MONTANA STATE UNIVERSITY

Validity of a stationary cycling protocol for tracking changes in uphill cycling time-trial performance
by Owen Frederic Murphy

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Health and Human Development
Montana State University
© Copyright by Owen Frederic Murphy (2001)

Abstract:
The ability to accurately quantify and predict endurance performance is imperative when assessing the effects of training interventions, dietary regimes, equipment modifications, and/or alterations in athlete position or technique. Due to the invasive nature and excessive costs of traditional physiological and biomechanical determinants of endurance performance, performance-based determinants of performance have received considerable attention.

The purpose of the present study was to determine the reliability and validity of a scaling-derived cycle ergometer protocol (SDP) for tracking changes in uphill time-trial (TT) cycling performance. Phase I of the study assessed the reliability of the Scaling Derived Protocol (SDP) via the administration of three SDP protocols within a ten-day time period. Phase II of the study determined the ability of the SDP to track longitudinal changes in uphill time-trial performance. Local competitive cyclists participated in either two or three testing periods separated by a minimum of ten weeks (May, July, September 2001). Each testing period consisted of an outdoor uphill TT (5-km, 8% grade) followed within ten days by a laboratory-based SDP. Longitudinal inter-trial changes in average TT speed (m/s) were compared with longitudinal inter-trial changes in SDP time-to-exhaustion (TTE, min) and relative SDP maximal power output (Wmax, watts kg^-1).

There were no significant Phase I inter-trial differences for time-to-exhaustion, maximum power output, maximum heart rate, and relative VO2max. Intraclass correlation coefficients were high for all variables, ranging from R = 0.933 to 0.992. Single-score reliability was also high, with correlation coefficients ranging from R(k=i)= 0.823 to 0.977.

High correlations were also observed between Phase II inter-trial changes in average TT speed and relative WMax (r = 0.70 to 0.94). Correlations between inter-trial changes in average TT speed and TTE were generally lower and more variable (r = 0.26 to 0.87). These results suggest that changes in relative SDP Wmax can reliably, validly, and practically track changes in uphill TT performance.

VALIDITY OF A STATIONARY CYCLING PROTOCOL FOR TRACKING CHANGES IN UPHILL CYCLING TIME-TRIAL PERFORMANCE

by

Owen Frederic Murphy

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Health and Human Development

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2001

N378
M9559

APPROVAL

of a thesis submitted by

Owen Frederic Murphy

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

Dr. Deborah King — Deborah L. King (Signature) — 11/19/01 Date

Approved for the Department of Health and Human Development

Dr. Ellen Kreighbaum — Ellen Kreighbaum (Signature) — 11/19/01 Date

Approved for the College of Graduate Studies

Dr. Bruce R. McLeod — Bruce R. McLeod (Signature) — 11-29-01 Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment for the requirements for a master's degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library.

If I have indicated my intention to copyright this thesis by including a copyright notice page, copying is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for permission for extended quotation from or reproduction of this thesis in whole or parts may be granted only by the copyright holder.

Signature Owen Murphy

Date NOVEMBER 19, 2001

ACKNOWLEDGEMENTS

I wish to express my appreciation to Drs. Deborah King and Dan Heil for their time, expertise, and inspiration. Thanks also to the poor saps who unwittingly volunteered to take part in my study – may you never have to do another max test again! My family and friends were very supportive throughout this process and I am indebted to them all. I am especially grateful for the patience and unconditional love of Natalie and Isabelle, thank you both.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The ability to accurately quantify and predict endurance performance is imperative when assessing the effects of training interventions, dietary regimes, equipment modifications, and/or alterations in athlete position or technique. Due to the invasive nature and excessive costs of traditional physiological and biomechanical determinants of endurance performance, performance-based determinants of performance have received considerable attention.

The purpose of the present study was to determine the reliability and validity of a scaling-derived cycle ergometer protocol (SDP) for tracking changes in uphill time-trial (TT) cycling performance. Phase I of the study assessed the reliability of the Scaling Derived Protocol (SDP) via the administration of three SDP protocols within a ten-day time period. Phase II of the study determined the ability of the SDP to track longitudinal changes in uphill time-trial performance. Local competitive cyclists participated in either two or three testing periods separated by a minimum of ten weeks (May, July, September 2001). Each testing period consisted of an outdoor uphill TT (5-km, 8% grade) followed within ten days by a laboratory-based SDP. Longitudinal inter-trial changes in average TT speed (m/s) were compared with longitudinal inter-trial changes in SDP time-to-exhaustion (TTE, min) and relative SDP maximal power output (W_{MAX}, watts·kg^{-1}).

There were no significant Phase I inter-trial differences for time-to-exhaustion, maximum power output, maximum heart rate, and relative VO_{2MAX}. Intraclass correlation coefficients were high for all variables, ranging from R = 0.933 to 0.992. Single-score reliability was also high, with correlation coefficients ranging from $R_{(k=1)}$ = 0.823 to 0.977.

High correlations were also observed between Phase II inter-trial changes in average TT speed and relative W_{MAX} (r = 0.70 to 0.94). Correlations between inter-trial changes in average TT speed and TTE were generally lower and more variable (r = 0.26 to 0.87). These results suggest that changes in relative SDP W_{MAX} can reliably, validly, and practically track changes in uphill TT performance.

