Santtu Isoniemi

Validity and Reliability Analysis of Cooper's 12-Minute Run on a curved non-motorized treadmill

Bachelor's Degree in Sports and Leisure Management

Autumn 2022

KAMK • University of Applied Sciences

Abstract

Author: Isoniemi Santtu

Title of the Publication: Validity and Reliability Analysis of Cooper's 12-Minute Run on a curved non-motorized treadmill

Degree Title: Bachelor of Sports Studies, Sports and Leisure Management

Keywords: cooper test, curved non-motorized treadmill, overground, runBEAT, vo2max

The purpose of this thesis was to assess the validity and reliability of Cooper 12-minute running test (12MR) on the curved non-motorized treadmill (cNMT). This was conducted by comparing the Cooper 12MR performed on an indoor track and on the cNMT. The commission for this thesis came from CSE Entertainment. Testing on the cNMT was performed with the runBEAT -software developed by the commissioning party (CSE Entertainment).

Eight participants (24 years ± 7 years) took part in the testing process and performed two familiarization trials on the cNMT as well as completed the Cooper 12MR on the cNMT and on an indoor track. Grouping was implemented by a matched pair design and a randomized block design to reduce potential variables. The results were measured, recorded, and later analysed with PSPP and Excel 2022. To test the significance of the results, a two-tailed t-test with paired samples was conducted together with Pearson's correlation. Additionally, the results and their means were compared against one another.

This thesis had three primary research questions. The first problem focused on the validity and reliability of the 12MR on cNMT. The second problem was to discover if there would be any differences regarding the challenge and physical demands between the tests, and the third problem was to determine whether cNMT together with runBEAT could be used as proper testing devices for the Cooper Test.

The results indicated a positive correlation (0,95) and are statistically significant. The participants achieved greater Vo_2 max estimations on track compared to the cNMT (23,9 ± 13.7%), which could indicate the more strenuous nature of the cNMT. These results are similar to previous findings of higher oxygen uptake on the cNMT. Further research could be conducted with a larger sample and the sample could be compared with direct VO_2 max testing to see how well the results correlate to actual VO_2 max by validating the test against a breath-to-breath analysis.

Contents

Appendices

1 Introduction

Endurance as a basic physical characteristic can be defined as the body's ability to resist fatigue during physical exertion. It is particularly affected by the condition of the respiratory and circulatory systems, muscle metabolism and the functioning of the nervous system. The most common way to determine endurance characteristics is to measure maximum oxygen carrying capacity by direct or indirect methods. (Keskinen, Häkkinen, & Kallinen, 2004)

Running is an exercise where the overall energy cost is depended upon the number of biomechanical, physiological and environmental factors. The place and surface where the individual runs or the running modality, is an environmental factor. The running modality can affect both biomechanical and physiological factors and therefore have an influence of individual's performance during the activity. According to Wee, Von Heimburg and Van den Tillaar (2016) Jones and Doust (1996) stated that between a motorized treadmill and running on a track, the latter has been found more strenuous due to the resistance on a track, surface, visual cues when moving, or overall, how familiar is the individual with the modality. (Wee et al., 2016)

The Cooper test is an indirect method VO_2 max. Cooper's own study reported a strong correlation between the direct VO_2 max and the VO_2 max estimation in a 12-minute test (Cooper, 1968). This high positive correlation implied that the indirect test had a strong correlation when compared with a direct testing method. Other studies have shown similar results within the correlation of the tests, supporting a high reliability. (Aitchison, Amjad, Corbett, Grant, & Wilson, 1995; Penry et al., 2011; Bandyopadhyay, 2015).

The primary purpose of this thesis is to assess the validity and reliability of the Cooper 12-minute running test performed on a curved non-motorized treadmill (cNMT for short) while using runBEAT -software. The commissioning party was interested to obtain data on how strenuous the 12MR is on the cNMT compared to 12MR performed on an indoor track. Earlier, the commissioning party had been given an estimation from the manufacturer of the treadmill. This thesis aims to provide proper results with theory supporting them.

The commission for this thesis came from CSE Entertainment. This thesis aims to provide results and analysis that have a theoretical basis supporting them, which then benefits the

commissioning party in the sports and fitness market. For the author the thesis has learning effects, giving the opportunity to gain new knowledge as well as at the same time apply the knowledge in practice. In addition, the author is able to utilize their acquired knowledge and skills: demonstration of a fundamental knowledge of anatomy and physiology, physical activity and fitness testing, management of planning, execution, and evaluation of extensive modes and modules.

A quantitative research method was chosen for this thesis. The author chose to use the quantitative research approach to obtain validity and reliability of the study. The quantitative study will be conducted using the measurements and results gained through the tests. Moreover, the sample, data gathering, and data analysis process will be explained.

The research problems of this thesis are:

1. Is the Cooper's 12-Minute Run valid and reliable on the cNMT?

2. Are there differences regarding the challenge and physical demand between the tests?

3. Can runBEAT and cNMT be used as proper testing devices for the Cooper Test?

2 The commissioning party

CSE Entertainment (CSE Simulation Oy) is a digital exercise games (exergaming) company based in Kajaani, established in 2012. The company's strengths include international expertise in game development, programming, hardware and electronics design, hardware interface design and the combination of these areas of expertise. (CSE Entertainment, 2021.)

Exergames are videogames that combine playing and exercising, therefore requiring the player to be physically active. (Kari, 2017, 5.) Exergames can alter the training's difficulty and volume through its virtual environment. Training in the virtual environment can add meaningful variety. Furthermore, it can be used as a training aid to support a sport or trait specific training by choosing suitable games. Moreover, it can be used solely for training or testing purposes.

The company's products include the Wall-franchise (iWall, tapWall, rehabWall) and the BEAT-franchise (runBEAT, groupBEAT, cycloBEAT). The company's aim is to inspire people of all ages to have fun while exercising thanks to their products. Exergaming industry has a rapidly growing market as many industries are looking for services that deliver experiences as well as engage customers. CSE's primary customer groups are fitness and sports centres, rehabilitation institutions, educational institutions, activity parks, tourism operators and public facilities. (CSE Entertainment, 2021.)

The company is constantly growing and has been nominated in Deloitte's Technology Fast 500 EMEA list, which consists of 500 fastest growing technology companies in Europe, Middle East and Africa. The company aims to establish their position as a significant international provider within exergaming products. The company's own staff develops sports games, both in terms of software and hardware. Several global companies are used as component suppliers. The company has supplied products to more than 40 countries. All products are manufactured in Finland. (CSE Entertainment, 2021.)

runBEAT is one of the products in CSE's BEAT-franchise. It utilizes a curved non-motorized treadmill which is designed and manufactured by DRAX. Treadmills are wirelessly connected to the runBEAT-system and therefore give the freedom on where to place treadmills. From two to eight

treadmills can be connected to one runBEAT system. The total space required for two treadmills is approximately 9 m^2. (CSE Entertainment, 2021)

runBEAT has five main modes, which are Free Run, Exercise, Fitness games, Competition and Fitness Test. The Free Run allows the player to run in a virtual environment of their choosing without time or distance limit. The player can start Free Run either by choosing the option manually from the menu or simply by just walking on the treadmill. A player can join in and leave at any given time even. Free Run allows the player to keep track of their time, distance, pace and heart rate. Currently, the player can choose between Greece, City and Forest. (CSE Entertainment ,2021)

Exercise mode includes warm up, interval and cooldown exercises. In exercise mode, the user is guided through the selected exercise by speed or heart rate. In a speed guided exercise mode, a short calibration will be done in the beginning. Based on the calibration, the user will have an individual exercise profile, where the user tries to match current speed to target speed, which changes several times during the workout. The user will have constant feedback during and at the end of the exercise. In the heart rate guided exercise, the user will set a maximum heart rate value to runBEAT software. Based on the set maximum heart rate, the user will have an individual program, where the user will see their current heart rate and target heart rate, which will change several times during the workout. (CSE Entertainment, 2021)

The Fitness game mode includes different games, where the treadmill is used as a game controller. Avatars are moving on the screen based on the treadmill's speed. The Competition mode allows the players to compete in a set distance race, ranging from 50 meters to 5000 meters. The players have a ten second window to join by running on the treadmill. After joining successfully, the game will countdown from three starting the race to the finish line. As the runBEAT provides instant feedback on the screen, the player knows if they have to run faster to win. (CSE Entertainment, 2021)

The Test mode consists of two tests: UKK 6-minute walk test as well as the Cooper 12-minute running test. Each mode can be played either alone or up to eight players. Before initiating any of the tests, the game presents a screen with information about the tests as well as possible health cautions. After each test, the player receives their test results according to their age group and sex. The average values of these charts for different age groups have been compiled from several different studies and other sources. The purpose of the charts is to provide an indication

of the player's fitness level. (CSE Entertainment) On the Cooper Test, the results show the treadmill distance as well as the real-world estimation. This estimation is approximately a 20% increase on the treadmill distance, imitating the 20% more strenuous nature of the curved non-motorized treadmill promised by the manufacturer. To correlate this, runBEAT multiplies the result with a 1.2 multiplier to imitate overground running. (CSE Entertainment, 2021)

The player can set Pace Runners by adjusting their desired distance covered in the test. For instance, in 12MRT the player sets one Pace Runner's total distance as 3000m. The Pace Runner will run at a pace that will cover 3000 meters in the 12-minute time limit. Ultimately, the player can set up to eight different Pace Runners with different distances; therefore, making it easier to achieve the preferred goal.

runBEAT uses a treadmill manufactured by DRAX. The Draxfit+ is a curved non-motorized treadmill where the individual themselves adjusts the speed by their stride and effort. The curved design together with the lower center position allows the individual to increase speed with bigger steps, and lower speed with smaller steps. In 2018, Schoenmakers and Reed published a study which suggested that self-powered treadmills can burn 30% more calories than a traditional treadmill. ("DRAX", 2020)

3 Theoretical background

The following chapters present the essential information, terminology and theories based on this thesis. Moreover, it introduces keywords such as curved non-motorized treadmill for example. On later occasions, they are abbreviated. For instance, the curved non-motorized treadmill is abbreviated as cNMT mostly but on some occasions (based on authors) it can be referred as NMCT.

