3-30-2015 12:00 AM

Factors Associated with Success in PARE Testing Among RCMP Officers

Robert Andre Seguin, *The University of Western Ontario*

Supervisor: Dr. Angie Mandich, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Health and Rehabilitation Sciences

FACTORS ASSOCIATED WITH SUCCESS IN PARE TESTING AMONG RCMP OFFICERS

(Thesis format: integrated article)

by

Robert A. Séguin

Graduate Program in Health and Rehabilitation Sciences,

Faculty of Health Sciences

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

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

The purpose of this integrated article dissertation was to examine the predictive factors for success in the RCMP's Physical Abilities Requirement Evaluation (PARE) in a retrospective observational study of 13, 709 unique records and a divisional subset of 620 for study two and three. Study one assessed the relative predictive power of the pursuit and body control times, while including covariates of height, weight, age. Significant ($p<$ 0.05) and equally strong effects were for pursuit log odds (LO) of 2.95% CI [2.49, 3.11], and body control time LO of 2.80, 95% CI [2.51, 3.14] with a weak predictor $\sqrt{Age}$, LO of 0.53, 95% CI [0.38, 0.72]. Not significant were height, weight, and sex with 99 % modeling accuracy.

Study two compared sex and performance factors on six repeated PARE pursuit circuit laps for pacing for both divisional data (535 men, 85 women) and 61 age and BMI matched male/female pairs. Results divisional data: significant strong performance (pass/fail) effects $F(1,616)=288.3$, $p< .00$, partial $\eta^2 = .32$ but weak sex (male/female) effects $F(1,616) = 27.2$, $p = .03$, partial $\eta^2 = .01$, interaction was significant, $F(1,616)=50.7$, $p< .01$, but weak, partial $\eta^2 = 0.014$. Repeat laps were significant, $F(3.7, 229)=195.1$, $p<.01$, with strong effects, partial $\eta^2=0.24$; performance*laps interaction was significant $F(3.7, 229)=4.5$, $p = .02$, with weak effects, partial $\eta^2 = .007$. Significant repeat lap contrasts were lap 1-2, strong effects and lap 2-3, lap 3-4, lap 4-5 weak effects, and not significant was lap 5-6. Results matched pairs data supported significant strong performance effects, $F(1,118)=90.9$, $p<.000$, partial $\square^2=.44$ and weak sex effects, $F(1,118)=13.5$, $p<.00$, partial $\square^2=.10$ not clinically significant. Significant repeat laps contrasts: laps 1-2, strong effect, lap 3-4 and lap 4-5, weak effect, lap 2-3 and lap 5-6 contrast was not significant. Men and women officers paced PARE repeat laps with slight ordinal interaction at lap three and six.

The purpose of study three was to assess self-reported physical activity (PA) frequency and intensity as potential additional significant predictors of PARE success. PA frequency of 3.5 day/week, and intensity of 2.2 of 3.0, did not support additional predictors. There appears to be insufficient PA to affect a maximal test.

Additional self-reported mode and dimensions of PA might increase PA predictability.

Keywords: PARE, Skill-related testing, Police, Fitness for duty, pursuit, body control, health promotion

Acknowledgments

My first acknowledgment is to my family who lost time with me during this course of study. Drake, Joshua, Rory, Felicia and Carolann -- thanks. Second to my supervisor Angela Mandich, who was more than patient with me and stayed positive during a challenging time. To Dr. Johnson and Dr. Irwin whose encouragement and guidance were always productive -- thanks. To my personal mentors: John Nuvoloni, Tracy Gedies, Carol and Claire Shaw, whose personal encouragement and positive outlook helped me get through the dark parts to finish. I owe you all a personal debt of friendship and thanks. Finally to Sgt. Sills and Officer in Charge Derek Matchett, your permission to use RCMP data was deeply appreciated and let the RCMP continue to be an organization of excellence and a strong community member sharing its best practices and findings with the general community.

Table of Contents

List of Tables

List of Figures

Appendices

Chapter 1

Introduction to Predictive Factors Associated with Physical Abilities Requirement Evaluation Success

Police abilities testing (Bonneau, 2001) during the Royal Canadian Mounted Police (RCMP) peak recruitment period reported close to 50,000 Physical Abilities Requirement Evaluation (PARE) when used for all applicants, cadets, incumbent (serving) officers, and college students in preparatory programs (Girard 2012). Police physical fitness for duty testing for all uses listed above is composed of two distinct and complimentary aspects. First, is the preceding periodic occupational medical clearances that not only report police job duty clearances but also report restrictions for duty as described in *Police Health* (Trottier & Brown, 1994). The medical clearances process includes and is identical for both police applicants and in-service officers for medical risk pretest screening for later testing and training. Second, medically cleared officers and applicants continue on to mandatory PARE testing (Bonneau, 2001). Cadets conduct three PARE tests during at depot training and in-service officers conduct PARE periodically with other related police skills recertification (Shell, 2003) as a best practice. The expectation for new cadets and in-service officers equally is for public safety, fellow officer safety and personal health and safety (Trottier & Brown, 1994).

In our modern pluralistic society, both male and female police officers must equally be fit for duty and be able to respond on demand for any given call out. The best descriptors of the majority of police officer shifts are routine and sedentary in nature. As well, officers described the timing of potentially maximal effort callouts during a shift as random (Anderson, Plecas & Segger, 2001). Police officers are high fit during structured police academy training (Knapik, Spiess, Swedler, Grier, Hauret, Yoder & Jones, 2011). However, after years of service, an officer’s physical fitness declines below comparative age matched population or inmate groups (Pollock, Gellman, Price, & Kent, 1977). The decline is in part due to a loss of aerobic power associated with aging for both active (slow loss) versus nonactive (accelerated loss) individuals (Akerman, Heckman, McKelvie, 2104; Fleg, Morrell, Bos, Brant, Talbot, Wright & Lakatta, 2005; Jackson, 1995; Hawkins & Wiswell, 2000; Shephard, 1998; Wilson & Tanaka, 2000). Therefor in-service training programming should consist of testing and remedial training (Major Cities Chiefs of Police, 2005; Shell, 2003). This is especially important for those on the lower end of the fitness test frequency distribution of scores seen in Figure 1.1.

This research narrative will examine in three parts the predictors of success in fitness testing for federal police officers. This first chapter will introduce the issues and research purposes of the following three integrated articles, each being a separate chapter. A final implications chapter will summarize and review the implications of the research findings. Some material will repeat in the general introduction and specific focused research articles, and final implications chapter. Part one of this introduction will focus on the physical demands and tasks analysis of operational policing. Part two will focus on the need for fitness programming, the state of the art of Police abilities testing and the legal aspects of bona fide occupational requirements. Part three will focus on the evolution of the fitness testing with human factors that have shaped fitness testing to date followed by the evolution of PARE used by the federal police. Finally, a brief literature review of the predictors of success in PARE will precede the research purposes of these integrated articles. Examined below are the issues of policing as a physically demanding profession with its essential tasks and demands analysis.

1.1 Physically Demanding Profession

Operational police officers, both male and female, need to be able to exert maximal efforts and meet the high physical demands of a critical incident in public safety 911 callouts (Anderson et al., 2001; Avery, Landon, Nutting & Maxwell, 1982, 1992; Bonneau & Brown, 1995; Trottier & Brown, 1994). Police physical demands analysis studies have revealed critical incidences that require moderate to maximal efforts (Anderson & Plecas, 2007; Anderson et al., 2001; Bonneau, 2001; Farenholtz & Rhodes, 1990). However, much of routine municipal and federal enforcement police shifts are reported and characterized as 50% sitting (373 min), over two hours standing, and one hour walking thus being characterized as largely sedentary of a low level demand (Anderson et al., 2001). Post-shift reports, interviews, and direct observations indicate critical incident callouts are interspersed into shifts at unpredictable times requiring policing competencies to be current and ready for deployment at any time (Anderson & Plecas, 2007; Anderson et al., 2001; Bonneau & Brown, 1995; Boyce, 2009; Farenholtz & Rhodes, 1990). Outlined below are the core tasks and the associated demands analysis of policing.

1.2 Task and Demands Analysis Methodology

Physical abilities tests used for competency screening take their discrete test items from the research tasks and demands analysis (Anderson et al., 2001; Deakin et al., 2001; Payne & Harvey, 2010). Task analysis reports usually cite the "most important, physically demanding" (Jamnik, Thomas, Shaw, & Gledhill, 2010, p. 45) and infrequent but critically occurring jobs tasks (Payne & Harvey, 2010). Task analysis research for municipal and federal police work has used both qualitative and quantitative inquiry methods (Anderson et al., 2001; Birzer & Craig, 1996; Payne & Harvey, 2010). Historically, police and security task analysis research include the following list of activities: (a) direct observations during ride-alongs, (b) post-shift interviews, (c) Delphi survey of important tasks, (d) focus groups with officers and subject matter expert opinions, (e) incident reports, and (f) follow-up test item validation opinion surveys by large police populations and subject matter experts (Anderson & Plecas, 2007; Anderson et al., 2001; Deakin, Smith, Pelot & Weber, 2001; Gledhill, Bonneau & Salmon, 2001; Jamnik, Thomas, Shaw, & Gledhill, 2010a). Seen in Osborn's (1976) methodology for the Los Angeles County Sheriff's Department was an original approach to skill-related test development. The methodology included "(1) questionnaire design, application, cross-validation and evaluation, (2) design and construction of the physical agility test site, and (3) establishment of time parameters through testing" (Osborn, 1976, p. 43). Current quantitative research steps always include large-scale officer opinion validation. Survey ratings and post-test debriefing include all essential, critical, frequent, and demanding tasks (Anderson & Plecas, 2008; Jamnik et al., 2010a). These research steps produced the short list of essential, frequent, or critical but infrequent tasks from the movement domains translated to discrete test items of policing abilities testing. The findings of this methodology`s core policing tasks are described next.

1.3 Task Analysis: Essential Movements

The task analysis studies have documented a consensus of core or essential movements. Osborn (1976) had reported 12 core movements including “balancing, carrying, climbing, crawling, dragging, hitting or kicking, jumping, lifting, pushing, pulling, running and wrestling” (p. 43). Maher (1994) reported additional elements of swimming, balanced walking, and crawling. Birzer and Craig’s (1996) surveys also validated the balancing and crawling items with 15 core movements of: kneeling, stooping/squatting, sitting, standing, walking, jogging/ running, jumping, lifting/carrying, pushing, and pulling. Bonneau (2001) reported core movements in the RCMP task analysis including: walking, running, climbing and descending stairs or hills, jumping, vaulting, lifting, carrying, pushing, pulling, dragging and balancing.

The research literature holds ratings by a large numbers of officers for both the essential demands and the core movements for federal, provincial policing and corrections (Anderson & Plecas, 2007; Anderson & Plecas, 2008; Anderson et al., 2001; Bonneau 1996, 2001; Farenholtz & Rhodes, 1986, 1990; Gledhill & Shaw, 1995; Lagasse, 1989; Jamnik et al., 2010a; Marchand, Thoden, Reed, & McNeely, 1995; Wilmore & Davis, 1979).

Anderson et al. (2001) validated these core movements for municipal policing in British Columbia, while eliminating non-essential tasks such as crawling and lifting over the shoulders, but did not address swimming. The Quebec provincial policing task analysis conducted by Bard et al., (1985) at Laval University, as reviewed in the Lagasse`s (1989) English translation, corroborates the ten basic movement demands of policing through extensive officer validation within Quebec. Anderson and Plecas (2007, 2008) validated both the reported core movements of federal police work, and the core discrete test items of the PARE used to test those movements. The authors used a stratified survey method for all RCMP police divisions across Canada. Divisional samples of officer's ratings were cross-validated with subject matter expert ratings. This produced the PARE discrete task item validation survey (Anderson & Plecas, 2008). Reviewed below are the physical demands analysis of these core movements used in police duties simulation testing.

1.4 Physiological Demands Levels

Direct measures of oxygen consumption during rich job simulation field-testing defines the physiological demands of policing (Reed, 1992 in Bonneau, 2001) Critical incident reporting documents the physical demands analysis of: distances covered, obstacles encountered, masses pushed and pulled, lifted, and carried (Anderson et al., 2001; Farrenholtz & Rhodes, 1990). Human factors data of police clients arrested have included race (which has been controversial), age, height, weight, sex, and the condition and the manner in which they were arrested (Avery et al., 1982; Farenholtz & Rhodes, 1986; Maher, 1984; Osborn, 1976). Reviewed next are the physiological demands findings of job simulations.

Reed (1996) in an unpublished report included in Bonneau's (2001) *The Development of PARE: An Evolution,* compared directly measured oxygen demands in treadmill running in laboratory testing versus in field PARE testing during the PARE foot pursuit and body control apprehension scenario. Maximal aerobic power (MAP) is a well know exercise physiology measure. MAP captured in real time is directly measured peak oxygen consumption difference between inspired and expired gas fractions per minute, written as $\dot{V}O_2$. When this value peaks and plateau during incremental exercise protocols it is taken to represent maximal capacity of the cardiovascular pulmonary systems efficiency in delivering oxygenated blood to the working muscles with working muscles extracting oxygen during work, and returning a diminished percentile of oxygen to exhaled gases.

MAP is marked as a peak consumption capacity ($\dot{V}O_{2\,max}$) of the additional exercise demands plus basal metabolism demand. Resting metabolic demands are usually equivalent to 3.5 ml of oxygen per kilogram body weight (ml/kg/min), termed one metabolic equivalent, or 1 MET in short form. MAP is measured during progressive incremental bouts of 1 to 3 minute duration per stage protocols. MAP is defined as a plateauing of minute volume of oxygen consumed. MAP is also associated with peak heart rate. MAP is also marked as the bodies power output on an ergometer associated with maximum cardiovascular effort plateauing $\dot{V}O_2$ stage (Astrand & Randahl, 1986; Brooks & Fahey, 1986; Gore, 2000).

Reed (2001) used a progressive incremental running test on a motorized treadmill for laboratory determination of MAP for a pool of 35 volunteer RCMP officers (27 male, 8 female). Using the TEEM portable oxygen analysis technology for MAP determination Reed (1996) assessed $\dot{V}O_2$ consumption elicited during PARE as a field test. The combined male and female subject characteristics mean were height, 1.77 cm; weight 80.8 kg; and age 37.6 years (range 25-56 years). Maximum heart rates achieved for the combined group in the MAP lab testing were 186 beats per minute (bpm). Laboratory MAP in $\dot{V}O_2$ in L/min (± SE) for males was 4.6 L/min (± 0.12) for absolute $\dot{V}O_2$ and relative to body weight, was $\dot{V}O_2$ 54.1 ml/kg/min (± 1.29). For females, absolute $\dot{V}O_2$s was 3.0 L/min (± 0.18) and relative to body weight values was 49.9 ml/kg/min (± 2.75). During PARE field testing peak $\dot{V}O_2$ s was 2.88 L/min (± 0.12) and 1.88 L/min (± 0.18) or 33.7 ml/kg/min (± 1.37) and 31.1 ml/kg/min (± 3.63) in relative to body weight units for 20 male and 5 female subjects, respectively. These reported levels were equal for sex groups relative oxygen consumption (corrected for body weight) elicited during field tests at 33.7 versus 31.1 ml/kg/min or about slightly less than 10 METs. The $\dot{V}O_2$ cost of police work reviewed by Reed (1996) contrasted slightly with Gledhill (1992) and Avery et al. (1992) who reported values ranging from 2.5 to 3.5 L/min. The field test elicited heart rates were 179 bpm for males and 186 bpm for females. The relative 2/3 peak oxygen consumption recorded (66%) during PARE field testing in relation to MAP values from the laboratory treadmill tests was postulated by Reed (1996) to be due to significant anaerobic metabolic pathway contributions to muscular energetics not compensated for by slower aerobic pathways as fully described by Brooks and Fahey (1986) in *Human Bioenergetics*. Rhodes and Farenholtz (1992) also found health-related field capacity fitness tests measures did not relate to police abilities scores and attributed it movement abilities not measured by those tests.

The difference in PARE scores for this small sample group of men and women were clinically significant. Men's scores were 214 s (3:34 min:s) a good pass versus women's score which were 246 s (4:06 min:s) – a slight fail on average, compared to the RCMP physical employment standard (PES) for cadet graduation. In contrast to defining police work with maximal or submaximal cardiovascular demands. Sharkey and Davis (2008) also defined police from a manual materials handling point of view. They cited policing as requiring hard heavy work, with heavy being greater than 50 lb mass pushed, pulled, lifted and/or being carried.

In summary, the physical demands of operational policing are largely sedentary shifts punctuated with unpredicted critical incidents callouts for any officer. The work as elicited in field-testing documents a maximal cardiovascular demand with heavy manual handling of loads. Therefore, police officers have to prepare to deal physically with their clients and be ready at a moment's notice. However, police officer fitness levels have not always been equal to the demands or their clients (Collingwood, 2005). Suspicions are that because of the lack of physical activity or exercise during a shift, especially for police of older age groups, with many years of service compared to population and client age matched groups, fitness is lower.

1.5 Police Fitness Levels

Initially in their careers, police officers' physical fitness is high during police academy training (Knapik et al., 2011). Wilmore and Davis (1979) state "it is apparent that while the average officer graduates from the academy in excellent physical condition, the normal sedentary nature of the officer's job leads to a rapid deterioration in physical fitness" (p. 37). Knapik et al. (2011) have documented high levels of police physical fitness in training academies for the Federal Bureau of Investigation (FBI). Likewise, the RCMP have documented high fitness levels for cadets at the point of Canadian federal cadet graduation. They pass both published health-related fitness component benchmarks as well as skill-related test like PARE. The passing prevalence is at the 99th percentile and 95th percentile pass rates for men and women's officer groups (Skolney 2010, personal communications).

However, after years of service, there are reports that police officers' physical fitness levels are below average population levels. Research intervention studies and physiological monitoring surveys has shown police officers of a previous generation (Frank & Anderson, 1994; Pollock et al., 1977; Wilmore & Davis, 1979) and the current generation are less fit than inmates and similar age-matched groups in the general population (Quigley, 2008; Strating, Bakker, Dijkstra, Lemmink & Groothoff, 2010). Police physical fitness testing has found officers to be either meeting or being below average in fitness levels, in particular older officers and women officers (Bissett, Bissett, & Snell, 2011; Pollock et al., 1977; Shephard, 1997; Wilmore & Davis, 1979). Hoffman (1996) cites fitness levels for health-related fitness constructs for police officers, in the 25,000 plus records of the Cooper Institute. Police fitness levels are posited as "aerobic fitness (35th percentile), body composition (40th percentile) and abdominal strength (40th percentile)" (p. 1) Pollock et al. (1977) found 100 inmates were "in better physical condition than police officers" with "higher working capacity, and cardiorespiratory endurance" (p. 45). Middle-aged police officers, 36 to 52 years of age, were "below average in working capacity, cardiorespiratory fitness, and body composition" scores (Pollock et al., 1977, p. 45). Bissett et al. (2011) confirm Pollock et al.'s (1977) findings in the previous generation of police officers with today's generation as they relate to decreasing fitness levels over time. Bissett et al. (2011) cites results for younger officers who were average for all health-related fitness construct scores; however, older officers were below average for health-related constructs of flexibility and endurance. Strating et al. (2010) confirmed lower fitness scores for older, female, and administrative duties groups in Dutch police officers, when using skill-related fitness testing. Boyce et al. (2008) conducted a 12.5 year follow-up on an urban police service using a large sample size and reported a significant increase in obesity rates equally for male and female officers ($p < .05$) with years of service.

In Canada, in the federal police academy, graduates as a group have lower PARE scores depicting higher fitness levels than the incumbent officer populations. Pilot work in this lab discovered that a frequency distribution of 3 years of PARE scores demonstrated a near normal distribution of scores. The frequency distribution data summary for older serving officers displayed a wider distribution of scores than for younger academy cadets. Many of the older officer's scores exceeded the PES of 4:00 min:s. A right skew of serving officer data is seen in the frequency distribution of PARE scores. The right skew of scores represents slow and low effort participatory scores of walking or jogging the PARE. The frequency distribution also demonstrated a mode score for cadets on PARE, for males and females, at 3:00 and 3:40 (min:s), respectively (personal communication Skolney, 2012). In comparison, mode scores for incumbent male and female officers, were 3:40 and 3:58 (min:s), see Figure 1. That the mode for female officers was under the 4:00 min standard is an important point for boni fide occupational requirements (BFORs) methodology when using the incumbent police population, or a large sample thereof (Jackson & Wilson, 2013) for normative referencing for PES (Payne & Harvey, 2010; Shephard & Bonneau, 2003).