CHAPTER ONE

INTRODUCTION

The year 2000 saw the rise to dominance some of the world's greatest cyclists. Johan Museeuw celebrated his third victory at Paris-Roubaix, soloing the last 40-km in a stunning display of brute strength and courage. Lance Armstrong won his second consecutive Tour de France, dominating the high-mountain stages and individual time-trials. Leontien Zijlaard was the revelation of the Sidney Olympic Games, winning an unprecedented three gold medals in women's road and track cycling events. Finally, Miguel Martinez became the first mountain biker in history to achieve the overall World Cup, Olympic gold, and the World Championships in the same year.

Such incredible performances (time and time again) beg the question: What sets these champions apart? Are they gifted with extraordinary lung capacities, lactate thresholds, or blood hematocrits? Do they possess the ideal strength-to-weight ratio? Have they developed superior technique or economy of movement? Do they command a greater strength of mind, will to suffer, or desire to win?

In truth, all of these factors contribute to successful endurance performance – cycling or other. Sport scientists today, however, are searching for a *single, easily quantifiable* determinant of performance that could be used to identify, track, and train the best endurance athletes in the world. More importantly, perhaps, coaches at the grassroots level are in need of a practical and affordable method with which to identify youth talent in a variety of endurance-based sports. Only then will it be possible to nurture the youths of today into the champions of tomorrow.

Historical Background

In 1923, Hill and Lupton were the first sport scientists to note that the world's best endurance athletes also happened to possess the world's highest $\dot{V}O_{2MAX}$ values when expressed relative to their body mass. It was subsequently proposed (and widely accepted) that relative $\dot{V}O_{2MAX}$ could be used as a primary determinant of endurance performance (Astrand and Rodahl, 1986; Costill, 1967; Foster, 1978; Hagberg and Coyle, 1983; Saltin and Astrand, 1967). More recently, physiologists have also proposed blood lactate transition thresholds (Farrell et al., 1979; Kumagai et al., 1982; Powers et al., 1983; Yoshida et al., 1987) and economy of movement (Basset and Howley, 2000; Costill and Winrow, 1970; Daniels, 1974; Dengel et al., 1989) as viable determinants of endurance performance.

Despite their strong association with successful endurance performance, however, physiological determinants ($\dot{V}O_{2MAX}$, lactate threshold, and economy of movement) are inherently expensive, invasive, and unavailable to the general public. "Indeed, few coaches and athletes have the ready access to laboratory facilities capable of measuring $\dot{V}O_{2MAX}$, as such a test requires expensive equipment, skilled personnel, and is time-consuming" (Hawley and Noakes, 1992, pp. 79). Given these drawbacks, physiological determinants of endurance performance are rendered impractical means with which to screen youth talent and/or repeatedly monitor an endurance athlete's progress.

Fortunately, several contemporary cycling-related studies have indicated the importance of average or peak muscular power as a valid and relatively affordable determinant of endurance performance. In such investigations, power-measuring devices (e.g. Computrainer™, SRM™, Tune Power Tap™) are used to measure a cyclist's power output in watts. The results have been stunning. For instance, Palmer et al. (1996), observed a correlation of r = -0.999 between average power output and both 20-km and 40-km cycling time-trial performance! Bishop et al. (1998) and Coyle et al. (1991) reported correlations of r = -0.81 and r = -0.88, respectively, between average power output and actual 40-km time-trial performance. In a study of sub-elite male cyclists, Balmer et al. (2000) found a correlation of r = -0.99 ($p < 0.001$) between peak power output (PPO) and 16.1-km time-trial performance. Lastly, Bishop et al. (1998), Lindsay et al. (1996), and Schabort et al. (2000) observed high correlations (r = -0.81 to -0.91) between PPO and 20-km to 40-km cycling time-trial performance.

In an effort to improve upon the validity, reliability, and theoretical foundation of PPO protocols, Heil et al. (2000) introduced a scaling derived protocol (SDP) based on the energy requirements of uphill cycling as influenced by body mass. Preliminary results indicate that SDP performance (time-to-exhaustion) correlates extremely well with actual uphill time-trial performance (r = 0.97 to 0.98), as does relative SDP peak power output (watts·kg^{-1}; r = 0.92 to 0.97). The SDP is easy to administer and completely objective because gravity affects all cyclists in the same manner, regardless of age, gender, or technique. In addition, the design of the SDP allows it to serve as a $\dot{V}O_{2MAX}$ test for cyclists when appropriate.

Given the impressive correlations between SDP performance and uphill time-trial performance, the current investigation was undertaken to determine the ability of the SDP to track *changes* in uphill time-trial performance. Though such ability may seem a rather simple and straightforward proficiency, the author knows of no existing protocol so apt.

Statement of Purpose

The primary purpose of this study was to determine the ability of the Scaling Derived Protocol (SDP) to validly track changes in uphill cycling time-trial performance. A secondary purpose of this study was to assess the test-retest reliability of the SDP.

Significance of Study

The results from this study will be of value to cyclists, coaches, and sport scientists as they attempt assess the effects of various training interventions, dietary regimes, equipment modifications, and/or alterations in athlete position and technique. This should ultimately lead to more effective cycling training. The portability, affordability, and practicality of the SDP also bodes well for local youth talent identification programs for cycling. Additionally, the potential exists to extend the principles of the SDP to other endurance-based sports such as running and cross-country skiing.

