that reported for highly skilled males and females for the standard YYIR1. Additionally, the men outperformed the females in a variety of outcome measures on the 5x6 Second RSA Cycle test including peak power for visit 1 [t(12) = -2.424, p=0.034, two-tailed] and visit 2 [t(12) = -3.946, p==0.002, two-tailed], total power for visit 1 [t(12) = -2.470, p=0.031] and visit 2 [t(12) = -3.629, p=0.004, two-tailed] and total work for visit 1 [t(12) = -3.237, p=0.008, two-tailed) and visit 2 [t(12) = -3.998, p=0.002, two-tailed). There was no significant difference between males and females for Work decrement for visit 1 [t(12) = 0.341, p= 0.739, two-tailed) or visit 2 [t(12) = -0.047, p=0.963, two-tailed)

Figure 5. Gender Comparison for YYIR1-GK and Total Work.

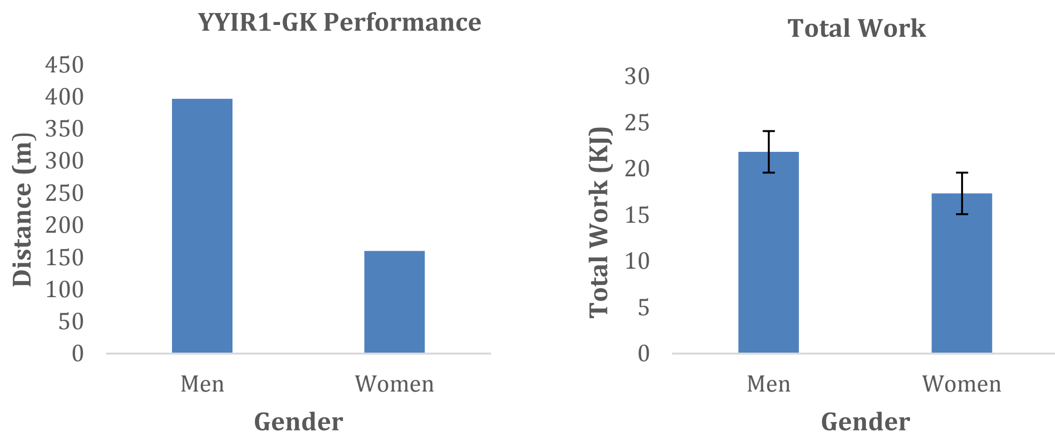


Figure 6. Gender Comparison for Peak Power and Total Power.

