

**A CRITICAL APPRAISAL OF THE MULTI-STAGE SHUTTLE RUN
AGAINST THE ASTRAND VO₂MAX TREADMILL PROTOCOL FOR
DETERMINATION OF AEROBIC FITNESS AMONGST DURHAM
AND DARLINGTON FIRE FIGHTER RECRUITS**

**Sub Officer Gary Bankhead
MAY 2001**

Acknowledgment

I would like to thank the Lecturers and Staff of New College, Durham for their assistance, advice and use of facilities; in particular Claire Orange for her guidance on this project. Special thanks must also go to the 'lads' from Recruit's Course 156, not only for their very hard work but also for their enthusiasm and attitude throughout the 4 weeks testing period. Additional thanks also to the Brigade Training School Instructors for their help with the assessments.

ABSTRACT

The aim of this study was quite simply to try and identify how accurate the Multi-stage shuttle test actually is compared to the most accurate method of evaluating aerobic capacity which is; the continuous, incremental maximal oxygen uptake (VO_2max) assessment, using a treadmill. Using all male recruit fire fighters ($n = 14$) testing took place over a 4-week period at New College, Durham. Performed, in random order, the Multistage shuttle test and an Astrand maximal treadmill protocol of increasing elevation, which elicited VO_2max .

Compared with direct measurements of VO_2max , the multistage shuttle test significantly under predicted the VO_2max of the whole group of recruits by a remarkable **11.87%**. The correlation (r) between VO_2max determined by the multistage shuttle test and by the treadmill protocol was 0.51 ($P < 1.38$). A stronger more positive correlation was identified with HR_{max} results which was 0.922 ($P < 0.06$) with no significant differences between the two test identified this indicates that total work rate were similar in both tests.

The study highlighted the need for further analysis of the accuracy of the Multistage shuttle test particularly as it is currently being used as an entry-level test into County Durham and Darlington Fire and Rescue Brigade.

CONTENTS

SECTION	PAGE
0. Glossary	5
1. Background	6
2. Introduction	8
3. Review of literature	11
4. Study Design and Method	14
5. Procedures	16
6. Results	18
7. Discussion	23
8. Recommendations	25
9. Bibliography	26
11. Appendices	27
• A Preliminary Questionnaire	27
• B Information sheet	28
• C Subject Data and Results	29
• D Subject Data and Results – Physical characteristics of subjects	30
• E Subject Data and Results – maximal heart rates	31
• F Table of predicting maximum oxygen uptake values - MSFT	32
• G Shuttle level achieved scorecard	33
• H Pre-Exercise Questionnaire	34

GLOSSARY

ml.kg ⁻¹ min ⁻¹	=	V02 max units-millilitres of oxygen per kilogram of body weight per minute.
NCD	=	New College, Durham.
n	=	Sample or group size.
ATT	=	Astrand treadmill test.
TTE	=	Time to exhaustion.
VO ₂ max	=	Maximum Oxygen uptake (pronounced 'vee oh two max')
CDDFRB	=	County Durham and Darlington Fire and Rescue Brigade
MSFT	=	Multistage Fitness Test (Shuttle Run)
UMTT	=	University of Montreal Track test
GXT	=	Graded Exercise Testing
IWG	=	The Implementation Working Group on Training Strategy
CST	=	Chester Step Test
CACFOA	=	Chief and Assistant Chief Fire Officers Association

1. Background

- 1.1 The Fire Service (Appointments and Promotions) regulations 1978 as amended require that each applicant seeking appointment to both the whole time and retained Fire Service:

Shall have satisfied a duly qualified medical practitioner selected by the fire authority that;

1. His / her lungs are assessed by the measurement of the forced expiratory volume in one second (FEV^{1.0}) + the forced vital capacity (FVC), and;
2. His / her general level of fitness, including his / her aerobic capacity (VO₂max), as measured by a step test.

Are as such as to enable him / her to undertake fire-fighting duties.

- 1.2 Because of formal and informal challenges to Brigades regarding the current statutory fitness tests, and the problem of recruit wastage during the initial training period, there has been growing concerns within the Chief and Assistant Chief Fire Officers Association (CACFOA) about the appropriateness of the current selection procedure. These matters led CACFOA (personnel and Training Committee) to form a working Group to investigate and evaluate current guidance and the practical application of the recruit selection procedures.

At its inaugural meeting in September 1997, it was decided that one of its main aims would be to:

Examine current guidance on recruit selection testing and in particular that contained Within Fire Service Circular 7/1997, the latest Home Office guidance on the Subject.

- 1.3 In order to obtain a clearer picture of the way in which Brigades were applying selection tests it was decided to survey all Brigades by way of a confidential questionnaire;

Statutory tests. Brigades were asked to specify which tests they were using to assess the physical fitness of applicants. These responses concerned those statutory tests prescribed by the Fire Service (Appointments and Promotions) regulations 1978 as amended.

1. **Step test** – Out of 48 responses 43 Brigades utilised the test but only 43 of these applied it to the recommended standard of the original promulgated test or the variant of the Chester Step Test.
2. **Non-Statutory tests** – Brigades were also asked to comment on non-statutory fitness tests, which were utilised. This revealed that 31 Brigades used the ***Multi-stage Fitness test***. In quantitative terms the responses indicated that in some instances this test was used as a replacement for the step test while some Brigades used it in addition to the step test. The comments received suggested dissatisfaction with the reliability, accuracy and methodology of the application of the step test as opposed to the clearer indication given by the Multi-stage Fitness test.

1.4 Whilst the origin of the working group began under the auspice of the CACFOA personnel and training committee it was quickly realised that its work would have wider ranging implications and would require consultation with a multitude of interested organisations. Therefore in July 1998, following a presentation of its initial findings to the working group of the Joint Strategic Committee on Personnel. It was decided to ask the Implementation Working Group on training strategy (IWG) to formally accept this report and approve its supported progression to the validation period with adoption across the British Fire Service.

1.5 **Physical Fitness Tests**

The IWG identified that the existing statutory test was the Chester Step Test (CST) and research has shown a number of problems with the test with examples of challenges to some Brigades citing examples of extremely fit individuals being seemingly incapable of reaching the required standard. Many Brigades had already introduced the Multi-Stage Fitness Test (MSFT) as a secondary means of assessing candidates for aerobic fitness, and research by the working group has concluded that there is justification for changing the testing modality from the CST to the MSFT particularly on the grounds of accuracy.

