UNIVERSITY OF ALBERTA

THE SENSITIVITY OF TWO PREDICTIVE TESTS TO CHANGES IN MAXIMAL OXYGEN CONSUMPTION FOLLOWING A TRAINING PROGRAM

BY

KEN ROSS

A THESIS

SUBMITTED TO

THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF PHYSICAL EDUCATION AND SPORT STUDIES

EDMONTON, ALBERTA

SPRING 1990

ISBN 0-315-60280-5

DEDICATION

To my wife, Janet, for her love, support, sacrifice and patience.

ABSTRACT

To study the sensitivity of predictive tests to changes in maximal oxygen consumption (VO_2 max), twenty-seven sedentary female subjects, age 20 to 29 years, were tested on each of the Canadian Aerobic Fitness Test (CAFT), the Astrand submaximal cycle test (ASTRAND), and a direct test of VO_2 max on a cycle ergometer (DIRECT) on three occasions. The first two occasions were separated by a one-week control period. The third occasion followed seven weeks of stationary cycle training (n=19) or control (n=8).

Relationships between the predictive tests and the DIRECT test on each occasion were analyzed to determine concurrent validity. Relationships between occasions one and two were analyzed for each test to determine test stability.

The training group showed a small (5%) but significant ($p<.05$) increase in VO_2 max after the training program. Stability coefficients of the tests were as follows: CAFT, .933; ASTRAND, .633; DIRECT, .903. Concurrent validity coefficients of the training group with DIRECT on each occasion, were: CAFT, .715, .735, .636; Astrand, .579, .637, .589. Significant ($p<.05$) differences between means were found between CAFT and DIRECT on all occasions. ASTRAND means were significantly different from DIRECT on occasions one and two but not on occasion three.

The stability of CAFT was considered to be spuriously high due to mathematical constraints on the regression equation and the omission of heart rate as a regression factor. The apparent instability of ASTRAND may have been due to habituation. The limited validity of the CAFT may be attributable to problems with the regression equation. Also discussed was the possibility that adaptation of heart rate and maximal oxygen consumption may follow different time courses.

It was concluded that because the relationships between the predictors and directly measured VO_2 max were weak, the use of CAFT or ASTRAND to predict VO_2 max for individuals is not justified. It was suggested that CAFT could be improved by including a heart rate related variable in the regression equation and that ASTRAND might be improved by basing the prediction on multiple regression.

The suitability of VO_2max as the criterion for aerobic fitness was questioned and it was suggested that investigations into the relationships between heart rate, oxygen uptake kinetics and anaerobic threshold may yield more valid information about the status of an individual's aerobic capabilities.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Maximal oxygen consumption (VO_2 max) is regarded as the reference standard for cardiovascular fitness (Mitchell et al, 1957; Shephard et al 1968). However, direct tests of VO_2 max are not always appropriate or accessible. They require skilled technicians, sophisticated instruments and cooperative, motivated subjects. To enable a wider population to have access to fitness testing, a number of tests have been developed to predict VO_2 max from submaximal exercise. These tests are generally based on extrapolation of submaximal measures of heart rate to predicted maximal heart rate, using the linear relationships between power output, heart rate and oxygen consumption (Astrand & Ryhming, 1954; Rowell et al, 1964). Two such tests which are widely used in Canada are the Astrand cycle test (ASTRAND) and the Canadian Aerobic Fitness Test (CAFT).

The ASTRAND test, (Astrand & Ryhming, 1954) is a six minute, single load test performed on a cycle ergometer. An empirically derived nomogram uses heart rate and power output to predict maximal oxygen consumption.

The CAFT was originally developed as a motivational and educational tool (Jette, 1977), but has since been used to predict VO_2 max (Jette et al, 1976), prescribe exercise (Jette, 1977), and, as part of the Canada Fitness Survey, to determine the cardiovascular fitness of Canadians (Shephard, 1986). The test is performed by stepping up and down a set of double 20.3 cm steps at a specified pace, in a series of stages. VO_2 max is predicted on the basis of a multiple regression equation.

Both of these predictive tests have been criticized. Wyndham et al (1959) and Rowell et al (1964) criticized the Astrand procedure based on their findings that oxygen consumption and heart rate reach the asymptote to their respective maxima at different rates. The Astrand test has also been found to systematically underpredict VO_2 max (Kasch, 1984) and to have correlations of about .6 to .7 with VO_2 max (Kasch, 1984; deVries and Klafs, 1965; Glassford et al, 1965). Astrand (1977) has stated that the test has an error of ± 10% for well-trained individuals and ± 15% for moderately-trained individuals. The CAFT has been criticized as unreliable (Morgan et al, 1984), subject to habituation (Thomas,1988), and having a mathematically restricted range of prediction (Leger, 1984). The original validation study has been criticized for having a small, heterogeneous validation group which may have resulted in a poor multiple regression equation (Bonen et al, 1977).

A variety of modifications have recently been made to the CAFT (Thomas, 1988) which may improve its reliability and validity.

Both of these tests are commonly used to assess the cardiovascular fitness of individuals, to prescribe exercise, and to monitor training improvement. Inaccuracy in assessment may result in incorrect classification, misinterpretation of training efficacy, or an inappropriate exercise prescription. While monitoring is potentially motivational, it is also conceivable that a test which does not accurately reflect training changes may actually be demoralizing.

The purpose of this study was to determine the sensitivity of the Astrand and CAFT to changes in VO_2 max following a training program. Sensitivity is defined, for this study, as the ability of the tests to discriminate between the trained and untrained state. Intuitively, this implies that the test must be accurate and, in turn, accuracy implies reliability and validity. It also implies that the validity must not change with changes in training status.

Reliability is a measure of consistency and has a variety of forms. The form required here is called stability, which Kerlinger (1983) describes as a measure of the consistency of scores obtained without regard for the truth of those scores. A stable test will yield the same, or very similar, results on repeated measurements.

Validity is closely related to reliability. However, it is a fundamental tenet that validity is a property of the inference and not of the measurement. The inference of interest in this study is the prediction of VO_2 max for an individual, before and after training, using a directly measured VO_2 max as the criterion of the prediction. This type of validity is commonly referred to as concurrent validity.

The specific objectives of this study were to determine the stability of the Astrand test and the CAFT test and to determine the concurrent validity of those tests before and after a training program.

METHODS

Design: The design of the study is illustrated in Figure 1. Random assignment to training (TG) or control (CG) groups was made prior to testing. Subjects were informed of their assignment on completion of the screening process, which included the first Astrand test. Subjects were then tested on the CAFT and a direct measure of VO_2 max on a cycle ergometer (DIRECT). All three tests were repeated one week later to determine the stability of each test. Subjects then repeated all tests after completion of a seven-week period of training or control. Each set of three tests was completed within seven days with at least 24 hours between successive tests within each set. As the Astrand test was included as part of the screening, it was always the first one performed. The other two were performed in arbitrary order. Subjects were asked to keep to their normal routine between tests.

Figure 1. Design of the Study.

PRE-STUDY	WEEK 1	WEEK 2	WEEK 3 - 9	WEEK 10
1. Recruitment	1. Assignment to Group	1. ASTRAND	Training or	1. ASTRAND
2. Orientation	2. ASTRAND and blood pressure	2. CAFT or DIRECT	Control	2. CAFT or DIRECT
3. PAR-Q	3. Subjects notified of assignment	3. DIRECT or CAFT		3. DIRECT or CAFT
4. Medical clearance	4. CAFT or DIRECT			
5. Informed consent	5. DIRECT or CAFT			

Subjects: Sedentary female subjects aged 20 - 29 years were recruited from the staff and student population at the University of Alberta. Sedentary was operationally defined as physically inactive for at least six months immediately preceding the study. This group was chosen as it was considered to be representative of a large fraction of fitness class participants and, also, was readily accessible. The age range was also consistent with one of the CAFT age categories, and sedentary subjects were considered to have good potential for improvement in VO_2 max.

After an orientation to the laboratory and to test procedures, prospective subjects were required to complete a Physical Activity Readiness Questionnaire (PAR-Q, Appendix 4), obtain medical clearance and give written, informed consent. Blood pressure responses to exercise were monitored during the first Astrand test. No subjects were excluded due to abnormal blood pressure response.