BFORs are legal discriminators for job function competencies, and as such they have to be brought to the job. One example is medical physical abilities of policing (Saw & Gledhill, 1995). See Wagner-Wiscotzki (2005) for a review of bring-to versus learn-on the job competencies. Pilot wok has shown that the mode decreases per 5-year age groups at a regression slope of 1 sec per year, or 10 s per decade, and 30 s per 30-year career period. This is suspected to be due to the demonstrated drop in MAP at 10% per decade (Hawkins & Wiswell, 2003) attributed to the progressive decline in functional capacity of the cardiovascular system with aging without training (Lemura, Dunvilad & Mookerjee, 2000). Staffing an officer in a protective unit requires the officer to attain and maintain a sub 4:00 min PARE score at any age. Protective units must recertify yearly. In contrast, general duty officers must recertify every three years. General duty officers do not have transfer consequences if they do not maintain sub 4:00 PARE. This is in contrast to those cadets who ran hard to complete the PARE and graduate to win their job. PARE use by recruiting units for pre-selection screening and PARE use by training units for in service periodic skills certification both have to address police fitness issues amongst other competencies (RCMP, 2010). Examined next are recruit and in-service occupational physical abilities testing, hereafter termed fitness testing.

1.6 Policing Competencies

Finding personnel capable of the physical demands of policing is only one of several issues facing policing recruiting units. Police services differentiate between bring-to-the-job applicant competencies versus developed on-the-job competencies (Wagner-Wisotzki, 2005). The provincial police Constable Selection System in Ontario (Gledhill & Shaw, 1995) defined the essential “bring to the job” competencies to include “medical/physical abilities” (p. 1). The other bring-to-the-job competencies were “analytical thinking, self-confidence, communications skills, flexibility, valuing diversity, self-control, relationship building, [and] achievement orientation” (Wagner-Wisotzki, 2005, p. 13). In contrast, on the job learning includes “information seeking, concern for safety, assertiveness, initiative, cooperation, negotiation/facilitation, work organization, community-service orientation, and commitment to learning” (Wagner-Wisotzki, 2005, p. 13).

1.7 In-service Police Fitness Programming

Without mandatory health and fitness promotion programming for incumbent officers, it is difficult to maintain high levels of physical fitness over the course of police officers’ career (Bissett et al., 2011; Blewett, Briley & Montgomery, 1993; Boni, 2004; Boyce, Jones, Lloyd, & Boone, 2008; Shephard & Bonneau, 2003; Winterhalder, 1993). Collingwood (1988a, 1988b, 2004) and Shell (2003) both describe the historical direction that fitness leadership programs for police services have taken to promote physical fitness for police duty with or without a Masters-level qualified subject matter expert. When police services are seen as too small to afford a qualified science-based fitness program leader, alternatives include using existing personnel trained as physical fitness coordinators/testers for “a recruit program or an in-service program” (Collingwood, 1988b, p. 28).

The Cooper Institute has a long-standing tradition of supplying police officers peer-to-peer fitness testing courses with normative data for interpreting age- and sex-based health-related fitness capacity test results (Hoffman & Collingwood, 2005). The authors describe the “ultimate goal that officers participate fully, developing and maintaining physical fitness throughout their career” (p. 28). Shell (2003) suggests a best practice is to pair physical fitness training with other police skills recertification, commonly termed ‘use of force training’ (e.g., defensive baton, pepper (OC) spray, carotid control, handcuffing and take down techniques, car stops and urban interiors arrest scenarios).

Surveys of common police health, fitness, and wellness promotion programs by the Major Cities Chiefs of Police in 2008 has documented the generic descriptions of these program strategies as: health; wellness and prevention; managing injury and illness; physical fitness for duty; and reintegration after military deployment (Major Cities Chiefs/National Executive Institute, 2008).

1.8 State of the Art Police Fitness for Duty Appraisal

A description of the state of the art appraisal of police fitness for duty based on abilities testing should include the current legal definition of the essentials of a BFOR test. The fitness test for provincial corrections officers (FITCO) is an example of a modern test that checks adverse impact for and reasonable accommodation of women in the test development stages. Zumbo (2001) described adverse impact as legal term representing disparate group failure rates in testing. Getting the initial lower pass rate group to score a pass percentage above 80% of the initial higher pass rate group is termed reasonable accommodation in fitness testing. Jamnik et al. (2010) provided several months of physical training for women’s groups failing on the initial testing session, to achieve the pass rate of > 80 % of the men's pass rate in subsequent testing. The failure of one group to pass above 80% of the other group pass rate, sometimes called the 4/5 rule, is termed in human rights tribunals and BFOR test validity literature as adverse impact. Next, the legal aspects of tests are discussed.

1.9 Process and Legal Basis of BFOR tests

The drivers for change in the United States (US) courts was new legislation and human rights challenges, including the Americans with Disabilities Act, the Discrimination in Employment Act, the Civil Rights Acts, and the *Uniform Guidelines* of the Equal Employment Opportunities Commission (Collingwood, 1995). The drivers for change in Canada have been similar policies and laws under the Human Rights tribunals and Supreme Court rulings on BFOR cases like the Meoirin case (Eid & Geh, 2001).

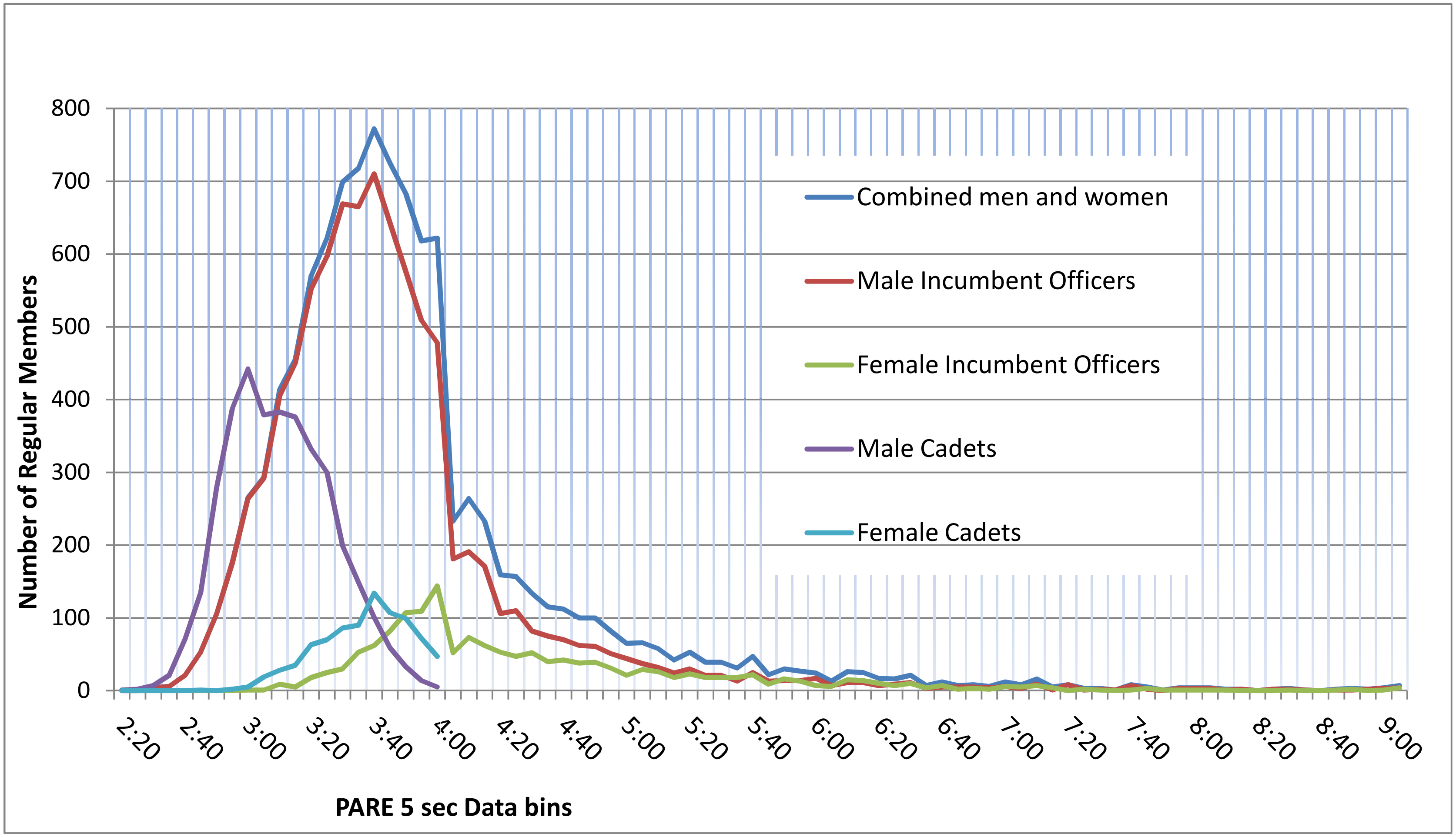
In Canadian law, the Supreme Court precedent setting case was the Meoirin case (Eid & Geh, 2001) which outlined the legal basis of a BFOR test. The BFOR test discrete items and overall test processes must be (a) "rationally connected to the performance of the job", (b) administered in "an honest and good faith belief", (c) with "the standard [that] is reasonably necessary to the accomplishment of the legitimate work-related purpose" (Eid & Geh, 2001, p. 55). The authors describe the dual aspects of objective and subjective elements of a BFOR discrimination that are defensible. The conditions (a) and (b) above are the subjective elements of a BFOR discrimination where (c) above is described as the objective element of a BFOR discrimination (Zumbo, 2001). Eid and Geh (2001) also cite failure to do the job must result in reasonable accommodations of the person with limited abilities up to a level of undue hardship on the organization. However, Gillis and Darby (2001) realize that "human rights and the duty to accommodate cannot, as a legal response, overcome all barriers to full participation in some jobs" (p. 86). These pressures above have directly affected ongoing BFOR test development, evaluation, and use and acceptance of fitness test PES use (Deakin, Smith, Pelot, & Weber, 2001; Eid & Geh, 2001; Gledhill, Bonneau & Solmon, 2001). When establishing and applying BFOR tests for any reasons human rights factors of group characteristics must be considered (Gledhill et al., 2001). Reasonable accommodations for adversely impacted groups must always be possible, as shown in the state of the art corrections test, the FITCO (Jamnik, et al., 2010c).

The Meoirin case findings are the political, cultural, and moral values desired by Messick (1996) for construct test validity. A BFOR test, granted when a human rights challenge is won by the test developer, or a BFOR defensible tests, are now “in wide spread use” (Birzer & Craig, 1996, p. 93) in most policing and corrections recruitment (Wagner-Wisotzki, 2005; Jackson, 1994). In-service promotion, placement, and periodic recertification of competencies requires the use of BFOR tests (Dunsmore & Hunter, 2001; Shell, 2003). This is partly because of the insufficiency of pre-employment and ongoing occupational medicals in police health without the complement of an abilities test (Gledhill, & Shaw, 1995; Trottier & Brown, 1994).

The changes in fitness testing appears to be driven by human rights and labour law. Changes in the courts, include, the Civil Rights Act, the American with Disabilities Act, and the Uniform Guidelines jointly from the US Department of Labour and the Equal Employment Opportunities Commission (Avery et al., 1992; Bonneau & Brown, 1995; Eid & Geh, 2001; Collingwood, 1995; Jackson, 1994; Jamnik, Thomas, Gledhill, 2010, Maher, 1994).

The cumulative effect of legislation has been to restrict or prohibit discrimination on prohibited grounds (Eid & Geh, 2001) and to examine adverse impact after the Meoirin decision through ongoing test evaluation and standards adjustments (Avery, et al., 1992; Gledhill, Bonneau & Solmon, 2001; Hogan & Quigley, 1986; Zumbo, 2001).

Figure1.1 Force wide PARE in a 3-year compliance cycle for 2008-2011



Note: Force wide PARE data for 10, 321 incumbent officers: Males (8,786), Females (1,536) and Male cadets (3,675) and Female cadets (859). PARE scores are normatively distributed with a long right skew for participatory PARE scores, as outliers. Data is from a large sample of the population of single best PARE results. Cadet's scores terminate at the physical employment standard (PES) of 4:00 min:s, which is their cadet graduation standard.

Historically the development of a specific occupational fitness test flowed out of the task analysis as previously described. Common state of the art methodological steps for test development have been outlined in the 2000 consensus conference proceeding, by editors Gledhill, Bonneau, and Salmon (2001). Gledhill et al. (2001) have described the research and implementation template process in 12 steps: (a) establish and justify the need for a BFOR, (b) create a project management team, (c) conduct a job familiarization analysis, (d) conduct a physical demands analysis, (e) determine the short list of essential tasks, (f) characterize the tasks, (g) develop and (h) standardize a test protocol, (i) establish scientific accuracy for validity, (j) develop performance standards and assess adverse impact, and (k) implement and (l) ongoing review and evaluation.

The 2000 BFOR consensus conference proceedings also reviewed content versus construct validity methodology (Deakin et al., 2001), the legal human rights basis of BFORs (Eid & Geh, 2001), with a methodology review by Zumbo (2001), and accommodations until undue hardships arguments -- all citing applications and critiques of the processes and law. The current state of the art BFOR defensible test, developed and published in peer reviewed academic articles, is that of Jamnik et al. (2010a, b, c) mentioned above. It is the most current example of a policing or corrections BFOR testing to use the traditional mixed research methods of direct observations, interviewing, focus groups, and follow-up worker survey validation.

FITCO was recently designed for provincial correctional officers in Ontario (see Appendix 1, Figure 5). The authors' demonstrated that failing groups can be reasonably accommodated with test familiarity, a training program, and sufficient time to achieve training enhancement – all resulting in passing in a timely fashion at rates greater than 80% for all who initially failed (Jamnik et al, 2010c). Presently, in a state of the art BFOR test, the task analysis and test development processes includes both final adverse impact (Zumbo, 2001) and reasonable accommodations assessment periods for all protected groups, as outlined in industry standard current BFOR test development published templates (Gledhill et al., 2001) and limited academic literature (Jamnik, Thomas, Scott, & Gledhill, 2010). Described below are the three main historical moments of occupational fitness testing methodology that have brought us to the state of the art fitness testing.

1.10 Historical Moments of Occupational Fitness Testing

Occupational fitness testing methods and the related human factors variables used in testing has evolved over what Denzin and Lincoln (2005) termed ‘moments’ or historical periods of methodological emergence and succession. Denzin and Lincoln (2005) defined the evolution of qualitative research philosophy, through seven moments, and described each successive historical period as being dominated by an emerging philosophical development, with a corresponding methodology, while a predecessor philosophy and methodology declined. Similarly, Police abilities testing methodology is in the process of development of its third historical period, of an emerging dominant methodology, namely skill-related police abilities fitness testing, while the previous methodologies are in decline. In particular, continued persistent supporters of health-related fitness testing, with age and sex separate physical employment standards (PES)s, await court rulings of fairness in gender separate standards (Collingwood, 1995).

1.11 First Moment: Height and Weight

The first historical moment of police abilities testing, used only for recruiting men, was the use of human factors, namely the applicants’ height and weight (Avery et al., 1992; Bonneau & Brown, 1985; Hogan & Quigley, 1986; Jackson, 1994; Shephard & Bonneau, 2003). This construct was considered necessary to successfully grapple and wrestle with suspects (Bonneau & Brown, 1995; Wilson & Bracci, 1982). Mahar (1984) also described minimum height and weight as related to what police officers needed for “assaults on police officers. Ability to see over crowds, fences, and so on. Ability to use equipment, such as driving cars, and firing a gun. Ability to gain respect and psychological advantage” (p. 173). Today, human factors of height, weight, age, and physiological capacity of workers, are still used in determining push and pull ability in the area of manual material handling, ergonomics, and other occupational biomechanics areas (Armstrong & Young, 2010; Ayoub & McDaniel, 1974; Boocock, Haslam, Lemon, & Thorpe, 2006; Chaffin, Herrin, Lee & Waiker, 1991; Chaffin & Resnick, 1995; Ciriello & Snook, 1991; DARCOR and ERGO web, 2103; Knapik & Marras, 2009; Kroemer, 1971). Explicit in policing scenario testing protocols are fair standardized testing practices controlling for subject height and weight factors during discrete item movement demands as potential unwanted moderators of test performance scores (RCMP, 2010). The mediators of test success are: testing coaching, fair and equitable equipment design that mitigates or avoids irrelevant task easiness and or difficulty, while measuring the main construct of physical abilities (Messick, 1995) without disparate group responses (Zumbo, 2001).

The initial PES height and weight assessment methodology was replaced with second moment health-related and third moment skill-related fitness testing. The first skill-related testing was described by Wilmore and Davies (1979) as “selection of state traffic officers work” (p. 33). Thompson’s (2010) definition of health-related or skill-related fitness testing are below.

1.12 Second Moment: Health-Related Constructs

The earliest second moment approach, described as health-related fitness testing, focused on constructs of muscular strength and endurance as the most relevant police job competencies (Avery et al., 1992). Thompson’s (2010) definitions of fitness testing included an expanded list of health-related constructs of “cardiovascular endurance, body composition, muscular strength, muscular endurance, and flexibility” (p. 3). The constructs were most commonly measured in a discontinuous format with discrete fitness tests items for a specific region of the body or the whole body, interpreted via age and sex normative tables and summarized into a fitness profile score (Hoffman & Collingwood, 1995; Shephard & Bonneau, 2002).

In the context of health-related testing, a battery of relevant fitness constructs are measured by a set of discrete test items and are scored in a summative fashion. Shephard and Bonneau (2003) defined the weighted composite scores of health-related testing as "a global score that summarizes performance on all test items" (p. 269). Deakin et al. (2001) described this type of test as a construct valid testing process with fitness tests performance supposedly highly correlated with job simulation performances. Shephard and Bonneau (2003) cite step-wise or canonical regression analysis and or principle component analysis, as the statistical steps taken to relate the global scores to policing performance abilities on the job. The later steps were considered valid if fitness scores were highly correlated with job simulations of the frequent, demanding, or critical tasks of police work described in the section above.

In health-related capacity testing, interpretation of results was via separate age and sex scores in quintile groupings (Hoffman & Collingwood, 2005, Sheskin, 2007). A generation of data supports health-related police fitness for duty testing with published and proprietary normative tables (Hoffman & Collingwood, 2005; Wagner-Wisotzki, 2005). These tables, and the health-related component fitness testing at the for-hire stage, have been retired by both the federal police with PARE in 1992 and by the Ontario Ministry's with the Physical Readiness for Police (PREP) in 1995 (Ministry of Solicitor General and Correctional Services, 2002).

There are advantages and disadvantages to using either a single pass standard or sex adjusted different pass standards. Shephard and Bonneau (2003) offer a perceived limitation to normative one score testing based on 50% of the mean of serving officers suggested by Wilson and Bracci (1982). The authors saw a single for hire target as "unlikely that women recruited on such a basis would be able to meet the demands of policing" (p. 275). The difference in average fitness for 20 to 50 years olds is cited by Shephard & Bonneau (2003) as the reasons against one score normative referencing. However, if normative referencing of job minimum duties are completely and accurately defined, and reflected in the abilities test (Avery et al., 1992) one standard and one cut score is transparent to the courts, the officers and the public (Deakin et al., 2001).

Health-related fitness testing methods are still used today for in-service or incumbent police and security health promotion programming across Canada. They are a used on a voluntary or self-determination basis. These programs are typically an incentive /recognition awards program for officers' physical fitness for duty. Universal and mandatory health promotion skill-related testing at the federal police level has mandatory medical and environmental workplace health and safety programs. Thus Green and Kreuter's (1999) description of a three part mandatory health promotion program, with medical surveillance, occupational health and safety, as well as mandatory health promotion fits the federal policing organizational culture. Unionized opposition has delayed the mandatory skill-related testing implementation at the provincial or municipal level for incumbent police abilities testing (Claire Shaw, personal communications).

1.13 Third Moment: Skill-Related Constructs

The third moment of occupational fitness testing emerged with the advent of rich job simulation skill-related police abilities testing (Farenholtz & Rhodes, 1986, Osborn 1976; Wilmore and Davis, 1979). The skill-related approach to fitness testing included two levels of test constructs use (Deakin et al., 2001). The first level construct of skill-related testing is the discrete test item critical or essential police movement tasks for constructs of "speed, coordination, reaction time, agility, balance and power" (Anderson & Plecas, 2008; Thompson et al., 2010, p.3). The second level construct is that of a critical incident police call out flow replication of "getting to a problem, dealing with a problem, and removing a problem" (Anderson et al. 2001, p. 8; Bonneau, 2001; Farenholtz & Rhodes, 1990). Each scenario segment has discrete test construct items integrated in two continuous timed circuit segments (Wagner-Wisotzki, 2005, p. 49-50). This multi-station single-test job simulation of police callouts cuts out the intermediate steps of a complex mathematical treatment of a global score of a battery of separate health-related test items correlated with a job simulation, described in health-related testing.