Hypotheses

1. Phase I: The Scaling Derived Protocol (SDP) is a reliable measure of cycling performance. Intraclass correlation coefficients (ICC) were assessed via the administration of three SDP protocols to local cyclists within a 10-day period.

H_o: ICC < 0.80

H_a: ICC ≥ 0.80

2. Phase II: The Scaling Derived Protocol (SDP) is a valid means of tracking changes in uphill cycling time-trial performance (TT).

H_o: $\rho = 0$ (No correlation exists between changes in SDP performance and changes in TT performance)

H_a: $\rho \neq 0$ (Correlation exists between changes in SDP performance and TT performance)

Limitations

1) Phase II: The time-trial course was three miles long and featured a 500-meter section of level grade approximately two-thirds of the way to the finish. Considering the available course options in the Bozeman area, however, the chosen venue was the most suitable to the purposes of this study.

2) Phase II: Weather conditions (namely wind and air temperature) varied between each of the three uphill time trials (May, July, and, September 2001).

3) Phase II: Weight categories (and consequent SDP resistance levels) were fixed for the entire study period, regardless of fluctuations in subjects' weight.

Delimitations

1) Phase I/II: The scope of this study was delimited to competitive and recreational cyclists in the Bozeman, Montana area.

2) Phase II: Due to health-related contraindications to maximal intensity exercise, this study was delimited to "low and moderate risk individuals" (as defined by ACSM's Guidelines for Exercise Testing and Prescription, 6th ed., 2000). In the the current study, "moderate risk individuals" refers to several male subjects over the age of 45 years.

Operational Definitions

Cadence:	The rate of pedaling, measured in revolutions per minute of one foot, e.g. pedal frequency or pedaling rate (Borysewicz, 1985).
Change-score:	An index of Phase II inter-trial change calculated by subtracting the performance score of one trial (pre) from that of another (post); the absolute difference between trials.
Computrainer™:	A commercially available stationary trainer that allows the application of variable resistance (in watts) to the rear wheel of a bicycle.
Economy of Movement:	The oxygen cost of exercising at a standard, submaximal velocity or power output.
Endurance Performance:	The ability to maintain a relatively high intensity of exercise (>80% max) for extended periods of time (>1 hour).
Lactate Threshold:	The balance point between lactate entry and removal from the blood.
Maximal Oxygen Consumption:	The highest rate at which ambient oxygen can be taken up and utilized by the body during maximal intensity exercise.
Power Output:	Watts produced per kilogram of body mass on a cycle ergometer; measured at the crank, rear hub, or tire/trainer interface.
Singe-Score Reliability:	The reliability of a performance protocol given only one test administration.
Time-Trial:	A race against the clock in which individual riders start at set intervals and cannot give aid or receive it from others on the course.

CHAPTER TWO

REVIEW OF LITERATURE

Introduction

The ability to accurately quantify and predict endurance performance has become an increasingly important focus in recent decades. Ever since Hill and Lupton (1923) demonstrated a positive correlation between maximal oxygen uptake ($\dot{V}O_{2MAX}$) and middle-distance running performance, researchers have sought to identify additional physiological and biomechanical factors associated with success in endurance events. Commonly referred to as "determinants of endurance performance", these factors are used to evaluate the effects of training strategies (Hickson et al., 1988; Lindsay et al., 1996; Westgarth-Taylor et al., 1997; Weston, et al., 1997), nutritional interventions (Coggan and Coyle, 1987; El-Sayed et al., 1997; Palmer et al., 1998), alterations in body position or bicycle geometry (Heil, 1997; Heil et al., 1997; Heil et al., 1995), and equipment changes (MacRae et al., 2000), as well as predict the performance capacity of endurance athletes (Balmer et al., 2000; Hawley and Noakes, 1992; Heil et al., 2000; Kuipers et al., 1985; Miller and Manfredi, 1987).

However, due to the inability of a single physiological or biomechanical factor to determine and/or predict endurance performance in homogeneous populations, a number of researchers have proposed the use of performance-based laboratory protocols. This review will examine the most common physiological and performance-based determinants of endurance performance, with particular respect to the validity and/or practicality of each. Other factors related to endurance performance, such as biomechanics, psychological preparation, and nutritional status will not be discussed.

Physiological Determinants of Endurance Performance

Maximal oxygen consumption ($\dot{V}O_{2MAX}$) has traditionally been used as *the* criterion measure of endurance performance (Astrand and Rodahl, 1986; Bassett and Howley, 2000; Costill, 1967; Coyle et al., 1988; Craig, et al., 1993; Foster, 1978; Hagberg and Coyle, 1983; Hill and Lupton, 1923; Noakes, 1988; Saltin and Astrand, 1967; Sleivert and Rowlands, 1996). However, several investigators have recently suggested that physiological parameters measured during submaximal exercise may provide better predictors of endurance performance than $\dot{V}O_{2MAX}$. These parameters include blood lactate transition thresholds (Barlow et al., 1985; Coyle, et al., 1991; Craig, et al., 1993; Farrell, et al. 1979; Hagberg and Coyle, 1983; Hopkins and McKenzie, 1994; Jacobs, 1986; Loftin and Warren, 1994; Miller and Manfredi, 1987; Powers et al., 1983; Schabort et al., 2000; Sjodin, 1981; Yoshida et al., 1987) and economy of movement (Conley and Krahenbahl,1980; Costill et al., 1985; Powers et al.,1983; Schabort et al., 2000).