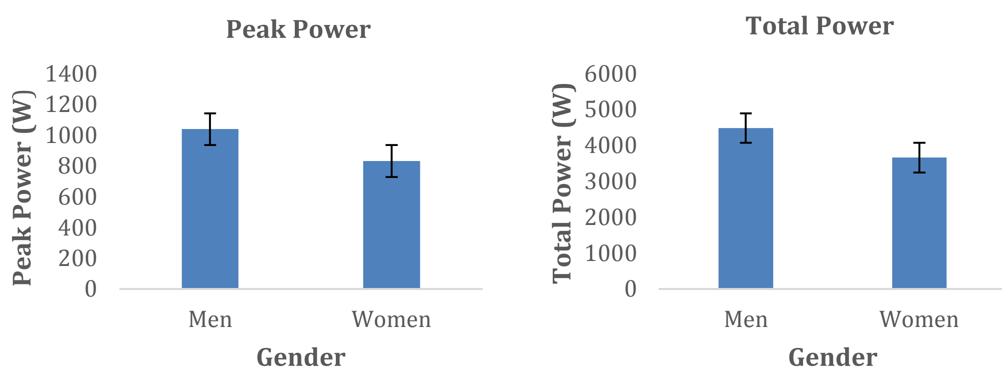


Figure 7. Gender Comparison for Decrement of Power and Work

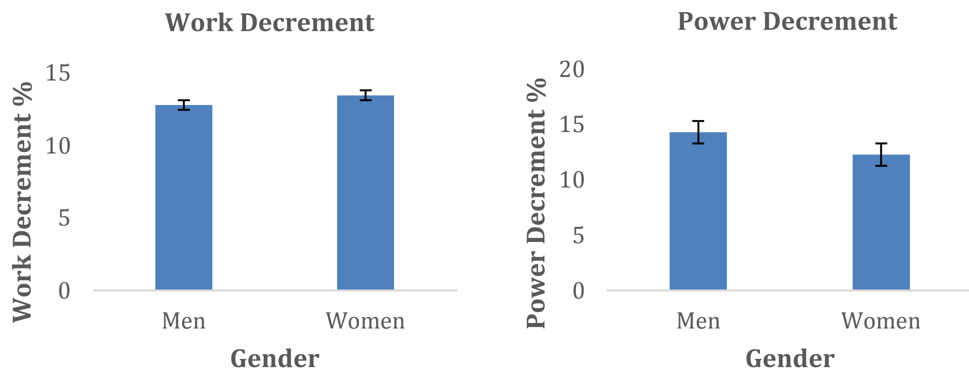


Figure 8. Heart Rate Responses for Each Test. The heart rate for the YYIR1-GK in terms of test duration and the heart rate for the 5x6 Second RSA Cycle test by sprint number.

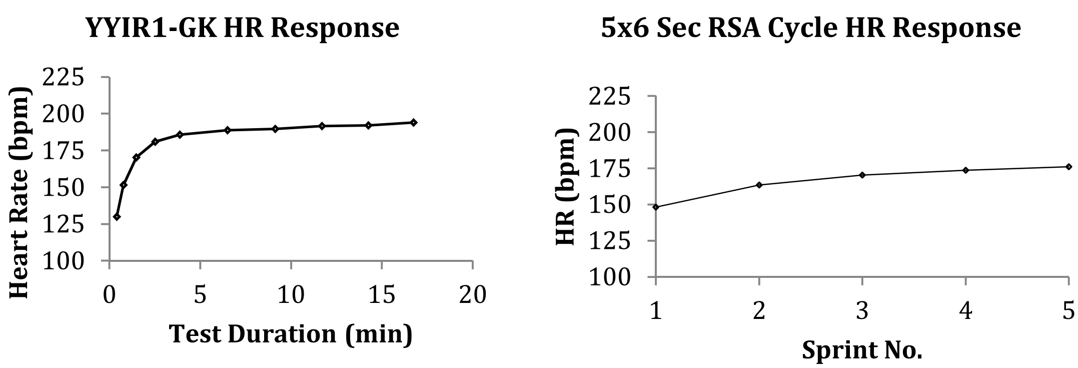
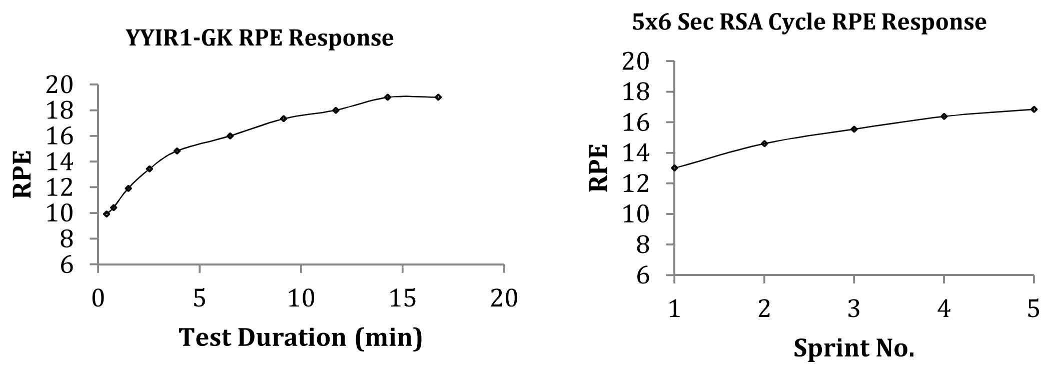


Figure 9. Rating of Perceived Exertion for Each Test. RPE for the YYIR1-GK in terms of test duration and RPE for the 5x6 Second RSA Cycle test by sprint number.