3.1 VO_2 max

The overall performance of the respiratory and circulatory system or the efficiency of the aerobic energy metabolism is best described by maximum oxygen uptake (VO_2 max). Maximum aerobic power is reliable, widely accepted indicator of the cardiovascular performance, when it is expressed as oxygen consumption per kilogram body weight per minute (ml x kg^{-1} x min^{-1}). Maximal oxygen consumption describes the body's ability to transport oxygen for the working muscles and the body's ability to utilize it for the aerobic energy production during maximal strenuous activity involving large muscle groups. (Keskinen et al., 2004)

VO2max is a value that measures the efficiency with which the individual can utilize oxygen while exerting themselves. The more oxygen an individual can use during high level exercise, the more energy they are able produce. This can be assessed directly by collecting expired gases in a controlled laboratory setting or indirectly through field tests. Indirect tests produce more variable results than the direct tests as they give an estimation of VO_2 max from the performance of a physical task rather than directly measuring oxygen uptake. (Penry, Wilcox, & Yun, 2011)

In practice, maximum oxygen uptake is the individual level reached by the circulatory and respiratory systems of the body. VO_2 max represents the maximal ability to utilize oxygen in litres of blood per minute, while ml/kg/min reflects the effect of body weight on the maximum value. Factors affecting VO_2 max are age, gender, training status, altitude, the number of working muscles, load model of the test, duration of the load. (Keskinen et al., 2004)

According to Penry et al. (2011), VO_2 max is measured most accurately and reliably through direct laboratory methods. This is done by collecting expired gases in a supervised laboratory setting.

Consequently, the direct testing is costly and time-consuming, not suitable for testing bigger groups. The most reliable way to assess the individual's VO_2max is by conducting a direct test in a laboratory setting therefore negating the possible measurement variables. Field tests are more practical way to assess a bigger group of individual's fitness as they do not require expensive laboratory setting. Furthermore, the field tests can be performed by multiple individuals at the same time, therefore being less time consuming. Direct laboratory tests require a laboratory setting, specialized staff as well as standardized measurement conditions. (Penry et al., 2011)

Due to this, different tests have been created to measure VO_2 max. The most important distinction between the tests being whether the test is performed in a laboratory environment (direct method) or in a non-laboratory environment (indirect method) (Cooper, 1968; Penry et al., 2011) These indirect, field tests generate more varying results than the direct laboratory variants as they estimate the VO_2 max rather than directly measure the participant's VO_2 max. Furthermore, as the field tests are not conducted in a controlled laboratory environment, the possibility for typical errors can arise. (Penry et al., 2011)

A maximal test predicts the maximal oxygen uptake better than a sub-maximal test. However, testing to the point of exhaustion puts a greater strain on the cardiovascular system than a sub-maximal test. Therefore, it is best suited as a testing method for young, people with the lowest possible risk of cardiac complications and for those who have had a medical check-up and been found to be healthy. (Keskinen et al., 2004)

3.2 Physiology of VO_2 max

During an endurance performance or training, the physiological demands of the human body increase. The human body requires energy and oxygen amongst other molecules (McArdle, Katch, & Katch, 2007). That energy is ATP, Adenosine triphosphate. When the workload of the muscles increase, more ATP will be expended and must be replaced in order to continue. (McArdle et al., 2007)

The intramuscular stores of ATP and PCr (immediate energy system) supply the energy for exercise during brief, intensive bouts (such as 100-meter sprints, lifting heavy weights). The majority of the energy is produced by anaerobic (short-term energy system) reactions of glycolysis during

less intense exercise that lasts longer (1 to 2 minutes). The aerobic system (long-term energy system) takes over as soon as exercise lasts more than several minutes. (McArdle et al., 2007) All energy systems are used continuously, however the contribution of each energy system to the ATP production is dependent of the needs of the muscles (McArdle et al., 2007).

The slow-twitch (ST) or type 1 muscle fiber, produces energy primarily through aerobic pathways. The capacity for ATP resynthesis is connected to an individual's VO_2 max levels. This means that type 1 muscle fibers produce ATP, during aerobic exercise, such as attempting a Cooper's Test in this scenario (McArdle et al., 2007). The majority of bodily functions require the molecule ATP, which can be obtained from through several chemical energy sources such as nutrition (food). The ATP molecule stores chemical energy that the type 1 muscle fibers can use to execute work. (McArdle et al., 2007)

The maximal point at which aerobic ATP re-synthesis can occur as well as the level of metabolic and physical loading at which oxygen consumption plateaus despite additional increase in exercise intensity signifies the maximal oxygen consumption, VO_2 max. (McArdle et al., 2007). Continuing this will only utilize energy from glycolysis, resulting lactate accumulation and ultimately soon the runner to be exhausted and unable to continue.

Oxygen has a vital part in the synthesis of ATP (Reece et al., 2014). When oxidative phosphorylation occurs, oxygen acts as an electron acceptor, making the molecule electronegative. This characteristic is crucial for generating a difference in concentration between two mitochondrial membranes. The ATP synthase protein complex drives protons from the mitochondrial intermembrane gap back into the matrix of the mitochondria, where they are phosphorylated to form ATP from adenosine diphosphate (ADP). (Reece et al., 2014)

The VO_2 max provides a quantitative measure of an individual's capacity for aerobic ATP resynthesis. Therefore, VO_2 max is a significant determining factor of how well an individual can maintain high-intensity performance for longer than 4 minutes (McArdle et al., 2007). Women on average achieve 15-30% lower VO_2 max scores compared to men (McArdle et al., 2007). The differences have been typically justified with the differences in body composition and haemoglobin concentration. Furthermore, the normal physical activity levels can differ between average woman and average man. McArdle et al. (2007) continues, that almost 70% of differences in VO_2 max scores among individuals are explained by the body composition (McArdle et al., 2007).

3.3 Cooper test – the indirect VO_2 max test

The Cooper 12-minute running test in an indirect method to assess individual's maximal oxygen consumption (VO_2 max). Cooper's own study reported a correlation of 0,897 between a direct VO_2 max and the VO_2 max estimation in a 12-minute test (Cooper, 1968). Other studies have shown similar results within the reliability, where the coefficients were 0,92, 0,93 and 0,96 (Aitchison, Amjad, Corbett, Grant, & Wilson, 1995; Penry et al., 2011; Bandyopadhyay, 2015). Cooper's own repeated test for the subjects brought a good estimate of VO_2max with little to no training effect as the scores were highly similar with the previous test. (Cooper, 1968)

Before Cooper's test, a method to estimate VO_2 max was Balke's field test. In 1963, Balke developed a 15-minute running test in order to assess soldiers' aerobic endurance. The average energy consumption or oxygen consumption were calculated based upon the time and distance they covered. Balke discovered that the values were similar with the VO_2 max values and created a prediction formula that could be used to predict the Vo_2 max within the 15-minute running test. Cooper's purpose was to develop a test to estimate VO_2 max with accuracy from a 12MR by modifying the Balke protocol and search for a correlation between a 12-minute run and individuals' VO_2 max and with those results create a predictive method of VO_2 max using a standardized 12-minute run. Cooper (1968, 1970, 1977) shortened the time to 12 minute and presented an alternative 1,5 miles running test for the 12-minute running test. (Keskinen et al., 2004)

The Cooper 12-minute running test is one of the most popular field tests to estimate VO_2max (Bandyopadhyay, 2015). Moreover, the test determines the individual's cardiovascular fitness and can be used as a relevant method of assessing changes in fitness lever over time (Cooper, 1968). Similar findings were found my Penry et al. (2011) and Aitchison et al. (1995).

Cooper's own study had a wide range of age groups as the range was from 17 to 52 years, average age being 22 years (Cooper, 1968). This thesis is performed similar to test that Cooper (1968); however, the subjects' age range is not as wide as in Cooper's subjects in 1968. However, the mean age will be similar. Other studies related to 12MR have used the following age groups: 21,8 ± 3,6 (Penry et al., 2011), 22,85±1,69 for females and 22,8±1,7 for males (Bandyopadhyay, 2015), 22,1 ± 2,4 (Aitchison et al., 1995). As it can be seen, most of the tests were performed on age groups of roughly 21,8-22,8 ± standard deviation. This is the age range on this research as well.

On few instances, older and younger age groups have been studied: 34,5 ± 1,9 (Alvero-Cruz, Giráldez García, & Carnero, 2016), 18,4 ± 1,54 (Das, 2013).

As mentioned, Cooper's Test is an indirect test, and it requires no laboratory equipment. It is low-cost and less time consuming to conduct that direct VO_2 max testing since the test administrator can conduct the test for multiple participants at the same time (Aitchison et al., 1995; Penry et al., 2011), however since it is a maximal VO_2 max test and strenuous by its nature it can lead to injuries. The test requires a high level of motivation to test the actual maximum level since it could otherwise influence the results (Cooper, 1968; Aitchison et al., 1995; Penry et al., 2011). Aitchison et al. (1995) continue that low motivation may affect differences between direct and indirect VO_2 max testing. Moreover, Cooper 12MR may be less accurate on population with weaker cardiorespiratory fitness level (Penry et al., 2011).

3.4 Non-motorized treadmill

Motorized treadmills are commonly used and considered as a legitimate alternative when it comes to measuring overground running performance. According to Jones and Doust (1996), a 1% gradient increase on the treadmill would measure between the traditional running modality by achieving a strongly correlated oxygen update (VO_2) measures. Benefits of treadmills is that the rate of works remains constant since the participants cannot continue if the belt goes too fast and therefore the work rate cannot be maintained (Wilmore & Costill, 1994). Furthermore, it is easier to track and analyse for the test administer as the participant stays in the same place during the test (Keskinen et al., 2004).