This was substantiated in that every predictive test of aerobic capacity has a Standard Estimated Error (SEE). The SEE of the MSFT (approx. 0.9%) being far less than the CST (approx. 9-10%). This is due to the maximal nature of the MSFT versus the sub-maximal nature of the CST stepping Protocol, and the fact that the more maximal the test, the greater the accuracy in terms of prediction.

It was with no surprise therefore that recommendation 4 of the IWG on point of entry selection tests (March 2000), states “*Having regard to the above information the working group recommend that the MSFT is introduced, with a required level set at level 9: shuttle 6 (equivalent to 45 ml/kg/min) as a replacement for the CST for all applicants to the fire service.*”

2. Introduction

- 2.1 There appears to be no widely accepted definition of fitness, nor any agreement as to what specific components should comprise a fitness evaluation. Over 50 years ago Steinhaus, evidently viewing fitness from the perspective of the physiologist, defined it as distance from death, a description somewhat like that of many in the medical profession, who tend to regard physical fitness as absence of disease. Willgoose 1973, perhaps slightly more accurately, has defined it as “a capacity for sustained physical activity”. A more appropriate and universal definition of physical fitness from a health perspective, is probably the one found in Mosby’s Medical and Nursing Dictionary, which defines physical fitness as “the ability to carry out daily tasks with alertness and vigour, without undue fatigue, and with enough reserve to meet emergencies or to enjoy leisure time pursuits”

The systems that are responsible for the delivery of oxygen to the respiring tissues are the cardiovascular system and the respiratory system, which utilise the blood to transport the respiratory gases for aerobic performance. Both of these systems contribute largely towards the final VO_2 max score, and undergo several physiological adaptations following training of an aerobic nature.

2.2 Aerobic Fitness Assessments.

Fitness per se can be defined as the capability of an individual to effectively perform his or her function or job. Fitness must consist of a number of elements such as mental (cognitive) ability, strength, endurance, flexibility, speed and health. Although the priority of these elements will vary, it is quite clear that the required fitness for sedentary jobs such as secretarial work will be very different to that of highly active jobs such as those embodied by the emergency services. Endurance or aerobic fitness is commonly the most significant element of an individual’s overall fitness with regards to highly active occupations. Consequently, the estimation or measurement of aerobic capability is a highly desired requisite within many industries in which it is recognised that the employee’s effectiveness is directly dependant upon his/her fitness.

Aerobic fitness has been measured by means of direct measurement of maximal oxygen uptake for a number of decades (Taylor et al., 1955). It was recognised early on that endurance capability was strongly related to the capability to take in and use oxygen (Astrand and Ryhming, 1954; Taylor et al., 1955). The better an individual was at acquiring and using oxygen the more impressive athletic performance tended to be. In due course, therefore, methods of measuring oxygen uptake during maximal exercise (maximal exercise being the point at which we all use oxygen to our personal best) were developed culminating in the evolution of what is probably the most accurate method of evaluating aerobic capacity at present, that is the continuous, incremental maximal oxygen uptake (VO_2 max) assessment, using a treadmill.

2.3 **Estimated or Predictive Maximal Oxygen Uptake Assessment**

The performance of a direct VO₂max assessment requires a certain degree of expertise from the administrator of the test, expensive equipment, a suitable location and can only be undertaken on one subject at a time. Furthermore as it is a maximal test there is a small risk to the subject, a risk that increases with the reduction of fitness of the subject i.e. those with questionable fitness are most at risk of developing a medical complication brought about by undertaking such a test. Considerations such as these have therefore lead to the development of methods, which do not constrain the administrator as much and/or which pose less risk to the subject. These methods of assessment can be categorised according to their design and purpose. All are predictive tests in that they estimate or 'predict' the subject's maximum oxygen uptake. Of these tests many are submaximal, in that they do not require the subject to exercise to maximum, and some are 'maximal' and thus do not necessitate the subject to exercise to maximum.

Of these two forms of test the maximal assessment naturally poses the greatest risk to the subject, however, the hazards involved are minimal in healthy subjects, but are principally concerned with stressing the cardiopulmonary systems to a high intensity. Fatigue during maximal exercise involving large muscle masses primarily results from an inadequate cardiac output and tissue perfusion (Skinner et al., 1990). The risks associated with such testing are minimal for healthy subjects under the age of 40 (Shepard, 1984). The likelihood of severe complications such as myocardial infarction or death have been shown to be 0.000008% in a survey of 500,000 properly controlled maximal exercise tests (Fardy et al., 1988).

- 2.4 The principal advantage of the maximum assessment is that because the subject performs maximally an indication of true capability is obtained. Furthermore, some methods of maximum assessment (in this case, the shuttle run) enable a large number of subjects to be assessed simultaneously, thus providing a high degree of cost effectiveness and time management, factors that often prove important under practical circumstances. In the case of the treadmill testing, however, as there are only two treadmill machines available at New College, Durham (NCD), assessments will be done in pairs.

- 2.5 Both the treadmill and Multi-stage shuttle tests can be used as maximal tests, The shuttle maximal test enables between 1 and 15 subjects to run at incrementally faster speeds until they are no longer able to maintain the required pace and are forced to stop. The speed (level) at which an individual stops is related to his/her maximum oxygen uptake. A table of shuttle run levels and associated VO₂max values provides the measure of fitness of the subject (appendix G).
- 2.6 The treadmill is also a maximal test. The treadmill console has a built in *Astrand Treadmill test* (ATT), which however, can only test one subject at a time. Again running at incrementally steeper gradients, until they are no longer able to maintain the pace and subsequently stop by hitting the treadmill cut out button. The treadmill console displays the VO₂max of the subject based on duration, heart rate and gradient achieved up to the point of exhaustion. Again this will provide a measure of the fitness of the subject, based on HR values.
- 2.7 The aim of this study is quite simply to try and identify how accurate the Multi-stage shuttle test actually is compared to the most accurate method of evaluating aerobic capacity at present that is the continuous, incremental maximal oxygen uptake (VO₂max) assessment, using a treadmill. As well as identify any possible consequences for the recruit selection procedure currently employed by CDDFRB.