Foot pursuit and resistive body control segments are represented by one real-time test completion time score. Job simulations are popularly accepted as the reflections of police demands (Anderson & Plecas, 2008; Deakin et al., 2001; Mahar, 1984; Shephard & Bonneau, 2003; Payne & Harvey, 2010). Although Messick (1996) believes all test validity is construct validity, job simulation testing has what is termed high face or content validity, not requiring subject matter expertise in exercise sciences or statistics to see the relevance of the discrete test items to the job. With that in mind, the rich job simulation testing has been described as being highly accepted by many groups (Deakin et al., 2001), including labour, the judiciary (Avery et al., 1992; Mahar, 1984), and federal police officer populations (Anderson & Plecas, 2008). Mahar opens his 1984 article with a quote from a judge stating: "Surely, it is difficult to imagine a more accurate way of testing ability to scale a 6 foot wall then to scale one (Hardy versus Stumpf, 1978)" (p. 173).

1.14 Hybrid Tests

Hybrid tests usually include both content- and construct-valid discrete test items (usually in a timed circuit), followed by a health-related fitness capacity discrete test item that was perceived to be missing in the initial test constructs. To date this approach has been seen in a negative tension in officer post test debrief critique, with officers citing a missing important construct that the initial skill-related test simulation did not adequately measure (Deakin et al., 2001; Gledhill & Shaw, 1995). An example of a test, originally missing an important construct and adding in a discrete item into the hybrid test version, to assess that construct, is the PREP (Gledhill & Shaw, 1996). PREP was missing a whole body full cardiovascular demand effort in its short two-minute test. PREP is now composed of a timed pursuit, body control, and arm simulation segment, plus an after the fact circuit added health-related construct of a demanding shuttle run, namely the Leger-Boucher 20 m shuttle run for maximal aerobic speed endurance, see Appendix A.

1.15 Current BFOR Tests

Current BFOR tests used in policing and corrections are: (a) the Police Officers Physical Abilities Test (POPAT, Farenholtz & Rhodes, 1986); (b) the PREP for Ontario Police and the PREP-A for Alberta provincial and municipal police recruits; (c) the FITCO for Ontario provincial corrections (Jamnik et al., 2010) which is a hybrid test similar to PREP, and (d) the PARE (Trottier & Brown, 1994), the latter being the only pure skill-related test. The PREP is a provincial and municipal policing hybrid test with both skill-related circuit of discrete items and with a health-related construct added, for example the Leger-Boucher 20 m shuttle run for aerobic power prediction (Gledhill & Shaw, 1995; Wagner-Wisotzski, 2005). See Appendix A for complete graphical comparison of the layouts of each test pursuit and body control simulations.

1.16 Controversies

Research of police task and demands analysis and BFOR test development are historically beset with controversy. Some of the issues are: organizational rights and undue hardships when working with BFORs (Gillis & Darby, 2001), measurement and methods validity (Avery et al.,1992; Jackson, 1994; Lonsway, 2003; Shephard, 2013; Zieky & Perie, 2004; Zumbo, 2001), safe supervision (Shephard & Bonneau, 2002), gender equity and critiques of the use of a masculine metric PES that may adversely impact women officers in a supposedly gender neutral test (Birzer & Craig, 1996; Jackson & Wilson, 2013; Lonsway, 2003; Prenzler, 1996; Shephard & Bonneau, 2003).

BFOR testing currently in use in policing is sporadically reported in published academic journals (Jamnik et al., 2001). Mostly police and corrections test development reports are found in unpublished government reports and archives for federal and provincial police in Canada (PARE by Bonneau, 2001). English provincial police reports covering Ontario can be found in the Ontario Police College Archives (PREP by Gledhill & Shaw, 1996). The Quebec provincial police test are stored in the Quebec national police training centre (le TAP by Lagasse, 1989). In contrast the current state of the art Ontario provincial corrections test, the FITCO, updated and replaced the proposed test by Marchand, Reed, Thoden and McNeely (1995).FITRCO is published in three article by Jamnik et al. in the CSEP peer reviewed academic journal (i.e., FITCO) (2010a,b,c).

Avery, Nutting, and Landon (1992a) report the validity of any test construct for content or face validity rests principally with the jurisdiction of use. They must assure the following points are considered: (a) the relevant physical duties captured by job analysis, (b) the sample of workers large enough and of appropriate qualified representatives of the job, (c) the test events matched, without over or under emphasis of relevant aspect of job, (d) the test items the same as on the job tasks without unethically stressing a worker, (i.e., correct number of stairs, height of wall, notwithstanding no one can ethically replicate real fear conditions), (e) the representative of non-standardized duty where the tasks take are of highly variant situations, like resistive arrest ground fighting, and f) the test items sufficiently represent all of the complexity of multi-variant job, (f) the tests has infrequent but critical important tasks, like firearms use. Messick (1996) described some instances of inexact measurements of a police job discrete item as either irrelevant task difficulty or irrelevant task easiness. Bonneau's (2001) personal communication on the PARE validation rebuttal to Avery, Nutting, and Landon's (1992) critiques were that PARE was never meant to deal with ground fighting or firearms, as this was not bring-to-the job but on-the-job training skill sets. Likewise for incumbents, firearms training should be separate from fitness testing scenarios.

In health-related fitness testing the pass / fail or cut score criterion references were originally drawn from representative handpicked groups of officers (Metivier, Gauthier, & Gaboriault, 1982). Later, with Farenholtz and Rhodes's (1986, 1990) in POPAT and COPAT testing the most demanding police clients' averages were used for cut score total test time. The resistance levels that could be achieved by young females in the validation group was represented as the criterion reference level for skill-related testing push pull resistance (Farenholtz, personal communications). There have been many controversies associated with using different norms for male and female officers in health-related fitness testing and many legislative pressures to avoid discrimination against protected human rights grounds of age, sex, race, and disability status, etc. (Avery et al., 1992; Hoffman & Collingwood, 2005; Jackson, 1994; Lonsway, 2003; Mahar, 1984; Prenzler, 1996; Shephard & Bonneau, 2003). However, the above shows male and female differences where always involved in tester developers trying to set fair cut-scores.

Using age and sex separate scoring for a sex neutral job demands is seen by some authors simply as sexism (Thomas & Means, 2000, Gillis & Darby, 2000). However, critical feminist researchers, some of whom are police officers themselves, view male sample group's results that determine PES cut-scores as a masculine-centered metric. The above is seen as limiting women's entry and advancement in policing (Andrews & Risher, 2006; Gaims, Falkemburg & Gamino, 1993; Lean & Durand, 2002; Lonsway, 2003; Prenzler, 1996). The use of small female sample groups to date to validate the original Canadian Police Officer Physical Abilities Test (POPAT) or the PARE tests (Bonneau 2001) or the PREP A (Alberta) has not been addressed in peer reviewed published journals by the authors of these tests. Potential sex differences and gender equity in passing major BFOR tests is unclear (Shephard & Bonneau, 2002) with one exception (Jamnik et al, 2010a). Jamnik et al., (2010) tested women in their BFOR test development stage, addressing lower scores by women in the initial test phase by using training as the reasonable accommodation that had final post training testing numbers how now adverse sex impact was avoided in final pass fail percentages. Also not addressed is the relationship between height and weight and age of an officer on their success rates in a BFOR PES tests. To date, not enough women, and not enough physically fit women, have been used in either predictive or validation studies on PARE (Bonneau, 2001; Farenholtz & Rhodes, 1990; Stanish, Wood & Campagna, 1999). Avery et al. (1992) describe the inferential approach for construct valid physical ability scores as mediating factors for job physical ability determinations (as previously described), and advocated for more construct validity approaches. They suggest supervisors report on physical ability ratings of officers to accompany physical testing scores for evaluations. However, based on the focus group findings of Bissett et al. (2011), which examined the importance of physical abilities in policing, as perceived by chiefs and supervisors, it was purported that such a task would likely be viewed as unrealistic in a policing organizational culture. Their opinion is that most in-service police officers have 'average' to 'above average' performance reviews. Further they conjecture that most of these officers have not completed a physical abilities test, nor could all complete one at the same 'average' to 'above average' level. However the approach of supervisors rating performance has been used in stat of the art cutcore determination with current fire and search and rescue with various methods, reviewed elsewhere by Rogers, Docherty and Petersen (2014).

Before the clear divisions of testing into either health-related or skill-related testing described by Thompson (2010) above, Metivier et al. (1982) reported creating a police service recruiting standard based on a set of ability constructs via expensive and complex laboratory and field testing processes. An example historically, before they adopted the PREP, was the Ottawa Police Service which provided a hand-picked group of male officers to the researchers to come up with test and cut-scores for this large urban Canadian police service. The volunteers were described as the "best all round physically, psychologically and most efficient" officers (p. 2). The authors conducted controlled laboratory testing for health-related fitness standards for constructs of: physique, muscular strength, power and endurance, joint flexibility and visual and auditory reaction time" (p. 2). Discrete item measurements included: physique through densitometry for percentage body fat; maximum repeat chin-ups, dips, and push-ups for muscular endurance of the arms; dynamometry for back, legs, and grip strength; sit and reach for flexibility of the hamstrings and low back muscle groups; a Margaria stair climb for dynamic leg power; standing broad jump for static leg power; and reaction time for auditory and visual stimulus as measured through a "performance analyzer 631" (Metivier et al., 1982, p. 2). This battery of test scores developed for one police service's specific small homogenous group, is a familiar challenge to physical agility test development seen in the PARE, COPAT and POPAT (Farenholtz & Rhodes, 1984, 1986), as well as in other programs developed in the late 1980s and early 1990s. This process, described above, had cut scores for a recruiting pre-entry test that can be critiqued as cost-prohibitive without being widespread or representative of all ages and sex group differentiation of officers for adverse impact analysis (Tinsley, 2000; Zumbo, 2001). Further, most small police services cannot afford a science-trained subject matter expert, let alone complex laboratory services to develop criterion referenced recruiting PES tests (Collingwood, 1988b; Shell, 2003).

Mass participant testing in groups using common health-related fitness field tests are seen as: less cost prohibitive, less equipment intensive, and highly portable alternatives. However, some health-related field test measures have been critiqued as having high standard error levels (Anderson & Plecas, 1999). In rare cases, the 1.5 mile run have resulted in deaths in fitness testing in Ontario (Deakin, Smith, Pelot, Stevenson, & Wolf, 1994).

Jackson (1984) and Maher (1984) reviewed controversial issues of BFOR test development and implementation challenges. The authors described height, weight, and age discrimination in testing that demonstrated what Zumbo (2001) described as disparate group impact. Mahar (1984) reported court challenges citing that not all-physical abilities testing has valid content elements that are reflective of important elements of the job, a point also cited by Avery (1992). Mahar (1984) also pointed out that scaling walls, pushing and pulling, lifting, and carrying heavy objects may require upper body strength and hence be adversely impacting one sex, namely women. The authors also commented that physical ability testing can be adversely impacting older workers and the over 30 years of age restriction of taking a BFOR test should be lifted.

Controversy in physical fitness test interpretation as a sexed issue comes from several areas. Scores in job simulation tests are compared to a policing client demographic group representing the most frequent and difficult arrests of the most able bodied clients. These fairly homogenous male client groups average test time scores have been used as the criterion reference for cut scores for policing tests as outlined in Farenholtz and Rhode's (1990) Canadian standards for police article and Bonneau's (2001) *Evolution of PARE* article.

1.17 The PARE

The Major Cities Chiefs of Police Toronto Conference after surveying most police services only reported one mandatory incumbent fitness testing program in North America, in Utah, and as such has missed one of the largest scale mandatory programs of for officers from hire to retirement, namely the Canadian federal police's *PARE test.* The RCMP have instituted mandatory attendance in-service police abilities testing in Canada. Originally it was paired with the periodic occupational medical on a biannual basis from 1997 to 2006 (Trottier & Brown, 1994), and later from 2007 to 2014 with the operational skills maintenance scenario recertification (Shell, 2003). In contrast the provincial police instituted medical / physical abilities testing at the recruit applicant level with scenario testing in 2002 (Gledhill & Shaw, 1996) with the PREP. This was complemented with a voluntary health-related fitness testing program for incumbents, mostly due to municipal and city police union resistance to mandatory testing for in-service officers with recruit BFOR tests (Green & Kreuter, 2005). Birzer and Craig's (1996) survey report show a large number of police services use recruiting agility selection tests. The 1979 Police Foundations Survey results showed 78% of municipal and 91% of US States reported a physical test use for applicants.

Periodic police officer operational skills recertification, with physical fitness programming is needed throughout an officer's career to remediate police fitness for duty as close to PES as possible. The question is, are PES standards the minimum necessary to perform the job (Collingwood, 1988a; Gledhill, Bonneau and Solmon, 2001). Wilson and Bracci (1982) clearly articulated the concept of mandatory incumbent testing and its relationship to PES standards that are still central to the issues we face today: *"if police officers currently employed are unable to pass job related agility tests required for hiring, how valid are the statements that these tests are necessary for satisfactory job performance" (p. 41). And further: if he[she] fails, there are two possible assumptions: the officer is unfit for duty – or the pre-employment standard is not essential to perform the police tasks*" (Wilson & Bracci, 1982, p. 41).

Fitness testing has used PES to graduate cadets to probationary cadet field training in the RCMP since 1992. Incumbent fitness testing, using the PARE job scenario with the current recruit PES and has now become almost universal amongst the federal police. Its genesis was the Wenger and Gaul incumbent officer PARE testing pilot project in 1992. With additional Certified Exercise Physiologists (CEP) hired in 1997/98, to enhance divisional compliance, PARE testing increased divisional compliance percentages amongst serving officers steadily above early voluntary self-determination programming The exception was the key area of protective policing positions where a change in policy requiring mandatory PARE at PSE 4:00 was the driver. Now PARE can be considered universally applied in all federal police divisions, with attendance being mandatory at PARE health promotion programming, although in test effort is mostly self-determined. The resulting divisional and national compliance vary from 50 to 80 percent compliance (RCMP, 2010).

The federal police in Canada are alone in using the same PES equally for recruits, as a BFOR, as well as for incumbent officers health promotion (Dunsmore & Hunter, 2001, Shell, 2003). Gaul and Wenger's (1992) conducted pilot work in the feasibility of mandatory health promotion PARE testing. Since that time PARE testing has evolved to become nearly universal for officer participation levels (Bonneau, 2001). The June 2014 compliance report cited a 70% compliance rate among federal police across Canada (Vincent 2014, personal communications). This has resulted in large data sets being available and captured covering the period since 2006 to 2014. Now that a generation of federal RCMP officers have completed periodic fitness testing (since 1997), a retrospective observational study of fitness testing records for predictors of fitness testing success is timely.

Previous fitness testing methodologies have all evolved as a result of multiple societal forces. The analysis of human factors, such as height, weight, age, and gender, many of which are protected human rights grounds, should be used to inform us and help avoid discrimination in our BFOR test studies. Similarly, the current two major constructs of the skill-related police abilities testing, should also undergo ongoing test evaluation for validity. Here validity includes Messick's (1996) social, moral, and political aspects of test use, not just test item construct development.

In federal policing, occupational fitness assessment is one of three parts of a mandatory health promotion program as described by Green and Kreuter (2005). It is delivered along with: (a) the legislated worksite health and safety committees, and (b) environmental safety audits and (c) periodic occupational medicals health surveillance; the latter as seen in police, fire, and ambulance services (Plat, Frings-Dresen, & Sluiter, 2011). Presently, this is delivered as a mandated best practice with other mandatory formative periodic police skills recertification as described by Shell (2003).

In federal policing, the periodical medical screening for occupational health in police work for in-service personal was originally published in *Police Health* by Trottier and Brown (1994). Officer occupational health exam and physical abilities test layouts were equally considered in the original publication of *Police Health*. The four objectives of periodic occupational health exams were described as assessing: (a) the individual's ability to do the job; (b) any threat to the safety of the public brought about by the interaction of the tasks of employment with disease or disability; (c) any threat to the safety of a co-worker; and, (d) any threat to the safety of the individual police officer" (Trottier & Brown, 1994, p. 40). The federal police occupational medicals have always played the role of gatekeeper and natural complement to abilities testing. The additional function of the periodic medical was established in June 2005, before mass testing began in earnest. Thus it became a universal comprehensive pre-screening for high or very high risk cardiovascular status, or equivalent risk, before the potentially maximally demanding PARE testing sessions -- as a best practice. In *Police Health* this was stated as the "physician who decides that the member suffered a condition and cannot or should not do the PARE is saying the individual cannot or should not do [operational] police work" (Trottier & Brown, 1994). See form 3986 in Appendix D.

The occupational medical examinations, termed periodic health assessments (PHAs) for federal police, have undergone numerous changes in 33 different iterations for question changes. In contrast, since its initial publication in *Police Health (*Trottier & Brown, 1994) the PARE underwent one human rights challenge by Genest resulting in one small change to the PARE protocol. The six foot running broad jump mat was shortened to a five foot mat (Bonneau, 2001). However, ongoing PARE equipment validity and reliability has seen major changes to the body control simulator test equipment, not to change to resistance level of the original protocol but to meet it reliably. The purpose was to reduce irrelevant task easiness or difficulty and the potential corresponding disparate group failures for shorter, taller, lighter, or heavier person's (Messick, 1995; Zumbo, 2001). The current day version of PARE, originally published in *Police Health* (Trottier & Brown, 1984) is in the unpublished government report titled *The Development of PARE: An Evolution* by Bonneau (2001). The current PARE Protocol and PARE Course Training Standard (CTS) approved by RCMP Learning and Development is published internally to accredited remote PARE sites and resides in the Fitness & Lifestyle archived files in Ottawa (PARE CTS, June 2014). Currently, certified PARE sites across Canada, from coast to coast to coast, use the document *The PARE: June 2014 Protocol* as their standardized instructions on equipment, test procedures, qualified personnel certification, and standard records reporting forms.

Bonneau (2001) documents the emergence of the PARE job simulation test from the original Canadian police test, the POPAT (Farenholtz & Rhodes, 1986, 1990; Trottier & Brown, 1994). POPAT acted as a genesis for the PARE and was based on common police task and demands analysis research studies. The PARE test development is examined in the next section.

1.18 The PARE Standard

The PARE standard has been part of RCMP human resources' practices for both recruiting and ongoing training for police officers for over a generation now. Since the early 1990s, all applicants entering the RCMP federal police training academy, called Depot, in Regina Saskatchewan, have had to pass PARE at the initial applicant standard of 4:45 min:s. After passing level one, the applicant's profile is considered for further processing and progress to enter depot. Applicants failing PARE at 4:45 min:s in the first three days of depot, after previously passing in their respective recruiting areas in the previous 6 months, are given additional attempts to meet the standard, although ultimately repeated failures are sent home. As graduates from the RCMP Depot training cadets completing the 24 week training program must pass PARE at 4:00 min:s with a slightly heavier push pull resistance setting (80 versus 70 lb) and a heavier bag carry section (100 versus 80 lb). PES of graduation is now 4:00 min:s rather than the initial applicant's 4:45 min:s. See Appendix A for PARE test layout. More recently, all general duty and protective program incumbent officers are scheduled for operational skills recertification which includes the PARE. Protective policing units currently require PARE 4:00 as the unit criteria to enter or stay in those federal public protective police units. Centralized record keeping of universal PARE testing data entry was facilitated by Human Resources Management Information Systems (HRMIS) in 2006 which in part facilitate staffing protective policing units with officer with PARE < 4:01 min:s.

1.19 PARE Testing Evolution

The current form of mandatory occupational testing has evolved from height and weight related standards for male applicants only (Bonneau & Brown, 1995; Maher, 1984), to health-related fitness testing interpreted separately for age and sex groups, to skill-related job scenarios interpreted for a pass or fail based on average police client scores (Farenhotlz & Rhodes, 1986). In Canada in the late 1980s, the Canadian Standardized Test of Fitness (CSTFA) was used as the health-related fitness test. Applicants were screened for a position in cadet training, with interpretations of scores based on sex and age normative tables. As mentioned above both applicant and incumbent police officer testing with PARE started in the early 1990's.