Discussion

This study examined whether a goalkeeper-specific adaptation to the standard YYIR1 would prove to be reliable when tested on two separate occasions one week apart. Further, this study examined whether performance on the YYIR1-GK would correlate with performance measures on a lab-based cycle test of repeated sprint ability. Finally, group comparisons were made to determine whether there was a similar gender gap in performance on the goalkeeper adaptation as that reported for the standard YYIR1. Overall, the YYIR1-GK exhibited strong test-retest reliability when tested one week apart with collegiate goalkeepers and the men performed significantly more distance than the women but there was little correlation between performance on the YYIR1-GK and the 5x6 Second RSA Cycle test.

The primary goal of this study was to establish the reliability of a goalkeeper-specific test of intermittent fitness and athleticism. To accomplish this, NCAA goalkeepers were recruited to come in for two visits exactly one week apart. In order to keep each visit as consistent as possible, several measures were taken. Before the first visit they were prompted to eat and hydrate as though they would prior to a match and record the nutritional information on diet logs that were provided to them. They were then asked to replicate this diet as closely as possible leading up to the second visit. They were also asked to not participate in intense exercise within 24 hours of each visit. The internal environment of the lab and gym were tracked in terms of temperature, humidity, and barometric pressure and compared for consistency. The participant was weighed at the beginning of each visit to compare weight fluctuations and they only drank out of a provided water bottle that had been measured for volume. The volume of water consumed was measured during each visit and recorded. The warm up performed prior to each testing of the YYIR1-GK was a standardized soccer warm up that was kept as identical as possible from visit 1 to visit 2 in terms of intensity and duration. There were no significant differences in participant weight or volume of water consumed as measured by paired t tests. In terms of internal environment, there were not significant differences in any of the measures except for the humidity of the gym. The warm up was shown to not be significantly different as well when comparing heart rate responses and ratings of perceived exertion during each phase. These findings indicate that the testing conditions were similar between visits 1 and 2.

The YYIR1-GK showed strong test-retest reliability as measured by Pearson's correlation, Intraclass Correlation Coefficient, and Coefficient of Variation. Correlation strength was stronger than what has been reported for the standard YYIR1 with recreational athletes (Thomas et al., 2006) and comparable to that reported with elite footballers (Krustrup et al., 2003). The strong reliability is slightly surprising considering the complexity of executing the series of four high intensity actions each bout compared to the standard test. While the standard YYIR1 involves twenty meter linear runs back and forth, the YYIR1-GK required goalkeepers to perform a succession of four actions including side shuffles in each direction, back pedaling, and forwards runs of four meters each. Goalkeepers often reported that the margin of error felt very small at the faster stages and that even a slight delay in transition between actions resulted in a failed attempt. Overall, the test exhibited strong test-retest reliability with collegiate soccer goalkeepers.

Reliability was also measured for the 5x6 Second RSA Cycle test. To maximize consistency for each session, recovery time between the YYIR1-GK and the 5x6 Second RSA was replicated between visit 1 and visit 2. Additionally, the cycle ergometer was set-up to the participant's comfort on visit 1 but the handlebar height, distance from the handlebar, and seat height measurements were recorded on visit 1 and the ergometer was set-up identically for visit 2. The warm up on the cycle ergometer prior to the 5x6 Second RSA Cycle test was also kept identical from visit 1 and visit 2 in terms of workload. There were no significant differences in heart rate responses during the warm up between visits.

This test of repeated sprint ability was shown to have moderately strong test-retest reliability in several outcome measurements such as total work, total power, and peak power but there was very weak reliability for work and power decrement. The significant difference in results using the decrement scores is likely due to several factors.

First, the subjects were not homogenous in terms of cycling experience. Some of the goalkeepers reported regular cycling as part of their training while others rarely cycled in any form. Additionally, there seemed to be a pacing effect with the repeated sprints on this test. The goalkeepers often varied significantly from visit 1 to visit 2 in terms of the drop off in power and work after the first sprint. There is the possibility that the goalkeepers initiated a pacing strategy where submaximal effort was given on certain sprints so that they could save energy for the final sprints. This was observed in several instances where the strongest sprints were the fourth of fifth, when fatigue should be setting in. This would affect the consistency of the decrement scores significantly. In the future with this test, it is imperative that the tester encourages maximal effort on each one of the five sprints to reduce any pacing effect the subject may incorporate. When comparing the two tests, the reliability was stronger in the YYIR1-GK, it can be performed simultaneously with the rest of the team in the field, and incorporates actions more specific to the goalkeeper position. The YYIR1-GK therefore is the more useful test of intermittent fitness for goalkeepers.