3. Review of literature

- 3.1 To aid understanding of the results from this study it is helpful to consider firstly the derivation of each of the predictive methods used on the 14 Recruits and secondly the findings of other researchers investigating related areas in particular those using the same fitness assessment methods as those employed in this study.
- 3.2 For example, in 1997 ST Claire, Broomhead, Lambert and Hawley from the University of Cape Town Medical School, Sport Science Institute of South Africa, conducted a study to assess the validity of the MSFT in predicting maximal oxygen uptake (VO_2max) in athletes trained in sports with different physical demands. 10 male long distance runners and 10 male squash players performed in random order, over a 14-day period the MSFT and a maximal treadmill protocol of increasing elevation, which elicited VO_2max . Compared with direct measurements of VO_2max , the shuttle run test significantly under predicted the VO_2max of the athletes as a whole ($n = 20$; $P < 0.01$). The correlation (r) between VO_2max determined by the shuttle run test and by the treadmill protocol was 0.67 ($P < 0.01$) for the athletes as a whole.

As the subjects represented athletes trained specifically in sports with different physical demands, this may explain why the correlations between the shuttle test and direct measurement of VO_2max are less robust than those reported in previous studies. These results also indicated that there are sport-specific differences when predicting VO_2max from the MSFT.

3.3 The Multistage Fitness Test

The precursor to the MSFT was the University of Montreal Track Test (UMTT) developed by Leger and Boucher (1980) as a means to estimate maximum oxygen uptake simultaneously for groups of individuals. The Test was conceived to predict aerobic power for “school children, healthy adults attending fitness classes and athletes performing sports with frequent stops and starts...” (Leger and Lambert, 1982). Although a very good relationship was noted between direct treadmill measurements of maximum oxygen uptake and the UMTT estimate ($r = 0.96$) and means were virtually identical (61.4 and $61.5 \text{ ml.kg}^{-1}\text{min}^{-1}$ respectively) the validation population was quite narrow. The primary validation pilot-studies involved 3 groups of college students with group mean ages of 20, 22 and 35 and of which only one was female. In a further study of the UMTT the reported reliability was good (means of 54.1 and $54.2 \text{ ml.kg}^{-1}\text{min}^{-1}$ for two sets of measurements from 60 subjects within 10 days of each other), however, the male/female ratio was not balanced (11 female, 49 male) and the mean ages of the sub-groups examined were 22.7 and 33.8 years (Leger and Boucher, 1980).

- 3.4 Leger and Lambert (1982) modified the UMTT, calling it the Multistage fitness Test, and further assessed its potential as a VO_2max predictor using (1 subjects, recording a significant correlation of $r = 0.84$ and a standard error of estimate of $5.4 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$ (Leger and Lambert, 1982). The stages were 2 minutes long, however, and when combined with small increments of speed lead to a long test for many subjects. The length of this version of the shuttle run may have potentially lead to some stopping earlier than they might otherwise have done if the overall assessment had been shorter.
- 3.5 The shuttle run as it is used today results from the modifications made by Ramsbottom, Brewer and Williams in 1998 who took a shortened version (1min Stages) and validated it against direct treadmill VO_2max measurements using 74 subjects (36 men, 38 women) between the ages of 19 and 36, predominantly from the student population. The correlations between maximal speed attained in the shuttle run, the direct VO_2max measures and the time taken to run 5km were significant, $r = 0.92$ and -0.94 respectively. The derivation of tables by the authors to show values of VO_2max in combination with corresponded Shuttle run speeds provided a simple guide for field assessment of aerobic fitness.

A number of studies comparing the results of the shuttle run to those of direct VO_2max measures show although the Shuttle run estimates are significantly related to direct measures (correlation coefficients of between 0.61 to 0.93), they are often slightly lower (Grant et al., 1995; Naughton et al., 1995) have been reported. Authors have also qualified shuttle run results as 'under predicting the VO_2max values of runners whilst accurately predicting those of squash players' (St Clair Gibson et al., 1998) or at 75% of the Shuttle run assessments undertaken were lower than direct VO_2max measurements (Sproul et al., 1993).

3.6 **Treadmill Testing**

Of the Treadmill graded exercise testing (GXT) protocols, the Balke walking test has been widely used. The protocol calls for a constant speed of 3.4 mph throughout the test. The grade is 0% for the first 2 min. The grade is then raised to 2.5% and further increased 1% per minute thereafter until the reaches their limit of tolerance. For low fit subjects, protocol elicits valid VO_2max values. However, for, the test duration is very long. Also, these subjects are required to walk at grades above 20% during the last few minutes of the test and complain of severe local discomfort in the lower back and calf muscles, which may limit their ability to achieve maximal work rates. McArdle, Katch, and Pechar compared VO_2max values measured using the Balke walking treadmill test to those measured using running treadmill tests in reasonably fit male subjects. They found that the Balke walking test yielded VO_2max values that were approximately 5% lower than those obtained on the running tests. The Balke test could be modified to use a faster walking speed, which would likely result in increased VO_2max values for fitter subjects.

- 3.6 A typical running treadmill protocol is that of Maksud and Coutts. It begins with the subject running at 6 mph on the level (0% grade) for 2 min. Thereafter, the grade is increased by 2.5% each 2 min; the speed is constant throughout the test.
- 3.7 Each of the protocols described are continuous. Some GXT protocols are discontinuous. An example is the treadmill test of Mitchell, Sproule and Chapman. The test begins with subjects walking at 3 mph for 10 min at a 10% grade. Then the subject rests for 10 min. next the subject runs for 2.5 min at 6 mph up a 2.5% grade. After another 10-min period, the subject runs for 2.5 min at 6 mph up a 5% grade. This procedure (rest 2.5% grade increase) is continued until the subject reaches their limit of tolerance.

The VO_2max obtained using a continuous protocol is the same as that obtained using a discontinuous protocol. McArdle, Katch, and Pechar compared the continuous Maksud and Coutts treadmill protocol to the discontinuous Mitchell, Sproule and Chapman treadmill protocol in 15 college-aged males. The mean VO_2max values for the continuous and discontinuous protocols were very similar, 4.109 and 4.145 $\text{L}\cdot\text{min}^{-1}$, respectively. What differed markedly was the test duration: 12.3 min for the continuous test and 67.3 min for the discontinuous test. Both subjects and the investigators preferred the continuous test for its shorter duration.

3.8 Maximum oxygen uptake (VO_2max)

- 3.9 Probably one of the most often-evaluated fitness tests parameters is **maximum oxygen uptake (VO_2max)**, which has long been considered the ‘gold standard’ for the evaluation of cardio-respiratory fitness (Evetts et al. 2000). Direct assessment of this variable is preferred. Many tests have, however, been devised to estimate maximum oxygen uptake, or general cardiovascular condition, using indirect measures purported to correlate well with direct assessment. Such tests have included time taken to run a specific distance, distance run in a specific time, or heart rate response either during, or recovering from, standardised work.
- 3.10 There are four parameters of aerobic fitness: aerobic power, or maximal oxygen uptake (VO_2max); work efficiency; time constants for VO_2 Kinetics; and the lactate threshold. Maximal oxygen uptake is an important parameter of human fitness because it represents the upper limit of aerobic exercise tolerance. Endurance activities are performed at some fraction of VO_2max ; if VO_2max is abnormally low, then the level of endurance performance is necessarily constrained (Adams, 1994).