The number of studies is fewer when the non-motorized treadmills are compared. These treadmills require the individual to drive the belt instead of a motor. Therefore, it has been suggested that non-motorized treadmills have a greater resemblance to overground running when comparing to treadmills with a motor (Wee et al., 2016). Wee et al. (2016) continue that the non-motorized treadmill is closer to track running compared to motorized treadmill as the participant determines the speed of the belt on the non-motorized treadmill with each step just as they would with track running. As in track locomotion, a non-motorized treadmill belt requires the leg to be manually pushed during each stride. Contrastingly, due to the high intrinsic resistance of the

treadmill belt on non-motorized treadmills, the power needed to move the belt increases with speed. (Wee et al., 2016)

Curved non-motorized treadmill, as its name states, is designed with a curved platform distinctive from the non-curved variants. The curved treadmill is designed so that the individual has an ability to run without restrictions such as a harness. Moreover, the curved design allows the individual to reach full speed by using similar running techniques as they would on a track (Gonzalez et al., 2013). According to Gonzalez et al. (2013), the individual cannot reach full speed on flat treadmills immediately and they must conquer a resistance to start running.

Smoliga et al. (2015) had previously compared physiological and perceptual responses while walking and running on a NMCT and a MT with same velocities. They found ultimately that the energy expenditure was higher on both instances on a NMCT. Oxygen uptake was 70% greater while walking and 43% while running at 8km/h respectively. However, these differences between the modalities could be explained as there were no familiarization trials before testing. (Wee et al., 2016)

Wee et al. (2016) continued upon this by comparing perceptual (Rate of Perceived Exertion) and physiological responses (oxygen uptake, heart rate and blood lactate concentration) when running on track, MT and NMCT. The results were that both physiological and perceptual values were higher on NMCT than compared to track or MT heart rate being 22% higher and oxygen uptake 37% higher. One of the reasons for greater perceptual and physiological responses on the NMCT might be the treadmill's curve as the contact area is higher due to the angle it creates since this increases the force the participants must use to maintain a constant speed. (Smoliga et al., 2015)

Penry et al. (2011) noticed that the variance due to physiological factors or issues with the familiarity of the testing modality (familiarization with the cNMT) is likely amplified with participants with lower aerobic fitness and motivation. This is since during the test (Cooper Test), the participant must pace themselves over the duration of the test rather than running for set distance. Moreover, similar findings of the importance of participant's motivation were found by Cooper (1986).

As everyone might not be familiar with the running modality, the cNMT, conducting familiarization trials should be done prior to testing. It has been suggested that a minimum of two familiarization trials, separated with at least 48 hours of each other, should be implemented prior to

testing on the cNMT to improve reliability (Wee et al., 2016). These trials would therefore eliminate or minimize the possible systematic error caused by the likelihood of a learning effect. Similar findings related to familiarization trials were found by Lakomy (1987), Hopker et al., 2009 to reduce the potential learning effect (as cited in Gonzalez et al., 2013).

To reduce any error caused by the training effect in VO_2 max, tests should be conducted within a 6-week period. (Penry et al., 2011) Penry et al. (2011) state that according to Carter et al. (2000), a 6-week training period of intense endurance training could increase VO_2 max value by 3-4%. Nevertheless, the participants were requested to retain their current activity for the testing period.

Previous studies of similar cNMT (Woodway Curve XL) by Woodway have shown good validity and reliability when assessing the VO_2 max, endurance efforts, sprint as well as recurrent sprint intervals. (Schoenmakers & Reed, 2018) Moreover, Schoenmakers and Reed (2018) added the findings of higher perceptual and physiological demands and increased cadence while running on the cNMT.

4 Research problems

The purpose of this thesis is to assess the validity and reliability of the Cooper 12-minute running test performed on a curved non-motorized treadmill while using runBeat -software. By assessing its validity and reliability, the possibility to assess how strenuous it is compared to traditional overground Cooper test is possible. This is implemented by comparing the results from tests (distance covered in the 12-minute time limit and furthermore with the VO_2 max estimation that can be measured from the distance).

The aims of this thesis are to obtain results and an analysis regarding the tests that have theoretical basis supporting them, which then benefits the commissioning party in the sports and fitness market. The author is able to utilize their acquired knowledge and skills. Moreover. they can demonstrate their fundamental knowledge of anatomy and physiology, physical activity and fitness testing as well as management of planning, execution, and evaluation of extensive modes and modules.

The findings of the study could contribute to the society by adding further knowledge regarding the reliability of Cooper 12-minute running test, specifically on the curved non-motorized treadmill. Especially since there was no other research where a test was conducted on a cNMT. In other words, the study could indicate whether or not the test could be reliably repeated with comparable results. In practice, bringing a new, less weather and running surface dependent option for conducting the Cooper test. Other option is by inspecting a possible difference between running modalities if there would be a distinction between results. This discussion could possibly be of use for coaches, trainers, instructors, sports teachers, future research and naturally for the commissioning party.

The thesis is important for the sake of the professionality of the author as the subject supports the author's specialization towards fitness testing, exergames and health promoting. Moreover, this thesis is valuable for the commissioning party as they can use this result in developing the software more as well as for marketing purposes. The aim within the Sports and Leisure Management in Kajaani University of Applied Sciences is work-life orientation as the university gives the student capabilities to become a sport professional.

Research problems are:

1. Is the Cooper's 12-Minute Run valid and reliable on the cNMT and on runBeat?
2. Are there differences regarding the challenge and physical demand between the tests?
3. Can runBeat and cNMT be used as proper testing devices for the Cooper Test?

The hypothesis the author has based on previous studies (Schoenmakers & Reed, 2018; Wee et al., 2016; Smoliga et al., 2015) is that the Cooper Test could be more strenuous and physically demanding than when performed overground. Schoenmakers and Reed (2018) found that the average oxygen uptake was 32,5% higher on the cNMT. Wee et al. (2016) reported a 37% higher Vo_2max while testing 1000-meter running test. Whereas Smoliga et al. (2015) observed a 43% increase while running at 8km/h and a 70% increase while walking on a non-motorized tread-mill.

5 Research methods

Quantitative research method provides projectable research outcome. It is a process of collecting and analysing numerical data. It can be utilized to find patterns, averages, variables, test causal relationships and generalize results to a wide population. Within the quantitative research, the researcher employs a sample which represents the chosen group or population. This method can be typically observed and objectively measured and analysed as it produces numerical data. (Gratton & Jones, 2010) This goes hand in hand within this study as the author himself will not be participating in tests, rather they will objectively observe, measure and analyse the tests and results. Moreover, as the variables (distance and the estimation of VO_2 max) are directly measurable within values and numbers, which can then be statistically analysed. Consequently, determining whether the variables are similar to one another. Therefore, making it more beneficial for this study than the qualitative research methods. In this study, beforementioned variables (VO_2 max estimation and distance) were chosen as the main values to be observed and analysed.

Reliability is a benchmark in quantitative research. It commonly implies to the consistency of the acquired results. Test-retest reliability is when the test would provide the same measurements and results if repeated at later time. Therefore, proving the stability and consistency of the test. (Gratton & Jones, 2010; Atkinson, 2012) Test-retest is the extent that will be used for this research as this method has been used in previous tests regarding the reliability of the Cooper 12-minute running test (Penry et al., 2011; Bandyopadhyay, 2015) Within this thesis, the reliability is assessed by comparing the results gotten from the two tests, are they close to one another. Hence the testing environment and testing aspects will be made as close to one another, naturally the testing modality will be different within the tests. Moreover, to minimize training effect, the testing period is done within a one or two weeks (Penry et al., 2011).

Second part is related to the validity. Validity refers to is the test (method) measuring what it is set to measure and therefore, are the conclusions gathered from the measurements valid. (Gratton & Jones, 2010) Within this thesis, the validity is assessed by verifying that do both tests measure the maximum oxygen uptake (VO_2 max).

Concepts of reliability and validity are alike and can therefore work in unison. The desired outcome is that the research is both valid and reliable. In this case, the measurement has been

identified correctly to reflect the phenomenon, consequently resulting in the same outcome. In other cases, the research can be either “valid but not reliable,” “reliable but not valid” or in the worst case “not valid and not reliable.” (Gratton & Jones, 2010)

This research compares the same test, 12MR on two different running modalities: indoor track and a curved non-motorized treadmill. As the purpose is to compare these tests with one another, Pearson’s correlation was used to determine the correlation between the tests. Moreover, a two-tailed paired samples t-test was utilized in order to determine whether there are statistically significant differences in the mean of the results and to reject or accept the null or alternative hypothesis. (Taanila, 2019) The means will then be compared against one another.

Before testing or familiarization trials, certain pre-test measures were done. These included performing a screening of health risks via a health questionnaire (See Appendix 1), obtaining an informed consent as well as information regarding previous Cooper Tests from the participants through the Cooper Familiarity document (See Appendix 2). Participants will be fully informed verbally and in written form of the purposes and potential risks of the study.

5.1 Participants

Fourteen participants (mean age = 22.9 ± 2 years), which included four women and ten men volunteered for this study. However, due to illnesses or other personal reasons, six participants dropped. This left the total number of participants at eight (mean age = 24.1 years ± 7 years), which included one woman and seven men. The mean height was 176,5 ± 5,81 cm and the mean weight was 73 ± 7,50 kg. Five of the participants were familiar and had completed the Cooper 12MR previously. These previous results ranged from 2950 to 3150 meters. Whereas only three participants had previous experience of the cNMT. Three participants rated their own fitness level as excellent. Four assessed their own fitness level as good and one participant rated their own fitness level as average. Table 1 presents these aforementioned characteristics.