When comparing the two tests, there proved to be no significant correlations between performance on the YYIR1-GK and various outcome measures from the 5x6 Second RSA Cycle test as measured by Pearson's correlations. There were moderately strong correlations between the distance covered on the YYIR1-GK and cycle outcomes such as total work, total power, and peak power, but none of these correlations were statistically significant. Work decrement and power decrement had extremely weak correlation scores, indicating that there is little to no relationship between those outcomes and YYIR1-GK performance. Despite previous literature showing correlations with intermittent sport performance and performance on the 5x6 Second RSA Cycle test, it is recommended that a test using actions more common to the goalkeeper be used in the future. Many of the goalkeepers reported a limiting factor of performance on the YYIR1-GK to be their explosiveness off of the starting line and when transitioning between actions. Therefore, lower extremity power may play a role in performance on this test and in goalkeeper performance in match play. Comparing results on the YYIR1-GK with common field tests of power development and athleticism such as the countermeasure jump, or vertical jump could show stronger correlations as they are actions more common to the goalkeeper position than cycling sprints.

Lastly, there were significant gender differences in performance on both of the tests as measured by independent t tests. The male goalkeepers performed significantly more distance on the YYIR1-GK than the female goalkeepers. The gap between male and female performance was even greater on the YYIR1-GK than that reported in the standard YYIR1 (Mujika et al., 2009). Males typically commented that the reason they eventually failed to maintain the pace of the test was the inability to recover between bouts of activity while the females often reported that the speed required for each actions was too great to maintain the pace. The males were covering average cumulative distances that were around 2.5 fold greater than the females and were likely experiencing a greater metabolic demand from the test than the female goalkeepers. The average duration of the testing for females was only about 3.5 minutes. This could indicate that the test should be adjusted for females. This could be done by shortening the four meter distance to allow for more stages to be successfully completed. Additionally, males significantly outperformed the female goalkeepers on various outcomes of the 5x6 Second RSA cycle test including total power, total work and peak power. However, there was no significant difference in work or power decrement between males and females.

The study is limited by a few different factors. First, the sample size is relatively small for between group comparisons due to the difficulty of recruiting large numbers of skilled goalkeepers. However, the effect size was large enough between male and female performances on many measures that there is still sufficient statistical power when comparing the groups. Since there was some evidence of correlation between several outcomes on the 5x6 Second RSA Cycle test and performance on the YYIR1-GK, it is possible that a larger sample size could lead to significant results in that regard. There also was not great compliance with filling out the diet logs. Participants were prompted to follow provided instructions for filling out diet information for three days prior to each visit but they either did not comply or provided insufficient detail in many of their food choices. A more detailed analysis of nutrition could benefit the reliability of this test.

Performance was also likely impacted by the time of testing. The goalkeepers were out of season and just beginning training for the upcoming season. Results may have been different had the goalkeepers performed the test in-season or at the end of their preseason when they are closer to peak fitness levels. Performing the test on a gym surface rather than in the field with cleats likely influenced results as well. A gym was chosen for the YYIR1-GK to minimize variation in the testing environment between visits but performance could potentially be improved by completing the test on a soccer field while wearing cleats. Additionally, many of the goalkeepers reported that they perform cross-over steps in match play when covering distances as great as four meters. The YYIR1-GK protocol limits the goalkeeper's lateral movements to side shuffling for consistency. For the goalkeepers who are more comfortable performing a crossover step for lateral movements, this could impact performance. Potential adjustments to the test should be considered for allowing crossover steps as well.

In terms of the cycling sprint test, the 5x6 Second RSA cycle test was always completed after the YYIR1-GK which could affect reliability measurements of the cycle test. Despite the effort to keep testing sessions as identical as possible, performance on the YYIR1-GK had slight fluctuations from visit 1 to visit 2 which could impact performance on the 5x6 Second RSA Cycle Test. The 5x6 Second RSA Cycle test still exhibited some evidence of moderate to strong test-retest reliability with the collegiate goalkeepers. Finally, this was not a homogenous sample in terms of cycling experience. Some of the goalkeepers reported regular cycling activity while others performed no cycling at all. This could prove to be a confounding factor for performance on the 5x6 Second RSA Cycle test.