Maximal Oxygen Uptake

Maximal oxygen uptake ($\dot{V}O_{2MAX}$) is considered a primary determinant of endurance performance, generally because it sets the upper limit for steady-state oxygen consumption (Astrand and Rodahl, 1986; Costill, 1967; Craig, et al., 1993; Evans et al., 1995; Foster, 1978; Hill and Lupton, 1923). It is defined as the highest rate at which ambient oxygen can be taken up and utilized by the body during maximal intensity exercise. Provided the absence of pulmonary disease, $\dot{V}O_{2MAX}$ is the best indicator of an individual's cardiovascular fitness level (Brooks et al, 1996).

Support for the use of $\dot{V}O_{2MAX}$ as a determinant of endurance performance is based on the fact that most endurance athletes have relatively high $\dot{V}O_{2MAX}$ values (Costill et al., 1973; Coyle, et al., 1988, 1991; Saltin and Astrand, 1967; Sleivert and Rowlands, 1996), and relatively strong correlations have been reported between $\dot{V}O_{2MAX}$ and endurance performance (Costill, 1967; Farrell et al., 1979; Foster et al., 1978; Sleivert and Rowlands, 1996). In contrast, a number of researchers have found that $\dot{V}O_{2MAX}$ is inadequate at predicting endurance performance within a homogeneous group of well-trained endurance athletes (Conley and Krahenbuhl, 1980; Coyle et al., 1988; Hagberg et al., 1979; Lindsey et al., 1996). "This [paradox] seems to indicate that a high $\dot{V}O_{2MAX}$ capacity is necessary to compete in endurance events, but $\dot{V}O_{2MAX}$ alone cannot be used to discriminate between highly trained endurance athletes of similar $\dot{V}O_{2MAX}$" (Barlow et al., 1985, pp. 194). In such cases, lactate threshold, economy, and/or performance during specific laboratory protocols have proven better indicators of endurance performance than $\dot{V}O_{2MAX}$ (Allen et al., 1985; Bishop et al, 1998; Conconi et al., 1982; Conley and Krahenbuhl, 1980; Craig et al., 1993; Farrell et al., 1979; Heil et al., 2000; Weltman et al., 1990).

Lactate Threshold

Lactate threshold (LT) and/or anaerobic threshold is defined as the balance point between lactate entry and removal from the blood. It is often represented by an athlete's $\dot{V}O_2$ (ml/kg/min) at LT, or their corresponding power output on a cycle ergometer (e.g. 300 watts). The use of LT as a determinant of endurance performance is supported by the following observations: 1) during endurance competition most athletes perform at some percentage of their $\dot{V}O_{2MAX}$ rather than at their absolute maximum (Costill and Fox, 1969; Coyle et al., 1988; Farrell et al., 1979); 2) athletes with a relatively higher LT exhibit superior endurance performance (Coyle et al., 1988; Farrell et al., 1979; Schabort et al., 2000; Sjodin and Jacobs, 1982); 3) a delayed LT should result in decreased glycogen depletion and increased time-to-exhaustion since elevated lactate levels are associated with increased glycogen utilization (a recognized limitation to endurance performance), (Costill et al., 1971a, 1971b; Coyle, 1988). Moreover, in studies that have correlated LT and $\dot{V}O_{2MAX}$ with endurance performance, LT has been more strongly related, both in trained subjects (Farrell et al., 1979; Kumagai et al., 1982; Powers et al., 1983) and untrained subjects (Yoshida et al., 1987).

However, the concept and application of lactate and/or anaerobic thresholds are the targets of considerable debate. Despite 60 years of blood lactate research, there is still a great deal of controversy surrounding the explanation of lactate and/or anaerobic thresholds and the methods used to identify them. In fact, points of contention probably outnumber points of agreement (Bourdon, 2000).

The principal point of contention involves the method by which blood lactate transition thresholds are defined – fixed blood lactate concentrations or individualized lactate thresholds. Fixed blood lactate concentrations ranging from 2 to 4 mmol/L have been used by a number of investigators in order to minimize problems associated with detecting inflections in the blood lactate response curve (Allen et al, 1985; Hagberg et al, 1986; Heck et al., 1985; Kindermann et al., 1979; Sjodin, 1983). However, fixed blood lactate concentrations may be influenced by an individual's nutritional and training status. For example, $\dot{V}O_2$ obtained at the onset of blood lactate accumulation was significantly lower after a high carbohydrate diet than after a low carbohydrate diet (Yoshida, 1984).

Stegmann et al. (1981) reported that steady state blood lactate concentrations can vary widely among athletes, thus lending additional support to the use of individualized lactate thresholds. However, the detection of inflection points in individualized blood lactate curves is a complicated and controversial matter in and of its own. Several interpretational methods exist, including log-log transformations (Beaver, et al, 1985), rates of metabolite accumulation (Thoden, 1991), tangential techniques (Cheng et al., 1992), and subjective visual inspections (Yoshida et al., 1985), but none has proven overly superior.