Most municipal police services where union have strong political sway, they resist universal incumbent fitness testing (Green & Kreuter, 2005), as a political issue, and focus on voluntary callouts for voluntary fitness recognition awards, like the Ontario Police Fitness Award. Bissett, Bissett and Snell, (2012) report survey results of most officers “indicated strong support of mandatory physical agility and fitness requirements for new recruits but not incumbents” (p. 1) and union positions reflect majority officer views. This presents a challenge when using fitness testing as health promotion for assessing low fit populations whom are unlikely to come out and fail a fitness recognition test (Wilson & Bracci, 1982). Second, the original PARE pilot projects were challenged by small sample sizes and low fit women participation and not enough older officers. Student studies of PARE predictors had the dual problem of low fit women volunteers as sample subjects, for reason not reported, and very small sample size. This has not helped provide predictive studies with sufficient power (Bonneau, 2001; Stanish et al., 1999). Finally, early PARE testing was fraught with equipment, protocol, standardization and medical clearance issues (Gaul & Wenger, 1992), limiting or casting doubt on any generalization of their findings to today’s highly standardized PARE test results.

The original studies to examine PARE predictors were challenged by small sample sizes and insufficient failures for one sex group so as not to be able to use logistic regression prediction with a dichotomous outcome variable of pass or fail (Stanish, Wood, & Campagna 1999). In a similar manner, most recruit selection systems are critiqued for not having universally applied the same testing for older officers on the job (Sharkey & Davis, 2008; Shephard & Bonneau, 2002). Second, only a small sample of studies focusing on health-related fitness capacity testing constructs were found to compare to passing scores on skill-related testing like the PARE or POPAT (Farenholtz & Rhodes, 1992; Stanish, Wood, & Campagna, 1999). Third, the initial PARE testing of incumbent officers’ pilot study had many and various standardization and protocol compliance issues such as slippery floors, and dysfunctional models of push pull machines, with artificially high officer medical clearance / restrictions issues (Gaul & Wenger, 1992). Officers would call in sick to avoid testing, when they were seen by a physician to avoid PARE testing, but without any police duty restrictions. Current policy has a clearance for PARE scenario training the same as a clearance for police work. They are mutually inclusive (RCMP, 2010).

Predictors of performance on the RCMP PARE is now possible from a large database of incumbent police officer scores encompassing some 30,000 cumulative data records. The data records can yield approximately 14,000 unique best PARE scores, one per officer when sorted, for fastest times with lower times deleted. Therefore, a retrospective observational study of a generation of incumbent RCMP officer's skill-related fitness tests can potentially create a quantitative prediction model of success for significant human factors from all three historical periods of fitness testing. Within these studies, human factors from previous testing should be evaluated as covariates, or in the cases of sex differences, as group factors.

1.20 Theory: Metabolism, Fatigue and Human Bioenergetics

Total maximal power output in an all-out or sustained high level effort demonstrates sustained power drops resulting from the lag of the three dominant bioenergetics systems to provide peak muscular power in children (Wells, Selvadurai & Tein, 2009), and in athletes, and is well documented in anaerobic Wingate testing for sports and performance (Gore, 2000; MacDougal, Wenger & Green, 1991). The transition from one dominant muscular metabolic energy system pathway to another, as the dominant source of energy, predicts whole body muscular power drops for tests of near maximum efforts, if the test is over a time course of two to three minutes (Brooks & Fahey, 1986; Wells, 2009). This bioenergetics energy systems theory sets the expectations for sustained performance time pacing decrements predictions and informs the current studies.

Wells, Selvadurai, and Tein (2009) review the three major metabolic pathways of human bioenergetics for muscular energy production. Use of the three pathways are time and intensity dependent. The three system are described as: (a) the aerobic oxidative systems for "longer duration activities of low to moderate intensities", (b) the aerobic glycolytic system for "short to moderate duration activities of higher intensity", and (c) the higher energy phosphate system for "short duration activities of high intensity" (Wells et al., 2009, p.83). Sport testing using Wingate-type tests of maximal or near maximal efforts that extend beyond the 30 seconds can show the rate of power drop theorized from each dominant energy delivery system, and has traditionally shown sustained power drops (Bar-Or, 1987).

Supplies of energy for resynthesized adenosine triphosphate (ATP) for use in high power muscle work is supplied quickly from free phosphate (Pi) and creatine phosphate (CP) plus available adenosine diphosphate (ADP), from the original breakdown of ATP in the muscle cytosol. Energy is liberated from CP stores via breakdown into creatine and Pi, with the energy supplying ATP resynthesis in a quick process. This system dominates for about 10 seconds at peak power and is usually spent at about 30 s.

Muscular fatigue is described by Brooks and Fahey (1986) as both central and peripheral and is described as “task specific, and its causes are multifocal” (p. 701). First fatigue could be local to the muscle as in ATP, CP depletion from high intensity exercise, where CP stores are insufficient to resynthesis ADP and Pi “CP depletion leads to muscle fatigue” where ATP stores are also in “depletion” (p. 704).

Simultaneous to this alactic process above, within muscle anaerobic glycolysis of six carbon sugars into three carbon lactic acid provides peak energy at about 20 s to 90 s. After 90 s, lower power but peak aerobic power processing of three carbon sugars and oxygen through the final electron transport in muscle cell mitochondria powers muscle cells ATP resynthesis. This process is slower, with lower rates of power for prolonged periods of time. It reaches equal production of power rates, with anaerobic power production, in all out test at 90 s or a little later for fixed paced but not maximal full out effort (Gore, 2000). Gore (2000) states “it is important to realize that energy delivery is achieved through a sequentially overlapping involvement of all three energy systems” (p. 45). The energy system contributions to a pacing effort, they have “relative contributions [that] vary over time in a coordinated metabolic response to the demands of exercise” (Gore, 2000, p. 45). Anaerobic contributions therefor bridge the gap between slower developing and lower aerobic power delivery as anaerobic energy production lags from peak levels by some times over 90 s to three minutes depending on the intensity of the external demand (Gore, 2000). “The minimum duration to exhaust anaerobic contributions to energy production" is seen by Gore (2000) in constant intensity test "as 120 s" (p. 47). Peak combined power output is shown as dropping, in a greater relative proportion to later power drops, in the first few seconds of exercise in a constant effort test (Gore 2000).

Fatigue in the anaerobic system for a 2 to 5 minute event is not from glycogen substrate depletion, as moderate levels of exercise can be maintained for hours (Brooks and Fahey, 1986). Also discounted are blood sugar depletion in short time bouts of exercise. In short high intensity bouts of exercise, the dissociated H^+ ion from lactic acid and other metabolic accumulation from rapid glycolytic breakdown of muscle glycogen can drop pH and inhibit the rate limiting step of glycolysis by inhibiting Phosphofructokinase (PFK). H^+ may also displace Ca^{2+} from "troponin thereby interfering with muscle contraction" (Brook & Fahey, 1986, p. **7**07). Other effects of elevated $H^{+,}$ lowered pH is to "inhibit the combination of O_2 with hemoglobin" (Brooks & Fahey, 1986, p. 707). Brooks and Fahey (1986) suggest NMR studies give weight to CP depletion over lactic acid accumulation as source of metabolic fatigue. Further the sequestration of Ca^{2+} by the mitochondria may uncouple oxidative phosphorylation of ADP to ATP. Inadequate oxygen may cause both CP levels to drop and lactic acid and hydrogen ions to accumulate. A rapid muscular O_2 stores depletion without sufficient central circulatory repletion in either low oxygen environments (hypoxia of altitude) or in diminished circulatory transport, like anemia, may limit oxygen hemoglobin affinity. Low or slow circulatory transport to muscles, due to mechanical limitations of blood flow or other causes, like certain tuck postures, may limit timely repletion of ATP by oxidative phosphorylation in high muscular work demands, limiting exercise to a reduced pace, till full circulation is restored – making mechanical posture another possible cause of fatigue. Previous training status overload without adequate recovery could see this in normally high fit high metabolic capacity person. This is not likely in a high fit person working at a low relative percentage of their maximum, but may also be present in a low fit person working, even for a couple of minutes at VO_{2max} levels (Brooks & Fahey, 1986).

Central nervous system fatigue could be implicated in any of the “proper functioning of receptors, CNS [central nervous system]¸integrating centres, sensory cortex, spinal cord, α-motor nerons, γ loop, motor ends plates” (Brooks & Fahey, 1986, p. 709). Depletion of Ca^{2+} cycling or other neural transmitters, may slow muscle function and power output.

Abbiss and Laursen (2008) defined the “distribution of work, or pattern of energy expenditure” as “pacing or pacing strategy” (p. 240). The issue here is how participants “regulate their work output in order to optimize overall performance” (Abbiss & Laursen, 2008, p, 240). The term “pacing more accurately refers to performance times or velocities” (Abbiss & Laursen, 2008, p. 240) and in the case of police circuit test time per lap of a circuit. The “regulation of pacing is largely determined by the ability to resist fatigue” (Abbiss & Laursen, 2008, p. 240). The difference between all out short duration sprints of less than 60 s, is contrasted by the authors with > 2 min paced endurance performance events. Strategies used include, negative, all out, positive and parabolic shaped and variable pacing strategies. Each of these strategies are described below from Abbiss and Laursen’s 2008 review article.

Negative pacing, usually called a negative split in track and field, see the second half of a distance run faster than the first half. The negative split strategy is seen as an increase in motor unit recruitment near the end of performance using any stores of anaerobic energy supply left, finishing just before metabolic accumulations limit performance during the final surge or sprint. All-out pacing is common for short duration sprint events, of less than 30 – 60 s, with the optimal distance cited by Keller (1974) as 291 m. where the cost of breaking inertia is beast paid early and lesser submaximal effort is used to hold a slight slowing pace as long as possible (Abbiss & Laursen, 2008). Holding an all-out pacing effort for greater 45 seconds has been found in Wingate anaerobic power tests to be impossible (Gore, 2000) with everyone slowing down after 10 s. Positive pacing is defined as when a performer slowly and gradually declines throughout a race, in contrast to negative pacing. This strategy is reportedly responsible for demonstrating the highest: fractional utilization of maximal aerobic power (VO_2 max.), the highest post-blood lactate concentrations, the highest respiratory gas exchange ratios, and the highest perceived exertions. – compared with even split pacing. Self-selected pacing in distance events tends to be of this nature. Even pacing is self-evident, and most often observed in the longer hours plus events. Parabolic-shaped pacing efforts represent high initial paces, followed by slower intermediate or middle event pacing, finishing with an event-ending surge of effort, as in a negative split. Performance profiling during distance events has seen athletes “progressively reduce speed during an endurance trial but tend to increase speed during the latter portion of the event” (Abbiss & Laursen, 2008, p. 245). Commenting on the usual J shaped faster start and finish of parabolic shaped pacing, Abbiss and Laursen comment that “the choice of pacing strategy does not appear to be dictated by changes in any one physiological system, but instead may be influenced via a complex system of integrated feedback from a number of sources, including prior experience and anticipated duration” (p. 246). Variable pacing is most common with varying external environmental conditions, as changing resistance during an event, like varying wind speeds, and terrain incline changes over a route, where it is considered an optimal strategy to vary pacing accordingly. Physiological monitoring has shown this to produce more accumulated metabolic fatigue agents than constant pacing strategies (Abbiss & Laursen, 2008).

1.21 Fatigue and Power Selection

Weir, Beck, and Housh (2006) states that power output as pace selection and the level of fatigue are responses solicited from a maximal effort that are driven by higher order cognitive functions in order to maintain homeostasis, when fatigue is explained by the central governor model. Running Canada’s website (http://runningmagazine.ca/) call one of these strategies the teleoanticiaptory effects of looking to the end of the run and anticipating how much of a pace can be sustained without completely failing with fatigue. This is critiqued by Weir et al. (2006) in comparing equal power outputs in 30, 33 and 36 s Wingate tests were teleoanticipation would have seen dropped power outputs in the longer tests. The authors cite Bigland, Ritchie, and Woods’ (1984) operational definition of fatigue as “an exercise-induced reduction in the ability to exert muscular force or power, regardless of whether or not the task can be sustained” (Wier et al., p. 574).

1.22 Human Factors and Ergonomics

Parallel to Kuruganti and Rickards' (2004) statement of "fit the job to the worker rather than fitting the worker to the job" (p. 455), human factors engineering and ergonomics can be used in PARE testing instruction to: (a) fit the BFOR test equipment to the worker, and (b) use ergonomic principles and human factors findings to reflect a more valid and simplified test construct for success. Common clinical teaching practices for PARE instruction that facilitate employment equity physical literacy instruction uses the ergonomics of manual material handling cues and optimal push pull ergonomic concepts. Every individual completing a PARE session has individualized instruction to compensate for human factors of height, weight, to advise optimal core and back muscle recruitments for safety and ergonomic factors for the push, pull, lift, and carry sections of PARE (DARCOR and ERGO web, 2103). Five concepts drawn from ergonomics and biomechanics used as teaching individual teaching cues in the PARE are: (a) in the push there must be greater than 50% coefficient of friction with the floor-shoe contact (Anders et al., 1983; Boocock, Gaul & Wenger, 1992; Haslam, Lemon, & Thorpe, 2006; Kroemer, 1971); (b) optimal push handle stability (no orthogonal plane rotation of push handles); (c) optimal grip height levels, while leaning forward, should be level with standing iliac crest level (Chaffin, Herrin, Lee & Waiker, 1991; Ceriello & Snook, 1991; Seo & Armstrong, 2009; Seo, Armstrong & Young, 2010); (d) on pushing full forward lean with both legs back acting as primary movers, with feet well back of push handles, (Cnyrim, Merger & Maurer, 2009); (e) chest kept low for peak forces push posture and arms slightly bent at the elbows (Ayoub & McDaniel, 1974); and (f) core activation of abdominals flexors and back extensors, plus transverse abdominals, to avoid a standing up posture with unsafe hyperextension of the low back in an attempt to maintain force against the resistance stack (Knapik & Marras, 2009).

1.23 Physical Activity Assessment

Welk (2002) reviewed the assessment of physical activity for health and reviews the definition of physical activity as "any bodily movement produced by skeletal muscles that result in caloric expenditure" (p. 4). The exercise definition was seen in a more narrow term: "physical activity that is planned, structured, repetitive, and results in the improvement or maintenance one or more facets of physical fitness" (Welk, 2002, p. 4). Physical fitness is termed "a set of outcomes or traits that relate to the ability to perform physical activity" (Welk, 2002, p. 4). Welk (2002) addresses the need for large samples and avoidance of "considerable error in the assessment" of physical activity (p.6). The advantages and disadvantages of the major methods of various assessment methods were reviewed, namely self-reports, activity monitors, heart rate monitors, pedometers, and direct observation. Self-reporting through direct observation, as in the Anderson et al., (2001) police task analysis studies, was considered the gold standard for PA assessment, but was considered labour intensive, requiring funded observers, and a large portion of time. Cheaper and less labour intensive measures included using objective data gathered by equipment added to the human body. The limited quantitative data of physiological measures or mechanical factors, like heart rates recorded or step counted, sometimes lost the mode of activity and was then dependent on self-reported PA recall for type of movement. PA recall of a generic nature was considered less effective than episodic recalls (Welk, 2002).

1.24 Literature Review

A literature review was conducted in Embase, CINAHL, PubMed, and PsychInfo using the following search phrases with inclusive operant "OR" for the following terms: "physical fitness OR health OR exercise OR exertion". To link the many studies to just the target population of police, the exclusive operant "AND" was used to limit studies to inclusive key words "law enforcement" OR "police" populations. All searches were limited to English. Due to the scant amount of academic peer reviewed journal articles on BFOR tests, government publications were also gleaned for the topics areas of BFOR task analysis, BFOR test methodology, officer perspectives of mandatory police fitness assessments, and PARE score predictor studies at the Ontario Police College, the Canadian Police College Library and the RCMP Fitness and Lifestyle Archives of papers in Ottawa. Reference lists of articles were consulted for additional relevant publications from journals, conferences, government sources, and/or media releases.

Contrasts between physical activity levels and physical fitness outcomes have been reported in police populations academic literature for low back pain (Henweer, Picavet, Staes, Kiers, & Vanhees, 2011), cardiovascular disease (Franke & Anderson, 1994; McMurray, Ainsworth, Jarrell, Griggs, & Williams, 1998; Sassen, Cornelissen, Kiers, Wittink, Kok, & Vanhees, 2009; Sassen, Kok, Schaalma, Kiers, & Vanhees, 2010), and musculoskeletal injuries (Nabeel, Baker, McGrail, & Flottemesch, 2007).

1.25 Predicting Success: Group Factors

An example in the literature of a large police population study that looked at success and performance in a skill-related police fitness tests with contrasting groups, including sex groups, was the work of Strating et al. (2010). This large Dutch police officers population, (n = 7,000) study used a “police fitness test” similar to PARE. Their reported findings were a lower fitness level with several officer group contrasts including: younger (< 40 years), between 40 and 55 years, and older (> 55 years), male versus female, and operational policing versus administrative duties groups, with administrative split into investigating and all other duties groups. Women as a group, were 26 seconds slower than men, and this was reported as reported as statistically significant, with an effect size of 1.15 -- without an effect size statistic name. Officers with high BMIs and lower physical activity profiles, scored slower for both male and female officers. Officers in “core police tasks” had significantly faster times at a significance level of $p < .01$ than officers in the “reaming function group” (Strating et al., 2011, p. 257). Listed as significant for regression factors were age, BMI, functional work group, and hours of physical exercise, all significant at $p < .001$. Both sexes were assessed against a single sex neutral job test/task specific standard. This study showed a limited sex group analysis for skill-related tests, however it helps us in predicting differences for factors of age, BMI, type of police worker, and hours of PA.

1.26 PARE Incumbent Pilot Project

The RCMP pilot incumbent PARE testing and training project by Gaul and Wenger (1992) was conducted from 1990 to 1992 in several urban and rural divisional detachments in Canada. Results from initial test to the two year follow-up testing reported overall combined men and women changes for passing from initial pass rates of 71% to a final pass rate of 78%. One subgroup, women, had only a 14% initial pass rate, but additional testing opportunities during the two years, with some training support had this figure increase to 48%. Gaul and Wenger (1992) attributed most of the difficulty for women at the push pull station. The under 30 years of age subgroup combined male and female pass rate was 80% after two years. The 31 to 40 years age group had similar results with up to 75% passing. The 41 to 50 years age group had only 75% under 4:30 min:s compared to the 4:00 min:s pass mark. Insufficient over 50 years of age group study participant's data were available to provide group results. All sex and age groupings of incumbents officers were tested using pass as under the 4:00 (min:s) standard at various times over the two year period.

1.27 Predictors of Success in PARE

Predictors of success in the RCMP's PARE has limited academic literature. Stanish, Wood, and Campagna (1999) contrasted health-related fitness component test scores to PARE scores. "The purpose of this study was to identify valid and practical field tests of physical fitness that accurately classify successful and unsuccessful PARE performers" (Stanish et al., 1999). Fitness test constructs of aerobic power, anaerobic power, muscular strength and endurance, and body composition were contrasted with PARE pass scores (4:45) for 28 students and 20 applicants. A 3-variable model was found in multiple regression analyses to be predictive of PARE score success with a 70 lb bench press, standing long jump, and agility "explaining 79% of the variance" for men (Stanish et al. 1999, p. 2). The authors reported that a small percentage of male scores were failures (9%), making the "classification" of pass/fail problematic for their proposed logistic regression analysis which requires sufficient dichotomous variable data to fit the sigmoid curve model. However, female data had sufficient dichotomous data (pass versus fail) to attempt to determine a relationship between pass / failure and health related fitness variables. The authors reported two variables as predictive of PARE score success in the women's group: namely the 1.5 mile aerobic power endurance run and the agility test, with "93% overall classification accuracy" (Stanish et al., 1999, p. 2). It should be noted though that the PES standard used here was the initial applicant entry time of 4:45 min:s to the RCMP applicant process, not the Cadet graduation standard of 4:00 min:s. She did however suggest future research of "partitioning the PARE score into run and fight components may provide more precise information regarding the contribution of these components to successful PARE performance" (p. 675).

With the recent focus on gender and respect (Gender and Respect: the RCMP Action Plan, 2012) within the RCMP and parliamentary committees looking at gender differences in assessment and women's advancement, it is not only timely to conduct studies finding significant predictors of success, but to also include gender differences if any, in this analysis. No studies were found comparing skill-related third moment occupational fitness testing constructs of pursuit and body control times to human factors of height and weight, or BMI, age, or sex to total PARE scores or PARE pass/fail success.