Thirty subjects were randomly assigned to the TG (n=20) or CG (n=10). Before testing began, and prior to being told of their assignment, two CG subjects decided not to participate. Another CG subject withdrew after learning she was assigned to the CG. One subject from each of the TG and CG withdrew from the study, due to time constraints, before completing one set of tests. Two subjects were subsequently recruited as controls, resulting in 19 TG and 8 CG subjects. Subject characteristics are summarized in Table 1.

All tests: Subjects were asked to adhere to strict pretest behaviours: to refrain from strenuous physical activity for six hours before testing, and to have no food, tobacco, caffeine, alcohol or drugs for at least one and a half hours. Most subjects were tested at approximately the same time of day.

DIRECT: Subjects performed a continuous, incremental test to exhaustion to determine maximal oxygen consumption. The test was performed on a mechanically braked cycle ergometer (Monark, model 668). Subjects pedaled at a nominal cadence of 65 rpm. The initial resistance was 0.5 kiloponds (kp) and increments of 0.5 kp were applied every two minutes until the subject was unable or unwilling to continue. Initial power output was 32 W (nominal) with 32 W increments. Strong verbal encouragement was given

Table 1. Subject Characteristics at the Start of the Study.

Age (yr)	Height (cm)	Weight (kg)	$\dot{V}O_2$ max ($l \cdot min^{-1}$)
24 (3)	166 (7)	62.6 (8.8)	2.216 (0.300)

Values are means (SD). $\dot{V}O_2$max is the directly measured $\dot{V}O_2$ max on the first test occasion.

n=27 except for $\dot{V}O_2$ max (n=26)

throughout.

One ergometer was used for all subjects throughout the study. It was calibrated by suspending known weights at the beginning of each test session and periodically throughout the study. The zero point was visually inspected regularly. The ergometer did not require calibration adjustment at any point in the study.

Visual feedback of pedalling speed was provided to the subjects with a digital cyclometer. Actual pedal revolutions were counted each minute by an electronic digital counter connected to the crank and manually recorded.

Expired gases were analyzed with a Beckman Metabolic Measurement Cart (SensorMedics). The volume turbine was calibrated each test day with a 1.04 L syringe. Zero was set on the O_2 and CO_2 analyzers using compressed nitrogen at the start of each testing session. The span on the analyzers was calibrated with precision mixed gases (O_2, CO_2 and N_2) before and after each test. The data output of the metabolic cart was corrected for any analyzer drift during the test using the posttest calibration value as the criterion for correction.

Heart rate was recorded in the last 10 seconds of each minute on an electocardiograph (Birtcher, model 365; Cambridge Instruments VS4). To determine the proximity of the subject's heart rate to age-predicted maximum during the test, an estimate was made by measuring the span of two consecutive QRS complexes using a plastic ruler designed for that purpose.

CAFT: The CAFT protocol was as described in the CSTF Operations Manual (Fitness and Amateur Sport Canada, 1986) with modifications as adopted by the Canada Fitness survey for the 1988 longitudinal follow-up study of Canadian fitness.

These modifications were as follows:

1. Subjects stepped for a maximum of five rather than three stages.
2. Heart rate was measured during the test rather than in the recovery period.
3. A telemeter system (PE 3000, Polar Electro), rather than a stethescope, was used to determine heart rate.

4. The cadence-setting taped music had a voice-over which counted out the step sequence throughout the work interval.
5. The criterion for ending the test was heart rate exceeding 85% of age-predicted maximum (220-age) at the end of the stage.

Maximal oxygen consumption was predicted using the following regression equation:

$$VO_2 \max = 32 + 16(VO_2) - 0.17(\text{weight}) - 0.24(\text{age})$$

where VO_2 max is in ml $O_2 \cdot kg^{-1} \cdot min^{-1}$, VO_2 is the O_2 cost ($L \cdot min^{-1}$) of the last completed stage of stepping, weight is in kg and age in years (personal communication, S. Thomas, 1989). That value was converted to $L \cdot min^{-1}$ for data analysis. The equation was suggested by Thomas as the best of three which were derived from a new validation study which incorporated the changes outlined above.

Pretest resting heart rate below $100 \cdot min^{-1}$ was required to begin the test and subjects above that level rested until this value was attained. Subjects were taught the stepping sequence for the CAFT and allowed to practise a few cycles of the first stage. The subject's stepping rate was continuously monitored to ensure that it matched the cadence set by the tape. Tape cadences were timed to determine consistency with the nominal cadence for each stage of the test and were found to agree. Heart rate was monitored with a telemeter as described above and recorded at the end of each stage. Simultaneously, an ECG record was taken at the time between the warning signal and the record signal on the tape (5 sec). This was then measured over six consecutive beats to determine heart rate at that time. The two methods were found to be in very close agreement. The telemeter reading was used in the analysis because it was the method used in the Canada Fitness Survey.

ASTRAND: Subjects were tested on a Monark 668 mechanically braked cycle ergometer. The test was not started unless pretest resting heart rate was below $100 \cdot min^{-1}$. Subjects who initially exceeded that rate were reassured and sat quietly until that level was attained. Subjects warmed-up for four minutes at a resistance of 0.5 to 1.0 kp (29 - 58 W), depending on the size of the subject. A pedal frequency of 60 revolutions per minute was used throughout. Subjects then cycled a further six minutes at a power output set to elicit a steady-state heart rate of between 130 and 170 beats per minute. Steady-state was considered to be achieved if the heart rate in the last two minutes differed by five beats per minute or less. If the difference was more than five beats, the subject continued for an additional minute or two until the criterion was attained. In a very few cases, heart rate at the end of the test was outside the target range. In this case, the subject was rescheduled and an appropriate change to the load was made to elicit the required heart rate at the retest.

In the second series of tests the same test resistance was used in establishing stability. In the third series of tests, the same resistance was used if it was judged that the subject's heart rate would still fall within the target range as above. Otherwise, the subject was given a higher resistance than previously.

Heart rate was measured with an ECG. A five-second trace was taken during the last 10 seconds of each minute. The distance between six consecutive heart beats was measured and heart rate calculated. The paper speed of the ECG was checked frequently with hand timing. A digital cyclometer was used to provide visual feedback of pedal frequency to the subject during the test. Actual pedal frequency was directly determined every minute with an electronic counter. Actual pedal frequency was used in calculating the power output for the purpose of predicting VO_2 max. Resistance was monitored and adjusted as required to keep it at the intended value. Ergometer calibration was carried out as detailed previously.

Maximal oxygen consumption was predicted using a computer solution to the Astrand nomogram (Shephard, 1970). Steady-state heart rate at the end of the test was used in the prediction formula.

Training Program: Subjects trained on stationary cycles three times per week, 30 minutes per session, over a period of seven weeks.

Initial intensity was set to elicit a heart rate of approximately $150 \cdot min^{-1}$. The subject's performance on the VO_2 max test was considered when prescribing training loads. Specifically, note was taken of the point where respiratory exchange ratio exceeded 1.00 on the assumption that levels exceeding this might not be sustainable. Subjects were allowed to select the combination of pedal frequency and resistance which they preferred to arrive at the prescribed heart rate level. Telemeters were worn by subjects to provide heart rate feedback to them. In the first two weeks, all subjects were encouraged to adopt an interval training approach alternating high and low power outputs but cycling continuously for 30 minutes.

Training sessions were supervised and each subject's heart rate, resistance, and pedal frequency were recorded each day. Intensity was progressively increased over this period through a combination of increased pace and resistance, according to the subject's preference. Training heart rates of 165 - 180 were typical in the final week of the study. In the first two weeks the 5-minute cycling warm-up was included in the 30-minute session; after that it was separate. Effectively this was an increase of 5 minutes duration. Random checks were made of the quality of training by using the PE3000's ability to sample heart rate every five seconds and generate a graph of heart rate during the training session.