In conclusion, the YYIR1-GK exhibited strong test-retest reliability with collegiate goalkeepers. Additionally, the 5x6 Second RSA Cycle test showed moderate reliability in several outcome measures. Finally, the male goalkeepers significantly outperformed the females on the YYIR1-GK and several outcomes of the 5x6 Second RSA Cycle test. This study sets the foundation for future studies with the YYIR1-GK. Future studies can proceed in a few different directions with the YYIR1-GK, including a direct comparison between performances on the standard YYIR1 as well as the YYIR1-GK to see if goalkeepers have similar performance or if they are as consistent on the standard test as the goalkeeper adaptation. Additionally, the standard YYIR1 has shown to be able to differentiate between skill levels of soccer field players (Mujika et al., 2009), so determining if there is a difference in performance between NCAA division level goalkeepers could offer some validity to the YYIR1-GK. Additionally, future studies should track performance on the YYIR1-GK throughout the season to see if the test shows seasonal or training-induced changes. If the test proves to be sensitive to training, it could be used to help guide training programs and also monitor the physical preparedness of goalkeepers throughout the season. Additionally, the vast difference in performance between men and women could indicate that an adjustment to the YYIR1-GK may be beneficial for female goalkeepers in terms of the distance they must cover. Studies evaluating the reliability of a modified YYIR1-GK for females should be considered.

Another future direction could observe correlations between performance on the YYIR1-GK and performance on common field measures of lower extremity power such as vertical jumps. Additionally, anthropometric measures such as leg length or lower extremity muscle mass could be a factor in performance on this test and should be correlated to performance as well. Finally, future studies should measure performance on the test in the field to better replicate the conditions of match play.

CHAPTER V

DISCUSSION

The motivation behind this study was to evaluate the reliability of a goalkeeper-specific adaptation to the popular Yo Yo Intermittent Recovery Test Level 1. The YYIR1-GK required participants to perform a series of four actions of four meters each in timing with a standardized audio file. The actions were more specific to the goalkeeper position than the standard test and could be used by coaches and trainers to evaluate the intermittent fitness and athleticism of their goalkeepers to guide training programs. Additionally, this study wanted to determine if there is any correlation between performance on the YYIR1-GK and performance on a laboratory-based cycling test of repeated sprint ability, the 5x6 Second RSA Cycle test.

Participants completed two visits exactly two weeks apart during which they completed both tests. Each visit was standardized and was kept as identical to each other as possible. This included prompting each participant to eat as similarly as possible leading up to each visit, standardized warm ups with similar intensity and duration, and analysis of factors such as lab and gym internal environment, water consumed, ratings of perceived exertion, and heart rate responses between each session. The YYIR1-GK was proven to have strong test-retest reliability with similar results as reported with the standard YYIR1. Additionally, the 5x6 Second RSA Cycle test exhibited moderately strong test-retest reliability in several outcome measures. However, there were no significant correlations between performances on each test. When looking at each test individually, there were significant gender differences in performance. Male goalkeepers performed significantly more distance on the YYIR1-GK than the females. The males also performed significantly greater total work, total power, and peak power on the 5x6 Second RSA Cycle test.

Overall, this study contributes to the literature by showing strong test-retest reliability for a goalkeeper-specific fitness test. There have been very few attempts at establishing reliable and convenient fitness tests for goalkeepers, as most studies have either involved tests that were cognitive, reactive or technical in nature omore suited for the field players. Establishing a test of intermittent fitness that uses actions common to the goalkeeper position could be very beneficial for soccer coaches and training staffs. This is also one of the only studies to have included female goalkeepers for fitness testing purposes.