Table 1 Descriptive characteristics for the measured sample

Description	Mean ± SD
Age (years)	24.13 ± 7.34
Height (cm)	176.5 ± 5.81
Weight (kg)	73 ± 7.50
Experience with 12MR (participants)	5
Experience with cNMT (participants)	3
Own fitness rating (excellent)	3
Own fitness rating (good)	4
Own fitness rating (average)	1

Participants were all sports students of Kajaani University of Applied Sciences. The participants were contacted via e-mail, in-class announcements as well as word-of-mouth advertising. Individuals of all fitness level were encouraged to participate if they did not possess more than one risk factor for cardiovascular disease. Participant's consent for the study will be obtained through a written informed consent form, which includes all necessary information. This form, alongside test procedures will be presented to acquire the Investigating/Research permission from Kajaani University of Applied Sciences (See on Appendix 1).

Prior to choosing university students as the sample, the author attempted to contact representatives of various sport clubs to gather participants. Due to not finding suitable time slots, the author decided to change the sample group to sport students at Kajaani University of Applied Sciences. The mean age was attempted to keep same the as within previous studies related to Cooper Test as well as general usage of non-motorized treadmill, meaning the studies that have been referred to earlier.

Sampling was done through serendipity – by taking anyone who volunteered as long as they hit the criterion of not having any illnesses, diseases, injuries or risk factors that could have influenced the tests or harmed the participants. Furthermore, as participants were university students who studied sports, meaning they had similar mean age group as Cooper (1968), Aitchison et al. (1995), Penry et al. (2011) and Bandyopadhyay (2015). Moreover, the sports studies could mean that the participants would be accustomed to running and fitness testing, making them an appropriate target group.

5.2 Study design and procedure

Prior conducting the tests, the author conducted a literature review regarding previously conducted studies as well as familiarizing within the topics. This review consistent of various scientific articles, books and articles found on the Internet and in the library. After the recruitment, the participants filled out the beforementioned Health Questionnaire and Cooper Familiarity form (See Appendix 1 and Appendix 2), based upon these, the participants were divided into two groups. Both groups completed two familiarization trials and two 12MR (one on the cNMT and one on the indoor track). After this, the data was recorded and analysed.

5.3 Measures

The participants were divided into two groups: group A would first perform 12MR on cNMT and then on the indoor track. Whereas group B would first perform 12MR on indoor track and then on the cNMT. The counterbalanced order aids to avoid bias and minimize the learning effect (Bandyopadhyay, 2014). This was to ensure any possible learning effect even though Cooper (1968) found that repeated tests brought highly comparable results. Moreover, the setting was chosen as both the author as well as the commission party were interested if there would be any major differences with the order. Indoor track testing was performed in Pallohalli Kajaani, and cNMT testing was conducted at the commissioning party's premises.

Grouping was done by utilizing a matched pair design according to the participant's previous Cooper test familiarity and previous record. The person with the best Cooper result was placed in Group A and the person with the second best, was placed in Group B and so on. ("Matched Pairs Design: Definition", 2021; (Choueiry n.d) Moreover, a randomized block design was used in the design. This method was chosen as it could remove gender as a potential source of variability and as a potential confounding variable as men and woman are physiologically different and therefore react differently as well. ("Randomized Block Design: Definition", 2021; Choueiry n.d) As with this method, the groups and matched pairs will be done so that they have equal or as

close to equal as possible number of women and men. However, due to the low number of participants, a perfect balance could not be established.

Between each test would be at least 48 hours, and preferably 72 hours to attain an adequate recovery. Prior to any testing, all participants completed 2 familiarization trials on the cNMT with 48 hours between the session to minimize the learning effect as well as improve reliability within the tests. (Gonzalez et al., 2013; Wee et al., 2016)

5.4 Overground procedures

Participants were instructed to cover as many laps as possible on a 300-meter indoor track by running, walking, or combining both within the 12-minute test period. Emphasis was on running as they were encouraged to give their best effort, to pace themselves, to speed up if they felt it was possible or slow down if they felt the need. The preferred track size for Cooper Test is 400 meters to minimize turning but 300-meter track works as well. Due to the size of the facility the test was held in, the 300-meter track was utilized. The track was measured with an open reel measurement tape by the test administer as well as by a test assistant a total of three times. Afterwards, cones were placed to indicate the track for participants. The ground was flat, and the test administrator had constant view of all participants during the entire test.

Participants could choose their own pace, preferably running but walking was allowed whenever the participant would feel the need. If necessary, participant could walk, rest or stop the whole test altogether. The participants were allowed to quit the test at any given time. Furthermore, the test administrator had the right to stop the test if they felt it necessary. Participants were told not to race against one another rather give their own best effort.

The tester counted the completed laps during the 12-minute test period by marking them down. During the test, the tester gave updates on regarding the elapsed time and distance. Every three minutes (three, six and nine minutes), called this to the participants as well as gave verbal encouragement to the participants. Moreover, the tester gave estimations of the participants result according to their current pace and what the result would be if they would keep that same pace. Though these steady pace estimations were rough estimations as the track was 300 meters instead of the usual 400 meters. During the last minute, participants were encouraged to give their

last-minute sprint. After 12 minutes had passed, the participants were told to stop and stay at their spot until the test administer got there to mark the result.

The total distance is measured by each full round the participant completed multiplied by the distance of the course added by the distance covered on the unfinished. The distance ran was measured to the nearest meter. Afterwards, to get the VO_2 max estimation, the running distance was converted by using Cooper's standardized equation derived from the original article by Cooper (1968) VO_2 max = (35.971 x distance in miles) – 11.288 or VO_2 max = (22.351 x distance in kilometres) – 11.288. (Cooper, 1968) In this case, the latter version was used for the sake of convenience.

Instructions and protocols were aimed to keep similar as possible within the testing modalities. For instance, the warm-up procedure was similar, only exception being instead of running on the track, the participants would run on the cNMT. Everything else will be attempted to keep the same as for an example, the rest of the warm-up (dynamic stretches) were the same. Instructions were similar, only exception being "running on the treadmill" rather than "running on the track." The warm-up consisted of five to ten minutes of jogging together with dynamic stretches, see Appendix 3 for the specific warm-up exercises. Cooldown consisted of five to ten minutes of slow jogging and walking, progressively lowering the speed as well as static stretches, see Appendix 3 for the specific cooldown exercises.

5.5 Curved non-motorized treadmill procedures

The participants ran on a curved non-motorized treadmill Draxfit +. Only three participants of the eight participants had experience from the cNMT. Therefore, the need for familiarization trials was emphasized. During the familiarization trials, preliminary instructions were given, a warm-up consisting of dynamic stretches as well as cooldown consisting of static stretches were held. The familiarization itself entailed getting on the treadmill, walking, jogging, running, practicing the Cooper pace with a 6-minute running test and getting off the treadmill. This was then repeated on the second familiarization session. Sufficient time for recovery was given between the two familiarization trials and the two test sessions.

On the cNMT 12MR, the participants were instructed to cover as much distance as possible while running the treadmill by running, walking, or combining both within the 12-minute test period. Same ground rules as with the track field Cooper test. The author set eight pace runners for the participants helping to pace themselves and maintain the desired running pace. The pacemakers' pace was set according to information gained from the Cooper Familiarity Form (see Appendix 2) as well as according to the participants' wishes.

After the 12-minute test period, the participants were told to stop running on the treadmill. No calculations were needed here as after the test, the program presented everyone's results as well as the chart to compare those results.

5.6 Procedures

All participants completed a total of two familiarization trials and two field tests, 12MR on an indoor track and 12MR on cNMT. Participants rested 72 hours between each test and trial. Familiarization and the tests were completed with two weeks. Participants were instructed to avoid any strenuous exercise a day before each test, participants nutrition was not monitored. Participants were instructed to avoid usage of alcohol, tobacco and caffeine for at least three hours before each test. Participants were instructed to arrive in their exercise attire, as well as avoid eating any big meals three hours before the test.

Testing on the cNMT was conducted on two different days, whereas testing on the track was done on a singular day with two sessions. Group A started on Monday with the cNMT and finished on Thursday with the track, whereas Group B started on Thursday with the track and finished on Sunday with the cNMT. Tests were done in groups of three to five participants, depending on the modality. Tests were done around the same time of day – 3 to 4pm as this was found to be the easiest option for the participants. Each participants wore same shoes and similar clothing for both tests.

Moreover, before testing a warm-up was conducted. Warm-up helps prepare the body for aerobic activity. It raises the body temperature and increases the blood flow, which then allows more oxygen to the muscles. It may additionally help reduce muscle soreness and lessen the risk of injury. The warm-up itself consisted of aerobic activity such as walking, jogging and running.

Additionally, dynamic stretches will be included. The warm-up procedure was 10-15 minutes as suggested by Keskinen et al. (2004). After the testing, a proper cooldown was done by doing aerobic activity with lower intensity and static stretches. This was done for five minutes as recommended by Keskinen et al. (2004).

5.7 Statistical analysis

For this thesis, the program PSPP -program (GNU PSPP version 1.4.1) as well as the Microsoft Excel 2022 were used to analyse the results. The results were illustrated by various tables and graphs such as box plots and histograms. Graphs were created with Microsoft Excel 2022 -program. The reliability and validity of Cooper 12-minute running test will be conducted by comparing results gained from the two tests – Vo2max estimation and the distance covered. These results were compared by their means and standard deviation, the difference in means as well as the differences individually within the participants.

Data was analysed using PSPP (GNU PSPP version 1.4.1) and Microsoft Excel 2022 and are presented as mean ± SD. Moreover, lower and upper quartiles are presented with the median. Differences in Vo2max estimation and distance covered were compared between Cooper cNMT and Cooper overground using paired samples t-test. Correlation between the test were tested with Pearson's correlation coefficient.

Sig. = p. Significance level represents; how high the risk is that the sample difference or dependence is due to chance. If p-value is less than 0,05, is the result nearly significant. P-value that is less than 0,01 is considered significant and a p-value that is less than 0,001 is considered highly significant (Heikkilä, 2004).