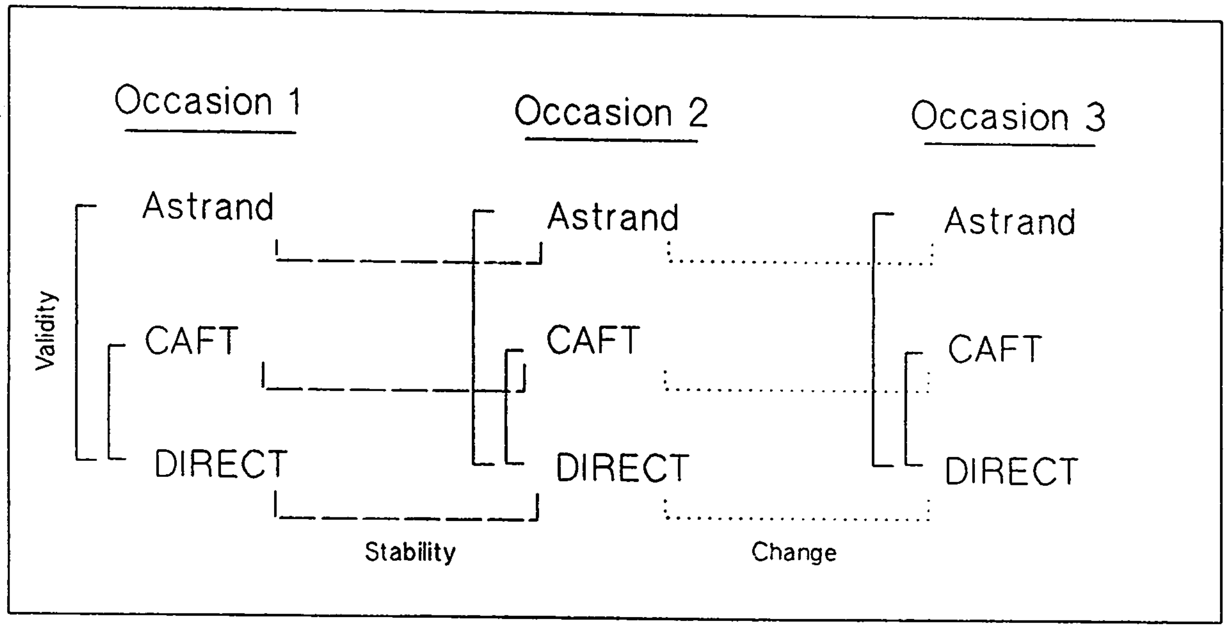
Statistical Analysis: Of the 27 subjects included in the study, there were four with incomplete data.

Two CG subjects began regular exercise prior to the retest and were not retested. Two TG subjects each had one VO_2 max in which the data did not meet criteria for inclusion (one test 1, one test 2). Because retests were not possible, their data was excluded from those analyses requiring complete data.

The analyses performed are illustrated in Figure 2. The following rules were followed in determining which data would be included in a particular analysis. For stability analyses, all subjects with data for both tests were included (see Table 2). For concurrent validity, only TG subjects with complete data sets were included (n=17). This allowed comparison of the validity coefficients on each occasion. In testing differences between means, only complete data sets were used (TG: n=17; CG: n=6).

Mean differences, stability coefficients and validity coefficients were calculated using a repeated measures ANOVA (University of Alberta, Department of Educational Research Services, ANOV14). Stability coefficients are intra-class reliabilities (single measure, unadjusted) and validity coefficients are inter-class reliabilities (single measure, adjusted). Significance was accepted at alpha=.05.

Fig. 2. Statistical analyses of the study.



Concurrent validity was determined by comparing the directly determined $\dot{V}O_2$ max with the predicted $\dot{V}O_2$ max on the same occasion. Stability was determined by comparing the values obtained on the same test over the first two occasions. Change was tested over occasions 2 and 3 (before and after training) on the same test.

RESULTS

Stability: Table 2 shows mean VO_2 max values for the first two test occasions and the stability coefficients. Only the Astrand test prediction was significantly different between the two occasions.

Validity: The concurrent validity coefficients of the predictors are presented in Table 3. Mean predicted VO_2 max on the CAFT test was significantly different from mean DIRECT on all occasions. Mean predicted VO_2 max for the Astrand tests were significantly different from mean direct VO_2 max on the first two occasions but not on the third.

Change: Mean VO_2 max values for the TG and CG on all tests are presented in Table 4, and graphically shown in Figure 3 and Figure 4. The training group showed a statistically significant increase from occasion 2 to occasion 3 on all three tests. The magnitude of that increase appeared to be largest when measured by the Astrand test. The CG change between occasions was not statistically significant.

Table 2. Mean $\dot{V}O_2$ max (± SD) and stability coefficients for the three tests.

	$\dot{V}O_2$ max ($L \cdot min^{-1}$)				
	Occasion 1	Occasion 2	p(F)	n	stability
ASTRAND	1.829 (.357)	1.998 (.405)	.006	27	.633
CAFT	2.492 (.218)	2.513 (.243)	.194	27	.933
DIRECT	2.217 (.307)	2.232 (.333)	.6071	25	.903

Stability coefficients were single measure, unadjusted reliabilities, obtained with dependent ANOVA.

Table 3. Mean $\dot{V}O_2max$ ($L \cdot min^{-1} \pm SD$) and concurrent validity coefficients (val) for TG subjects on all occasions.

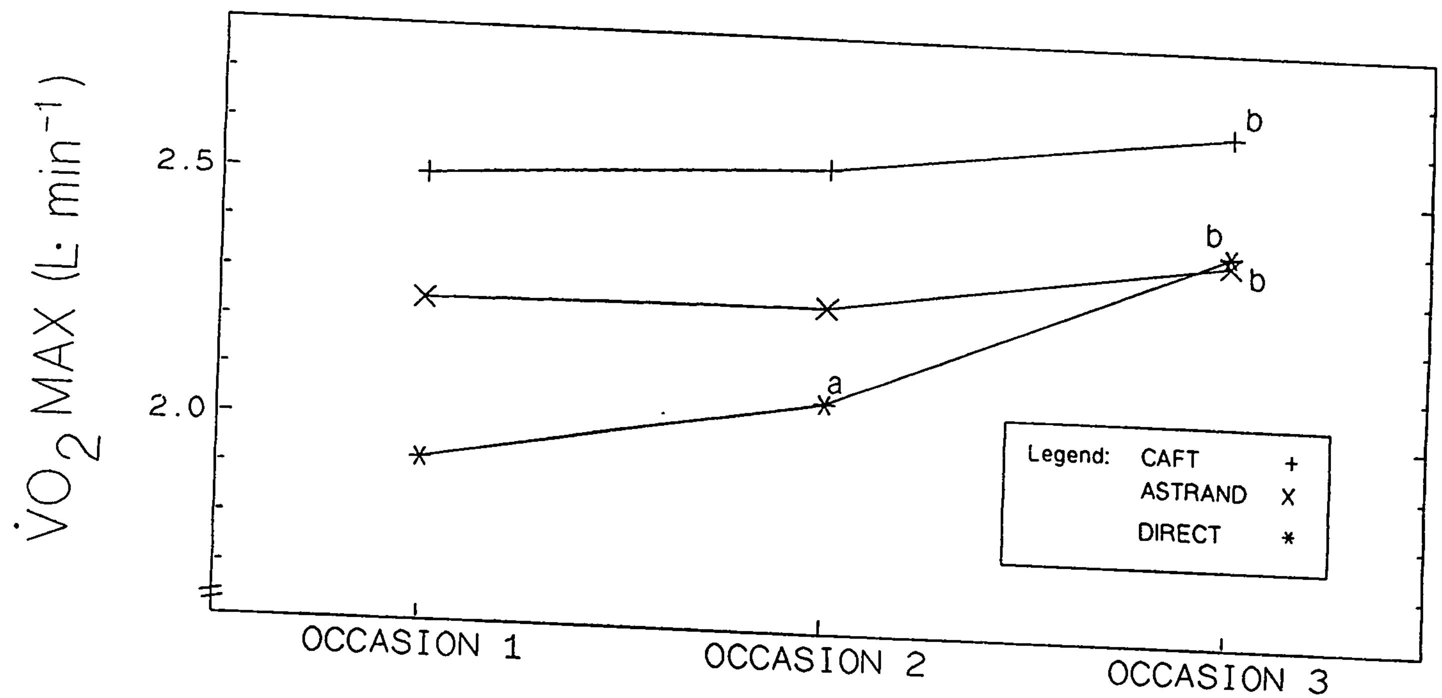
	Occasion 1		Occasion 2		Occasion 3	
	$\dot{V}O_2max$	val	$\dot{V}O_2max$	val	$\dot{V}O_2max$	val
ASTRAND	1.918 (.359)	.579	2.055 (.346)	.637	2.387 (.505)	.589
CAFT	2.506 (.237)	.715	2.534 (.278)	.735	2.629 (.326)	.636
DIRECT	2.244 (.337)		2.253 (.344)		2.369 (.344)	

Coefficients reported are single measure, adjusted for means, obtained with dependent ANOVA. Only subjects with complete data were included (n=17).