- 3.11 The highest oxygen value achieved during the VO_2 max tests represents the maximal oxygen consumption. This peak is most likely to occur at the final minute of the progressive exercise test. Because the peak oxygen consumption value may vary with the mode of exercise and limiting symptoms, the term peak oxygen consumption denotes the highest oxygen consumption for a specific type of exercise, and the term maximal oxygen consumption denotes the single highest possible oxygen consumption elicited among the different modes of exercise. For example, peak oxygen consumption values are similar for treadmill and step tests for most people, but they are about 5%-10% higher on a treadmill than on a bicycle ergometer

The physical rationale for sub maximal cycle ergometer, as with the physical rationale for sub maximal treadmill protocols are both very similar. Indeed the highest oxygen value achieved during the VO_2 max tests represents the maximal oxygen consumption, (Maximal VO_2 values are approximately 10% higher when measured during treadmill running, compared to cycle ergometry).

4. Study Design and Method

4.1 Subjects

14 subjects participated in the two assessments. Of these all were male, aged between 19 and 31. Of the 14 subjects, 7 were aged 26 years or less and 7 were aged over 26 years. All of the subjects were Recruit Fire Fighters serving with CDDFRB each with similar levels of fitness of an intermediate standard. They had all previously completed a physical fitness assessment in order to pass through the various stages of the selection process in December 2000; this included a MSFT assessment requiring a pass level of 9.6. All of the subjects were Caucasian.

Table 1 Physical characteristics of subjects

Values are Mean \pm SD			
	All (Males)	Males 26 and Under	Males Over 26
n =	14	7	7
Age (yr)	25.5 ± 3.91	22.5 ± 8.83	28.4 ± 1.39
Weight (kg)	77.8 ± 8.33	76.8 ± 9.5	78.7 ± 7.61
Height (cm)	179.6 ± 5.42	182 ± 5.29	177.2 ± 4.78

4.2 The design of the study required the subjects complete two fitness assessments (Multistage shuttle run test and Astrand treadmill test) over a four-week period. The Shuttle run was to be assessed on the third week as this would not prejudice any individuals who may have had the benefit of extra training; each assessment was to be conducted on the same day of the week and at the same time. The subjects were requested to continue with their usual exercise programme (if any), and not to embark on any additional exercises if they did not normally exercise on a regular basis

The Astrand treadmill test and Shuttle run tests were both conducted by myself assisted by CDDFRB Training Instructors

4.3 Randomisation

Due to difficulties with subject scheduling, all testing was only possible on a Wednesday afternoon. This would require 3 weeks in total with the MSFT testing taking place in the middle week. However 2 subjects picked up slight injuries prior to the first weeks test and 1 other subject's results for the Astrand treadmill test was lost due to an electrical fault immediately on completion of the test. Therefore a fourth week was required to complete the tests on all 14 subjects. The first assessment was an Astrand test for 43% of subjects; the second assessment was a MSFT for 85% of subjects. The third assessment was an Astrand Test for 50% of subjects. The fourth assessment was a MSFT for 15% and an Astrand test for 7% of subjects.

4.4 **Statistical analysis**

The t-test (one-tailed version, assuming the difference between the two means lies in one direction only) was used to detect significant differences in $VO_2\text{max}$ and HR_{max} between the shuttle run tests and the treadmill protocol. Pearson's product moment correlation coefficient was used to determine relationships between $VO_2\text{ max}$ and HR_{max} in the shuttle runs and treadmill protocol. Repeatability and agreement between variables derived from the multistage shuttle run test was determined using the procedure described by Bland and Altman (1986).

The maximum R-squared (R^2) method is used in figures 1, 2, 3, to estimate the degree of meaningfulness in terms of the amount of common association (or amount of variance held in common) between the dependant and independent variables. A multiple regression method in which the so-called best of all possible one-variable models is selected, as is the best two-variable model, the best three-variable model, and so on until a predetermined criterion the ends the calculation is reached.

5. Procedures

5.1 The Astrand, Constant Speed, Progressive Grade Treadmill Test Protocol

The Astrand Treadmill Test (ATT) was conducted at the NCD Sports centre using a *Technogym Treadmill run race HC1200* incorporating a built in ATT. Two subjects at a time were tested, on the two identical treadmills. All of the subjects were allowed to stretch and loosen up the leg muscles for 5 minutes prior to the test. The subject's were fitted with a heart rate monitor that was worn around the chest under the clothing and next to the skin thus ensuring adequate contact; the heart rate monitor was compatible with the treadmill and gave a real time indication of the subject's heart rate. The test required that before the start of the test specific data had to be fed into the treadmill console, namely the weight and age of the subject were entered into the treadmill console, The speed at which the test was to be carried out at, i.e. 13 km/hour and the type of test to be carried out i.e. the Astrand

- 5.2 The test started with the subjects mounting the treadmill and having been shown the emergency stop button and methods in which to dismount safely, commenced the test with a 5-minute warm up which was carried out at walking pace. After 5 minutes the subject's heart rate was recorded and the speed of the treadmill was increased to the subjects chosen running speed (13 km/hour) and the gradient was increased from 0 to 2.5%. After a further 3 minutes the process was repeated with the heart rate again recorded and the gradient increased again by 2.5% to 5%. The time period between increases in gradient was then reduced to every 2 minutes; subjects were instructed to continue until exhaustion. All subjects were given equal verbal encouragement both from the assessor and from their fellow recruits who were watching the tests this was influenced by the Brigades policy on building teamwork. The encouragement from the assessor related directly to breathing techniques and references to how well they were doing.

This process was further repeated until the subject reached exhaustion at this point the treadmill was stopped, the final gradient of the treadmill, the finish time and the heart rate of the subject were recorded. The treadmill console displayed the maximum heart rate, VO_{2max} , L/min VO_2 and the Mets scores achieved of the subject based on the duration, heart rate and gradient reached up to the point of exhaustion.