6 Results

Originally, 14 participants volunteered for the study. However, in the end, eight participants took part in the familiarization trials and actual tests. Six participants dropped through the study process due to getting ill or other personal reasons. Eight participants (Age mean ± SD; 24 years ± 7 years) completed two familiarization trials as well as the Cooper Test on the cNMT and on track. Five participants had previous experience from the Cooper Test, previous results ranging from 2950-3150 meters and only (See characteristics on Table 1).

The results for the VO_2 max estimations and the Cooper Test distance of participants are presented with a box plot (see Figure 1 and Figure 3) as well as with a bar graph (see Figure 2 and Figure 4). Within these can be seen the minimum and maximum values as well as the lower and upper quartile, median as well as the mean values. Each participant achieved a higher score (covered distance and VO_2 max estimation) on overground. Possibly implying that the cNMT was more strenuous than overground.

The mean for the VO_2 max estimation on the track was 49,85 ml/kg/min (standard deviation = 7,84 ml/kg/min n=8) was greater than the mean of the VO_2 max estimation on the cNMT 40,24 ml/kg/min (SD = 9,18 ml/kg/min n=8). By dividing the mean of VO_2 max on the track (49,85 ml/kg/min) with the mean of VO_2 max on cNMT (40,24 ml/kg/min), it was found that, the participants achieved approximately 23,88% greater VO_2 max while running on the track. This could therefore mean that the cNMT was physically more demanding, which caused the differences between the tests.

The lower quartile (Q1) represents the value under which 25% of data points are when the data is arranged in ascending order. The upper quartile (Q3) is the value under which 75% of the data points are when the data is arranged in ascending order. The median is also known as the second quartile (Q2). The interquartile range is the difference between Q3 and Q1. Q1 was found with equation ¼ * n(8) and then connecting the correct value from the ascending order. Q2 followed the formula of 2/4* n (8) and Q3 was determined with the formula of ¾*n(8) (Taanila, 2019).

Comparing the lower and upper quartiles within the VO2max found that the lower quartile (Q1) on track was 42.35 ml/kg/min whereas on cNMT it was 30.44 ml/kg/min. Median (Q2) was 45.7

ml/kg/min on the track and on the cNMT it was 38.51 ml/kg/min. The upper quartile (Q3) however was 57.37 ml/kg/min on the track and 49.51 ml/kg/min on the cNMT. These differences can also be seen on Figure 1 below.

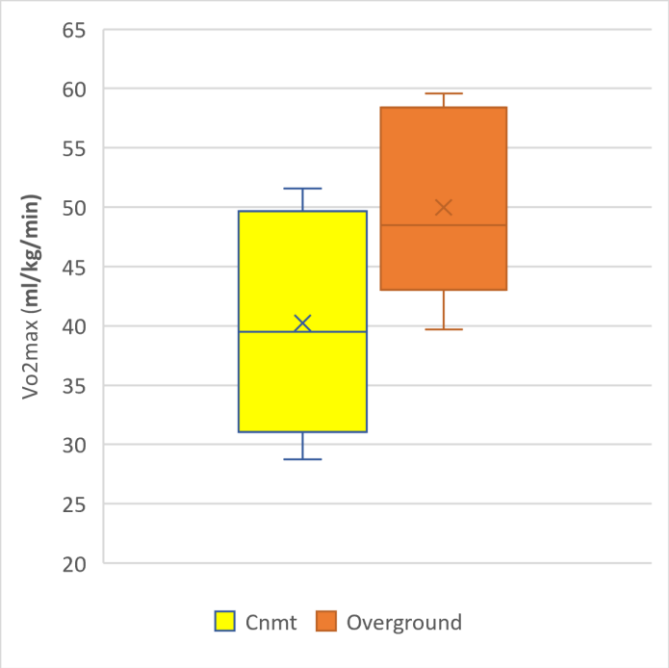
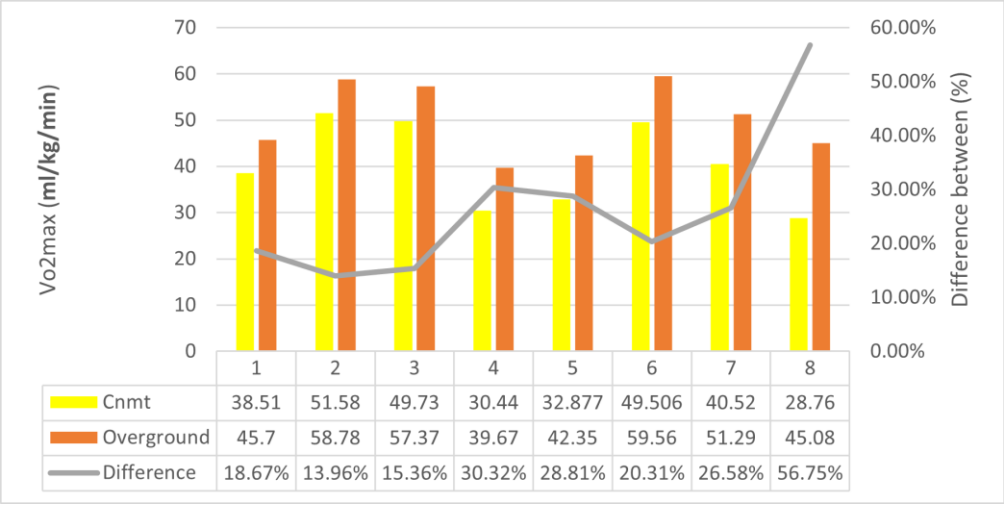


Figure 1 Box plot of participants' VO2max (ml/kg/min) comparison (n=8)



	1	2	3	4	5	6	7	8
Cnmt	38.51	51.58	49.73	30.44	32.877	49.506	40.52	28.76
Overground	45.7	58.78	57.37	39.67	42.35	59.56	51.29	45.08
Difference	18.67%	13.96%	15.36%	30.32%	28.81%	20.31%	26.58%	56.75%

Figure 2 Individual differences between VO2max estimations

The mean for the distance covered on the overground was 2742 meters (SD = 352 meters, n = 8) was higher than the mean of the distance covered on the cNMT 2306 (standard deviation = 412meters n = 8). Participants managed to cover approximately 18.91% more distance while running on track.

Comparing the lower and upper quartiles of the Cooper Distance found that the lower quartile (Q1) on track was 2400 meters whereas on cNMT it was 1867 meters. Median (Q2) was 2552 meters on the track and on the cNMT it was 2228 meters. The upper quartile (Q3) however was 3072 meters on the track and 2720 meters on the cNMT. These differences can also be seen on Figure 3 below.

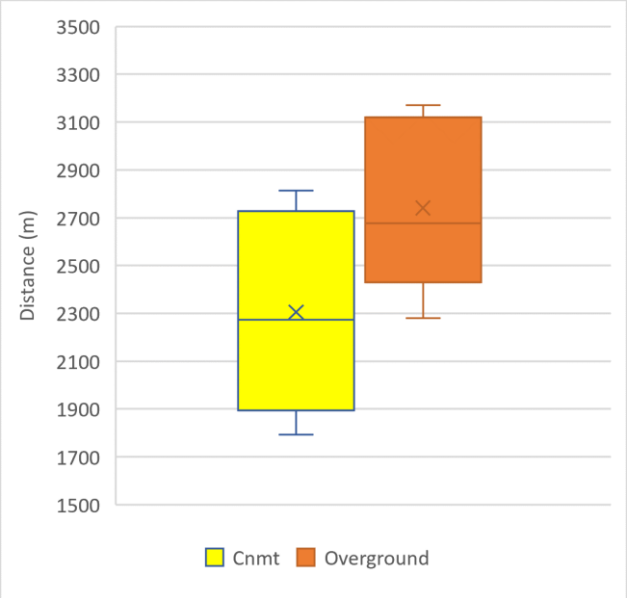
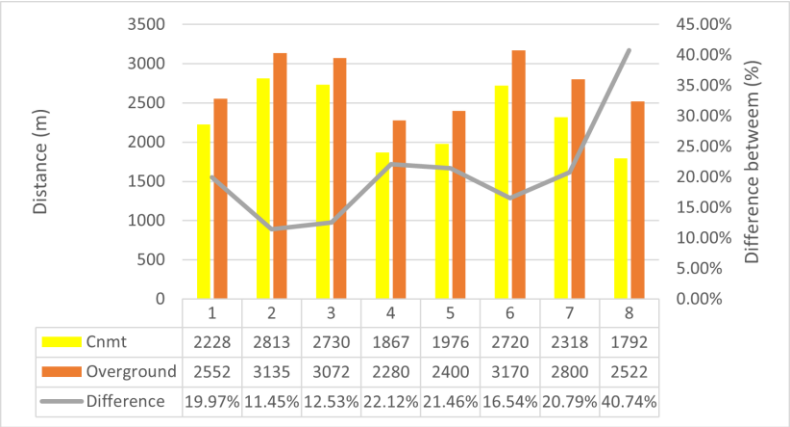


Figure 3 Box plot of participants' Cooper Distance (m) comparison (n=8)



	1	2	3	4	5	6	7	8
Cnmt	2228	2813	2730	1867	1976	2720	2318	1792
Overground	2552	3135	3072	2280	2400	3170	2800	2522
Difference	19.97%	11.45%	12.53%	22.12%	21.46%	16.54%	20.79%	40.74%

Figure 4 Individual differences between Cooper Test results

Paired samples t-test was chosen as measurements included the same group under two different conditions (cNMT and overground). The paired samples t-test indicates whether there are statistically significant differences in the mean of the results. (Taanila, 2019)

Null hypothesis would be that there would be no difference between mean in cNMT and Overground. Alternative hypothesis on the other hand, would be that there is a difference between mean in cNMT and Overground. The lower the p-value, the more support there is that the mean of the differences is significantly different from zero. (Taanila, 2019)

As the p-value is less than 0,001 (p = 0,000), can the result be seen statistically highly significant. Therefore, the null hypothesis can be rejected. Alternative hypothesis stays, meaning that there is a difference between the cNMT and overground.