1.28 Purpose

The overall purpose of this dissertation was to build a quantitative predictor model of factors from the three different moments of occupational fitness assessment to predict success in the RCMP's PARE test. Specifically, the purpose of study 1 was to assess the relative predictive power of the pursuit and body control segment times for PARE success and assess the relative difference across sex groups while controlling for human factors of height, weight, and age. The purpose of study 2 was to examine the effect of pacing in the six laps pursuit circuit on success in PARE for sex groups and pass or fail performance groups, controlling for height, weight, and age. The purpose of study 3 was to add self-reported physical activity frequency and intensity factors as potential predictors of PARE success to the quantitative prediction model.

1.29 References

Abbiss, C. R., & Laursen, P. B. (2008). Describing and understanding pacing strategies during athletic competition. *Sports Medicine*. 38(3): 239-252. doi: 0112-1642/08/003-0329

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M.L., Swartz, A. M., Strath, S. J., … Leon, A.S. (2000). Compendium of physical activities: An update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise*, *32*(9), S498-S516. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Anderson, G. S., & Plecas, D. (2007). *PARE Task Analysis*. RCMP Learning and Development Webpage, Ottawa, ON: Retrieved from RCMP internal Learning and Development intranet web page archives.

Anderson, G. S., & Plecas, D. (2008). *Police abilities testing: Revalidation of the task analysis*. Ottawa, ON: RCMP Learning and Development Archives.

Anderson, G. S., Plecas, D., & Segger, T. (2001). Police officer physical ability testing: Re-validating a selection criterion. *Policing: An International Journal of Police Strategies and Management, 24*, 8-31. Retrieved from http://dx.doi.org/10.1108/13639510110382232

Andrews, A., & Risher, J. (2006). What does that have to do with being a cop? Employment standards in law enforcement. *International Association of Chiefs of Police Conference,* Boston, MA. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=187913

Astrand, P., & Radahl, K. (1986). *Textbook of work physiology: Physiological basis of exercise*. McGraw Hill, New York.

Avery, R. D., Nutting, S. M., & Landon, T. E. (1992). Validation strategies in physical ability testing in police and fire settings. *Public Personnel Management*, *21*(3), 301-312. Retrieved from http://ppm.sagepub.com/

Avery, R. D., Landon, T. E., Nutting, S. M., & Maxwell, S. E. (1992). Development of physical ability tests for police officers: A construct validation approach. *Journal of Applied Physiology, 77*(6), 996-1009.

Bar-Or, O. (1987). The Wingate anaerobic test: An update on methodology, reliability and validity. *Sports Medicine, 4*, 381-394. doi:10.2165/00007256-198704060-00001

Bissett, D., Bissett, J., & Snell, C. (2012). Physical agility test and fitness standards: Perceptions of law enforcement officers. *Police Practice and Research: An International Journal, 13*(3), 208-223. doi.org/10.1080/15614263.2011.616142

Birzer, M. L., & Craig, D. E. (1996). Sex differences in police physical ability test performance. *American Journal of Police, 15(*2), 93-108. doi:10.1108/07358549610122494

Boocock, M. G., Haslam, R. A., Lemon, P., & Thorpe, S. (2006). Initial force and postural adaptations when pushing and pulling on floor surfaces with good land reduced resistance to slipping. *Ergonomics*, *49*(9), 801-882. doi:10.1080/00140130600562876

Boni, N. (2004). *Exercise and physical fitness: The impact on work outcomes, cognition, and psychological well-being for police*. Marden, South Australia: Australasian Centre for Policing Research.

Bonneau, J. (1996). *PARE: Standard summary presentation to the Canadian Human Rights Commission*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J. (2001): *The development of PARE: An evolution*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J., & Brown, J. (1995). Physical ability, fitness and police work. *Journal of Clinical Forensic Medicine*, *2*, 157-164. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Brown, J., & Trottier, A. (1994). *Police health: A physician's guide for the assessment of police officers*. Ottawa, ON: Canada Communications Group.

Brooks, G. A., & Fahey, T. D. (1986). *Exercise physiology: Human bioenergetics and its applications*. New York, NY: MacMillan.

Boyce, R., Jones, G., Lloyd, C., & Boone, E. (2008). A longitudinal observation of police: Body composition changes over 12 years with sex and race comparisons. *Journal of Exercise Physiology, 11*(6), 1-14. doi:10.1093/occmed/kql057

Boyce, R. W., Jones, G. R., Schendt, K. E., Llyod, C. L., & Boone, E.L. (2009). Longitudinal changes in strength of police officers with gender comparisons. *Journal of Strength and Conditioning Research, 23*(8), 2411-2418.

Canadian Society for Exercise Physiology (1986). *Canadian standardized test of fitness (CSTF): Operations manual (3 rd Ed.).* Ottawa, ON. Canada Communications Group.

Casson, E. J. (2000). Establishing BFOR and validated standard for vision. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Canadian Society for Exercise Physiology. (2012). *The Canadian Physical Activity, Fitness & Lifestyle Appraisal: CSEP's Plan for Healthy Active Living* (2nd Ed.). Ottawa, ON: Canadian Society for Exercise Physiology.

Collingwood, T. R. (1988a). Implementing programs and standards for law enforcement physical fitness. *The Police Chief,* April *55*(4*),* 20-24. Retrieved from Canadian Police College Library.

Collingwood, T. R. (1988b). Physical fitness leadership in law enforcement. *The Police Chief*, April *55*(4), 28-34. Retrieved from Canadian Police College Library Ottawa.

Collingwood, T. R. (1995, February). Physical fitness standards: Measuring job relatedness. *The Police Chief, 62,* 31-31. Retrieved from Canadian Police College Library Ottawa.

Collingwood, T. R. (2004, March). Underlying physical fitness factors for performing police officer physical tasks. *The Police Chief, 71*(3), 1-7. Retrieved from Canadian Police College Library, Ottawa.

Cynrim, C., Mergner, T., & Maurer, C. (2009). Potential role of force cues in human stance control. *Experimental Brain Research*, *194*(3), 410-433. doi:10.1007/s00221-009-1715-7

Darcor and Ergoweb. (2013). The ergonomics of manual material handling: Pushing and pulling tasks. *Sutton Javelin Communications for Darcor Limited*. Retrieved from http:/www.darcor.com/technical/ergonomics of manual material handling/

Deakin, J. M., Smith, J. T., Pelot, R. P., & Weber, C. L. (2000). Methodological considerations in the development of physical maintenance standards. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Deakin, J. M., Smith, J. T., Pelot, R. P., Weber, C. L., & Stevenson, J.M. (1994). *Development of a bona fide physical maintenance standard for CF and DND firefighters.* Kingston, ON: Ergonomics Research Group, Queens University. Retrieved from Ontario Police College Archives.

Denzin, N. K., & Lincoln, Y.S. (Eds.). (2005). *The Sage handbook of qualitative research* (3 rd Ed.). Sage, Thousand Oaks, CA.

Dunsmore, R. R., & Hunter, A. J. (2001). Maneuvering the obstacle course: Implementing incumbent physical fitness testing in the workplace. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Eid. E., & Geh, S. (2001). Challenges posed by the Supreme court of Canada in the Meiorin decision to employees in physically demanding occupations. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Equal Employment Opportunity Commission. (1978). *Uniform guideline on employee selection procedures*. Retrieved from www.uniformguidelines.com

Farenholtz, D. W., Rhodes, E. C. & Burrell, A. J. (1983). *Development of physical abilities test for correctional officers in British Columbia.* Unpublished paper retrieved from the Canadian Police College Archives.

Farenholtz, D. W., & Rhodes, J. M. (1992). *Comparison of the police officer's physical abilities test to selected laboratory and field tests of physical fitness.* New Westminster BC: Justice Institute of BC. Unpublished paper retrieved from the Ontario Police College.

Farenholtz, D. W., & Rhodes, E. C. (1986). *Police Officer Physical Abilities Study*. New Westminster BC: Justice Institute of BC. Unpublished retrieved from the Canadian Police College Library archives. .

Farenholtz, D. W., & Rhodes, E. C. (1990). Recommended Canadian standards for police physical abilities. *Canadian Police Journal, 14*(1), 37-49. Retrieved from http://www.questia.com/library/p90673/canadian-journal-of-police-and-security-services

Fleg, L., Morrell, C. H., Bos, A. G., Brant, L. J., Talbot, L. A., Wright, J. G., & Lakatta, E. G. Accelerated longitudinal decline of aerobic capacity in healthy adults. *Circulation*, *112*, 674-682. Retrieved from http://circ.ahajournals.org/cgi/content/full/ CIRCULATIONAHA.105.545459/DC1 doi: 10.1161/CIRCULATIONAHA.105.545459

Frank, W. D., & Anderson, D. F. (1994). Relationship between physical activity and risk factors for cardiovascular disease among law enforcement officers. *Journal of Occupational Medicine, 36*(10), 1127-1132. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Gaul, W., & Wenger, H. (1992). *RCMP Physical Abilities Requirement Evaluation demonstration project: Final report.* Victoria, BC: University of Victoria, School of Physical Education. Retrieved from RCMP Archives.

Gillis, A. D., & Darby, B. (2001). To serve and protect, post Meiorin an argument for due diligence is an argument for undue hardship: The liability labyrinth faced by the public safety in recruitment and retention of competent employees. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Girard, G. (2012). *Five-year PARE compliance report.* Ottawa, ON: RCMP Learning and Development internal communications.

Gledhill, N., & Shaw, C. (1996). *Report on the characterization, test construction and validation phase of the medical, physical, skills and abilities project.* Toronto, ON: Race Relations and Policing Unit of' the Ministry of the Solicitor General and Correctional Services.

Gledhill, N., & Bonneau, J. (2001). Objectives, processes, and consensus summary of the national forum on establishing Bona Fide Occupational Requirements. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Gledhill, N., Bonneau, J., & Salmon, A. (2001). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Gore, C. J. (Ed.). *Physiological tests for elite athletes*. Windsor, ON. Human Kinetics.

Green, L. W., & Kreuter, M. W. (2005). *Health promotion planning: An ecological and educational approach*. New York, NY: McGraw Hill.

Hawkins, S. A., & Wiswell, R. A. (2003). Rate and mechanism of maximal oxygen decline with aging. *Sports Medicine, 33*(12), 877-888. doi: 0112-1642/03/0012-0877/

Hoffman, B. (1996). How today's law-enforcement officers rank. *Joe Weider's muscle and fitness*, *76*, 232-235. Retrieved from http://www.muscleandfitness.com/news-and-features/athletes-and-celebrities/life-joe-weider

Hoffman, R., & Collingwood, T.R. (2005). *Fit for Duty: An officer's guide to total fitness.* (2 nd Ed). Champaign, Ill: Human Kinetics.

Hogan, J., & Quigley, A.M. (1986). Physical standards for employment and courts. *American Psychologist,* 41, 1193-1217. Retrieved from http://www.apa.org/pubs/journals/amp/

Howell, D.C. (1997). *Statistical Methods for Psychology*. Belmont, CA: Cengage Learning.

Jackson, A.S. (1994). Pre-employment physical evaluation. *Exercise and Sports Science Reviews, 22*(1), 53-90. Retrieved from http://journals.lww.com/acsm-essr/pages/default.aspx

Jackson, C. A., & Wilson, D. (2013). The gender-neutral time obstacle course: A valid test of police officers? *Occupational Medicine*, *63*, 479-484. doi:10.1093/occmed/kqt102

Jackson, A. S., Beard, E. F., Wier, L.T., Ross, R.M., Stutville, J.E., & Blair, S.N. (1995). Changes in aerobic power of men, ages 25-70 yr. *Medicine & Science Sports & Exercise*, *27*(1), 113-120.

Jamnik, V. K., Thomas, S. G., Shaw J.A., & Gledhill, N. (2010a). Identification and characterization of the critical physically demanding tasks encountered by correctional officers. *Applied Physiology, Nutrition and Metabolism*, *35*, 45-58. doi: 10.1139/H09-121

Jamnik, V. K., Thomas, S. G., Burr, J. F., & Gledhill, N. (2010b). Construction, validation, and derivation of performance standards for a fitness test for correctional officer applicants, *Applied Physiology, Nutrition, and Metabolism*, *35*, 59-70. doi: 10.1139/H09-122

Jamnik, V. K., Thomas, S. G. & Gledhill, N. (2010c). Applying the Meiorin decision requirements to the fitness test for correctional officer applicants: Examining adverse impact and accommodation. *Applied Physiology, Nutrition and Metabolism*, *35*, 71-81. doi: 10.1139/H09-123

Keith, T. Z. (2006). *Multiple regression and beyond*. New York, NY: Pearson.

Keller, J. B. (1974). Optimal velocity in a race. *American Mathematics Monthly. 81*(5): 474-80.

Kinnear, P. R., & Gray, C. D. (2010). *PASW 17: Statistics made simple*. New York, NY: Psychology Press.

Kleinbaum, D. G., & Klein, M. (2002). *Logistic regression: A self-learning text*. New York, NY, Springer.

Knapik, C. G., Spiess, A., Swedler, D., Grier, T., Hauret, K., Yoder, J, & Jones, B.H. (2011). Retrospective examination of injuries and physical fitness during Federal Bureau of Investigation new agent training. *Journal of Occupational Medicine and Toxicology*, *6*(26), 1-11. doi: 10.1080/00140139.2011.645887.

Knapik, C. G., & Marras, W. S. (2009). Spine loading at different lumbar levels during pushing and pulling. *Ergonomics*, *52*(1), 60-70. doi:10.1080/00140130802480828

Knapik, J. J., & Sharp, M. A. (1998). Task-specific and generalized physical training for improving manual-material handling capability. *International Journal of Industrial Ergonomics, 22*(3), 149-160.doi:http://dx.doi.org/10.1016/S0169-8141%2897%2900003-6

Koller, M., Gledhill, N., Payne, N., Jamnik, R., Fisher, K., Gordon, C., & Halligan, J. (1999). Evaluation of a new force meter for assessing upper body strength. *Medicine & Science in Sports and Exercise*, *31*, S114. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Kuruganti, U., & Rickards, J. (2004). The role of human factors engineering in establishing occupational fitness standards. *International Journal of Industrial Engineering*, *34*(6), 451-457. doi: .1016/j.ergon.2004.05.004

Lagasse, P. P. (1989). *Assessment of Physical Standards for Admission in the Police Forces*. [Élaboration des normes physiques d'admission aux d'agents de la paix]. Unpublished manuscript, Faculty of Physical Education, Laval University, Quebec City, Canada.

Legault, J. P. (2000). Establishing a Bona Fide Occupational Requirement for occupational requiring a hearing abilities. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the consensus forum on establishing BONA FIDE requirements for physically demanding occupations*. Toronto, ON: York University.

Lemura, L. M., Von Dunvillard, S. P., & Mookerjee, S. (2000). The effects of physical training of functional capacity in adults: Ages 46-90: A meta-analysis. *Journal of Sorts Medicine and Physical Fitness, 40*, 1-10.

Lonsway, K. A. (2003). Tearing down the wall: Problems with consistency, validity, and adverse impact of physical agility test in police selection. *Police Quarterly*, *6*, 237-276. doi: 10.1177/1098611103254314/

Maher, P. T. (1984). Police physical ability tests: can they ever be valid. *Public Personnel Management,* 32(2), 173-178. Retrieved from http://ppm.sagepub.com/

MacDougal, J. D., Wenger, H. A., & Green, H. G. (1991). *Physiological testing of the elite athlete*. Champaign, Ill: Human Kinetics.

Major Cities Chiefs of Police. (2008). Promoting health and wellness: Returning to full duty. Report of the Major Cities Chiefs and Federal Bureau of Investigation National Executive Institute. Retrieved from www.neiassociates.org/stoarge/PromotingHealth-fullduty.pdf

Marchand, R., Reed, A. Thoden, J. S., & McNeely, E. (1995). *Physical screening protocol for correctional officer applicants.* Toronto, ON: Ministry of the Solicitor General and Correctional Services.

McMurray, R. G., Ainsworth, B. E., Harrell, J. S., Griggs, T. R., & Williams, O. D. (1998). Is physical activity or aerobic power more influential on reducing cardiovascular disease risk factors? *Medicine & Science in Sports & Exercise,* 30(10), 1521-1529. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts*. Thousand Oaks, CA, Sage.

National Executive Institute Association. (2008). *Promoting health and wellness: Returning to full duty* Proceedings of the Toronto 2008 Conference of the Major Cities Chiefs and the Federal Bureau of Investigation. Retrieved from www.neiassociates.org/PromotingHealth-fullduty.pdf

Nabeel, I., Baker, B. A., McGrail, M. P., & Flottemesch, T. J. (2007). Correlation between physical activity, fitness, and musculoskeletal injuries in police officers. *Minnesota Medicine, 90*(9), 40-43. Retrieved from http://www.minnesotamedicine.com/Past-Issues/Past-Issues-2007/September-2007/Clinical-Nableel-September-2007

Nutbeam, D. (1998). *Health promotion glossary*. World Health Organization. Retrieved from http://www.who.int/hpr/NPH/docs/hp_glossary_end.pdf

Ontario Ministry of Public Safety and Security. (2002). *The Ontario constable selection system: Technical guide for users of the Physical Readiness Evaluation for Police the PREP*. Toronto, ON: Queens's Printers.

Osborn, G. D. (1976). Physical agility testing: Validating physical agility tests. *The Police Chief*, January, 45-50. Ottawa: ON. Retrieved from Canadian Police College Library.

Physical Education Canada. (2013). *What is physical literacy?* Retrieved from http:/phecanada.ca/programs/physical-literacy/what-physical-literacy.

Plat, M. J., Frings-Dresen, M. H. W., & Sluiter, J. K. (2011). A systematic review of job-specific workers' health surveillance activities for fire-fighting, ambulance, police and military personnel. *International Archives of Occupational Environmental Health*, *84*(8), 839-857. doi: 10.1007/s00420-011-0614-y

Police Fitness Personnel of Ontario. (2012). *Certified Fitness Appraiser's Manual.* Aylmer, ON: Ontario Police College.

Pollock, M. L., Gellman, L. R., Price, C. L., & Kent, D. A. (1977). *Physical fitness program for law enforcement officer: A manual for police administrators.* Report of the International Chiefs of Police. Retrieved from the Canadian Police College.

Royal Canadian Mounted Police. (2010). *Physical Abilities Requirement Evaluation (PARE) Course Training Standard (CTS).* Ottawa, ON: RCMP Learning and Development.

Royal Canadian Mounted Police. (2012). Gender and Respect: The RCMP Action Plan *National Program Evaluation Services Royal Canadian Mounted Police*. Ottawa, ON: Retrieved from RCMP at http://www.rcmp-grc.gc.ca/gba-eces/action/gba-eces-action-eng.pdf

Reed, A. (1998). The oxygen cost of running the Physical Ability Requirement Evaluation. In Bonneau (2001). *The development of PARE: An evolution*. Ottawa, ON: RCMP Learning and Development Archives.

Sassen, B., Cornelissen, V. A., Kiers, H., Wittink, H., Kok, G., & Vanhees, L. (2009). Physical fitness matters more than physical activity in controlling cardiovascular disease risk factors. *European Journal of Cardiovascular Prevention and Rehabilitation*, *16*, 677-673. doi: 10.1097/HJR.0b013e3283312e94

Sassen, B., Kok, G., Schaalma, H., Kiers H., & Vanhees, L. (2010). Cardiovascular risk profile: Cross-sectional analysis of motivational determinants, physical fitness and physical activity. *BMC Public Health, 10*(592), 1-10. Retrieved from http://www.biomedcentral/.com/1471/2458/10/592.

Seo, N. J., and Armstrong, T. J. (2009). Biomechanical analysis for handle stability during maximum push and pull exertions. *Ergonomics*, *52*(12), 1568-1575. doi:10.1080/00140130903287999

Seo, N. J., Armstrong, T. J., and Young, G. J. (2010). Effects of handle orientation, gloves, handle friction and elbow posture on maximum horizontal pull and push forces. *Ergonomics*, *53*(1), 92-101. doi:10.1080/00140130903389035

Sharkey, B. J., & Davis, P. O. (2008). *Hard work: Defining physical work performances,* Windsor, ON: Human Kinetics.

Shields, M., Tremblay, M. S., Laviolette, M., Craig, C. L., Gorber, S. C. (2010). Fitness of Canadian adults: Results from the 2007-2009 Canadian health measures survey. *Statistics Canada Health Reports, 21(*1), 1-15. Catalogue no 82-003.