5.3 The Multistage Fitness Test Protocol (MSFT)

The multistage fitness test (MSFT) entailed the subject running on a flat surface between two markers 20m apart. The MSFT was carried out inside the gymnasium at NCD. The pace was set by an audio-tape which provided auditory indications of pace requiring the subject to pivot and turn at one end of the 20m track when each tone was heard. The starting pace was $8.5\text{km}\cdot\text{h}^{-1}$ and increased by $0.5\text{km}\cdot\text{h}^{-1}$ per level, each level lasted one minute and consisted of 9 to 16 'tones' or 'shuttles'. Level one was the slowest pace allowing 9 seconds in which to run each 20m shuttle. The fastest pace occurred at level 21.

- 5.4 The subjects were divided at random into two groups of 7, this allowed for sufficient space between each person during the test, as the width of the gymnasium was restricted. The subjects ran for as long as possible until they could no longer maintain the pace set by the tape. This was quantified by their inability to reach a marker in time with a tone twice in succession. The other group of 7 subjects helped record the heart rate of the subjects at various stages of the test (Prior to, level 9.6 and at the end) as well as the level and number of shuttles achieved (appendix G). Again during the final stages of the test all subjects were given verbal encouragement from their fellow recruits, this in many cases seemed to 'push' the subjects that little bit harder during the final few seconds of the test. The level and number achieved was then found on a table provided (appendix F) with the audiotape from which the relevant VO₂max values could be read.

6. Results

6.1 Comparison of means

The VO₂max means for all groups for each method of assessment are presented in table 2. The mean VO₂max reached for the MSFT was 54.57 ±4.62 (mean ±s) for the 26 and under subjects and 50.57 ±1.71 for the over 26 subjects as a whole, all subjects reached 52.57 ±3.94. The mean VO₂max reached for the ATT was 61.85 ±1.86 (mean ±s) for the 26 and under subjects and 58 ±2.51 for the over 26 subjects, as a whole, all subjects reached 59.92 ±2.92. The mean shuttle level reached during the MSFT was 12.28 ±1.39 (mean ±s) for the 26 and under subjects and 10.97±1.24 for the over 26 subjects, as a whole all subjects reached 11.62 ±0.63.

Table 2 Maximum oxygen uptake (ml kg⁻¹ min⁻¹) as predicted from the multistage shuttle run and as measured during the treadmill protocol (mean ±s), and MSFT shuttle level reached.

Values are Mean + SD			
	26 and under	Over 26	All
Shuttle Run (MSFT)	54.57 ±4.62	50.57 ±1.71	52.57 ±3.94
Treadmill (ATT)	61.85 ±1.86	58 ±2.51	59.92 ±2.92
MSFT level reached	12.28 ±1.39	10.97±1.24	11.62 ±0.63.

6.2 Examination of Relationships

The correlation between the predictive and direct assessments of VO₂ max according to groups are presented in table 3

Table 3

Pearson's product moment correlation coefficient of VO₂ max Results	
All subjects	0.51
26 and under	0.31
Over 26	0.23

The relationship between the MSFT and ATT assessments of VO₂ max gives a moderately positive correlation for all subjects $r = 0.51$ ($P < 1.38$), but no relationship for the over 26 age group $r = 0.31$ ($P < 0.002$), and again a no relationship for the 26 and under age group $r = 0.23$ ($P < 0.0001$)

Table 4 Maximal heart rates (beats min⁻¹) attained by the subjects during the MSFT and ATT protocols (mean \pm s).

Values are Mean + SD			
	26 and under	Over 26	All
Shuttle Run (MSFT)	186.71 \pm 10.49	189.71 \pm 6.26	188.21 \pm 8.45
Treadmill (ATT)	185 \pm 9.71	188.57 \pm 8.16	186.78 \pm 8.81

6.3 The HR_{max} attained during the shuttle run tests was slightly higher than that obtained during the Treadmill tests (P<0.06) for all subjects. With no significant differences in HR_{max} for groups 26 and under and over 26.

6.4 The correlation between the MSFT and ATT assessments of HR_{max} according to groups are presented in table 5

Table 5

Pearson's product moment correlation coefficient of HR_{max} Results	
All subjects	0.922
26 and under	0.931
Over 26	0.936

The relationship between the MSFT and ATT assessments of HR_{max} gives a strong positive correlation for all subject and both age groups.

Figure 1 (a) Relationship between VO₂max as determined by the ATT and the MSFT in all subjects showing a moderately positive correlation (r = 0.51). The black line represents the Regression line of best fit (R² method) between MSFT and ATT VO₂max results.

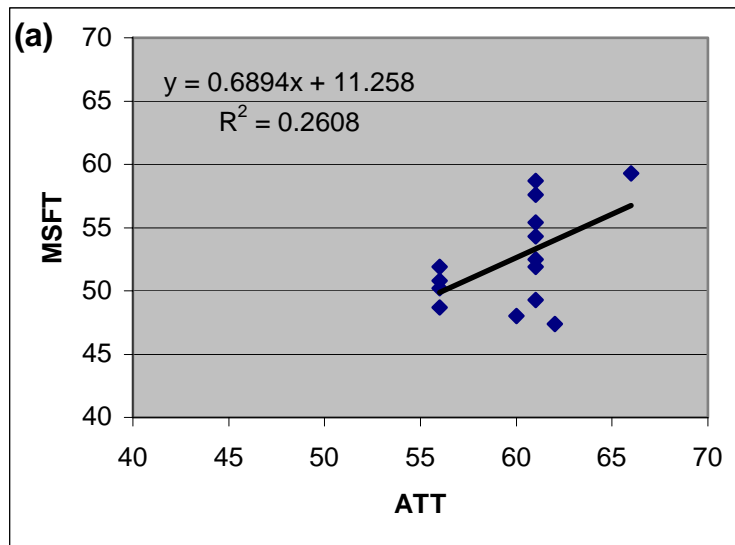


Figure 2 (b) Relationship between VO₂max as determined by the ATT and the MSFT in subjects 26 and under showing a slightly positive correlation (r = 0.31). The black line represents the Regression line of best fit (R² method) between MSFT and ATT VO₂max results.