Table 4. $\dot{V}O_2$ max ($L \cdot min^{-1}$ ±SD) values obtained by the Training Group (TG; n=17) and the Control Group (CG; n=6) on all tests on all occasions.

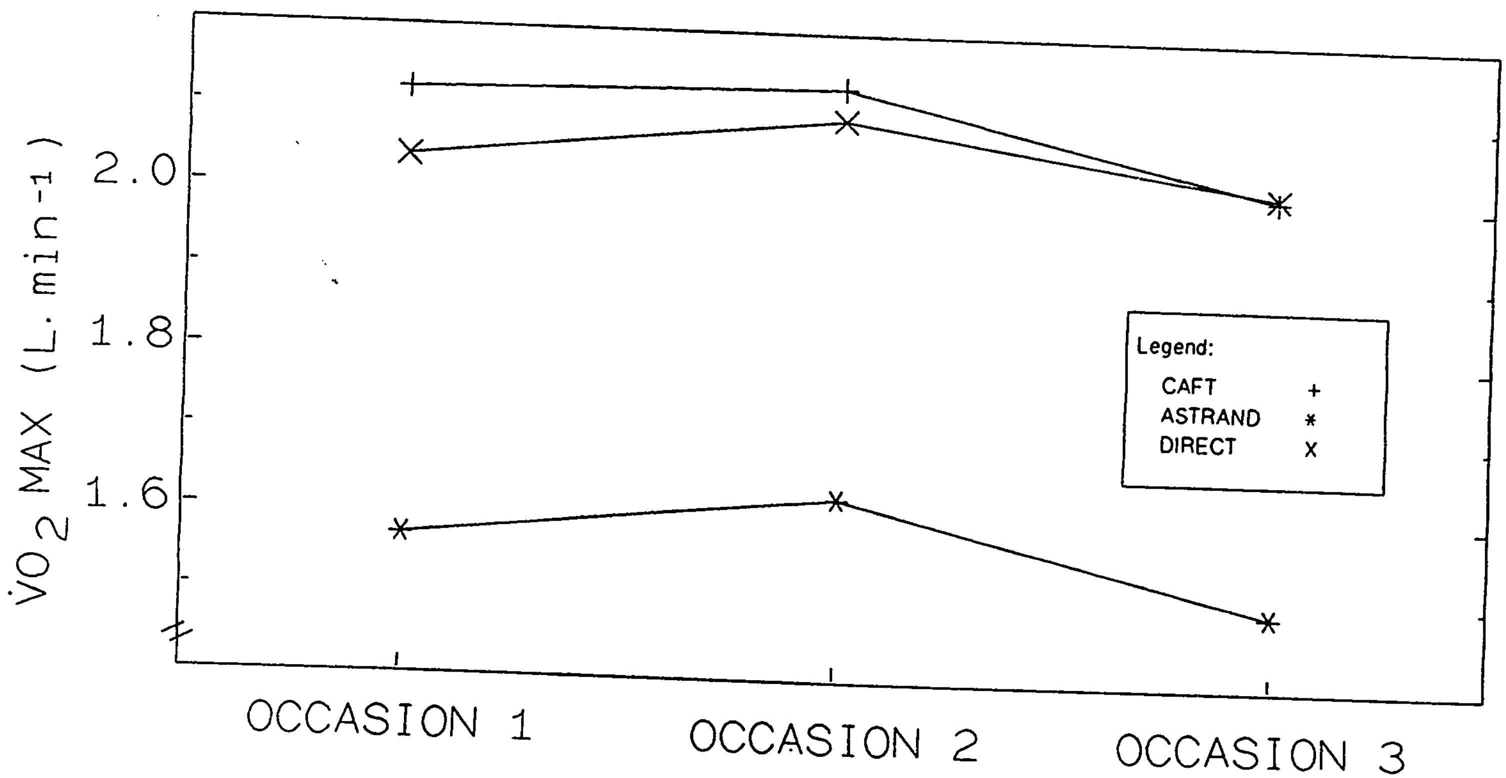
	Occasion 1		Occasion 2		Occasion 3	
	TG	CG	TG	CG	TG	CG
ASTRAND	1.918 (.359)	1.567 (.310)	2.055 (.346)	1.620 (.283)	2.387 (.505)	1.487 (.156)
CAFT	2.506 (.237)	2.121 (.178)	2.534 (.278)	2.132 (.111)	2.629 (.326)	2.007 (.079)
DIRECT	2.244 (.337)	2.038 (.243)	2.253 (.344)	2.093 (.331)	2.369 (.344)	2.011 (.356)

Figure 3. $\dot{V}O_2$ max values for the Training Group (n=17) on all tests on all occasions.



[a]significantly different from occasion 1 ($p<.05$)
[b]significantly different from occasion 2 ($p<.05$)

Figure 4. $\dot{V}O_2$ max values for the Control Group (n=6) on all tests on all occasions.



DISCUSSION

At the outset, it should be noted that the results of this study were obtained under well-controlled, laboratory conditions. If the conditions under which the predictive tests are less than ideal, the obtained results can only be less stable and less valid.

Stability: For a test to be sensitive, it must first be stable. To the extent that it is not, true changes in what is measured will be masked by the instability.

The variance between consecutive tests may be partitioned into systematic variance, biological variability and technological or method variance (Katch, 1982, Mitchell, 1957). It is assumed that there should be no systematic variance because no treatment has taken place. All of the variance should be due to random, or error, variance, which is comprised of biological and method error.

The stability of the DIRECT test is consistent with values reported in other investigations (Katch et al, 1982; Mitchell et al, 1957).

Based on these results, the CAFT test appears to be even more stable than the DIRECT test. The high stability for the CAFT is likely spurious. In the development of the regression equation, consideration was given to minimizing any potential habituation effect, specifically, changes in heart rate. (Thomas, 1988). Therefore, the regression equation does not include heart rate. The only influence that heart rate can have on the outcome of the test is in determining when the test is stopped. This, in turn, will determine the final stage of stepping, which is a factor in the regression equation. However, for a given individual, the prediction can only take on five discrete values, corresponding to the five stages of the test. Thus, unless a subject's heart rate is different enough, on a second test, to cross a stage threshold the same VO_2 max will be predicted.

The Astrand test was the only one which appeared to be unstable. This test was the first one done by all subjects and included the screening for blood pressure. Predicted VO_2 max was significantly greater on occasion two than occasion one. The difference may have been due to habituation to the laboratory situation. If the subject was more anxious on the first occasion, and heart rate was elevated, the effect would be a lower predicted VO_2 max. Certainly there were many subjects with heart rates which were initially well above $100 \cdot min^{-1}$. Even though no tests were begun until heart rate was below this level it is possible that anxiety may have affected them. A habituation effect, if real, is of great importance in the applied setting. A client undergoing a fitness assessment will probably have only one test, and the effect of habituation will be unknown. No further effect was noticed for the CG subjects on the third test, implying that the predicted increase in VO_2 max for the TG was not likely due to habituation.

Validity: The validity coefficients for both predictive tests are low. With values in the region of .7, only 50% of the variance is common to the predicted and directly measured VO_2 max. This is a substantial commonality, but insufficient to warrant the use of the predicted VO_2 max in individual cases. There appears to be no difference between the coefficients obtained before and after training. However, the mean predicted VO_2 max for the TG on ASTRAND was not significantly different from VO_2 max on the DIRECT test after training, yet was significantly different before training.