Shell, D. E. (2003). Exercise science and law enforcement. *Strength and Conditioning Journal*, *25*(3), 52-57. Retrieved from the Canadian Police College Library.

Shephard, R., & Bonneau, J. (2002). Assuring gender equity in recruitment standards for police officers. *Canadian Journal of Applied Physiology*, *27*(3), 263-295.

Shephard, R., & Bonneau, J. (2003). Supervision of occupational fitness assessments. *Canadian Journal of Applied Physiology, 28*(2), 225-239

Shephard, R. (2013). A brief critique of physiological employment standards for physically demanding public safety occupations. *Health & Fitness Journal of Canada, 6*(4), 1-8.

Sommerfeld, H., Dunlop, B., Neary, J. P. (1998). Heart rate response during the RCMP PARE test [Abstract]. *Canadian Journal of Applied Physiology*, 23(5), 509.

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Research Methods and Statistics*, *19*, 1099-1104. doi: 10.1111/j.1553-2712.2011.01185.x

Stanish, H. I., (1993). Comparison of the Physical Ability Requirement Evaluation to measures of physical fitness. [Abstract]. *Canadian Journal of Applied Physiology, 18:*436 Retrieved from http://www.nrcresearchpress.com/journal/apnm

Stanish H., & Campagna P, (1993).Comparison of the Physical Ability Requirement Evaluation simulated fight component to measures of muscular strength, endurance and power. Ottawa, RCMP Fitness & lifestyle Archives of an unpublished manuscript.

Stanish, H .I., Wood, T. M., & Campagna, P. (1999). Prediction of performance on the RCMP Physical Abilities Requirement Evaluation. *Journal of Occupational and Environmental Medicine, 41*(8), 669-677. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Strating, M., Bakker, R. H., Dijkstra, G. J., Lemmink, K. A. P.M., & Groothoff, J. W. (2010). A job related fitness test for the Dutch police. *Occupational Medicine, 60*, 255-260. doi: 10.1093/occmed/kqq060

Tinsley, P. N. (2000). *Assessing discrimination in a police recruit assessment center*. Doctoral dissertation, University of British Columbia). Retrieved from http://scholar.google.ca/scholar?q=+assessing+discrimination+in+a+police+recruit+Tinsley+2000&btnG=&hl=en&as_sdt=0%2C5

Thoden, H. C. (2002). *Testing for Normality*. New York, NY: Marcel Decker Inc.

Thomas & Means. (2000). *Use of gender and age adjusted physical performance standards in the law enforcement profession*. PRIMA conference position paper. Alymer ON: Retrieved from Ontario Police College Archives.

Thompson, W. R., Gordon, N. F. & Pescatello, L. S. (Eds.). (2009). *ACSM's: Guidelines for exercise testing and prescription, (8th Edition).* New York, NY: Lippincott Williams & Wilkins.

Trottier, A., & Brown, J. (1994). Occupational health in police work: A Canadian perspective. *Journal of Clinical Forensic Medicine*, *1*, 39-42. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Trottier, A., & Brown, J. (1994). Police health: *A physician's guide for the assessment of police officers.* Ottawa, ON: Canada Communications Group.

Wagner-Wisotzki, N. (2005). *Fitness and Lifestyle Management for Law Enforcement*, (2nd Ed.). Toronto, ON: Edmund Montgomery Publishers.

Welk, G. J. (2002). *Physical activity assessment in health-related research.* Windsor, ON: Human Kinetics.

Wenger, H. A., & Gaul, C. A. (1992). *RCMP Physical Abilities Requirement Evaluation demonstration project: August 1990-August 1992, Final Report.* Ottawa, ON: RCMP Health Services Archives.

Weir, J. P., Beck, T. W., Cramer, J. T., & Housh, T. J. (2006). Is fatigue all in your head? A critical review of the central governing model. *British Journal of Sports Medicine, 40*(7), 573-586. doi: 10.1136/bjsm.2005.023028

Wells, G. D., Selvaduria, H., & Tein, I. (2009). Bioenergetics provision of energy for muscular activity. *Paediatric Respiratory Reviews*, *10*, 83-90. doi: 10.1016/j.prrv.2009.04.005

Whitehead, M. (2007). Physical Literacy: Philosophical considerations in relation to developing a sense of self, universality and propositional knowledge. *Sport, Ethics and Philosophy, 1*(3), 281-298. doi: 10.1080/17511320701676916

Wilmore J. H., & Davis, J. A. (1979). Validation of a physical abilities test for the selection of state traffic officers. *Journal of Occupational Medicine*, *21*(1), 33-39. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Wilson, D., & Bracci, R. (1982). Police agility test: Another look at pre-employment screening. *Law and Order*, August, 36-42. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=85008

Wilson, T. M., & Tanaka, H. (2000). Meta-analysis of the age-associated decline in maximal aerobic capacity in men: relation to training status. *American Journal of Physiology and Heart Circulation*, *278*, H829-834.

World Health Organization. (1986). *Ottawa Charter for Health Promotion*. Retrieved from http://www.who.int/hpr/NPH/docs/ottawa_charter.hp.pdf

Zieky, M. J., & Livingston, A. (2004). *A Primer on Setting Cut Scores on Tests of Educational Achievement.* Citing material published elsewhere by Zieky & Perie (2004) in *Passing Scores: A manual for setting standards on educational and occupational Tests*. Retrieved from https://www.ets.org/Media/Research/pdf/Cut_Scores_Primer.pdf

Zumbo, B.D. (2001). Methodology and measurement matter in establishing a Bona Fide Occupational requirement for physically demanding occupations. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Chapter 2

The Relative Predictive Power of the Pursuit and Body Control Simulations and their Relationship Across Sex Groups

Policing is a physically demanding occupation requiring maximal efforts in unpredictably timed callouts for critical incidents (Anderson et al., 2001; Wilson & Bracci, 1982). As such, police officer physical fitness preparation for duty is an essential aspect of public and police officer safety (Bonneau & Brown, 1995; Trottier & Brown, 1994). The physical demands of police shifts between callouts are predominantly: walking, standing, and sitting, and are therefore, described as sedentary (Anderson et al. 2001; Anderson & Plecas, 2007; Wilson & Bracci, 1982). With years of policing, officer's physical fitness levels can fall from the high levels achieved as younger recruits (Boyce, Jones, Lloyd & Boone, 2008; Knapik et al., 2011; Strating, Bakker, Dijkstra, Lemmink & Groothoff, 2010; Wilmore & Davis, 1979). The declines in fitness levels for incumbent male and female officers are seen in Figure 1 depicting cadet's versus incumbent police officer's fitness scores. The high level of recruit fitness can be attributable to an almost universal use of Physical Employment Standards (PES) tests for minimal bring-to-the job ompetencies for cadet selection followed by quality PES for police academy graduates (Gledhill & Shaw, 1996; Knapik et al., 2011; Wagner-Wisotzki, 2005). Applying PES standards to incumbent offers provides an opportunity for health promotion programming for incumbent officers from hire to retire. However, these tests and standards have changed over time, with different focuses on different human factors, and different test constructs which are reviewed below.

2.1 Historical Moments of Occupational Fitness Testing

Denzin and Lincoln (2005) called the evolution of qualitative research philosophy, "moments" and described each successive moment in historical periods as dominated by an emerging philosophical development and a corresponding methodology, while a predecessor methodologies declined. Similarly, police occupational physical ability assessment is in the process of development of its third moment of an emerging dominant methodology, while the previous methodology declines, and not without some continued persistent supporters (Collingwood, 1988a, 1988b, 1988c, 1995, 2004).

2.2 First Moment: Height and Weight

The first moment in fitness testing for police was the use of human factors of height and weight standards, as a construct of what was needed for men to grapple and wrestle with suspects (Bonneau & Brown, 1995; Trottier & Brown, 1994; Wilson & Bracci, 1982). Today, human factors of height, weight, age, and physiological capacity of workers are used in determining pushing and pulling ability. This is seen in the area of occupational biomechanics instruction affecting manual material handling, ergonomics, and fitness for police and security personnel using body control simulators (Armstrong & Young, 2010; Ayoub & McDaniel, 1974; Boocock, Haslam, Lemon, & Thorpe, 2006; Chaffin, Herrin, Lee & Waiker, 1991; Chaffin & Resnick, 1995; Ciriello & Snook, 1991; DARCOR and ERGO web, 2103; Knapik & Marras, 2009; Kroemer, 1971). Implicit in policing movements and fair job simulation standardized testing is the controlling of potential mediators of test success to avoid irrelevant task difficulty or easiness (Messick, 1996) like height and weight factors that could result in disparate group responses (Zumbo, 2001). Police services started changing their police recruitment abilities testing from height and weight standards in the early 1980's (Anderson et al., 2001) to health-related fitness component testing (Hoffman & Collingwood, 2005) to skill-related testing in the 1990's (Farenholtz & Rhodes, 1986) as described below.

2.3 Second Moment: Health-related Fitness Assessments

The second moment in fitness testing was the construction and use of what Thompson (2010) defined as health-related fitness testing. The constructs most commonly measured by discrete test items, for the capability for either a specific region of the body or the whole body were for capacities for cardiovascular endurance, muscular strength, muscular endurance, flexibility, and body composition (Thompson, 2010). An example of a body region health-related test would be maximum push-ups capability for upper body strength endurance work capacity. In health-related testing, a battery of relevant constructs measured by a set of discrete item tests is summated and globally scored, acting as perceived mediators of job simulation ability (Shephard & Bonneau, 2003). All health-related score interpretations use separate age and sex normative percentile scores tables. These scores are generalized to physical fitness for duty status for all police duties (Shepherd & Bonneau, 2003). A generation of data has supported and been published as age and sex grouped normative tables for police health-related fitness for duty testing (Hoffman & Collingwood, 2005; Wagner-Wisotzki, 2005).

Deakin et al. (2001) described this type of test as a construct valid testing process with fitness tests performance supposedly highly correlated with job simulation performances. The later steps were considered valid if fitness scores were highly correlated with job simulations of the frequent, demanding, or critical tasks of police work. With regard to physical agility tests, Avery (1992) reports that the accurate reflection of police work was the predominant effort and preoccupation of test developers and researchers, as well as court reviewers. There have been controversy associated with using separate sex group normative references or one minimum cut score for both genders. Should there be gender specific and different pass cut scores for men and women, if there is one job related task?

2.4 Third Moment: Skill-Related Job Simulations

The third moment of physical occupational fitness testing, hereafter called skill-related testing, emerged with the advent of job testing scenarios. The discrete test construct items of essential police movements were determined by gold standard direct police physical movement observations and post shift officer reporting, and follow-up large-scale surveys, all triangulated for accuracy by researchers (Anderson et al., 2001; Farenholtz & Rhodes, 1990). Further, the tasks were integrated in the rich job simulations as a whole timed circuit. These tests cut out the intermediate steps of a complex mathematical treatment of a large number of test batteries, with only one real-time test score representing critical constructs easily and directly related to the job. They also allow a weak score on one section to be compensated for with a strength in another section (Deakin et al, 2001; Mahar, 1984). The general pattern of the police skill-related testing is that of a reflected critical incident callout of “getting to a problem, dealing with a problem, and removing a problem” (Anderson et al., 2001; Farenholtz & Rhodes, 1990). Furthermore, Thompson (2010) described these tests as using discrete test item constructs of: speed, reaction time, balance, agility, coordination, and power. Deakin et al. (2001) have described the skill-related job simulation testing as being highly accepted by many groups, including labour, police management, and even the judiciary system. Mahar opens his 1984 article with a quote from a judge stating: “Surely, it is difficult to imagine a more accurate way of testing ability to scale a 6 foot wall than to scale one (Hardy versus Stumpf, 1978)” (p. 173). A hybrid test is a combination of the above two test methods with discrete test items from both types of methods (Deakin et al., 2001).

With findings in large data samples (Strating et al., 2010) women as a group have lower or slower time scores than men in skill-related testing. With this in mind should there be one job related gender neutral minimum time score, low enough for women to get in without disparate group selection, but not too easy for men (Strating et al., 2010; Zumbo, 2001). There are many legislative pressures to avoid discrimination against protected human rights grounds of age, sex, race, and disability (Avery et al., 1992; Hoffman & Collingwood, 2005; Jackson, 1994; Lonsway, 2003; Mahar, 1984; Prenzler, 1996; Shephard & Bonneau, 2003).

2.5 The State of the Art of Fitness Tests

The genesis of the Canadian police and corrections skill-related tests were the Police Officers Physical Abilities Test (POPAT) and the Correctional Officers Physical Ability Test (COPAT) by Farenholtz and Rhodes (1986; see Appendix 1). Farenholtz and Rhodes (1992) reported very little correlation between health-related discrete field tests and the POPAT scores, for either maximal aerobic power or anaerobic components alone, although they suggested the two possibly working together to predict POPAT success. Conclusions about health-related fitness testing for POPAT included, low correlations between laboratory and filed fitness tests “suggest[ing] that the fight station is very specific and controlling a suspect by police officer is not dependent on how many push ups or pulls ups he [or she] is able to demonstrate” (Farenholtz & Rhodes, 1992, p. 7). Farenholtz and Rhodes (1992) concluded, “that POPAT being a valid, task specific, job related test consists of motor abilities and techniques, as much as generalized fitness parameters” (p. 1).

The Physical Abilities Requirement Evaluation (PARE) followed in the early 1990s, as a modified POPAT. The RCMP use of the PARE, as a modified POPAT, met the US Department of Labour’s *Uniform Guidelines* task analysis filter for test appropriation by another jurisdiction. Namely, do the discrete test items, distance run, weights lifted, etc. of the test appropriated equally apply to the job as described by your jurisdictions task analysis results. PARE was adopted as a modified POPAT (Bonneau, 2001) in 1990. Equally important, *Uniform Guidelines* methods questions include: do tests avoid adverse impact for protected groups when a cut score for pass or fail is applied? Zumbo (2001) defines adverse impact as “a legal term describing the situation in which group differences in task performance result in disproportionate examinee selection” (p. 42). Jamnik et al. (2010b) cautioned and demonstrated that one-time BFOR testing resulted in adverse impacts for one group (i.e., women), but could be reasonably accommodated by a sensible physical training period. Thereafter, follow-up test development steps with repeat testing showed that a greater than 80 % pass rate for initially failing groups of women can be achieved (i.e. 80% of female group equals the pass rate percentile of the male group).

2.6 The PARE Test

The current PARE test was originally measured and is now currently published in imperial units only. It has officers successfully navigate each of six laps by running and jumping a five foot simulated ditch mat (representing a water ditch or small hazard). Then they climb and descend a five-step staircase. This is followed by running and hurdling two 18-inch obstacles 10 feet apart. A vault is then conducted over a three-foot high rail, followed by a controlled landing. The subject then falls to the ground for a controlled chest or back fall, on alternate laps, and rise quickly to repeat another circuit. The total distance covered is roughly 1,180 feet (360 m). The original suggested time of completion for station one, the pursuit circuit was 2:30 min:s or 25 s per lap for the six laps. See figure 1 in Appendix 1 to compare flow for PARE,POPAT, PREP, and FITCO tests. In station two the body control segment, officers simulate the gross motor skills of body control of a subject by pushing and maintaining a weighted carriage, several inches above the ground with 80 lb (70 lbs for applicants), while controlling the rectangular steering wheel centered at belt height. After lifting the weight stack, by pushing and overcoming the resistance, the officer moves in a complete 180-degree arc. After completing six arcs, four controlled body falls to the ground, are completed, alternately to the chest and back. Then again, after lifting the resistance stack by pulling, the officer completes six 180-degree arcs, while pulling on the waist height triceps rope attachment. The optimal posture and biomechanics cues for pushing are: both feet back, use non-slippery shoes, suggested pretest standing belt height grip, legs fully extended acting as primary movers/pushers, complete core muscle activation, and never let go of the steel handles for safety. For the pull, the universal athletic stance of triple flexion of ankle, knee, and hip with the chest anterior to the hips and arms bent at the elbows (90-degree angle). Pulling is completed with a low to the ground hip height. The original suggested completion time of the body control station was 90 seconds.

2.7 The Skill-Related Fitness Test Performance Standard

Canadian police skill-related testing (i.e., POPAT, PARE, PREP) used a criterion reference PES that represents the average score solicited from young males, including inmate groups, on the test (Farenholtz & Rhodes, 1990). In the case of POPAT and then PARE, this score was validated with supplemental male group testing using a similar young male demographic group data, as representative of the most able, hence most demanding police arrest, as a normative criterion reference (Bonneau 1996, 2001; Farenholtz & Rhodes, 1990, Zieky & Perie, 2004). The use of male inmate’s times and similar homogenous groups average BFOR scores is over 25 years old and is the standard in both provincial and federal Canadian police and corrections skill-related fitness testing. The original suggested total PARE time standard for incumbent officers' PARE testing was the PES of 4:00 min:s for operational constables and corporals, or equivalent positions, with 4:30 min:s as a health promotion goal for other non-operational of older officers (Trottier & Brown, 1994).

2.8 Cut-score Determination

Shephard and Bonneau (2003) have reviewed the process of using an incumbent population of police officers' score distributions, in light of gender equity, and suggested a population-wide percentage be set as the standard for recruits. No published consensus has been reached regarding which percentage value should be used. More recent cut score methods have been reviewed by Rogers, Docherty and Petersen (2014). The existing methods for cut-score determination are first, that of normative data referencing the cut score in terms of + 1 or 2 SD above the mean of some criterion population that has been tested (Deakin, Pelot, Smith, Stevenson, Wolfe, Lee, Jaenen, et al.; 1996; Jamnik et al., 2010; Rogers et al., 2014; Shephard & Bonneau, 2002; Tipton, Milligan and Reilly, 2012). This is considered valid when the test components involves "job-related tasks and is a valid simulation of the essential elements of time-sensitive work" (p. 1751). Bonneau, (2001) based the PARE cut-score on a criterion group of young males, and two other groups of young males, with similar demographics as the arrest records of the criterion group. However, he used the mean of the groups, at PARE 3:57 min:s, rounded to 4:00 min:s, as the cut-score for depot graduation for cadets. This was a departure from the POPAT 4:15 min:s, but with the 10 rail jumps integrated into PARE, the new mean was used at 4:00 min:s. The 4:00 min:s mean of the criterion group coupled with + 2 SD above the mean + 44 s, leads to the initial PARE entrance standard for applicants at 4:45 min:s. Depot training weeks are used as reasonable accommodation time to shift to a higher standard without adverse impact. This cut score is determined without age and gender variant scores based on a one-job proficiency level (Rogers et al., 2014; Shephard and Bonneau, 2002).

The other cut score methods listed by Rodgers et al. (2014) were relative, contrasting or borderline group (Livingstone and Zieky 1982), absolute procedure (Nedelsky 1954) and its modifications (e.g. Maguire, Skakun, and Harley 1992) (p. 1751). Also listed were Angoff (1971) procedure and its modifications (Ricker 2006) and compromise (Hofstee, 1983; Fielding et al., 1996). The Angoff method was partly used to determine the newer RCMP Emergency Response Team PARE test by four RCMP DFLAs, Fiona Vincent, Sylvain Lemelin, Ken Ross and this author. As in Angoff rating methodology, several increasing speeds of the ERT PARE tests were filmed and rated by a subject matter expert panel. The final cut score was triangulated with an acceptable percentages of the existing population that would pass, after testing the entire RCMP ERT population to get normative reference numbers, added to supervisors rating of ERT PARE times – all to validate the cut score using both normative data and subject matter expert opinions.

Also listed as cut-score methodologies were item-mapping Bookmark method (Lewis, Mitzel, Green, & Patz, 1999). This method was used by the Canadian Forces to establish the Fire Fighter Fitness Maintenance Standard Evaluation (FF FME). In this method, faster and faster test times for firefighters were filmed and reviewed by a panel of judges. After each of the first two rounds decision were debated, then a final third round voting. After each round the range or cut-core dispersion dropped and a narrower range of scores was possible. Final determination used the final rated score + 2 standard errors of the mean (SEM) for the minimum necessary for the safe and efficient conduct of the job, as an inclusive step for borderline scores (Lewis, Mitzel, Green, and Patz 1999) of what can actually be done (Zieky & Perie, 2004).

Other cut score methods listed by Rogers et al. (2014) are Body of Work and its modifications (Kingston et al., 2001) and Analytic Judgment Method (Plake and Hambleton 2000, 2001). These are reviewed elsewhere and not pertinent to general duty federal police skill-related testing.