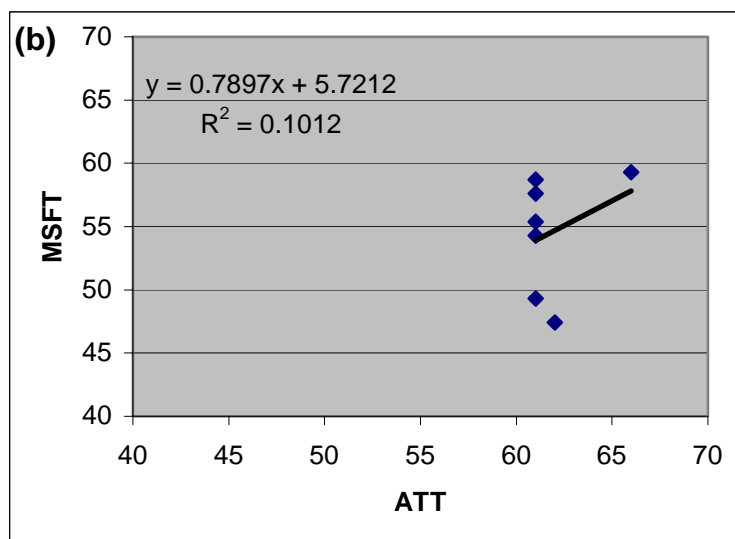
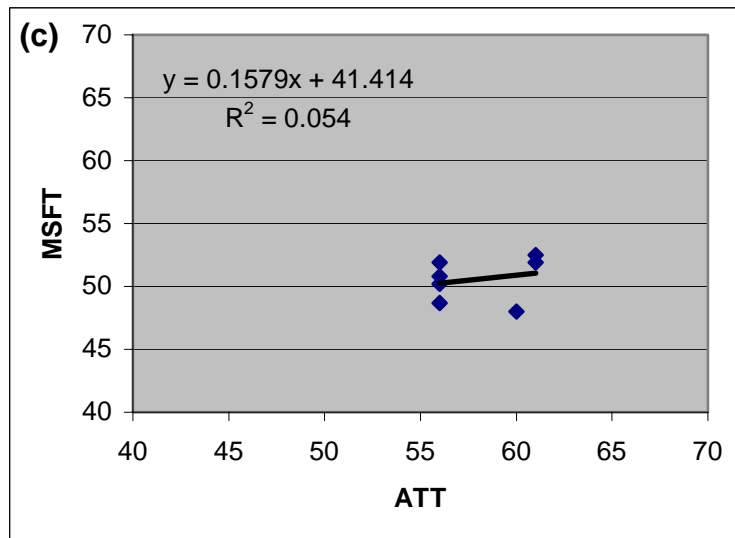


Figure 3 (c) Relationship between VO₂max as determined by the ATT and the MSFT in subjects over 26 showing a slightly positive correlation (r = 0.23). The black line represents the Regression line of best fit (R² method) between MSFT and ATT VO₂max results.



7. Discussion

- 7.1 The aim of this study was to assess the validity of the multistage shuttle test against the Astrand treadmill test in predicting maximal oxygen uptake (VO_{2max}) in CDDFRB recruit fire-fighters ($n = 14$). The recruits were aged between 19 and 31 ($25.5\text{yrs} \pm 3.91$ mean $\pm s$) and were all at similar intermediate levels of fitness. Over a 4-week period the recruits performed, in random order, the multistage shuttle test and a maximal treadmill protocol of increasing elevation, which elicited VO_{2max} .
- 7.2 Although it is recognised that a sample size of 28 subjects would yield more significant differences between derived measures. The limitations of the study were directly related to the size of the recruit course. However similarities with the gender and physical characteristics of the recruits as well as the narrow age difference will have had a positive effect on the level of accuracy of the results.
- 7.3 Conclusions from a study carried out by the King's College London (2000), which assessed methods of predicting maximal oxygen uptake in the London Fire Brigade identified that, the shuttle run if used to measure the fitness of a group would probably under estimate VO_{2max} particularly if a large group of women make up the sample. The study actually goes on to recommend that because of the reliability of the shuttle test a simple correction factor could be applied to 'elevate' the predictions thus improving accuracy.
- 7.4 The study in 1997 by ST Claire, Broomhead, Lambert and Hawley from the University of Cape Town Medical School, Sport Science Institute of South Africa, which conducted a study to assess the validity of the shuttle test in predicting maximal oxygen uptake (VO_{2max}) in athletes trained in sports with different physical demands, also concluded that there was an under prediction in VO_{2max} in the shuttle test. Suggesting that the type of event for which a subject is trained may influence the outcome of the test, and that the prediction of VO_{2max} from the shuttle test may be sport-specific.
- 7.5 From the results one point, which must be raised, are the VO_{2max} scores achieved by subject 6, who scored very well on the ATT (62) but very poorly on the MSFT (47.4), when questioned later on the reasons why said that "it was the lack of Brigade instructors present, and that their was no real reason to push to the limit, no motivation". The testing in his case was carried out with only 4 people present. Consequently had this subject gone on to reach the potentially higher score that was expected then this would again have increased the accuracy and validity of all results.

- 7.6 Another point which should be raised is the question of motivation and encouragement as previously mentioned, CDDFRB recruit courses are geared to teambuilding and supporting each other. Testing in weeks 1-3 was done in the presence of several fellow recruits as well as Brigade Training School Instructors. All individuals benefited from some form of verbal encouragement and Individuals also would have felt the need to prove their physical ability in front of their 'audience' as well as the competition factor of wanting to be the best. These points go to strengthen the reasons for subject 6's poor MSFT result as previously mentioned in 7.5 above. Overall these factors would have a negative effect on the final accuracy of VO₂max levels achieved.
- 7.7 An important finding from the results of the study into the CDDFRB recruits was that the multistage shuttle test significantly under predicted VO₂max for all subjects by 11.87% when compared to the Astrand Treadmill Test it seems therefore that sufficient evidence has been found to agree with both findings from the King's College study and the Sport Science Institute of South Africa study.
- 7.8 However one question that arises from all the testing is which test actually gave the more accurate results although both tests were maximal all results would be irrelevant if the accuracy of one test was in doubt. To this end all subjects were asked on completion of both test which test in their opinion was the more accurate, 100% of subjects agreed that the ATT was by far the more realistic test. The relationship between the MSFT and ATT assessments of HR_{max} ($r = 0.92$) gives a strong positive correlation for all subject and both age groups, this further strengthens the validity of the tests.
- 7.8 One huge benefit CDDFRB enjoys while using the MSFT as an entry level test for potential recruits is the ability to process huge numbers of applicants in a very short time scale. 2000 plus applications are not uncommon when fire-fighter vacancies arise, and some form of getting this number down to around 150 for a regional pool is therefore required. The MSFT does indeed play a large part in reducing these numbers down to an acceptable level requiring very little manpower to over see the process for such a high number of applicants, with many hundreds failing the tests. Perhaps therefore this is a key reason for using the test in the first place, CDDFRB know in advance that statistically sufficient will pass to move on to the next stage of physical testing.
- 7.9 Due however, to perhaps political pressure as well as equality and fairness issues CDDFRB are under pressure from various parties to improve statistically the number of female fire-fighters it currently employs. One big stumbling block in this area is the very high number of females who actually fail the MSFT, consequently leading to few or if any, going on to a successful recruitment. With the information identified in this report along with several other reports each indicating an under prediction in VO₂max scores, should CDDFRB therefore consider testing females with a maximal VO₂max treadmill test?