CAFT overestimations of VO_2 max were as high as 33%. A calculation of all possible predicted values for each stepping stage for each subject reveals limitations in the prediction range. The minimum possible predicted VO_2 max ($ml \cdot kg^{-1} \cdot min^{-1}$) for the lightest subject in this study was 40.2 $ml \cdot kg \cdot min^{-1}$. The actual VO_2 max of that subject was 31.0 $ml \cdot kg \cdot min^{-1}$., a 33% overestimation.

Comparison of group means gives another perspective on validity. The CAFT mean predicted VO_2 max was significantly different from the DIRECT mean on all three test occasions. The Astrand predicted mean VO_2 max was significantly different from the DIRECT mean on the first two occasions but not on the third. The CAFT regression equation was validated using a treadmill VO_2 max test. Treadmill tests have been shown to yield VO_2 max values about 5 - 10% higher than on the cycle ergometer. (Cardus, 1979). With this factored in, the CAFT means are quite close to the DIRECT means.

Change: The TG showed a small (5%), but significant change in VO_2 max (DIRECT) following the training program when compared with their own pretraining level. There was no significant change in VO_2 max (DIRECT) in the CG over the control period. Although the numbers in the CG were small, it would seem reasonable to infer that the observed changes were due to the training program.

The small measured change in VO_2 max is at variance with subjective observations of the progress of the subjects through the training. Subjects claimed to feel much fitter and although the relative difficulty of the training session did not appear to change over the course of training, much more work was being done in each session in the latter part of the training. Workloads were progressively increased, and subjects worked out faithfully. The workloads and rate of progression were similar to those which might be expected in a standard fitness class.

Sensitivity: The sensitivity of the test must be evaluated with respect to the precision demanded of the test. One perspective on this problem is that of statistical power analysis. The power to detect a significant difference is a function of the sample size, the effect size, and alpha, the acceptable probability of a type I error. Continuing the analogy, the problem of discriminating between the trained and untrained state for an individual has a sample size of one. Effect size is the ratio of the difference between tests and the standard deviation. In the case of the individual, the standard deviation of interest is the reliability of that individual on the test. To maximize effect size, the reliability must be maximized (minimizing the standard deviation) and the difference between tests must be maximized. If the training effect is subtle, it can only be seen if the reliability is very high.

Problems with the predictions:

The problems discussed in this section relate solely to the ability of the tests to predict VO_2 max.

The possibility that the correlations may have been attenuated due to a restricted range in the sample was considered. However, the selection of sedentary individuals did not result in a substantial restriction of the range of VO_2 max.

The basis for prediction of VO_2 max is that a linear relationship exists between heart rate and oxygen consumption. One of the problems with this, the differential rate of approach of heart rate and oxygen consumption to their respective asymptotes, has already been identified. However, in monitoring progress, it is even more important to consider the extent to which training-induced bradycardia is reflective of VO_2 max changes. For example, heart rate is subject to neural control (Stone et al, 1985) and unless the neural control of heart rate adapts with an identical time course with VO_2 max adaptation, it will not necessarily be a valid indicator of VO_2max adaptation.

A second problem with the tests lies in their lack of specificity. While VO_2 max may be an indicator of the maximal capabilities of the cardiovascular system, the expression of VO_2 max is clearly a function of the test mode (Cardus, 1979). To the extent that predictive tests are used to quantify fitness in a non-athletic population, they must be able to reflect changes in VO_2 max independent of the type of training which might be undertaken.

The CAFT regression equation is problematic. Currently, the factor for the completed stage is the average oxygen consumption of the validation group. Oxygen consumption for different individuals at a given stage will vary with their weight and their mechanical efficiency. This mean value should not be used in deriving the multiple regression equation. Equally important, there is currently no heart rate related variable in the equation. The prediction cannot possibly differentiate between two individuals who finish the same stage with different heart rates. The equation also uses only final stage data. As there are up to five stages of data available, a substantial amount of information is being disregarded.

Another problem may be in the use of three minutes stages. The implicit assumption in this is that steady state will be achieved in this time. If this is not the case, the recorded heart rate will be less than it would be at steady state for that work rate, possibly allowing the subject to proceed to another stage when steady state heart rate would have terminated the test. The test cannot differentiate between two subjects having the same heart rate at the end of the stage, even though one of the subjects may have attained steady state, and the other subject's heart rate is still rising.

The ASTRAND results suggest some kind of differential validity before and after cycle training, which makes monitoring of changes virtually impossible. One possible explanation lies in the fact that the training mode was the same as the test mode. As the subjects were sedentary, seven weeks of cycling likely improved their cycling skill. This in turn would improve their mechanical efficiency. By definition, this would mean that their actual power output would be decreased at the nominal power output of the test. That is, less unmeasured work is done after training, which would result in a lower oxygen requirement and, possibly, a lower heart rate. This would result in a prediction of an increased VO_2 max even if no increase had taken place.

The ASTRAND uses only one power output for its prediction, implicitly assuming the same heart rate - oxygen consumption - power output relationship for all subjects, in a given age range, with a given VO_2 max.

Improving the Predictions

The CAFT regression equation should be reworked. Instead of using the predicted oxygen consumption for the stage, the subject's weight and stepping cadence should be used as a power output variable. A heart rate related variable should be forced in to improve the sensitivity of the test to differences in fitness within a stage. By making use of all of the stages of the test, rather than just the last one, the actual slope of the heart rate response could be employed as a variable. With further research into heart rate kinetics, a variable such as time to steady state or half-time to steady state could possibly be used. This is entirely feasible with current technology. The Sport Tester PE3000 is able to record heart rate every five seconds.

The ASTRAND test might be improved by employing more than one stage and employing a multiple regression equation instead of the current empirical formula. Watt-pulse, the ratio of power output to heart rate, could be determined at two or three loads and used as a variable in the regression.

VO_2 max as Criterion

It is implicit in this study that VO_2 max is an appropriate criterion of aerobic fitness. However, if the CAFT and ASTRAND are found to be invalid with respect to that criterion, this does not invalidate the tests themselves. They may in fact be able to validly and sensitively measure other indices of oxidative metabolism, for example endurance or anaerobic threshold. Future work on these tests must first deal with the question of construct validity.

Even the sensitivity of direct measures of VO_2 max could be questioned. Biological variability in the direct measurement of VO_2 max has been estimated as 5.6%, on average, with college athletes (Katch et al, 1982) and 7.9% with well-trained men (Kuipers et al, 1985). It seems a reasonable hypothesis that untrained individuals would be even less reliable because they are not used to the intense effort and uncomfortable sensations attending an exhaustive test.

The DIRECT results indicated a small change in VO_2 max. Given that this is accurate, the disparity between this result and the subjectively noted improvements in the subjects' condition give additional cause to continue searching for a more appropriate indicator of aerobic fitness.

New Approaches

The simplest approach to monitoring training changes is to provide a standard, progressive exercise test and record heart rate. Progress will be seen over time in reduced heart rate and no reference to VO_2 max need be made.

At a slightly more sophisticated level, future research may be able to determine submaximal variables which more accurately reflect what we wish to know about changes in the body's capacity for aerobic metabolism. Relationships between submaximal heart rate, oxygen consumption and power output before and after training could be related to anaerobic threshold. Investigation of heart rate and oxygen uptake kinetics, with trained and untrained subjects, and with respect to the time course of adaptation, may yield variables which are more reflective of a state of training than of an absolute value such as VO_2 max.

From a mathematical approach, factor analysis could be employed in order to define general factors from the physiological measures described above. This may better define the somewhat blurry construct of endurance.

Conclusions

Though the relationships indicate that the predictive tests share a substantial amount of variance with the criterion, neither of the predictive tests shows a strong enough relationship with VO_2 max to warrant its use as a predictor for individuals.

The uses of CAFT and ASTRAND, in their current form, to predict VO_2 max in the context of fitness assessment, exercise prescription or monitoring progress is not justified.

Some suggestions have been made which may improve the predictions, but I believe that future work should not concentrate solely on VO_2 max.

REFERENCES

Astrand, P.-O., & I. Ryhming. (1954). A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. Journal of Applied Physiology; 7:218-221.