Some limitations to this study are as follows. First, compliance for filling out the nutritional logs was relatively poor. Some participants either did not fill out the logs as they were asked or did not follow the directions given to them. The result is not having a sufficient understanding of nutritional intake prior to each visit. Second, the sample size was rather small for many between group analyses. Since the study was limited to NCAA collegiate goalkeepers, the population to draw from was small. Additionally, the need to set aside two days in a week proved difficult for many collegiate athletes. There was however a large enough effect size to provide sufficient statistical power when comparing gender differences on test performance. Third, the 5x6 Second RSA Cycle test was performed the same day as the YYIR1-GK. The resulting test-retest reliability was likely affected to some degree by any differences in performance on the YYIR1-GK. Attempts were made to provide a standardized rest period between the tests that was sufficient to minimize any interference between the two but a study design where they are performed on different days would be ideal. Finally, there was a large degree of differences in cycling experience with this sample. Some regularly cycled while others cycled little to none. This could have affected results on the 5x6 Second RSA Cycle Test.

Future directions for this research would be to determine if the YYIR1-GK can differentiate between skill levels as the standard YYIR1 has been reported to (Mujika et al., 2009). It would also be beneficial to see if YYIR1-GK performance is sensitive to training, as has been shown with the standard test as well (Krustrup et al., 2003). If the higher skill level goalkeepers outperform those of lower level, it would offer some validity to the test. Additionally, if the test proves sensitive to training, it would allow coaches to monitor changes in intermittent fitness throughout the year and adjust training accordingly. Possible adjustments to the test such as shortening the distance for females or allowing crossover steps instead of side shuffles should be evaluated for reliability as well. It would also be beneficial to compare performance on the YYIR1-GK with field-based tests of lower extremity power such as vertical jumps, or with anthropometric measures that could make an impact on performance such as leg length or lower extremity muscle mass. Finally, future studies should try to incorporate larger sample sizes, and different populations of goalkeepers. This study included only a small number of NCAA collegiate goalkeepers, but it could be beneficial to determine how youth or recreational goalkeepers perform on the YYIR1-GK as well.

REFERENCES

Al-Hazzaa, H. M., Al-Muzaini, K. S., Al-Refaee, S. A., Sulaiman, M. A., Dafterdar, M. Y., Al-Ghamedi, A., & Al-Khuraiji, K. N. (2001). Aerobic and anaerobic power characteristics of Saudi elite soccer players. *Journal of Sports Medicine and Physical Fitness*.

Aziz, A. R., Mukherjee, S., Chia, M. Y. H., & Teh, K. C. (2008). Validity of the running repeated sprint ability test among playing positions and level of competitiveness in trained soccer players. *International journal of sports medicine*, *29*(10), 833-838.

Balsom, P. (1994). Evaluation of physical performance. *Football (soccer)*, 102-123.

Bangsbo, J. (1994). *Fitness training in football: a scientific approach*. August Krogh Inst., University of Copenhagen.

Bangsbo, J., Iaia, F. M., & Krustrup, P. (2008). The Yo-Yo intermittent recovery test. *Sports medicine*, *38*(1), 37-51.

Bangsbo, J., & Michalsik, L. (2002). Assessment of the physiological capacity of elite soccer players. *Science and football IV*, 53-62.

Bangsbo, J., Mohr, M., & Krustrup, P. (2006). Physical and metabolic demands of training and match-play in the elite football player. *Journal of sports sciences*, *24*(07), 665-674.

Bangsbo, J., Nørregaard, L., & Thorsoe, F. (1991). Activity profile of competition soccer. *Canadian journal of sport sciences= Journal canadien des sciences du sport*, *16*(2), 110-116.

Bishop, D., Girard, O., & Mendez-Villanueva, A. (2011). Repeated-sprint ability—Part II. *Sports Medicine*, *41*(9), 741-756.

Bishop, D., Spencer, M., Duffield, R., & Lawrence, S. (2001). The validity of a repeated sprint ability test. *Journal of Science and Medicine in Sport*, *4*(1), 19-29.

Carling, C., Bloomfield, J., Nelsen, L., & Reilly, T. (2008). The role of motion analysis in elite soccer. *Sports medicine*, *38*(10), 839-862.

Carling, C., Williams, A. M., & Reilly, T. (2005). *Handbook of soccer match analysis: A systematic approach to improving performance*. Psychology Press.

Davis, J. A., & Brewer, J. (1993). Applied physiology of female soccer players. *Sports Medicine*, *16*(3), 180-189.