The Appendix 4 shows the correlations between VO_2 max scores on overground and cNMT with each participant. Pearson's correlation r = 0,95 indicates a strong positive correlation on the participants VO_2 max scores on the track and on the cNMT.

The mean for the VO_2 max estimation on track was 49,85 ml/kg/min (SD= 7,84 ml/kg/min), n=8) was greater than the mean for the Vo_2 max estimation on the cNMT that was 40,24 ml/kg/min (SD = 9,18 ml/kg/min), n=8). The difference proved to be significant with the paired samples t-test:

t(7) = 9,21, p < 0,001, 2-tailed. The 95% confidence interval of the mean of the differences is 7,14 – 12,08. As seen on Appendix 4.

The mean for the Cooper Distance on track was 2741,38 meters (SD = 351,49 meters), n=8) was greater than the mean for the Cooper Distance on the cNMT that was 2305,50 meters (SD = 410,62 meters), n=8). The difference proved to be significant with the paired samples t-test: t(7) = 9,27, p < 0,001, 2-tailed. The 95% CI of the mean of the differences being 324,70 – 547, 05. As seen on Appendix 4.

Two-tailed test was chosen as this design is set to examine both sides of data whether a sample is greater than or less than the range of values. Furthermore, two-tailed test is used when it is not completely certain whether the mean of the differences is positive or negative. (Taanila, 2019) Typically, it is recommended to use t-test with a sample size of 30 (Taanila, 2019). However, de Winter (2013) presents that a paired t-test is acceptable and viable even with smaller samples (as small as N=2) and there would be no fundamental objections. However, the larger the sample size, the more reliable conclusions can be reached. (de Winter, 2013)

The correlation coefficient shows that there is a significant strong positive correlation between VO2max scores between overground and cNMT, [r(8) = 0,95, p = 0,000] and the result is statistically significant (Heikkilä, 2019). The correlation coefficient shows that there is a strong positive correlation between Cooper Distance covered between overground and cNMT, [r(8) = 0,95, p = 0,000] and the result is statistically significant (Heikkilä, 2019).

7 Discussion

The objectives for the study were to assess and determine the reliability and validity of estimated VO_2 max with Cooper's 12MR on cNMT on paired samples scenario. The results show a positive correlation of 0,95 and the result is statistically highly significant. (Heikkilä, 2019) In other words, the participants achieved higher scores (both VO_2 max estimation and covered distance) on track compared to the 12MR on cNMT. It shows that the cNMT test was physically more demanding for the participants as the scores were lower despite testing the same sample.

Cooper 12MR has yet to be validated on the cNMT but the results of the current study show a positive reliability correlation for the 20-30% more strenuous finding by Schoenmakers and Reed (2018) and could therefore be used for further studies in the area.

7.1 Results

The present study aligns with the previous studies pointing on the reliability and validity of the cNMT when testing a participant's VO_2 max. The study together with previous studies (Smoliga et al., 2015; Wee et al., 2016; Schoenmakers & Reed, 2018) indicate that cNMT is a more strenuous training modality.

Since the present study tested the correlation between Cooper 12MR on the track against the cNMT, further studies could be conducted by re-testing the cNMT test only. By doing the exact same test twice on the cNMT and comparing these results would then show how repeatable the cNMT 12MR is and whether any learning effect occurred within the cNMT.

Each participant had differences in their VO_2 max value between the tests ranging from 14% difference to 56% difference with the mean difference of 23%. The differences in the distance covered ranged from 11% to 40% difference with the mean of 21%.

The present study had similar result as Schoenmakers and Reed (2018), where the average oxygen uptake was 32.3 ± 4% higher on the cNMT. Within this study, the participants had 23,9 ± 13.7%

greater Vo2max on track compared to the cNMT, which can be a result of the more strenuous nature of the cNMT. Even higher oxygen uptake was found by Wee et al. (2016), where the Vo2max was 37% higher while testing 1000-meter run. Moreover, Smoliga et al. (2015) discovered a 70% increase when walking on NMCT and a 43% increase when running at 8 km/h.

Participants with previous experience from the Cooper Test performed the best and achieved the highest VO_2 max estimations on both tests. However, these individuals had no previous experience of the cNMT; showing that the running modality itself did not affect their results as drastically as with other participants in this scenario.

In terms of the quartiles on the VO_2 max estimations, it was found that the lower quartile (Q1) had a 39,1% difference alone between the tests. Interestingly, the median (Q2) had a 18,7% difference and the upper quartile (Q3) had a 15,9 % difference. This shows that the median and the upper quartile had smaller differences between the results than the lower quartile. The greater differences of the lower quartile could be explained as the results within the 12MR can be amplified or altered with individuals with lower aerobic fitness levels. (Penry et al., 2011) This could explain why the difference is greater on the lower quartile and smaller on the median and upper quartile.

Consequently, the quartiles on the Cooper distance showed that the Q1 had a 28,6% difference, whereas the Q2 and Q3 had a 13,2% and 12,9% difference respectively. As VO_2 max estimations are based upon these distances, similar findings can be seen here with the higher difference in distance within the lower quartile compared to the median and the upper quartile. However, the difference is smaller than with the VO_2 max estimations.

Grant et al. (1999) and Penry et al. (2011) mention the importance of motivation for performing maximal testing such as Cooper's 12MR and it can be applicable to the current study as well. Results may vary due to the motivation of the participants as they must pace themselves on the track and on the cNMT. There was a possibility that not every participant was motivated on the cNMT test, which could have caused some discrepancies or lower scores in general. With heart rate tracking, blood lactate concentration samples or direct VO_2 max testing, this could have been confirmed or denied. Varying results could also be explained by the new unfamiliar running modality that the majority of the participants were not accustomed to.

As mentioned, the mean difference between the VO_2 max scores was 23,9%. However, there was a participant who demonstrated a large score variability across tests whereas other participants did not, at least not to that level. The highest difference on the VO_2 max estimation between the tests was 56%. It could be possible due to the previously mentioned non physiological variables such as psychological factors or issues with the running modality as pointed out by Penry et al. 2011. Furthermore, as the score was 56% lower on the cNMT it may be due to the level of motivation and, therefore the participant did not exert themselves on a maximal level since they must pace themselves for a certain duration rather than running an exact distance. Indirect tests including a greater number of these inconsistent participants in a sample could potentially be expected to demonstrate lower reliability values.

Errors in scoring were likely to be very low on the cNMT, since the runBEAT -program tracked the whole test. On the track errors were low; however, there is always the possibility for a random error caused by the test administer when counting the laps and measuring the final lap distance. Instrumental error was identified low as the test administer was familiar with the measuring tape wheel and the instrument was functional. Since the track was measured and marked with cones manually, the test leader and the test assistant measured the track multiple times to obtain an even 300-meter lap. This meant, however, depending on how tightly or loosely the measuring wheel was turned; slight differences were possible to occur. The test leader did test this prior to the Cooper Test. Additionally, the test assistant tested and measured the track with the same distance providing a stabler and reliable track.

Participants were advised and told how to prepare for the tests (See Appendices 2). Other than that, there was no attempt to control the behavioural patterns of the participants prior and between the two tests. Factors such as alterations in sleep and diet may have caused variability into the data.

The paired t-samples test is known to be sensitive to the sample size and it is recommended for samples greater than 30. The larger the sample size is, the smaller the sampling error can potentially be. (Taanila, 2019). Therefore, the use of only eight participants may have had some effect on the results of this study.

The large loss in sample size affected the usefulness of the matched paired design together with the randomized block design to compare two groups with a balance of men and women as well

as having the baseline mean Cooper result as close as possible. Group B (track + cNMT) lost three participants during the process leaving the already small groups even smaller and imbalance.

7.2 Methods

An advantage that the cNMT has with runBEAT is its visual cues as they can help to maintain the pace as well as the motivation. (Jones & Doust, 1996; Edwards et al., 2017) During the test, the participant sees their avatar running in real time on the track with other participants, their distance, current pace, how much time is left, their speed as well as the average speed. By monitoring the visual cues, the participant knows whether they should increase their pace or keep it as it is.

When performing Cooper 12 Minute Running test, a 400-meter running track would be more optimal than the 300-meter that was used for the current study both in terms of the running economy and to guarantee that the track is 400 meters. However, no suitable 400-meter running track was available for this study due to the cold weather as well as the possibility of snow and ice, therefore the indoor 300-meter track was chosen. As both tests were performed indoors, the error caused by weather was minimized. However, small changes in temperature in the indoor track or at the Commissioning Party's premises could potentially have had a minor impact in the results.

Complications occurred during the recruitment of participants for the tests, especially when recruiting for maximal testing that would be done twice in a shorter notice. The initial goal was to recruit 15-20 subjects with no complications or health risks for testing. No other restrictions such as age, sex or fitness level were chosen as it was predicted that it could be difficult to recruit participants. After the recruitment process, 14 participants had volunteered and filled in the necessary documents. However, six participants dropped out before the testing had begun, leaving the total number of participants at eight. With the original participants, there was a better balance between men and women but as the number dropped, so did the balance.

Time was also a factor in decision-making; insufficient time to recruit participants and perform the tests had an impact on the decisions, especially in order to have the necessary 72 hours' recovery between the tests.

The optimal research would be to study validity and reliability of Cooper's 12 Minute Running test on the cNMT with a larger sample and compare it with direct VO2max testing, which was unfortunately not possible within the current study due to lack of resources and time.