A second major point of contention surrounding the concept of "lactate threshold" addresses the protocol(s) used to elicit blood lactate responses, collect blood samples, and analyze lactate concentrations. "Protocol-related factors such as the sampling site, workload duration, continuous versus discontinuous exercise bouts, and choice of ergometer can all affect the measurement of blood lactate response to incremental exercise" (Bourdon, 2000, pp. 56). The longer the workload duration, for instance, the lower the LT (Heck et al., 1985; McLellan 1987). In contrast, the longer the rest breaks during a discontinuous protocol, the higher the LT (Heck et al., 1985; Foster et al., 1995). A complete summary of the exercise protocols and sampling/analysis methods used to evaluate blood lactate concentrations is beyond the scope of this paper, but suffice it to say that a significant number of procedural, physiological, and environmental factors exist with the potential to alter the blood lactate-exercise intensity relationship.

Economy of Movement

There is a distinct linear relationship between submaximal power output and $\dot{V}O_2$ for each individual. However, considerable variation exists in just how much oxygen is required to exercise at a given intensity or power output. Several researchers have proposed this $\dot{V}O_2$ discrepancy can be attributed to inter-individual differences in economy of movement (Basset and Howley, 2000; Costill and Winrow, 1970; Daniels, 1974).

Economy of movement is defined as the oxygen cost (ml/kg/min) of exercising at a standard submaximal velocity or power output (Bassett and Howley, 2000; Farrell et al., 1979; Powers et al., 1983; Schabort et al., 2000). Intuitively, better economy results in better performance: a more economical athlete uses less oxygen at a standard velocity and theoretically is able to move faster and/or conserve more energy than a less economical athlete (Sleivert and Rowlands, 1996).

Research concerning the use of economy as a determinant of endurance performance first appeared in the literature during the early 1970s. Costill and Winrow (1970) suggested that variations in the performance of two middle-aged ultramarathon runners with similar $\dot{V}O_{2MAX}$ values could be attributed to individual differences in economy. Several years later, Daniels (1974) reported that economy of movement was responsible for nearly identical 2-mile run times (10:31 and 10:35) among two champion male runners with significantly different $\dot{V}O_{2MAX}$ values (> 10 ml/kg/min). In each case, the more economical runner was able to perform at a lower percentage of his or her relative $\dot{V}O_{2MAX}$, thus decreasing the power of $\dot{V}O_{2MAX}$ to predict overall performance.

Additionally, Conley and Krahenbahl (1980) observed a correlation of $r = 0.82$ between economy of movement and 10-km running performance in a study of highly trained male distance runners with similar $\dot{V}O_{2MAX}$ values. These data were later used by Basset and Howley (1997) to explain how running economy could account for a large portion of the variability in distance running performance. In other sports, Dengel et al. (1989) reported that the strongest predictor of triathlon performance was a combined measure of swimming, cycling, and running economy.

In direct contrast to the above researchers, Farrell et al. (1979) and Powers et al. (1983) reported a significant *lack* of correlation between economy and endurance performance. Working independently, Farrell and Powers reported poor correlations ($r = 0.49$ to 0.59) between running economy and distance running performance in homogeneous well-trained populations. "The failure to show a significant relationship between running economy and running performance… suggests that in some populations the individual differences in running economy at a standardized speed is not great, and that running economy may be of limited value in differentiating distance running performance" (Powers et al., 1983, pp. 182).

Noakes (1998) is another prominent physiologist hesitant to endorse economy of movement as a primary determinant of endurance performance. In rebuttal to the aforementioned claims presented by Basset and Howley (1997), Noakes cites large discrepancies in their study population – individual performances differed by 18 minutes (38%) in a 16-km race, and by 9 minutes (20%) in a 10-km race! However, "…in studies of athletes whose performances are more similar, neither $\dot{V}O_{2MAX}$ nor economy is a good predictor of running performance" (Noakes, 1998, pp.1392).

Performance-Based Protocols as Determinants of Endurance Performance

The primary purpose of performance-based protocols is to simulate the demands of real world competition in a controlled environment. Ideally, these protocols should be reliable, valid, and sensitive to small changes in an athlete's fitness level. These qualities are especially important when performing repeated measurements over a period of time, or when laboratory data is applied to the field to predict actual performance (Balmer, Davison, and Bird, 2000).

Traditional performance-based endurance protocols include time-to-exhaustion at a fixed workload or percentage of $\dot{V}O_{2MAX}$ (Acevedo and Goldfarb, 1989; Billat et al., 1988; Burnham et al., 1995; Coggan and Coyle, 1987; Coyle et al., 1988; Hickson et al., 1988; Jeukendrup et al., 1996; McLellan et al., 1995), time to completion of a fixed workload or distance (Acevedo and Goldfarb, 1989; Coyle et al., 1991; Farrell et al., 1979; Foster et al., 1993; Hickey et al., 1992; Jeukendrup et al., 1996; Lindsay et al., 1996; Loftin and Warren, 1994; Nichols et al., 1997; Palmer et al., 1995; Schabort et al., 1998; Williams and Cavanagh, 1987), average power output or power output at lactate threshold (Bishop, 1997; Bishop et al., 1998; Coyle et al., 1991; Davison et al., 2000; Hopkins and McKenzie, 1994; Nichols et al., 1997; Palmer et al., 1996; Smith et al., 1999), and peak power achieved during exercise (Balmer et al., 2000; Bishop et al., 1998; Hawley and Noakes, 1992; Lindsay et al., 1996; Morgan et al., 1989; Noakes, 1990; Scott and Houmard, 1994; Schabort et al., 2000).