Di Salvo, V., Baron, R., Tschan, H., Montero, F. C., Bachl, N., & Pigozzi, F. (2007). Performance characteristics according to playing position in elite soccer. *International journal of sports medicine*, *28*(03), 222-227.

Di Salvo, V., Benito, P. J., Calderon, F. J., Di Salvo, M., & Pigozzi, F. (2008). Activity profile of elite goalkeepers during football match-play. *Journal of Sports medicine and Physical fitness*, *48*(4), 443.

Borg, E., & Kaijser, L. (2006). A comparison between three rating scales for perceived exertion and two different work tests. *Scandinavian journal of medicine & science in sports*, *16*(1), 57-69.

Brožek, J., Grande, F., Anderson, J. T., & Keys, A. (1963). Densitometric analysis of body composition: revision of some quantitative assumptions. *Annals of the New York Academy of Sciences*, *110*(1), 113-140.

Fitzsimons, M., Dawson, B., Ward, D., & Wilkinson, A. (1993). Cycling and running tests of repeated sprint ability. *Australian Journal of Science and Medicine in Sport*, *25*, 82-82.

Girard, O., Mendez-Villanueva, A., & Bishop, D. (2011). Repeated-sprint ability—Part I. *Sports medicine*, *41*(8), 673-694.

Granier, P., Mercier, B., Mercier, J., Anselme, F., & Prefaut, C. (1995). Aerobic and anaerobic contribution to Wingate test performance in sprint and middle-distance runners. *European journal of applied physiology and occupational physiology*, *70*(1), 58-65.

Green, S., & Dawson, B. (1993). Measurement of anaerobic capacities in humans. *Sports Medicine*, *15*(5), 312-327.

Helgerud, J., Hoff, J., & Wisloff, U. (2002). Gender differences in strength and endurance of elite soccer players. *Science and Football IV*, 382.

Jackson, A. S., & Pollock, M. L. (1978). Generalized equations for predicting body density of men. *British journal of nutrition*, *40*(03), 497-504.

Junge, A., & Dvorak, J. (2004). Soccer injuries. *Sports medicine*, *34*(13), 929-938.

Kerwin, D. G., & Bray, K. (2006). Measuring and modelling the goalkeeper's diving envelope in a penalty kick. In *The Engineering of Sport 6* (pp. 321-326). Springer New York.

Knoop, M., Fernandez-Fernandez, J., & Ferrauti, A. (2013). Evaluation of a specific reaction and action speed test for the soccer goalkeeper. *The Journal of Strength & Conditioning Research*, *27*(8), 2141-2148.

Krustrup, P., Mohr, M., Amstrup, T., Rysgaard, T., Johansen, J., Steensberg, A., & Bangsbo, J. (2003). The yo-yo intermittent recovery test: physiological response, reliability, and validity. *Medicine and science in sports and exercise*, *35*(4), 697-705.

Krustrup, P., Mohr, M., Ellingsgaard, Helga & Bangsbo, J. (2005). Physical demands during an elite female soccer game: importance of training status. *Medicine and science in sports and exercise*, *37*(7), 1242.

Lüthje, P., & Nurmi, I. (2002). Fracture-dislocation of the tarsal navicular in a soccer player. *Scandinavian journal of medicine & science in sports*, *12*(4), 236-240.

Matsukura, K., & Asai, T. (2013, May). Area covered by diving actions performed by male college soccer goalkeepers. In *Science and Football VII: The Proceedings of the Seventh World Congress on Science and Football* (p. 249). Routledge.

McMorris, T., & Hauxwell, B. (1997). Improving anticipation of soccer goalkeepers using video observation. *Science and football III*, 290-294.

Mohr, M., Krustrup, P., & Bangsbo, J. (2003). Match performance of high-standard soccer players with special reference to development of fatigue. *Journal of sports sciences*, *21*(7), 519-528.

Mihalik, J. P., Myers, J. B., Sell, T. C., & Anish, E. J. (2005). Maxillofacial fractures and dental trauma in a high school soccer goalkeeper: a case report. *Journal of athletic training*, *40*(2), 116.

Mujika, I., Santisteban, J., Impellizzeri, F. M., & Castagna, C. (2009). Fitness determinants of success in men's and women's football. *Journal of Sports Sciences*, *27*(2), 107-114.