2.9 Incumbent Officer PARE Testing

Incumbent fitness for duty is also a priority of modern police services, and therefore some physically demanding services have extended the use of applicant physical testing to incumbent populations as described by Dunsmore and Hunter (2001) in the BFOR template consensus conference reporting manual. The RCMP PARE pilot project by Gaul and Wenger (1992) to test and train incumbent RCMP officers was conducted from 1990 to 1992 across several urban and rural divisional detachments. The pilot project used varying non-standardized push-pull equipment compared to today's highly standardized PARE testing. Results reported that the males' initial pass rate of 71% after the intervention and with additional attempts to pass over 2 years, increased to 78% two years later. Only 14% of women initially passed, with a two-year intervention change to 48% after additional attempts. Gaul and Wenger (1992) attributed most of the difficulty for women at the body control simulation station. The under 30 years of age group pass rate was 80% after two years. The 30 to 40 years age group had similar two-year results with 75% passing. The 41 to 50 years age group had only 75% under 4:30 min:s, although the goal was still 4:00 min:s. Insufficient numbers in the over 50 years age group were available to present results.

2.10 Literature Review

Success predictors in the skill-related testing for large sample sizes has been researched in part by Strating et al. (2010) for a large sample of Dutch officers. More directly related to PARE predictors of success, Stanish, Wood, and Campagna (1999) conducted a small student study.

2.11 Predictors of Success in Skill-Related Testing

Strating et al. (2010) researched a large sample of 7,000 Dutch police officers using skill-related fitness testing. They reported significant group difference for older, female, and non-operational police duty groups. A decrease in fitness level with aging for younger (< 40 years), between 40 and 55 years, and older (> 55 years) officers, and women's versus men’s groups, and operational policing versus administrative duties groups, were reported. Administrative groups were split into investigating and all other duties groups. They reported on the significant predictive effects of BMI, age, and hours of sports on passing a skill-related test scores. The test items included: foot pursuit running 226.5 m while climbing a 1.1 m obstacle, jumping over low obstacles, pushing three times and pulling two times a 200 kg cart over six m, lifting and carrying a five kg ball 18 times for three m, dragging a 48 kg dummy for five m in a timed circuit. Women as a group were 26 seconds slower than men. This was shown as statistically significant, at $p < 0.05$, reporting an effect size of 1.15. Officers with higher BMIs (> 30) and lower physical activity profiles, scored slower for both male and female officers. Officers in “core police tasks” had significantly faster times with at a p of < .01 than officers in the “remaining function group” (Strating et al., 2011, p. 257). Listed, as significant results for regressions ($p < .001$) were factors of age, BMI, functional work group, and hours of physical exercise. Both sexes were assessed against a single sex neutral job test/task specific standard.

Stanish, Wood, and Campagna (1999) contrasted health-related fitness component test scores to PARE scores “to identify valid and practical field tests of physical fitness that accurately classify successful and unsuccessful PARE performers” (Stanish et al., 1999, p. 667). Fitness test constructs of aerobic power, anaerobic power, muscular strength and endurance, and body composition were contrasted with PARE pass scores (4:45 min:s) for 28 students and 20 applicants. A three-variable model was found in multiple regression analysis to be predictive of PARE score success: a 70 lb bench press, standing long jump, and agility “explaining 79% of the variance” for men (Stanish et al. 1999, p. 2). The authors reported that a small percentage of male scores were failures (9%), making the “classification” of pass/fail problematic for their proposed logistic regression analysis which requires sufficient dichotomous variable data to fit the sigmoid curve model. However, female data had sufficient dichotomous data to determine passes and failures. The authors reported two variables as predictive of PARE score success in the women’s group: namely the 1.5-mile run and the agility test, with “93% overall classification accuracy" (Stanish et al., 1999, p. 2). The PES standard used by Stanish et al. (1999) was the initial RCMP applicant entry time of 4:45 min:s not the Cadet graduation employment standard of 4:00 min:s. Stanish et al. (1999) did however suggest future research of "partitioning the PARE score into run and fight components [that] may provide more precise information regarding the contribution of these components to successful PARE performance" (p. 675). The authors reported "performance scores on the PARE were found to be significantly (*p* < .01) correlated to percentage body fat, sum of 5 skin folds and average time to complete lap 1and 2" (p. 1). Additionally, negative correlations were found between both maximal aerobic power and height and weight and PARE scores. Reed's (1996) direct measures of PARE solicited oxygen consumption would explain the low maximal aerobic power (MAP) correlation, as although near maximum heart rates were solicited; only 66% of MAP oxygen consumption was evident in direct O_2 measured during field PARE testing, arguing for a significant anaerobic contribution. This evidence, plus other Rhodes & Farenholtz (1992) leads us to determine PARE test constructs are related to cardiovascular measures in part, as well as agility measures and strength measures, but no so much health-related fitness component capacity measures. The majority of the variance in PARE run performance time scores reported by Stanish et al. (1993) using multiple regression techniques was determined to be: maximal aerobic power, percent body fat, sum of 5 skin folds, and average lap time.

2.12 Purpose

The movement of police fitness testing from historical use of height and weight standards (Bonneau & Brown, 1995), to health-related testing with sex and age separate standards (Hoffman & Collingwood, 2005; Wagner-Wisotzki, 2003) to the current skill-related job simulation standards (Farenholtz & Rhodes, 1990) needs to be evaluated. One way of examining the historical movements in fitness testing would be to build a prediction model with factors from all three moments of fitness testing, with height and weight factors from moment one, health-related fitness capacity testing age and sex group factors from moment two, and skill-related performance times for pursuit and body control factors from moment three. Therefore, the purpose of this study is to compare the relative predictive power of the pursuit and body control segment of the PARE test to success, and any differences across sex groups while including covariates of height, weight, and age.

2.13 Ethical Clearance of use of Data

Written permission to use the RCMP national database, to conduct a doctoral study, was granted by the RCMP Fitness & Lifestyle Policy Centre, in January 2013. Permission to use divisional data was granted by front line supervisor and line officer in charge of Human Resources, in January 2013. All personal and work identifiers were removed and a unique coded research number was assigned to each data line. Only group data reporting was used to help ensure personal confidentiality. The RCMP granted publication of an easily retrievable electronic version of a doctoral dissertation using RCMP PARE data.

2.14 Records Download

All PARE data was retrieved from the Human Resources Management Information System (HRMIS). Captured electronic data of incumbent police officer PARE test results conducted by the RCMP *Fit for Life, Fit for Duty* program covered the 2006 to April 2013 period. These records were downloaded to a spreadsheet to facilitate duplicate removals after sorting for times selection of one unique best PARE test score per officer. Thus over 30,000 lines of PARE data were downloaded, with over half eliminated to obtain 14,401 one best (lowest) PARE time per officer. Data was uploading to SPSS version 21 for logistic regression analysis. Data lines uploaded represented 14,401 records. With outliers > + 2 SD removed the resulting data file held 13,709 unique PARE records. This represented a performance profile for a six-year period of a police population real-world testing program, without the slowest participation walking scores. The dependent outcome variable was coded pass = 0, fail = 1, based on the PES 4:00

min:s or 240 s, as the cut-score point. Sex coding was male = 0 and female = 1. PARE times were recorded to the nearest second with hand held stopwatches.

Logistic regression procedural steps were conducted as suggested by Fields (2013) and Stoltzfuz (2010): (a) influential outliers above 2 SD removed; (b) model building through the generation of a short list of significant predictor variables; (c) verification of linearity of the logit for significant short list independent predictor variables; with (d) multicollinearity diagnostics checks. Every significant predictor factor was checked for linearity of the logit with the predictor variable interaction, coded using the natural log of the variable interactions with the variable (ln*(variable)* variable)* as suggested by Fields (2013). Any non-linearity of a variable during multicollinearity checks had the variable transformed (Fields, 2103; Hilde, 2009). The factor age was transformed into the square root of age, to create a normalcy profile that could be used, after failing the first diagnostic check. With all factors entered; namely: height, weight, sex, square root of age, pursuit time, body control time, run*control(interaction term), mat penalty time, and stick penalty time, a complete separation of data at step nine was reported without a unique solution. Mat and stick penalty times were rejected due to high residuals (30+) and their disproportionate loading in failure test scores (80-20). Therefore, a reduced model was run without penalty times. A Chi-square statistic for maximum likelihood was assessed for each model building step with significance levels at $p < .05$, as suggested by Menard (2010) and Fields (2013).

2.15 Results

The descriptive statistics of the officers' age, height, weight, and BMI representing 13,709 unique PARE records, for the time 2006 to April 2013, are displayed in Table 2.1 for combined and separate male and female groups. The demographics of study one's data was: height 178.2 cm, weight 87.3 kg, BMI 27.4 and age 37.2 in years which were very close to Anderson et al.'s (2001) revalidation study incumbent sample with: height at 179 cm, weight at 84 kg, and age at 36 years. The female proportion of the 13,709 unique PARE data records was 2,217 / 13,709 of 15.5 %, while the odds were 2,217 /

11,582 or 18.4 %. The serving RCMP incumbent officers' population sex profile is close to 80:20 male/female. Anderson et al. (2001) had 19% female percentage in his 1992 pilot study sample.

The PARE pursuit, body control, mat, and stick penalty times for all groups are in Table 2.2. The complete reporting of the logistic regression analysis, as suggested by Fields (2013) includes the intercept and predictors coefficients, Wald statistic and corresponding significance values, log odds (LO), and ± 95% confidence intervals (CI), are all displayed in Table 2.3. The mean probability of passing with no predictors, as the null model first step, was equal to the population mean pass score of 78.8%. The baseline probability for the combined data of 13,709 records was 0.79 for passing PARE and 0.21 for failing PARE. The odds of failing PARE was 0.26 calculated numerically as 0.21 / 0.79 = 0.26. The initial logistic regression model analysis produced a unique solution with no reported variables until mat and stick penalties were removed. A repeat analysis with a performance only model of no penalties and > 2 SD PARE time outliers removed, provided a reported solution with a 99.6 percent prediction model rating.

The combined data set did support nearly equal pursuit times and body control LO, see table 2.3. Of the human factors covariates tested, only the square root of age was a significant predictor. The data set did not support statistical significance for height p= .71 and weight p= .18, BMI, p =.34, and sex p =.09, or the interaction of control*run, p=.74 in the model output of variables in the equation.

The prediction formula for the combined data set, female data set and men's data set are

combined $P(Y) = -243 + 1.022(run) + 1.031(control) - 0.643(\sqrt{age})$,

female $P(Y) = -209 + 0.873(run) + 0.895(control) - 0.410(\sqrt{age})$,

male $P\,(Y) = -258 + 1.091(run) + 1.097(control) - 0.832(\sqrt{age})$.

Table 2.1

Generational Demographics Data for the Retrospective Observational Study

	Variable	*M*	*SD*	*SE*	Variance
Combined	Age (years)	37.2	8.1	.07	65.0
(n = 13,709)	Height (cm)	178.2	8.0	.07	63.3
	Weight (kg)	87.3	14.6	.13	212.7
	BMI (kg/m^2)	27.4	3.6	.03	12.8
Female	Age (years)	35.2	6.8	.14	46.6
(n = 2,127)	Height (cm)	167.8	6.4	.13	40.7
	Weight (kg)	69.7	11.3	.24	127.9
	BMI (kg/m^2)	24.7	3.7	.08	13.4
Male	Age (years)	38.1	8.4	.08	70.8
(n = 11,582)	Height (cm)	180.2	6.6	.06	44.0
	Weight (kg)	91.2	13.0	.12	167.5
	BMI (kg/m^2)	28.1	3.6	.03	12.7

Note: *M* is the mean; *SD* is the standard deviation; *SE* is the standard error of the mean; BMI is body mass index as a ratio of body weight in kilograms per height in meters squared.

The probability changes from 0 to 1 for the pursuit and body control times, while controlling for age and one of the two performance factors, are shown in Figures 2.2, 2.3 and 2.4. In all three data sets the probability of passing PARE as a changing value for the continuous variables were calculated while holding one of the other factors constant at the mean level of that variable (Pampel, 2000). The change in probability ranging from 0 to 1, representing passing PARE to failing PARE, as clinically important outcomes, are demonstrated in Figures 2.1, 2.2, and 2.3. The unique shift point in probabilities from

Table 2.2

PARE Performance Scores for the National Data Set

Factor	Mean	*SD*	%	Mean	SD	%	Mean	SD	%
	Combined (n = 13,709)			Male (n = 11,582)			Female (n = 2,127)		
PARE (s)	221.9	32.0	100	217.0	*29.7*	100	248.7	30.7	100
Pursuit (s)	156.0	21.8	70	153.3	20.9	70	170.7	20.3	68
Control (s)	65.4	11.4	30	63.3	11.8	30	76.8	15.8	31
Mat (s)	0.5	2.2	--	0.4	1.9	--	1.1	3.3	1
Stick (s)	0.1	0.4	--	0.5	0.4	--	0.1	0.4	
Pass/Fail (%)	79%			84%			49%		
Odds Ratio	.79/.21 = 3.76			.84/.16 = 5.25			.49/.51= 0.96		

Note: *M* = mean; *SD* = standard deviation; % = percentage of total unique PARE data odds point for the men's data was: pursuit slightly < 163 s and for the body control slightly > 73 s. set, total test time (PARE), foot chase simulation circuit time (Pursuit), body control push and pull simulation time (Control); a five s penalty for touching the five foot mat (Mat), a two s penalty for knocking down an 18 in hurdle stick (Stick).

below 0.5 to greater than 0.5 probability of passing, which is equal to an odd of 1.0, is posited on the graphical summary of the probabilities calculations with an arrow. In Figure 2.2, the combined data 1.0 unity odds point for the pursuit at: slightly > 167 s, (2:47 min:s), and for the body control just slightly > 83 s. (1:23 min:s).In Figure 2.3 the unity odds points for women's data were pursuit at slightly < 164 s and body control at slightly < 71 s. In Figure 2.4 the unity odds point for men's data was: pursuit slightly < 163 and for body control slightly < 73 s. Probability predictions conversion from log odds to predictions follows the formula:

$P(\mathrm{Y}) = \exp^{(Bo+B1x1+b2X2+b3X3)} / 1+\exp^{(Bo+B1x1+B2X2+B3X3)}$ (Pampel, 2000, p. 16).

As suggested by Pampel (2000) the change in the continuous independent variable X are displayed as a probability prediction outcomes, while controlling for other linear regression factors using their mean values. This was accomplished by use of the above formula to create meaningful probability outcomes for the pursuit range times for the combined data, at 162 to 173 s, for the women's data at 157 to 170 s and for the men's data at 158 to 168 s. Similarly, meaningful probability outcomes for the body control range times were for the combined data was 78 to 89 s, for the women's data as 64 to 77 s and for the men's data was 68 to 79 s. In terms of probability changes ranges, the women's pursuit times were closer to the combined pursuit times.

2.16 Discussion

The main purpose of study one was to compare the relative predictive power of the pursuit and body control segments on PARE test success. Further, this study sought to examine whether the relationship held or varied across sex groups. This analysis also encompassed human factors from previous historical moments of fitness testing of height and weight, and age.

Gaul and Wenger's (1992) conducted an incumbent project of PARE testing. The overall best percentage pass rate for combined group of men and women was 72% versus the current retrospective observational study, which reports a pass rate of 79% for the combined group. The 1992 pilot study female pass rate was at 25% initially and reached 48% over 2 years. Converting the significant predictive factors into probabilities shows no adverse impact for the large women's group as passing PARE has a large 90"s percentile probability which could be presented as non-discriminatory to the human rights courts, which is explored below.

The current retrospective observational study has a female pass rate of 49% over 6 years. The current observational study's male pass rate was 84%, over 6 years. The Gaul and Wenger (1992) study male pass rate was 71 % and reached 78% by two years. Note Gaul

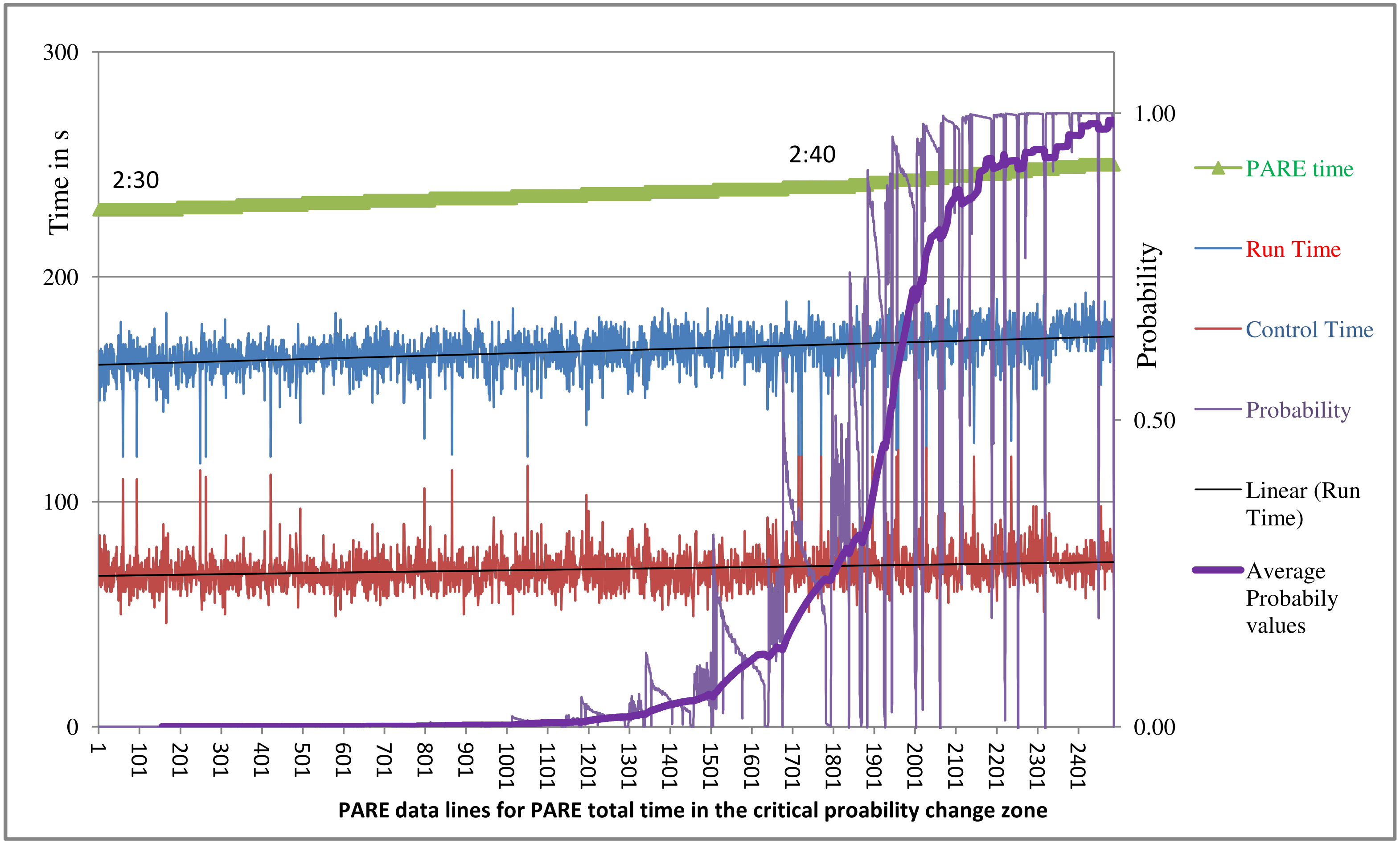
Table 2.3

Statistically Significant PARE Success Predictors for Initial and Final Model

Initial Model	Factor	β	*SE*	Wald	df	Sig.	Exp$^{(\beta)}$	95% CI [lower, upper]
Divisional Combined (n = 13,709)	Run	1.00	0.1	218.2	1	< .001	2.72	[2.38, 3.10]
	Control	.99	0.1	104.3	1	< .001	2.70	[2.23, 3.26]
	C X Run	.00	0.0	0.1	1	.74	1.0	[0.99, 1.00]
	Sex	.50	.29	2.9	1	.09	1.64	[0.93, 2.90]
	$\sqrt{Age}$	-.74	.18	17.3	1	< .001	.48	[0.34, 0.68]
	BMI	-.03	.03	.91	1	.34	.97	[0.92, 1.03]
Final Model	**Factor**	β	**SE**	**Wald**	**df**	**Sig.**	**Exp (β)**	**95% CI [lower, upper]**
Combined (n = 13,709)	Run	1.02	0.57	324.1	1	.01	2.78	[2.49, 3.11]
	Control	1.03	0.57	324.7	1	.01	2.80	[2.51, 3.14]
	$\sqrt{Age}$	-.64	0.16	16.0	1	.01	0.53	[0.38, 0.72]
	Intercept	-243						
Male (n = 11,582)	Run	1.09	0.7	229.6	1	.01	3.00	[2.59, 3.43]
	Control	1.10	0.7	228.2	1	.01	3.00	[2.60, 3.46]
	$\sqrt{Age}$	-0.83	0.3	17.4	1	.01	0.44	[0.29, 0.64]
	Intercept	258						
Female (n = 2,127)	Run	0.87	0.9	91.4	1	.01	2.40	[2.00, 2.86]
	Control	0.90	0.9	95.7	1	.01	2.45	[2.05, 2.92]
	$\sqrt{Age}$	-0.41	0.3	1.4	1	.01	0.66	[0.34, 1.31]
	Intercept	-209						