7.3 Validity and reliability of the thesis

Familiarization results could have been recorded to discover if there would have been a learning effect between the two familiarization sessions. However, other studies have pointed out that there has not been any learning effect between sessions (Schoenmakers & Reed, 2018). This study does not lack limitations, and the results could be biased by the intensity of the test, so it could be argued that the participants did not exercise at maximum level or with the same effort in both test occasions as mentioned previously. By utilizing HR or direct VO_2 max testing equipment, the intensity of aerobic exercise could have been easily confirmed.

7.4 Development of own competences

During the thesis process, I was able to combine new knowledge with my previously acquired experiences and knowledge. I learned plentiful new matters and developed myself professionally. Within the thesis process, taking the full responsibility and planning work became even stronger since all tasks were only my responsibility. I have learned how to effectively acquire literature in the field of sport and how to assess it critically. Within the thesis, I tried to compare different studies and, where possible, choose the most recent data. Through this, I demonstrated my ability to understand the importance of source criticism. Naturally, there are studies that are older than others, but they are essential on the subject.

The topic of the thesis supported my professional development since one of my main interests in the sports field is sports science and fitness testing. During the thesis process, I learnt how to conduct quantitative research, how to analyse the results and how to present the results in a meaningful way. There were some shortcomings in the study that could not be amended due to time, space and personnel constraints.

The competences of the Sports and Leisure Management programme consist of sport competence, well-being and health competence, pedagogical and sport didactic competence, social, managerial and entrepreneurial competences in sports, events and activity tourism (Kajaani University of Applied Sciences n.d). In this study, I believe that the competences of sport and physical activity were acquired satisfactorily. A large part of the study was the Cooper tests and familiarization trials, which mainly consisted of testing the participants and guiding through the process. Planning the training (including warm-up, cooldown and testing) requires knowledge of human physiology, anatomy and coaching skills. It is essential to understand why matters are done and what is intended to be achieved with them. In this study, in my opinion, it was comprehensively justified on the basis of previous research, studies and literature what was done and why was it done.

During the thesis process, I gained a lot of practical experience as a tester. I see that within that aspect, there was a good amount of development. Prior to this, I had some previous practical experience and practical training but mostly within courses by the university so the familiarization trials and tests themselves were useful for professional development on testing. I believe that by becoming more familiar with fitness testing, I will be able to work as a specialist in the field of sport. Furthermore, learning about fitness testing has developed my ability to plan goal-oriented training programmes, as fitness testing (including maximal endurance testing) is an important part of training planning in order to see results and how to proceed. Ultimately, the thesis process has developed my professional knowledge and I feel that I am able to use what I have learned in practice.

7.5 Ethical considerations

As the participants were students of Kajaani University of Applied Sciences, a research permit had to be obtained prior to testing. This permission, however, would not replace the participant's consent form for the study. KAMK has committed to comply Good scientific practice guidelines provided by Tutkimuseettinen neuvottelukunta (TENK). (Tutkimuseettinen Neuvottelukunta, 2012)

The guidelines aim to ensure the ethical acceptability and credibility of the research results. The good scientific practice includes adherence to practices recognised by the scientific community, due diligence, precision in research and honesty. (Tutkimuseettinen Neuvottelukunta, 2012) This study was conducted according to these guidelines. The research data was stored on the researcher's own computer or in a database with up-to-date firewalls and antivirus software. After the analysis of the research results, the data are transferred to a USB flash drive or external hard drive for archiving. If the data does not need to be archived, it can be destroyed. The results are presented as they are without alteration to ensure the honesty and credibility.

All participants volunteered by themselves, read and signed the informed consent, health questionnaire and written participant information form (one copy and one original). The documents contained all relevant information related to the tests including requirements of the tests (goal and aim of the test, how to perform the test), potential health risks, ethical and social considerations, benefits of the test, and usage of personal information. The original documents were kept safe, as to keep it confidential, and the documents were approved by KAMK.

Moreover, the personal information of the participants was kept anonymous. The participants were able to withdraw from the study at any given time. No recognizable personal information is present in the analysis, as the names have been changed into numbers to conform the participant's confidentiality.

Findings of this study could contribute to the society by adding further knowledge to the reliability of the Cooper 12-minute running test, especially on the cNMT, which is interesting as outdoor tests have the weather variable that can alter the test results. Therefore, the places where it is difficult to organize an outdoor Cooper Test with preferable weather conditions could benefit from an opportunity to reliably conduct the test with the cNMT, knowing that the weather or

terrain will not affect the result. In other words, the study could indicate whether Cooper 12MR could be reliably repeated with the same outcome. In addition, if the outcome would have differences, discussions could help to discover how to achieve or explain the score differences, similarly to Jones and Doust (1996) who found that 1% gradient increase on the treadmill would reflect between overground by achieving strongly correlated oxygen uptake measures.

8 Further Notes

The usage of precise equipment and laboratory settings would be beneficial to obtain the most precise data as possible. Therefore, bringing the possibility to accurately assess the estimated values from the tests with the direct test conducted within a laboratory setting. In this thesis, only indirect tests were compared with one another.

The previously mentioned study by Schoenmakers and Reed (2018) compared physiological and perceptual response of running on a curved non-motorized treadmill (cNMT) and it is currently one of the latest studies regarding the subject. Furthermore, this study is currently the reference point made by the manufacturer of the treadmill.

As the sample contained mostly men, differences between the genders remain unanswered. This could be investigated in the future as women generally have lower VO_2 max values compared to men. Reasons for the differences are the body composition as women have more fat mass, less fat-free mass and blood haemoglobin content that is lower with women and, therefore, they have less oxygen carrying capacity (Keskinen et al., 2004). Moreover, according to Edwards et al. (2017) female participants with a lower body mass may have a disadvantage on the cNMT due to the treadmill belt resistance.

9 Conclusion

Cooper's 12 MR performed on the cNMT showed a high reliability for estimating VO_2 max according to the manufacturer's (Drax) 20% more strenuous statement, which was found by Schoenmakers and Reed (2018). The results show that the prefixed 1.2 multiplier in the runBEAT software is comparable to this scenario. Systematic errors could be reduced by providing motivation to every participant to guarantee an optimal testing, and by familiarizing less-active and less-experienced participants with proper techniques and information to obtain optimal results.

Further research could be conducted with a larger sample and the sample could be compared with direct VO_2 max testing to see how well the results correlate to actual VO_2 max by validating the test against a maximal treadmill test with breath-to-breath analysis.

Bibliography

Aitchison, T., Amjad, A.M., Corbett, K., Grant, S., & Wilson, J. (1995). A comparison of methods of predicting maximum oxygen uptake. British Journal of Sports Medicine, 29 (3), 147-152

Alvero-Cruz, J., Giráldez García, M., & Carnero, E. (2017). Reliability and accuracy of Cooper's test in male long distance runners. Revista Andaluza De Medicina Del Deporte, 10(2), 60-63. doi: 10.1016/j.ramd.2016.03.001

Atkinson, M. (2012). Key Concepts in Sport & Exercise Research Methods. SAGE Publications Ltd.

Choueiry, G. (N.d.) Matched Pairs Design: An Introduction. Retrieved 9 October 2022, from https://quantifyinghealth.com/matched-pairs-design/

Cooper, K. (1968). A Means of Assessing Maximal Oxygen Intake. *JAMA*, *203*(3), 201. doi: 10.1001/jama.1968.03140030033008

CSE Entertainment. (2021). Retrieved 12 October 2021, from https://www.cse.fitness/en_US/

Das, B. (2013). Estimation of maximum oxygen uptake by evaluating cooper 12-min run test in female students of West Bengal, India. Journal Of Human Sport And Exercise, 8(4), 1008-1014. doi: 10.4100/jhse.2013.84.11

de Winter, J.C.F. (2013). "Using the Student's t-test with extremely small sample sizes," Practical Assessment, Research and Evaluation, 18:10, August, ISSN 1531-7714

DRAX. (2020). Retrieved 13 October 2021, from http://draxfit.com/en/product/draxfit/

Edwards, R., Tofari, P., Cormack, S., & Whyte, D. (2017). Non-motorized Treadmill Running Is Associated with Higher Cardiometabolic Demands Compared with Overground and Motorized Treadmill Running. Frontiers In Physiology, 8. doi: 10.3389/fphys.2017.00914

Gonzalez, A., Wells, A., Hoffman, J., Stout, J., Fragala, M., Mangine, G. et al. (2013). Reliability of the Woodway CurveTM Non-Motorized Treadmill for Assessing Anaerobic Performance. Journal of Sports Science and Medicine 12, 104-108

Gratton, C., & Jones, I. (2010). Research methods for sports studies (2nd ed.). London: Routledge.

Heikkilä, T. (2004). Tilastollinen tutkimus. Edita.

Jones, A., & Doust, J. (1996). A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. Journal of Sports Sciences, 14(4), 321-327. doi: 10.1080/02640419608727717

Kajaani University of Applied Sciences (n.d). Bachelor's Degree in Sports and Leisure Management -Opinto-opas. Retrieved 17 October 2022, from http://opinto-opas.kamk.fi/index.php/en/68146/en/68091

Kajaanin ammattikorkeakoulu. (n.d). Tutkimusluvat. Retrieved 12 November 2021, from https://www.kamk.fi/fi/Tutkimus-ja-kehitys/Tutkimusluvat

Kari, T. (2017). Exergaming Usage: Hedonic and Utilitarian Aspects.

Keskinen, K., Häkkinen, K., Kallinen, M., & Aho, J. (2004). Kuntotestauksen käsikirja. Helsinki: Liikuntatieteellinen seura.

Matched Pairs Design: Definition. (2021). Retrieved 4 November 2021, from https://stattrek.com/statistics/dictionary.aspx?definition=matched%20pairs%20design

McArdle, W. D., Katch, F. I., & Katch, V. L. (2007). Exercise Physiology: Energy, Nutrition, and Human Performance (6th ed.). Philadelphia: Lippincott Williams & Wilkins.