Time-to-Exhaustion at a Fixed Submaximal Workload

A continuous protocol at a submaximal workload until volitional exhaustion is the classical and most common determinants of endurance performance (Jeukendrup et al., 1996). This type of protocol is usually performed at exercise intensities ranging from 60% to 90% of $\dot{V}O_{2MAX}$, and has reportedly correlated well with endurance performance (Acevedo and Goldfarb, 1989; Billat et al., 1988; Burnham et al., 1995; Coggan and Coyle, 1987; Coyle et al., 1988; Hickson et al., 1988). However, a growing number of researchers have indicated this genre of protocol to be highly variable and unreliable (Jeukendrup et al., 1996; Krebs and Powers, 1989; McLellan et al., 1995).

For example, Krebs and Powers (1989) and McLellan et al. (1995) independently determined coefficients of variation (CV) ranging from 2.8% to 55.9% in untrained subjects cycling to fatigue at 80% $\dot{V}O_{2MAX}$, indicating the irreproducibility of submaximal tests to exhaustion. Jeukendrup et al. (1996) reported a CV ranging from 17-40% in well-trained cyclists riding to fatigue at approximately 85% $\dot{V}O_{2MAX}$. Additionally, Billat et al (1994) reported a CV of 25% when "sub-elite" subjects were asked to run to exhaustion at their relative lactate threshold.

The poor reproducibility of time-to-exhaustion (TTE) protocols is widely attributed to psychological factors. "Because [TTE] tests have an 'open end', psychological factors that might influence performance, such as motivation, monotony, and boredom, are more pronounced than in the time trial oriented approach where the 'endpoint is known'" (Jeukendrup et al., 1996, p. 269). TTE protocols are further criticized because they are time-consuming and fail to mimic normal competitive situations (Schabort et al., 1996).

Time to Completion of a Fixed Workload or Distance

In comparison to TTE protocols, endurance performance protocols in which athletes are asked to complete a fixed amount of work or distance in the shortest possible time are more reliable (Bishop, 1997; Hickey et al., 1992; Jeukendrup et al., 1996; Palmer et al., 1995; Schabort et al., 1996). Palmer et al. (1995) and Hickey et al. (1992), determined a CV ranging from only 1.0% to 3.1% for cyclists performing 20 to 40-km and 5 to 40-mile time-trials, respectively. Similarly, Bishop (1997) and Jeukendrup et al. (1996) determined a CV of 2.7% and 3.4%, respectively, when subjects were asked to complete as much work as possible in 60 minutes of cycling.

As well as being reliable, protocols of fixed duration or workload have also correlated well with actual endurance performance (Coyle et al., 1991; Palmer et al., 1992). In a study of elite male cyclists, Coyle et al. (1991) reported a correlation of $r = 0.93$ between time to completion of a 40-km laboratory time-trial and actual 40-km time-trial performance. Likewise, Palmer et al. (1995) reported a correlation of $r = 0.98$ between 20-km and 40-km laboratory time-trials and actual road racing performance.

Similar to TTE protocols, however, fixed duration/workload protocols are inherently time-consuming. For example, a 40-km time-trial may take up to 1.5 hours to complete for a sub-elite cyclist. This fact alone renders fixed duration/workload protocols impractical for situations that necessitate the testing of multiple subjects. Also, due to their time-consuming nature, prolonged fixed duration/workload protocols suffer the same negative psychological factors that affect TTE protocols (Hickey et al., 1992).

Measures of Power

Of all the existing determinants of endurance performance, none have proven so overwhelmingly reliable and valid as measures of power output ($watts \cdot kg^{-1}$). Power is, in essence, a manifestation of all possible physiological and biomechanical determinants of endurance performance. By definition, power is the capacity to perform a given amount of work in the shortest possible amount of time. Although power-measuring devices used to be relatively expensive, recent advances in technology have made the quantification of power (in watts, W) much more feasible for coaches and athletes alike.

It is important to note the existence of two different methods of quantifying power output: 1) average power output and 2) peak power output. The first studies to explore the relationship between power and endurance performance relied on measures of average power output (Bishop et al., 1998; Coyle et al., 1991; Davison et al., 2000; Hopkins and McKenzie, 1994; Nichols et al., 1997; Palmer et al., 1996; Smith et al., 1999). Incredibly, Palmer et al. (1996) observed a near perfect correlation ($r = -0.999$, $p < 0.001$) between average power output and both 20-km and 40-km cycling time-trial performance. Bishop et al. (1998) and Coyle et al. (1991) reported correlations of $r = -0.81$ and $r = -0.88$, respectively, between average power output during a one-hour laboratory time-trial and actual 40-km time-trial performance among experienced male and female cyclists and triathletes. Finally, Nichols et al. (1997) and Smith et al. (1999) independently reported correlations of $r = -0.77$ to -0.91 between power output at LT and 13.5-km to 20-km cycling time-trial performance.