Astrand, P. -O. & K. Rodahl. (1977). Textbook of Work Physiology (2nd ed.). McGraw Hill Book Company.

Bonen, A., J. Gardner, J. Primrose, R. Quigley, & D. Smith. (1977). An evaluation of the Canadian Home Fitness Test. Canadian Journal of Applied Sport Sciences; 2:133-136.

Cardus, D. (1979). Exercise testing: methods and uses. Exercise and Sports Science Reviews; 8:59-103.

de Vries, H.A., & C.E. Klafs. (1965). Prediction of maximal oxygen intake from submaximal tests. Journal of Sports Medicine and Physical Fitness; 5:207-214.

Fitness and Amateur Sport Canada. (1986). Canadian Standardized Test of Fitness, Operations Manual, 3rd edition. Minister of State, Fitness and Amateur Sport.

Glassford, R.G., G.H.Y. Baycroft, A.W. Sedgwick & R.B.J. Macnab. (1965). Comparison of maximal oxygen uptake values determined by predicted and actual methods. Journal of Applied Physiology; 20(3):509-513.

Jette, M. (1977). A calculator to predict maximal oxygen consumption for use with the Canadian Home Fitness Test. Canadian Journal of Public Health; 68:195-198.

Jette, M., J. Campbell, J. Mongeon, & R. Routhier. (1976). The Canadian Home Fitness Test as a predictor of aerobic capacity. Canadian Medical Association Journal; 114:680-682.

Kasch, F.W. (1984). The validity of the Astrand and Sjostrand submaximal tests. The Physician and Sportsmedicine; 12(8):47-51.

Katch, V.L., S.S. Sady, & P. Freedson. (1982). Biological variability in maximum aerobic power. Medicine and Science in Sports and Exercise; 14(1):21-25.

Kerlinger, F.N. (1973). Foundations of Behavioural Research (2nd ed.). Holt, Rinehart and Winston Inc.

Kuipers, H., F.T.J. Verstappen, H.A. Keizer, P. Geurten & G. van Kranenburg. (1985). Variability of aerobic performance in the laboratory and its physiologic correlates. International Journal of Sports Medicine; 6:197-201.

Leger, L. (1984). The Canadian Home Fitness Test revisited. CAHPER Journal; July-Aug.;6-8.

Mitchell, J.H., B.J. Sproule, & C.B. Chapman. (1958). The physiological meaning of the maximal oxygen intake test. Journal of Clinical Investigation; 37:538-547.

Morgan, K., A.O. Hughes, & R. Philipp. (1984). Reliability of a test of cardiovascular fitness. International Journal of Epidemiology; 13:32-37.

Rowell, L.B., H.L. Taylor and Y. Wang. (1964). Limitations to prediction of maximal oxygen intake. Journal of Applied Physiology; 19(5):919-927.

Shephard, R.J. (1986). Fitness of a Nation: Lessons from the Canada Fitness Survey. Karger AG, Basel, Switzerland.

Shephard, R.J., C. Allen, A.J.S. Benade, C.T.M. Davies, P.E. di Prampero, R. Hedman, J.E. Merriman, K. Myhre, & R. Simmons. (1968). Standardizatino of submaximal exercise tests. Bulletin of the World Health Organization; 38:765-775.

Thomas, S. (1988). CAFT Research Project Final Report.

Wyndham, C.H., N.B. Strydom, J.S. Maritz, J.F. Morrison, J. Peter & Z.U. Potgieter. (1958). Maximum oxygen intake and maximum heart rate during strenuous work. Journal of Applied Physiology; 14(6):927-936.

APPENDIX 1

REVIEW OF THE LITERATURE

A. Maximum Aerobic Power

1. Terminology and Relevance

Various terms have been used over the years, sometimes inappropriately, to describe the capacity of the human body for oxidative metabolism: maximum oxygen consumption, maximum oxygen intake, maximum oxygen uptake, maximum aerobic capacity, maximum aerobic power. There has been recognition, in recent years, that there are two aerobic parameters which determine the ability of the body to sustain prolonged physical activity (Whipp et al, 1981; Boulay et al, 1985). The first is maximum aerobic power (VO_2 max), defined as the maximum rate at which the whole body can consume oxygen (Thoden et al, 1982). The second parameter, less well defined, is the maximum aerobic capacity, which is considered to be related to the concept of anaerobic threshold (Whipp et al, 1981; Boulay et al, 1985). According to Whipp et al (1981), the anaerobic threshold sets the upper limit for oxygen consumption without a sustained arterial lactic acidosis.

Maximal aerobic power has been adopted as the reference standard for cardiorespiratory fitness (Shephard et al, 1968b; Mitchell et al, 1957). Although many factors potentially limit VO_2 max, a series of studies by Saltin (1985), provided strong evidence that the capacity of the heart to deliver oxygen was the limiting factor in whole body exercise.

2. Measurement of VO_2 max

Early protocols, such as that of Taylor et al (1955), were discontinuous, requiring subjects to exercise at progressively increasing loads on different days. These measurements were time consuming for the subject and the researcher, and had the added possibility that fitness might change while measurements were carried out (Shephard et al, 1968b). Continuous, incremental protocols overcome some of these drawbacks and comparisons between discontinuous and continuous test protocols yielded no difference in the VO_2 max measured (McArdle et al, 1973; Shephard et al, 1968b).

In a recent handbook of testing procedures, Thoden et al (1982) recommend that VO_2 max tests start with a warm up, have increments which are small enough to avoid large increases in lactate but large enough that the test is not unduly prolonged, and is as specific to the training mode as possible. The amount of active muscle mass is a determinant of VO_2 max (Taylor et al, 1955; Cardus, 1979) and it was suggested by Mitchell et al (1958), that VO_2 max should not be considered as absolute; rather, it is relative to the test conditions. Cycling generally yields VO_2 max values about 5-10% less than treadmill exercise (Cardus, 1979; Shephard et al, 1968b; McArdle et al, 1973). Despite this, Thoden et al (1982) suggest that non-specific tests are still able to give valid information regarding the central component of the oxygen transport system and the effects of training on this component.

3. Reliability of VO_2 max

Mitchell et al (1958) studied the reliability of VO_2 max measurement and reported test-retest correlation coefficients of 0.95. Katch et al (1982), in a more recent study with between eight and twenty-one repeated measures on five highly-trained athletes found a coefficient of variation of 5.6%, mainly due to biological rather than technological variability. However, to ensure reliability, standardization of the protocol is essential. Shephard et al (1968b; p.762) laid out some of the essentials: "measurements must not follow a heavy meal, previous exercise or exposure to heat... the subjects must be well motivated". Others emphasize "rigorous standardization" of the test, including: warm up, loading, (Taylor et al, 1955; p.78), and "rigid criteria for determining the point at which maximal oxygen intake is reached" (Mitchell et al, 1958; p.542).

4. Criteria for VO_2 max

The most valid criterion of VO_2 max is failure of oxygen consumption to rise with an increase in power output. However, this is not necessarily observed, depending on the size and duration of increments in the test. Taylor et al (1955) found that oxygen consumption increased by approximately 300 ml/min with 2.5% increments in grade during a treadmill test. He suggests that an increase of less than 150 ml/min is evidence that VO_2 max has been reached. This has been criticized by Wyndham et al (1959) who found that the rise in oxygen uptake was asymptotic at maximum levels. Rowell et al (1964) confirmed the work of Wyndham. The implication is that the end of linear oxygen uptake response with increasing power output is not sufficient evidence of VO_2 max, and further workloads should be attempted. Practically, this implies work to exhaustion in a progressive test. Further, it must be recognized that no absolute value should be adopted in ascertaining whether a test has yielded VO_2 max. The magnitude of the increase will be dependent on the magnitude of the power output increase. The changing ratio of VO2/P.O. should be a guide. It has been observed that oxygen consumption approaches its asymptote more slowly than heart rate does(Astrand & Ryhming, 1954; Wyndham et al, 1959). This has lead to a secondary criterion for assessing whether a test has yielded a "true" maximum: heart rate should be close to age-predicted maximum (220-age(yrs)) (Thoden et al, 1982). Thoden suggests also that a respiratory exchange ratio ("R": VCO2/VO2) exceeding 1.00 or a plasma lactate concentration exceeding 10 mM are indications of having reached VO2 max. These last two rely on the premise that anaerobic processes become increasingly important as VO2 max is approached and are the only methods of increasing power output once VO_2 max is reached.