Penry, J., Wilcox, A., & Yun, J. (2011). Validity and Reliability Analysis of Cooper's 12-Minute Run and the Multistage Shuttle Run in Healthy Adults. Journal Of Strength And Conditioning Research, 25(3), 597-605. doi: 10.1519/jsc.0b013e3181cc2423

Randomized Block Design: Definition. (2021). Retrieved 4 November 2021, from https://stattrek.com/statistics/dictionary.aspx?definition=randomized%20block%20design

Reece, J.B., Urry, L.A., Cain, M.L., Wasserman, S.A., Minorsky, P.V., Jackson, R.B., and Campbell, N.A. (2014). Campbell biology (10th edition). Pearson.

Schoenmakers, P. and Reed, K., 2018. The physiological and perceptual demands of running on a curved non-motorised treadmill: Implications for self-paced training. Journal of Science and Medicine in Sport, 21(12), pp.1293-1297.

Smoliga, J. M., Hegedus, E. J., & Ford, K. R. (2015). Increased physiologic intensity during walking and running on a non-motorized, curved treadmill. Physical Therapy in Sport, 16(3), 262-267. doi:10.1016/j.ptsp.2014.09.001

Suni, J., & Taulaniemi, A. (2012). Terveyskunnon testaus - Menetelmä terveysliikunnan edistämiseen (1st ed.). Helsinki: Sanoma Pro Oy.

Taanila, A. (2019, April). *Kahden riippuvan otoksen vertailu*. Akin menetelmäblogi. Retrieved 9 October 2022, from https://tilastoapu.wordpress.com/2012/02/14/kahden-riippuvan-otoksen-vertailu/.

Taanila, A. (2019, April). *Yläkvartiili – Akin menetelmäblogi*. Akin menetelmäblogi. Retrieved 8 October 2022, from https://tilastoapu.wordpress.com/tag/ylakvartiili/.

Tutkimuseettinen Neuvottelukunta. (2012). *Hyvä tieteellinen käytäntö ja sen loukkausepäilyjen käsitteleminen Suomessa*. Retrieved from https://tenk.fi/fi/tiedevilppi/hyva-tieteellinen-kaytanto-htk

Wee, V., Von Heimburg, E., & Van den Tillaar, R. (2016). Comparison of perceptual and physiological variables of running on a track, motorized treadmill, and non-motorized curved treadmill at increasing velocity. Acta Kinesiologiae Universitatis Tartuensis, 22, 20. doi: 10.12697/akut.2016.22.02

Wilmore, J., & Costill, D. (1994). Physiology of sport and exercise. Champaign, IL: Human Kinetics.

Appendices

Appendix 1 – Health Questionnaire and consent form

It is important that we know about your lifestyle and your level of physical activity before we test you. Please fill out the form carefully.

Name:_____ **DOB:**_____ **weight**____ **height**_____
Gender:_____

Symptoms over the past 6 months:	Yes	No	Unknown
1. Have you had chest pain?			
At rest?			
While exercising/moving?			
2. Have you had breathing difficulties?			
3. Have you experience light-headedness?			
4. Have you had experienced any feelings of arrhythmia?			
5. Do you have any pain that restricts movement/exercises?			
Where?			
6. Have you had any symptoms of stress or burn-out?			

Do you have/Have you had any of the following? (circle)

01 coronary artery disease	02 heart attack	03 high blood pressure	04 heart valve problems
05 stroke	06 cerebrovascular accident	07 heart arrythmia	08 pacemakers
09 myocarditis	10 deep vein thrombosis	11 other cardiovascular illness	12 chronic brochitis
13 pulmonary emphysema	14 asthma	15 other lung illness	16 allergies
17 thyroid dysfunction	18 diabetes	19 anemia	20 high cholesterol
21 rheumatism	22 joint problems	23 chronic back pain	24 stomach ulcer
25 hernia	26 inflamation of esophgus	27 tumor or cancer	28 recent surgery
29 mental health problems	30 recent trauma	31 low blood K or Mg	32 high eye pressure
33 vision/hearing problems	34 recent injury		

other illnesses or symptoms, what?:_____

Medications: Do you use some sort of medication on a regular basis or often? 1 No 2 Yes
What?:_____

Do you smoke? 1 No 2 Yes 3 I have quit (when?)

Do you use chewing tobacco? 1 No 2 Yes 3 I have quit (when?)

When have you last consumed alcohol?_____ How much?_____

Pregnancy/births: 1 I am pregnant, week_____ 2 I gave birth _____months / years ago

Fever, flu-like symptoms or otherwise unusual feeling of fatigue during the last two weeks:
1 No 2 Yes

Has there been any premature deaths related to cardiac problems? 1 No 2 Yes
Close relative? _____ At what age?_____
Do you have any diagnosed heart problems? _____

How often have you participated in physical activity within the last month?_____
Examples?_____

Evaluate your physical condition: 1 = weak, 2 = satisfactory. 3 = average, 4 = good, 5 = excellent

How often do you exercise/train during a week? _____

Consent

I have answered the questions truthfully to the best of my knowledge. I am familiar with the test procedures and will participate at my own risk. I understand the purpose of this testing and possible risks of participating in this testing as well as my rights and insurance responsibilities. I agree to participate in the testing and to follow the instructions given to me. I will not participate in the measurements if I have a cold/flu or fever or if I am recovering from illness or otherwise not feeling well. I understand that I may drop out/quit the testing at any time without and consequences. My data can be used for reporting in a manner that I cannot be identified from the presentation. Research data is stored on the researchers' own computer or in a database with up-to-date firewalls and antivirus software. After analysis of the research results, the data are transferred to a USB flash drive or external hard drive for archiving. If the data does not need to be archived, it can be destroyed.

In Kajaani _____2021 _____Phone:_____

Signature of client

In Kajaani _____2021 _____Phone:_____

Signature of tester

Instructions for tests

You will be participating in two (2) Cooper 12-minute running tests. Cooper's Test is a maximal aerobic fitness test. One of them will be performed on a track whereas the other one will be performed on a non-motorized treadmill at CSE Entertainment's premises. These tests (Cooper test) will give you an indication of your aerobic fitness

In addition to the two (2) tests, you will be performing two (2) brief familiarization trials on the non-motorized treadmill to familiarize on it. Between the tests, there will be at least 48 hours to recover.

Testing session will take roughly about 30-40 minutes including testing, warm-up and cooldown. The actual maximal testing takes 12 minutes.

Before the test starts, there is a warm-up. We will then take the test. It is important that you are healthy and uninjured, as mentioned earlier, so that you can complete the test with high intensity and motivation - i.e., that you can run for 12 minutes continuously to the best of your ability. The idea is not to compete against others, but it is important that you can complete the 12-minute run, at your own pace, with a good motivation.

During the test, the test proctor will announce intervals at 3, 6 and 9 minutes. This will also be announced at the last minute. When the test proctor yells "Stop" stop where you are and wait for the test proctor to come to you and mark your score. After the test, a cool-down will be performed.

Questions about the procedures used in the physical fitness tests are encouraged. If you have any questions or need additional information, please ask the test administrator to explain further.

How to Prepare*

- Avoid eating any big meals, drinking alcohol, caffeine, and smoking at least 3 hours before the test.
- Rest well. Avoid any strenuous activity a day before the test.
- Enough sleep the night before (6-8 hours)
- Use good, comfortable clothes that are suitable for sports.
- Remember to drink enough prior to the test to take care of the fluid balance. Also after the test.

*these guidelines are modified based upon guidelines by ACSM 2000

Appendix 2 – Cooper Test Familiarity form

In order to conduct the test and divide into groups, we'd like to know some background information. Please fill out the form carefully.

Cooper Test Familiarity

Name: Date of Birth:

1. Are you familiar with the Cooper Test? 1. Yes 2. No

2. If yes, have you performed a Cooper Test before? 1. Yes 2. No

3. If yes, have you performed a Cooper Test within the last 12 months? 1. Yes 2. No

If (3) Yes, what was your results in meters? _____

If (3) No, when was your previous test? _____

What was its result in meters? _____

What is your best result and when? _____

(if not above mentioned)*

If (2) Yes, where have performed the Cooper Test? (Can be multiple) 1 Track field 2 Treadmill
3 Non-motorized treadmill 4 Something else, what?
_____

Have you run or walked on a non-motorized treadmill before? 1. Yes 2. No

Figure 1 DRAXFIT+ curved-non motorized treadmill (this is the one we'll be using)

Appendix 3 – Warm-up and cooldown procedures

Warm-up

- 5-10 minutes of jogging/running
- Dynamic stretches:
    - Walking on toes, walking on heels
    - Frankenstein stretch
    - Leg swings (full range of motion)
        - (Front, lateral, opening, hurdles)
    - Hamstring lunges with reach and kick
    - Squat with rotational twist
    - Kick backs
    - High knees
- Individual stretches or last-minute jog before the test

Cooldown

- 5-10 minutes of slight jogging/walking
- Static stretches:
    - Quad stretch
    - Standing abductor stretch
    - Standing calf stretch
    - Butterfly stretch
    - Hamstring stretch
    - Hip flexors stretch
    - Half knee hamstring stretch

+Instruct the participants to perform further static stretches 2 hours after the test.

Appendix 4 – Paired Samples t-test

Paired Sample Statistics

		N	Mean	Std. Deviation	S.E. Mean
Pair 1	Vo2max on track	8	49.85	7.84	2.77
	Vo2max on cNMT	8	40.24	9.18	3.24

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Vo2max on track & Vo2max on cNMT	8	.952	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	S.E. Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Vo2max on track - Vo2max on cNMT	9.61	2.95	1.04	7.14	12.08	9.21	7	.000

Paired Sample Statistics

		N	Mean	Std. Deviation	S.E. Mean
Pair 1	Overground	8	2741.38	351.49	124.27
	cNMT	8	2305.50	410.62	145.18

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Overground & cNMT	8	.951	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	S.E. Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Overground - cNMT	435.88	132.99	47.02	324.70	547.05	9.27	7	.000