More recent studies of power have involved the measurement of peak power output (PPO) (Balmer et al., 2000; Bishop et al., 1998; Hawley and Noakes, 1992; Heil et al., 2000; Lindsay et al., 1996; Morgan et al., 1989; Noakes, 1990; Scott and Houmard, 1994; Schabort et al., 2000). In a study of sub-elite male cyclists, Balmer et al. (2000) reported a correlation of $r = -0.99$ ($p < 0.001$) between PPO and 16.1-km time-trial performance. Bishop et al. (1998), Lindsay et al. (1996), and Schabort et al. (2000) also observed high correlations ($r = -0.81$ to -0.91) between PPO and 20-km to 40-km cycling time-trial performance. In runners, the peak treadmill velocity achieved during an incremental test to exhaustion (a reciprocal measure of PPO) has proven to be a strong indicator of endurance performance as well (Morgan et al., 1989; Noakes, 1990; Scott and Houmard, 1994).

A new PPO protocol has been proposed by Heil et al. (2000). Based on the external power demand (W_D, W) of steep uphill cycling, Heil and coworkers developed a scaling derived cycling protocol (SDP) that correlates extremely well with actual uphill cycling time-trial performance ($r = 0.97$ to 0.98). The SDP "...utilizes findings from both correlational and body mass scaling studies in cycling to derive an ergometer test protocol that [can] accurately predict uphill cycling time-trial performance" (Heil et al., 2000, p. 3).

In general, W_D for outdoor cycling is considered to be the vector sum of rolling resistance (R_R), aerodynamic drag (R_D), and gravitational resistance (R_G) (Bassett et al., 1999, di Prampero et al., 1979, Olds et al., 1995). During steep uphill cycling, however, R_R and R_D are considered negligible and R_G is the primary resistance ($W_D \propto R_G$). As such, the SDP is completely objective (gravity affects all cyclists in the same manner, regardless of age, gender, or technique). The SDP is also easy to administer, practical, and cost-effective (necessary equipment < \$1000) for coaches and athletes outside the sports science laboratory environment. In addition, the design of the SDP as an incremental test to volitional exhaustion makes it possible for the SDP to serve as a $\dot{V}O_{2MAX}$ test for cyclists when appropriate.

Summary

Average and peak power output are the most valid determinants of endurance cycling performance. Consequently, the need to measure physiological variables such as $\dot{V}O_{2MAX}$, LT, and economy of movement is questionable: few coaches and athletes have the ready access to laboratory facilities capable of measuring $\dot{V}O_{2MAX}$, or the finances. Additionally, the affordability and availability of power-measuring devices has made performance protocols like the SDP increasingly attractive.

The SDP is a relatively new laboratory cycling protocol based the external power demand of steep uphill cycling. The SDP is a practical and cost-effective measure of peak power output and, therefore, uphill cycling time-trial performance. SDP time-to-exhaustion ($r = 0.97$ to 0.98) and W_{MAX} (watts per kilogram, $r = 0.92$ to 0.97) are highly correlated with actual uphill time trial performance, indicating that a laboratory protocol of relatively short duration (10-12 minutes) can predict exercise performance of much longer duration (15-60 minutes). Therefore, the SDP may serve as an effective marker for tracking changes in actual time-trial performance.

CHAPTER THREE

METHODOLOGY

Introduction

This investigation was conducted in two phases. Phase I of the study assessed the reliability of the Scaling Derived Protocol (SDP); Phase II of the study determined the ability of the SDP to track changes in uphill time-trial performance. In this chapter, the subject characteristics, data collection techniques, and required instrumentation are presented. Included in this chapter is a brief review of the rationale behind the SDP. Finally, the statistical procedures for the stated hypotheses of this study are documented.

The subjects recruited for both phases of the study consisted of a heterogeneous mix of male and female cyclists of varying age, athletic ability, and cycling experience. All subjects were informed of the nature of the investigation and written consent was obtained in accordance with the guidelines set forth by the Montana State University Human Subjects Committee (see Appendix A). Subjects were instructed to prepare for each of the laboratory and field tests as they would for a true competition. Subjects used their own racing bicycle mounted on a commercially available stationary trainer (Computrainer™ Pro, Model 8001, Racermate Inc., Seattle WA., USA) for all laboratory-based SDP tests.

Phase I

Subjects

Twenty-three cyclists from Montana State University and the local Bozeman community volunteered participation for Phase I of the study. Two subjects resigned due to general fatigue and acute knee pain, respectively. Physical characteristics and cycling experience of the remaining subjects (n = 21) are presented in Table 1.

Table 1. Subject Characteristics, Phase I (n=21). Repeated measures reported as Mean ± SD.

Subject ID#	Gender	Age (yrs)	Body Mass (kg)	Height (m)	Competition Level*
1	F	29	59.3 (0.6)	1.66	2
2	F	27	50.8 (0.4)	1.57	2
3	F	23	71.5 (1.0)	1.73	2
4	F	21	59.3 (0.2)	1.63	3
5	F	27	62.0 (0.5)	1.65	2
6	M	21	66.8 (0.3)	1.78	1
7	M	26	81.5 (0.8)	1.84	3
8	M	25	79.5 (1.0)	1.87	2
9	M	36	80.5 (1.8)	1.85	1
10	M	21	71.3 (0.3)	1.77	1
11	M	24	86.8 (0.1)	1.72	3
12	M	24	71.8 (0.5)	1.76	3
13	M	28	77.3 (0.5)	1.82	2
14	M	21	84.0 (0.0)	1.85	3
15	M	26	71.3 (0.7)	1.74	1
16	M	33	78.0 (0.2)	1.88	1
17	M	20	76.8 (0.1)	1.80	2
18	M	29	80.0 (1.0)	1.89	1
19	M	31	80.0 (0.6)	1.85	2
20	M	34	80.0 (0.6)	1.87	1
21	M	21	77.8 (0.2)	1.87	1
Ave (SD)		26 (4.7)	73.6 (9.3)	1.78 (0.09)	