8. Recommendations

- 8.1 If CDDFRB continues to use the MSFT for future entrance level testing for potential recruits the consideration must be given to the validity and accuracy of the results in particular its tendency to under predict VO₂max. The possible use of a simple correction factor as mentioned in 7.3 above should increase the number of applicants passing this stage of the entry procedure.
- 8.2 Consideration must be given to the use of a maximal VO₂max treadmill test for entry level testing, in particular for female applicants. As this will be a maximal test it is most probably the most accurate method of evaluating aerobic capacity at present, and will result in more females reaching recommendation 4 of the IWG on point of entry selection tests (March 2000), of 45 ml/kg/min. for applicants to the fire service.”
- 8.3 This is the first Physical fitness study of CDDFRB recruits of its kind and as such the data from the research should be used as an ongoing project to periodically re-test the physical levels of the recruit subjects involved, following CDDFRB’s implementation of its new Health and Fitness policy which is to be introduced in 2002. Further on going studies should be carried out on future recruits courses.
- 8.4 CDDFRB should initiate more detailed research into potential pass rates for females where aerobic capacity is evaluated on a maximal VO₂max treadmill test. This testing should be carried out on a much greater scale (n<28) and could be accomplished by testing females who attend open days with a view to finding out more about recruitment.

9. Bibliography

Adams, G.M. (1994). Aerobic Fitness Tests. *Exercise Physiology - Laboratory Manual*. 11-15, 87-146.

Astrand, P.O. and Rhyning, I (1954). A normogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. *Journal of applied Physiology*, 7, 218-221.

Baumgartner, TA and Jackson, AS. Evaluating Aerobic Fitness (VO₂ Max) for Adults. *Measurement for Evaluation*. 10, 271-302.

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

Evetts, N. Evetts, M. Ernsting, J. (2000). An assessment of the methods of predicting maximal oxygen uptake in the London Fire Brigade. *King's College, London*.

Grant, S. Corbett, K. Amjad, A. Wilson, J and Aitchison, T. (1995). A comparison of methods of predicting maximum oxygen uptake. *British journal of Sports Medicine*, 29, 77-152.

Home Office. (2000). Medical standards and guidance. *Fit for duty?* Part two 35-37.

Home Office. (2000). Physical fitness, *Implementation working group - point of entry selection tests*, 19-20 (March)

Home Office. (2000). Statutory Tests, *Implementation working group - point of entry selection tests*, 5-7 (March)

Leger, L. and Boucher, R. (1982). An indirect continuous running multistage field test – the University de Montreal Track Test. *Canadian Journal of applied Sports science*, 5, 77-84

Ramsbottom, R. Brewer, J. and Williams, C. (1988). A progressive shuttle run test to estimate maximal oxygen uptake. *British Journal of Sports Medicine*, 22, 141-144.

Shepard, R. (1984). Tests of maximal oxygen uptake: A critical review. *British Journal of Sports Medicine*, 1, 99-124.

Skinner, J. Baldini, F. and Gardener, A. (1990). Assessment of fitness exercise, *Human Kinetics*, Illinois

St Clair Gibson, A. Broomhead, S. Lambert, M.I. and Hawley, J.A. (1997). Prediction of maximal oxygen uptake from a 20-m shuttle run as measured directly in runners and squash players. *Journal of Sports Sciences*, (1998) 16, 331-335.

Taylor, H. Buskirk, E. and Henschel, A. (1955). Maximal oxygen uptake as an objective measure of cardio respiratory performance. *Journal of Applied Physiology*. 8, 73-77.

Astrand Treadmill and Shuttle Run VO₂max Assessments.
Preliminary Questionnaire

Number

1. Name _____
2. Age _____
3. Height_____Weight_____
4. Do you currently exercise? Yes_____No_____
5. If yes, how often and to what standard _____

6. Do you smoke? _____
7. How much Alcohol do you drink / week (units) _____
8. Have you ever had a VO₂max Assessment before? Y/N _____
9. If yes, how and when _____
10. What level do you think you will obtain in the shuttle runs? _____
11. Have you ever used a treadmill before? Y/N_____
12. Have you had any muscle or joint injuries in the last 6 months_____
- _____

VO₂max Assessment Information Sheet

Date of assessments: 7th, 14th, 21st and 28th February 2001

Time: 1530-1700 Hrs

Location: New College, Framwellgate Moor, Durham.

Dress Requirements:

1. Running shoes.
2. Shorts (not Leggings).
3. Short sleeved t-shirt.
4. Spare t-shirt
5. Towel.

Food and Beverages:

1. Do not consume a large lunch prior to the assessments,
2. Please refrain from drinking caffeine beverages (tea and Coffee) for two hours prior to the tests. Keep well hydrated leading up to the tests (water only).
3. Bring adequate water to rehydrate body after assessments.
4. Avoid any form of alcohol on the evening prior to the tests.
5. Do not smoke two hours prior to the tests.

For your information these tests are maximal tests, that is they push your bodies energy systems to the limit, the risk of feeling nauseous will be greatly increased if you do not follow the advice given on food and beverages !!!!!

Physical activity: Please refrain from Physical activity within 24 hours prior to your test (unless your instructors tell you otherwise).

Pre-exercise questionnaire: Please complete both the pre-exercise questionnaire and the Preliminary Questionnaire and bring them with you on the first day of assessments.

Confidentiality: All results will be treated with the strictest confidence. Each recruit will be allocated a candidate number and any reference made thereafter will be to that number and that number only.