B. Submaximal Predictive Tests

1. Justification

Continuous protocols and advances in data acquisition technology considerably shortened the time needed for testing VO_2 max. However, there are still drawbacks to this type of test. Subjects may be unwilling or unable to cope with the nausea, breathless, chest pain or quadriceps pain which are symptomatic of the maximal exertion required at the end of the test (Taylor et al, 1955; Shephard et al, 1968b). If monitoring VO_2 max is worthwhile, then providing a submaximal protocol will include a much larger part of the population. Submaximal tests are less reliant on expensive and sophisticated equipment and technical expertise, making them the instrument of choice in remote areas and more practical for the majority who do not have access to a laboratory.

2. Physiological Bases

The existence of a linear relationship between heart rate and oxygen consumption over a wide range is the basis for most predictive tests of VO_2 max (Astrand & Ryhming, 1954; Rowell et al, 1964). An assumption is also made that, within age groups, maximum heart rate is constant (Margaria et al, 1965), and that maximum heart rate declines with age at the rate of about 10 beats/min per decade (Davies, 1968). Submaximal oxygen consumption is inferred from the relatively constant mechanical efficiency of the given mode of exercise (Shephard et al, 1968a) and a linear projection is made using the slope of the HR/VO2 curve to maximum heart rate (age predicted).

3. Astrand Nomogram

In 1954, Astrand and Ryhming published a nomogram based on empirical studies of the relationship between heart rate and oxygen consumption in well trained young men and women. They determined that maximum heart rate was 195 beats per minute; that men, on average, reached 50% of their VO_2 max at a heart rate of 128 bpm, 70% VO2 max at 154 bpm, and 100%

VO_2 max at 195 bpm; women reached 50% of VO_2 max at a heart rate of 138 bpm, 70% VO_2 max at 174 bpm, and 100% VO2 max at 195 bpm. Subsequent studies determined age adjustment coefficients to account for the decline of maximum heart rate with age. The protocol has been modified by researchers attempting to improve its usefulness. The basic protocol is one stage at a fixed load, for six minutes. If the difference in heart rate in the last two minutes is less than or equal to 5 bpm then the last value is accepted as steady state. Otherwise, exercise continues until this criterion is reached. Heart rate and power output are located on the nomogram and VO_2 max read off. Shephard (1980) has derived a computer solution to the nomogram.

4. Canadian Aerobic Fitness Test (CAFT)

The CAFT, originally known as the Canadian Home Fitness Test, was developed in the mid 70's as a simple "motivational tool to promote cardio-respiratory fitness" (Jette, 1977; p.195). The test consists of stepping up and down a double 20.3 cm step at a rate which is set by a tape recording. Stages of the test are three minutes long. At the end of each stage, heart rate is measured and compared with a chart. If the heart rate is low enough, the subject progresses to the next stage which increments power output by increasing the stepping rate. The test was designed to demand 65-70% of average maximal aerobic power for a person ten years older than the subject in the first stage. Subsequent stages demand 65-70% of average VO_2 max for the same age group and an age group ten years younger, respectively. In development of the test, basal oxygen consumption was assumed to be 3 $ml \cdot kg^{-1} \cdot min^{-1}$; mechanical efficiency of stepping was assumed to be 16%; and energy yield was assumed to be 21 joules per litre O_2 (5 kcal/L). Step heights were based on Canadian home construction standards of 8" risers. (Bailey et al, 1976).

This test has been in wide use in Canada in promotional settings and in collecting data for the Canada Fitness Survey.

5. Validity of Predictive Tests

Wyndham et al (1959) challenged the notion that oxygen consumption and heart rate were linearly related up to maximum. Their data, subsequently confirmed in studies by Rowell et al (1964) demonstrated that both the heart rate and oxygen consumption curves were asymptotic near maximum levels. Further, they showed that the oxygen consumption curve approaches its asymptote more slowly than the heart rate curve; thus the linear relationship between the two breaks down at near-maximum levels. The result of such projections is systematically to underpredict VO_2 max.

Thoden et al (1982), stated that variations in maximum heart rates are also problematic. Davies (1968) also mentions this but states that it is a relatively small effect compared with the error of the asymptotes.

Rowell et al (1964) note that at submaximal levels of exercise, the heart rate can be influenced by emotional state, previous meal, degree of hydration, ambient temperature and posture. Davies (1968) found that the coefficient of variation of heart rate was larger at lower workloads, and redu to only 2% at high submaximal work rates.

From the outset, Astrand & Ryhming (1954) set limits on the generalizability of their nomogram. Still, a great amount of research has been devoted to validating and attempting to improve the test. Davies (1968) estimated the error of the prediction to be around 20%.

Specific criticisms of the CAFT have been made. The validation group, from which the regression equation was determined comprised only 59 subjects to cover all age and sex groups (Jette, 1976) and Bonen et al (1977) suggest that the width of age and fitness may have given a spuriously high multiple correlation. Leger has considered the regression equation and suggested that mathematical limits on the prediction systematically reduce the possible range of predicted VO_2 max.

Recently, modifications have been made to the test to deal with some of the criticisms. Previously, heart rate was measured by auscultation or palpation during the interval between stages. Now heart rate is being measured during exercise with a heart rate telemeter. Also, additional stages have been introduced to allow fitter individuals to attain a higher fraction of their VO_2 max.

C. Training

The training methods for improvement of VO_2 max in non-athletic populations has been outlined by the American College of Sports Medicine (1978). Training should be 3-5 days per week, at 50-85% of VO_2 max, 15-60 minutes of continuous aerobic activity. The mode of activity should be one which engages large muscle groups and is continuous, rhythmic and aerobic in nature. A variety of permutations of intensity, frequency and duration of training will produce adaptation. However, exercise which is too frequent or intense may lead to muscle soreness and injury. Increases up to 25% VO_2 max and more are reported, the highest adaptations occuring in the least fit individuals.

REFERENCES

American College of Sports Medicine. (1978). Position statement on the recommended quantity and quality of exercise for developing and maintaining fitness in healthy adults. Physician and Sports Medicine; 6:39-41.

Astrand, P.-O., & I. Ryhming. (1954). A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during submaximal work. Journal of Applied Physiology; 7:218-221.

Bailey, D.A., R.J. Shephard, and R.L. Mirwald. (1976). Validation of a self-administered home test of cardio-respiratory fitness. Canadian Journal of Applied Sport Sciences; 1:68-78.

Boulay, M.R., G. Lortie, J.A. Simoneau, P.Hamel, C. Leblanc, & C. Bouchard. (1985). Specificity of aerobic and anaerobic work capacities and powers. International Journal of Sports Medicine; 6:325-328.

Cardus, D. (1979). Exercise testing: methods and uses. Exercise and Sports Science Reviews; 8:59-103.

Davies, C.T.M. (1968). Limitations to the prediction of maximum oxygen intake from cardiac frequency measurements. Journal of Applied Physiology; 24(5):700-706.

Jette, M. (1975). An exercise prescription program for use in conjunction with the Canadian Home Fitness Test. Canadian Journal of Public Health; 66:461-464.

Jette, M. (1977). A calculator to predict maximal oxygen consumption for use with the Canadian Home Fitness Test. Canadian Journal of Public Health; 68:195-198.

Jette, M., J. Campbell, J. Mongeon, and R. Routhier. (1976). The Canadian Home Fitness Test as a predictor of aerobic capacity. Canadian Medical Association Journal; 114:680-682.

Kasch, F.W. (1984). The validity of the Astrand and Sjostrand submaximal tests. The Physician and Sportsmedicine; 12(8):47-51.

Katch, V.L., S.S. Sady, & P. Freedson. (1982). Biological variability in maximum aerobic power. Medicine and Science in Sports and Exercise; 14(1):21-25.

Leger, L. (1984). The Canadian Home Fitness Test revisited. CAHPER Journal. July-Aug.;6-8.