*Competition Level is reflective of the classification systems used by the United States Cycling Federation (USCF) and the National Collegiate Cycling Association (NCCA): 1 = USCF Category 1-2 and/or NCCA Category A; 2 = USCF Category 3 and/or NCCA Category B; 3 = USCF Category 4 and/or NCCA C-D.

Scaling Derived Protocol

Upon arrival at the laboratory, bicycle rear tire pressure was standardized at 100 p.s.i. using a track pump with tire gauge (Specialized, USA). Each subject's bicycle was fitted to a Computrainer™ stationary trainer (Computrainer™ Pro, Model 8001, Racermate Inc., Seattle WA., USA) and calibrated according to the manufacturer's recommended procedure (see Instrumentation, Pg. 29, for a complete description of the calibration process). Subject age was recorded to the nearest year, subject height (barefoot) was measured to the nearest centimeter, and subject mass (with cycling clothes, minus shoes) was measured to the nearest 0.25 kilogram. When anthropometric measurements were complete, subjects were permitted a subjectively adequate warm-up period (10-30 minutes).

The power output (in watts, W) for each stage of the SDP was based on research previously conducted in this laboratory. In brief review, Heil (1998) determined that the external power demand (W_D) of uphill cycling scales with body mass (M_B) to the +0.89 power (e.g. $W_D \propto M_B^{0.89}$). Thus, the relative energetic cost of uphill cycling at a given speed and grade is less for heavier cyclists than for lighter cyclists. Assuming that the net external power demand of uphill cycling is equivalent to a cyclist's metabolic power supply (Heil, 1998), Heil et al. (2000) modified the standard formula for power (Power = [average acceleration * mass] * average speed) to reflect the scaling relationship between body mass and uphill power demand:

$$\text{Power} = [g * M_B^{0.89} * \sin\theta)] * s$$

where "g" is the constant of gravitational acceleration (9.81 m/s/s), "θ" is the average inclination angle of the road surface, and "s" is the average forward ground speed of the cyclist.

Using this equation, the power required for each cyclist to ride up a 5° slope (θ) at 6.72 m/s (s) was calculated and designated as Stage 5 of the SDP. Stage 1 of the SDP was then calculated as 20% of the Stage 5 power output value, Stage 2 as 40%, Stage 3 as 60%, Stage 4 as 80%, Stage 5 as 100%, and each stage thereafter increased by 10% (see Appendix B for a complete outline of the first 14 stages for subjects weighing 40-100 kg). Due to wattage limitations imposed by the Computrainer™, it was necessary to approximate SDP power output to the nearest ±10W. End measures of the SDP included time-to-exhaustion (TTE, min), maximum power output (W_{MAX}, W/kg), and maximum oxygen uptake ($\dot{V}O_{2MAX}$, ml/kg/min).

SDP Subject Procedure

The subjects were allowed to select their own pedal cadence throughout the first five stages of the protocol, with the stipulation that their chosen pedal cadence be held constant (± 2 rpm) within a given stage. When the subjects reached Stage 6, however, the investigator fixed the cadence for the remainder of the test based upon the observed cadence at that moment. The subjects remained seated during the entire protocol and were verbally encouraged to continue until volitional exhaustion, or until the investigator noted a drop in cadence of $\geq$ 5 RPM for $\geq$ 5 seconds. Time to completion (e.g. time-to-exhaustion) of the SDP test was recorded to the nearest second.

Maximal power output (W_{MAX}) and maximal oxygen uptake ($\dot{V}O_{2MAX}$) were also determined during the SDP test. W_{MAX} was defined as the power output corresponding to the final completed stage of the SDP, plus a fraction of the power output maintained in the final incomplete stage (Kuipers et al., 1985):

$$W_{MAX} = (W_{com} + t/60) * \Delta W$$

where W_{com} is the power output of the final completed SDP stage, t is the amount of time (sec) completed in the final uncompleted stage, and ΔW is the power output of the final incomplete stage. Maximum oxygen consumption was calculated as the average of the three highest consecutive VO_2 measures during the SDP. The measurement of $\dot{V}O_{2MAX}$ was considered successful if two of the following three criteria were met: 1) A maximal heart rate within 10 beats of each subject's age predicted maximum heart rate (e.g. 220-age); 2) A maximum respiratory exchange ratio of ≥ 1.10 before the last successful $\dot{V}O_2$ measurement; 3) A plateau of $\dot{V}O_2$ evidenced by a change ≤ 2.1 ml/kg/min with an increase in power output.

In order to determine the test-retest reliability of the SDP, each subject performed three maximal SDP tests within a ten-day period. The tests were conducted at the same time of day and separated by a minimum of 48 hours. Subjects were asked to maintain their normal diet and training load throughout the 10-day testing period.