Results: The finalised results and all analysed data acquired will be available for any interested parties from June 2001 onwards

Appendix C

SUBJECT DATA & RESULTS

Subject	Age	Weight	Height	Maximal Oxygen Uptake (ml.kg ⁻¹ min ⁻¹)		Shuttle
				ATT	MSFT	
1	19	11	6.1	61	54.3	12.1
2	26	12.7	5.11	61	58.7	13.6
3	26	11	5.96	61	57.6	13.1
4	28	13	5.10	56	50.8	11.1
5	20	11.6	5.9	61	55.4	12.5
6	19	11.7	6.2	62	47.4	10.1
7	29	13	5.11	61	52.5	11.7
8	26	15.4	6.2	61	49.3	10.8
9	27	12.6	5.10	61	51.9	11.5
10	27	14.6	6.1	56	50.2	10.11
11	28	11	5.8	56	51.9	11.5
12	31	11	5.9	56	48.7	10.6
13	22	12	5.11	66	59.3	13.8
14	29	12	5.7	60	48.0	10.3

SUBJECT DATA & RESULTS - con't

Subject	Multi-Stage Fitness Test			Astrand Treadmill Test			
	HR - Start	HR - 9.6	HR - Fin	L/min VO ₂	Max-HR	Duration	Mets
1	92	190	200	4.209	193	13.16	17.4
2	88	166	180	4.697	175	13.03	17.4
3	76	165	168	4.209	170	13.54	17.4
4	76	180	186	4.648	184	13.32	16.0
5	79	190	194	4.331	193	12.13	17.4
6	88	181	184	4.526	181	14.01	17.4
7	86	177	184	5.185	184	12.35	17.4
8	98	188	190	5.978	194	12.40	17.4
9	89	192	197	4.758	198	12.25	17.4
10	98	191	191	5.208	189	11.08	16.0
11	96	194	197	4.032	194	11.42	16.0
12	93	192	192	3.864	196	11.36	16.0
13	85	180	191	8.052	189	14.48	18.8
14	93	174	181	4.536	175	12.19	16.0

Appendix D

Table 1 Physical characteristics of subjects

Values are Mean \pm SD			
	All (Males)	Males 26 and Under	Males Over 26
n =	14	7	7
Age (yr)	25.5 \pm 3.91	22.5 \pm 8.83	28.4 \pm 1.39
Weight (kg)	77.8 \pm 8.33	76.8 \pm 9.5	78.7 \pm 7.61
Height (cm)	179.6 \pm 5.42	182 \pm 5.29	177.2 \pm 4.78

Table 2 Maximum oxygen uptake ($\text{ml kg}^{-1} \text{min}^{-1}$) as predicted from the multistage shuttle run and as measured during the treadmill protocol (mean \pm s), and MSFT shuttle level reached.

Values are Mean \pm SD			
	26 and under	Over 26	All
Shuttle Run (MSFT)	54.57 \pm 4.62	50.57 \pm 1.71	52.58 \pm 3.94
Treadmill (ATT)	61.86 \pm 1.86	58 \pm 2.51	59.92 \pm 2.92
MSFT level reached	12.28 \pm 1.39	10.97 \pm 1.24	11.62 \pm 0.63.

Table 3

Pearson's product moment correlation coefficient of VO_2 max Results	
All subjects	0.51
26 and under	0.31
Over 26	0.23

The relationship between the MSFT and ATT assessments of VO_2 max gives a moderately positive correlation for all subjects, but only a slight positive correlation for the two age groups.

Appendix E

Table 4 Maximal heart rates (beats min⁻¹) attained by the subjects during the MSFT and ATT protocols (mean \pm s).

Values are Mean + SD			
	26 and under	Over 26	All
Shuttle Run (MSFT)	186.71 \pm 10.49	189.71 \pm 6.26	188.21 \pm 8.45
Treadmill (ATT)	186 \pm 9.71	188.58 \pm 8.16	7.6 \pm 8.81

The HR_{max} attained during the shuttle run tests was slightly higher than that obtained during the Treadmill tests (P<0.01) for all subjects. With no significant differences in HR_{max} for groups 26 and under and over 26.

Table 5

Pearson's product moment correlation coefficient of HR_{max} Results	
All subjects	0.922
26 and under	0.931
Over 26	0.936

The relationship between the MSFT and ATT assessments of HR_{max} gives a strong positive correlation for all subject and both age groups.

Appendix F

Table of predicting Maximum Oxygen uptake Values – MSFT

Level	Shuttle	Predicted VO ₂ max	Level	Shuttle	Predicted VO ₂ max	Level	Shuttle	Predicted VO ₂ max
4	2	26.8	10	2	47.4	16	2	68.0
4	4	27.6	10	4	48.0	16	4	68.5
4	6	28.3	10	6	48.7	16	6	69.0
4	9	29.5	10	8	49.3	16	8	69.5
			10	11	50.2	16	10	69.9
5	2	30.2				16	12	70.5
5	4	31.0	11	2	50.8	16	14	70.9
5	6	31.8	11	4	51.4			
5	9	32.9	11	6	51.9	17	2	71.4
			11	8	52.5	17	4	71.9
6	2	33.6	11	10	53.1	17	6	72.4
6	4	34.3	11	12	53.7	17	8	72.9
6	6	35.0				17	10	73.4
6	8	35.7	12	2	54.3	17	12	73.9
6	10	36.4	12	4	54.8	17	14	74.4
			12	6	55.4			
7	2	37.1	12	8	56.0	18	2	74.8
7	4	37.8	12	10	56.5	18	4	75.3
7	6	38.5	12	12	57.1	18	6	75.8
7	8	39.2				18	8	76.2
7	10	39.9	13	2	57.6	18	10	76.7
			13	4	58.2	18	12	77.2
8	2	40.5	13	6	58.7	18	15	77.9
8	4	41.1	13	8	59.3			
8	6	41.8	13	10	59.8	19	2	78.3
8	8	42.4	13	13	60.6	19	4	78.8
8	11	43.3				19	6	79.2
			14	2	61.1	19	8	79.7
9	2	43.9	14	4	61.7	19	10	80.2
9	4	44.5	14	6	62.2	19	12	80.6
9	6	45.2	14	8	62.7	19	15	81.3
9	8	45.8	14	10	63.2			
9	11	46.8	14	13	64.0	20	2	81.8
						20	4	82.2
			15	2	64.6	20	6	82.6
			15	4	65.1	20	8	83.0
			15	6	65.6	20	10	83.5
			15	8	66.2	20	12	83.9
			15	10	66.7	20	14	84.3
			15	13	67.5	20	15	84.8

Appendix G

Shuttle level achieved scorecard

Level Completed	Shuttle Number Completed
1	1234567
2	12345678
3	12345678
4	123456789
5	123456789
6	12345678910
7	12345678910
8	1234567891011
9	1234567891011
10	1234567891011
11	1234567891011112
12	1234567891011112
13	12345678910111213
14	12345678910111213
15	12345678910111213
16	1234567891011121314