Mulqueen, T., & Woitalla, M. (2011). *Complete Soccer Goalkeeper, The*. Human Kinetics.

Narayanan, V., Josty, I. C., & Dickson, W. A. (2000). Lime burns in a professional football goalkeeper—an unusual hazard. *Burns*, *26*(8), 754-756.

Padulo, J., Haddad, M., Ardigò, L. P., Chamari, K., & Pizzolato, F. (2015). High frequency performance analysis of professional soccer goalkeepers: A pilot study. *J Sports Med Phys Fitness*, *55*(6), 557-562.

Rampinini, E., Bishop, D., Marcora, S. M., Bravo, D. F., Sassi, R., & Impellizzeri, F. M. (2007). Validity of simple field tests as indicators of match-related physical performance in top-level professional soccer players. *International journal of sports medicine*, *28*(03), 228-235.

Rebelo-Gonçalves, R., Figueiredo, A. J., Coelho-e-Silva, M. J., & Tessitore, A. (2016). Assessment of Technical Skills in Young Soccer Goalkeepers: Reliability and Validity of Two Goalkeeper-Specific Tests. *Journal of Sports Science and Medicine*, *15*, 516-523.

Reilly, T. (2003). *Science and soccer*. Routledge.

Reilly, T. (2007). The science of training–soccer. *Oxon: Routledge*.

Reilly, T., Bangsbo, J., & Franks, A. (2000). Anthropometric and physiological predispositions for elite soccer. *Journal of sports sciences*, *18*(9), 669-683.

Reilly, T., & Bowen, T. (1984). Exertional costs of changes in directional modes of running. *Perceptual and Motor Skills*, *58*(1), 149-150.

Sainz De Baranda, P., Ortega, E., & Palao, J. M. (2008). Analysis of goalkeepers' defence in the World Cup in Korea and Japan in 2002. *European Journal of Sport Science*, *8*(3), 127-134.

Siri, W. E. (1961). Body composition from fluid spaces and density: analysis of methods. *Techniques for measuring body composition*, *61*, 223-44.

Sørensen, V., Ingvaldsen, R. P., & Whiting, H. T. A. (2001). The application of co-ordination dynamics to the analysis of discrete movements using table-tennis as a paradigm skill. *Biological cybernetics*, *85*(1), 27-38.

Spencer, M., Lawrence, S., Rechichi, C., Bishop, D., Dawson, B., & Goodman, C. (2004). Time–motion analysis of elite field hockey, with special reference to repeated-sprint activity. *Journal of sports sciences*, *22*(9), 843-850.

Sporis, G., Jukic, I., Ostojic, S. M., & Milanovic, D. (2009). Fitness profiling in soccer: physical and physiologic characteristics of elite players. *The Journal of Strength & Conditioning Research*, *23*(7), 1947-1953.

Stewart, A. D. (2002). Physiological Tests for Elite Athletes.

Stølen, T., Chamari, K., Castagna, C., & Wisløff, U. (2005). Physiology of soccer. *Sports medicine*, *35*(6), 501-536.

Svensson, M., & Drust, B. (2005). Testing soccer players. *Journal of Sports Sciences*, *23*(6), 601-618.

Thomas, A., Dawson, B., & Goodman, C. (2006). The Yo-Yo Test: Reliability and Association With a 20-m Shuttle Run and VO 2max. *International Journal of Sports Physiology & Performance*, *1*(2).

Williams, A. M., & Hodges, N. J. (2005). Practice, instruction and skill acquisition in soccer: Challenging tradition. *Journal of sports sciences*, *23*(6), 637-650.

Witte, T. H., & Wilson, A. M. (2004). Accuracy of non-differential GPS for the determination of speed over ground. *Journal of biomechanics*, *37*(12), 1891-1898.

Wood, G., & Wilson, M. R. (2010). A moving goalkeeper distracts penalty takers and impairs shooting accuracy. *Journal of Sports Sciences*, *28*(9), 937-946.

Ziv, G., & Lidor, R. (2011). Physical characteristics, physiological attributes, and on-field performances of soccer goalkeepers. *International journal of sports physiology and performance*, *6*(4), 509-524.