Margaria, R., P. Aghemo, & E. Rovelli. (1965). Indirect determination of maximal O_2 consumption in man. Journal of Applied Physiology; 20(5):1070-1073.

McArdle, W.D., F.I. Katch, & G.S. Pechar. (1973). Comparison of continuous and discontinuous treadmill and bicycle tests for max VO2. Medicine and Science in Sports; 5(3):156-160.

Mitchell, J.H., B.J. Sproule, & C.B. Chapman. (1958). The physiological meaning of the maximal oxygen intake test. Journal of Clinical Investigation; 37:538-547.

Morgan, K., A.O. Hughes, and R. Philipp. (1984). Reliability of a test of cardiovascular fitness. International Journal of Epidemiology; 13:32-37.

Rowell, L.B., H.L. Taylor and Y. Wang. (1964). Limitations to prediction of maximal oxygen intake. Journal of Applied Physiology; 19(5):919-927.

Sady, S.P., M.P. Savage, W.H. Thomson, & M.M. Petratis. (1983). The reliability of the VO_2-HR relation during graded treadmill exercise in prepubertal boys and adult men. Research Quarterly for Exercise and Sport; 54(3):302-304.

Saltin, B. (1985). Hemodynamic adaptations to exercise. American Journal of Cardiology; 55:42D-47D.

Shephard, R.J. (1970). Computer programs for solution of the Astrand nomogram and the calculation of body surface area. Journal of Sports Medicine and Physical Fitness; 10:206-210.

Shephard, R.J. (1981). Present views on the Canadian Home Fitness Test. Canadian Medical Association Journal; 124:875-879.

Shephard, R.J., D.A. Bailey, and R.L. Mirwald. (1976). Development of the Canadian Home Fitness Test. Canadian Medical Association Journal; 114:675-679.

Shephard, R.J., C. Allen, A.J.S. Benade, C.T.M. Davies, P.E. di Prampero, R. Hedman, J.E. Merriman, K. Myhre, and R. Simmons. (1968).Standardization of submaximal exercise tests.Bulletin of the World Health Organization; 38:765-775.

Shephard, R.J., C. Allen, A.J.S. Benade, C.T.M. Davies, P.E. di Prampero, R. Hedman, J.E. Merriman, K. Myhre, and R. Simmons. (1968). The maximum oxygen intake. An international reference standard of cardiorespiratory fitness. Bulletin of the World Health Organization; 38:757-764.

Taylor, H.L., E. Buskirk, & A. Henschel. (1955). Maximal oxygen intake as an objective measure of cardio-respiratory performance. Journal of Applied Physiology; 8:73-80.

Thoden, J.S., B.A. Wilson, and J.D. MacDougall. (1982). Testing Aerobic Power. In Physiological Testing of the Elite Athlete. J.D. MacDougall, H.A. Wenger, H.J. Green, eds. Canadian association of Sport Sciences.

Whipp, B.J., J.A. Davis, F.Torres, & K. Wasserman. (1981). A test to determine parameters of aerobic function during exercise. Journal of Applied Physiology; 50(1):217-221.

Wilmore, J.H., J.A. Davis, and A.C. Norton. (1976). An automated system for assessing metabolic and respiratory function during exercise. Journal of Applied Physiology; 40(4):619-624.

Wyndham, C.H., N.B. Strydom, J.S. Maritz, J.F. Morrison, J. Peter and Z.U. Potgieter. Maximum oxygen intake and maximum heart rate during strenuous work. Journal of Applied Physiology; 14(6):927-936.

APPENDIX 2

INFORMED CONSENT FOR EXERCISE TESTING AND TRAINING

Predictive Testing Study

Investigators: K.R. Ross, S.R. Petersen, Ph.D. and H.A. Quinney, Ph.D.

The purpose of th[illegible] [illegible]y is to determine whether some commonly used p[illegible] [illegible]ve tests accurately reflect changes in aerob[illegible] [illegible] resulting from training.

The following procedures will be carried out prior to the training program (twice) and after eight weeks of training.

1. Maximal Aerobic Power Test.
   Following a warm-up, subjects will pedal a cycle ergometer at progressively increasing workrates up to voluntary exhaustion. Subjects will breathe through a mouthpiece which directs expired air to a gas analysis instrument. Heart rate will be measured using an ECG.

2. Canadian Aerobic Fitness Test.
   Subjects will step up and down a double step in 3 minute stages at an age and sex specific rate. This rate increases every stage until either 5 stages have been completed or heart rate exceeds the limit for continuing.

3. Astrand Submaximal Cycle Test.
   Following a brief warm-up, subjects will cycle at submaximal work rates while heart rate is measured on an ECG.

4. Stationary cycle training.
   Subjects will train for eight weeks, three sessions per week (on different days), thirty minutes per session on a stationary bike at an individually prescribed exercise rate based on maintaining a heart rate of 150 beats per minute.

I agree to take part in this study. The test procedures have been described to me in detail and I understand that some of the tests are strenuous and require maximal effort. I have had an opportunity to ask questions about these procedures, including any risks of taking part in such tests. My participation is voluntary, and I understand that I am free to refuse any procedure or to withdraw from the study at any time.

For my part, I will participate in the study to my best ability. If I am assigned to the training group, I will train on the assigned days for the duration of the study; if I am assigned to the control group I will not train during the study. I have completed a Physical Activity Readiness Questionnaire and I will inform the investigators of any changes in my health status.

I consent to the use of my results for the study and for publication purposes. I understand that the investigators will collect, keep and publish data in a manner which assures confidentiality and respects my privacy.

________________________ ________________________

Name (print) Address

________________________ ________________________

Subject's Signature Phone

________________________ ________________________

Investigator's signature Date

NAME

DATE

PAR Q & YOU

PAR-Q is designed to help you help yourself. Many health benefits are associated with regular exercise, and the completion of PAR-Q is a sensible first step to take if you are planning to increase the amount of physical activity in your life

For most people physical activity should not pose any problem or hazard. PAR-Q has been designed to identify the small number of adults for whom physical activity might be inappropriate or those who should have medical advice concerning the type of activity most suitable for them

Common sense is your best guide in answering these few questions. Please read them carefully and check (√) the ☐ YES or ☐ NO opposite the question if it applies to you

YES	NO	
☐	☐	1 Has your doctor ever said you have heart trouble?
☐	☐	2 Do you frequently have pains in your heart and chest?
☐	☐	3 Do you often feel faint or have spells of severe dizziness?
☐	☐	4 Has a doctor ever said your blood pressure was too high?
☐	☐	5 Has your doctor ever told you that you have a bone or joint problem such as arthritis that has been aggravated by exercise, or might be made worse with exercise?
☐	☐	6. Is there a good physical reason not mentioned here why you should not follow an activity program even if you wanted to?
☐	☐	7 Are you over age 65 and not accustomed to vigorous exercise?

If You Answered

YES to one or more questions

If you have not recently done so, consult with your personal physician by telephone or in person BEFORE increasing your physical activity and/or taking a fitness appraisal. Tell your physician what questions you answered YES to on PAR-Q or present your PAR-Q copy

programs

After medical evaluation, seek advice from your physician as to your suitability for:

- unrestricted physical activity starting off easily and progressing gradually.
- restricted or supervised activity to meet your specific needs, at least on an initial basis. Check in your community for special programs or services

NO to all questions

If you answered PAR-Q accurately, you have reasonable assurance of your present suitability for

- A GRADUATED EXERCISE PROGRAM — a gradual increase in proper exercise promotes good fitness development while minimizing or eliminating discomfort
- A FITNESS APPRAISAL — the Canadian Standardized Test of Fitness (CSTF)

postpone

If you have a temporary minor illness, such as a common cold

* Developed by the British Columbia Ministry of Health. Conceptualized and critiqued by the Multidisciplinary Advisory Board on Exercise (MABE). Translation, reproduction and use in its entirety is encouraged. Modifications by written permission only. Not to be used for commercial advertising in order to solicit business from the public.
Reference: PAR-Q Validation Report, British Columbia Ministry of Health, 1978
* Produced by the British Columbia Ministry of Health and the Department of National Health & Welfare