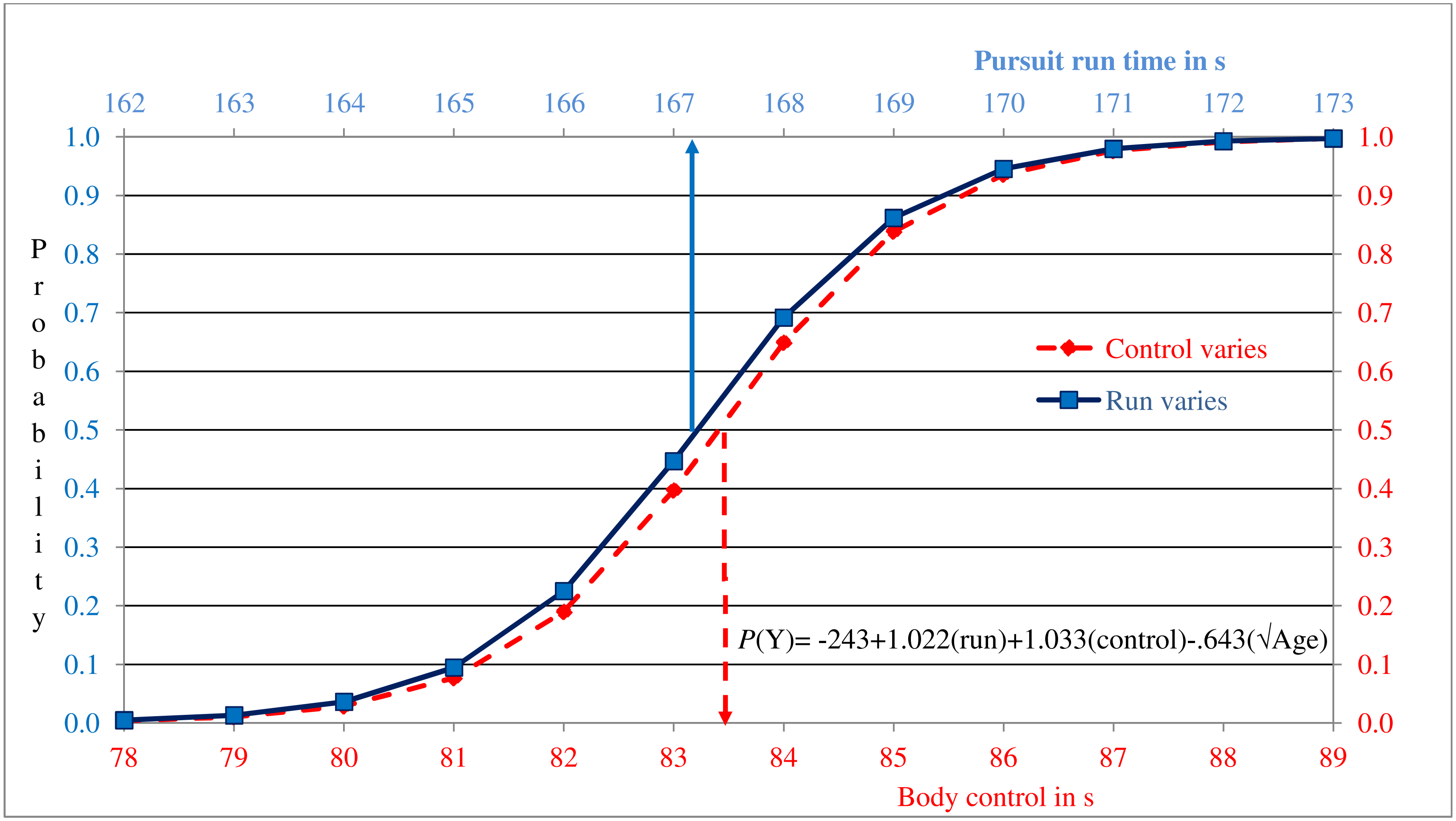
Note: statistical significance at $p < .05$; six lap pursuit segment (Run), body control push and pull (Control), transformed age variable that meets log-linearity assumptions of logistic regression covariate and dependent variable $\sqrt{\textbf{Age}}$ (Fields, 2103); degrees of freedom (df); regression coefficients of the logistic regression analysis (β); standard error of β (SE); the Wald statistic ration (β/SE); the log odds Exp$^{(\beta)}$s; confidence interval (CI).

Figure 2.1 Change in probabilities of failing from 0 to passing at 1 for all best PARE scores



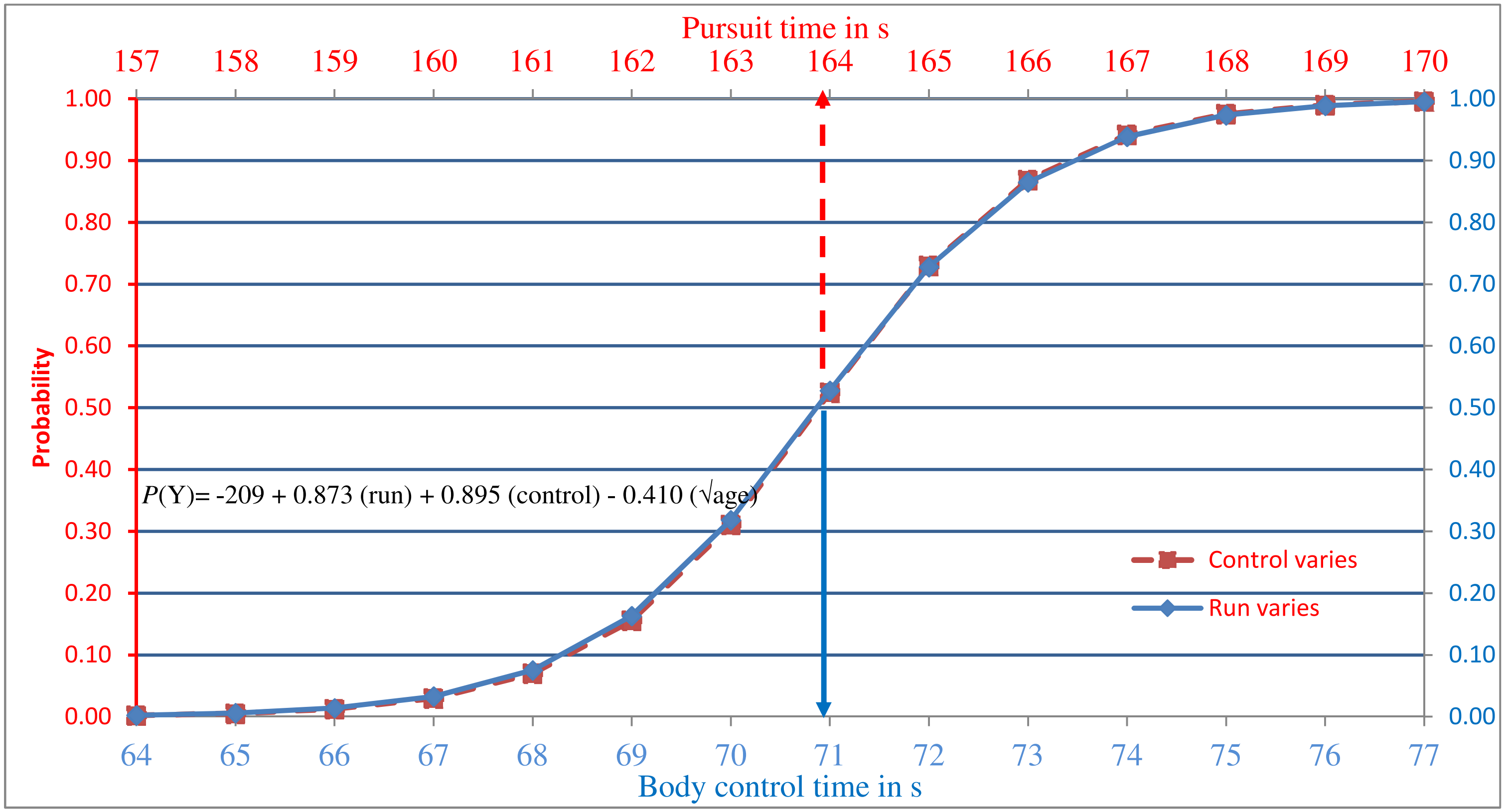
Note Change in probabilities from 0 to 1 for PARE from 230 s to 250 s, the green top line steps each represent a 1 s increase in PARE time. Noise on the curve is averaged showing the expected logistic regression sigmoid curve of probabilities changes with changes in pursuit, blue line and body control, red line. Graph shows only the clinically important transition zone.

Figure 2.2 Probability changes from 0 to 1 for combined male and female data.



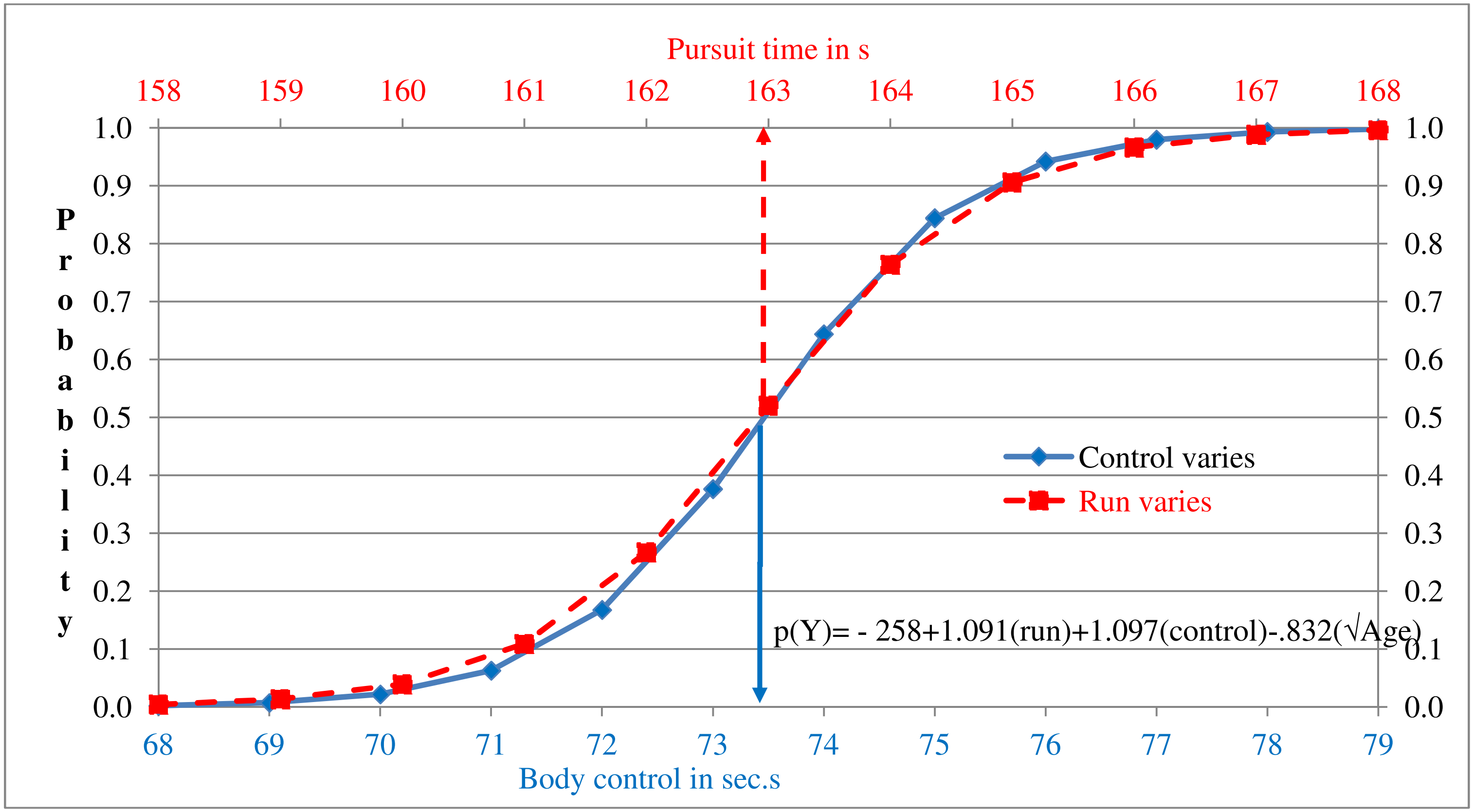
Note: Probability changes for combined data for critical probability zone of 0 to 1with clinically significant PARE times "using the mean of the dependent variable" …"as a starting point for changes" in "X" (Pampel, 2000, p29). Blue solid line: Age held constant at mean 37.2 years, body control held constant at mean of 65.4 s, while run time varies; red dashed line run time held constant at mean 156 s and age held constant at mean 37.2 years, while body control times vary.

Figure 2.3 Probability changes from 0 to 1 for female data



Note: Probability changes for combined data for critical probability zone of 0 to 1with clinically significant PARE times "using the mean of the dependent variable" …"as a starting point for changes" in "X" (Pampel, 2000, p29). Blue solid line: age held constant at mean 35.2 years, body control time held constant at mean 76.84 s, while run time varies; red dashed line: run time held constant at mean 170.7 s and age held constant at mean 35.2 years, while body control times vary.

Figure 2.4 Probability changes from 0 to 1 for male data.



Note: Probability changes for combined data for critical probability zone of 0 to 1 with clinically significant PARE times "using the mean of the dependent variable" …"as a starting point for changes" in "X" (Pampel, 2000, p29). Blue solid line: age held constant at mean 38.1 years, body control time held constant at mean 76.84 s, while run time varies; red dashed line: run time held constant at mean 170.7 s and age held constant at mean 38.1 years, while body control times vary.

and Wenger (1992) had very few women in their study samples and very few women officers under a time score of 4:30 min:s. Whereas the average for the women officers' PARE time in the retrospective observational study of 2,127 officers was 4:08 min:s, with a large percentage under 4:00 min:s. Comparisons between the two populations can be very limited in that they represent very different samples of police groups. The later study has 2,127 women officers whereas the earlier has very limited number of women, at about 37 who completed the project.

For the baseline probability of passing at .79 or failing at .21, Pampel (2000) gives the interpretation of this probability as, for each PARE test; only .21 of the times does a failure occur. For 13,709 PARE in the data supporting study one, 2,910 failures occurred. The odds of failing PARE was 0.26 calculated as .21 / .79 = 0.26. The data supported an initial logistic regression model that had a 78.8% null predictor model. If we guessed everyone passed we would be right 78.8% of the time. If we guessed everyone failed we would be right 21.2% of the time (Pampel, 2000). Male pass/fail proportions were .84 passed (9,755/11,582) and .16 failed (1,827/ 11,582). Male pass odds were .84/.16 = 4.0. Female pass/fail proportions were .49 (1,044/2,127) passed and .51 failed (1,083/2,127). Female pass odds were =.49/.51= .96. Final model contingency table prediction accuracies reported were for the combined model, 99.6%; for the men's data 97.8%; and for the women's data 98.2%.

Fields (2013) describes the logit of the outcome as "the natural log of odds of Y occurring" (p. 784). Pampel (2000) describes three properties of equal change in logit metrics, which have no intuitive metrics. One, they have no ceiling or floor value being P_i = 1or 0 is undefined. Two, logit changes are symmetrical about the unity point of .50 probability or logit of 1.0. Three probabilities as prediction outcome can vary in value for a given unit of the independent variable X. Changes in outcome probabilities are not proportionate to logit changes, especially as they approach 1 or 0. Pampel (2000) states the relationship between "the independent [predictor] variable[s] and the probabilities are nonlinear and non-additive" and "cannot be fully represented by a single coefficient" at any one point on the sigmoid curve (p. 23). Therefore, meaningful metrics are derived via determining a clinically significant range of probabilities and outputting their values, holding one value even and exploring the change in other values probabilities. The process described by Pampel (2000) is to reverse the original logit function (taking the natural log of variables) by exponentiation of the values for log odds of Y with base *e*. The probability statement for the outcome of Y, namely *P*(Y), is then the outcome from the logistic regression analysis. Figures 2.2, 2.3 and 2.4 refer, meaningful clinical times show changes in probabilities from 0 to 1, showing a change in just a few seconds, at the critical threshold of the logistic function can mean the differences between a 0 probability of failing to a 100% chance of failing $P(\mathrm{Y}) = 1.0$, the results of which are explored below.

The major finding of this retrospective observational study is that the combined PARE data set and the subset of men and women's data supports a prediction equation with nearly equal pursuit and body control log odds (LO). Near equality was seen for the comparisons within data sets for: combined data pursuit versus body control LOs, 2.78 vs 2.80; men's data subset pursuit versus body control LO, 3.0 vs 3.0; and women's data subset pursuit versus body control LO, 2.40 vs 2.45. This near equality was also seen for comparisons between: men's versus the women's data subset pursuit LO, 3.00 vs 2.78; for men's versus women's body control LO, 3.00 vs 2.80. The differences between and all data set and gendered subgroups are all within 1 % point when individual value LO were converted to probabilities, P_i. The LOs differences from highest value to lowest value equated to a very small differences in the percentiles, based on the conversion formula of :

$P_i = \exp^{(Bo+B1x1+b2X2+b3X3)} / 1{+}\exp^{(Bo+B1x1+B2X2+B3X3)}$ (Pampel, 2000, p. 16)

This conversion of significant predictors log odds to success probabilities, with men and women scoring both in the high 90's percentiles, matches real world pass rates for men and women, and is another proof of the gender-neutral PARE test without adverse impact, that could be presented to human rights courts.

An implication from these results may be that clinically important education focuses equally on each component of pursuit and body control regardless of the 70-30 % descriptive statistical times split in actual average time for pursuit versus body control (2:22 vs 1:56 min:s) portions of the PARE total time.

One of the strengths of the study is the large sample size representing a generation of skill-related fitness testing. Finally, in a time where gender equity is socially and politically important, the findings that male/female group differences were not a significant predictor of PARE success is noteworthy, to the public, the courts, and ourselves.

Another major finding was that first moment factors of height and weight were not significant predictor of PARE success. Nor did these factors change the near equality of pursuit and body control predictor value log odds. Since the age factor did not demonstrate a linear relationship with the natural log of the variable, as described by Fields (2013), the age scores were transformed into the square root of age. However, the transformed age factor (i.e., the square root of age) was linear with its natural log, a fundamental assumption of logistic regression and square root of age was a statistically significant predictor factor. Initial logistic regression analysis of the 14,401 PARE score data resulted in a 95% confidence interval for age as a factor centered around the odds value of 1.0 or a 50 - 50 probability, which is unity and clinically insignificant. However as scores > 2 SD outliers were removed the square root of age become statistically significant and somewhat predictive above the unity point, see table 2.3, LO of Exp $^{(\beta)}$ = 0.53.

Figure 2.1 depicts a graphical summary of the changing probability curve around the pass / fail point of 4:00 min:s. The widely changing probability curve depicts a “noisy” signal for probabilities of passing PARE as body control and run times increase. In the graphic the total PARE times steps one second at a time towards the pass / fail score of 240 s. When the probability line, is smoothed out to remove the noise-to-signal effect with an averaging function (excel 138 averaging function), we are left with a depiction of the logistic function sigmoid curve, as expected, in real worlds data. As can be seen in the three probability graphics, in Figure 2.2, 2.3 and 2.4, with three factors we must hold two variables constant to calculate the change in the third variable for probabilities outcome changes. The major changes in probabilities seem to occur at about 242 s in Figure 2.1, closely approximating the cut score pass / fail dummy coding at 240 s. The change in probabilities by about ten seconds in each of the pursuit or body control segments of Figure 2.2, 2.3 and 2.4 can be a clear parsimonious message for clinical education for PARE. The clinically significant message for exercise physiologists can be "just change your PARE run or body control score by about *10 seconds*" as you approach the PES and chance of passing rises quickly, to 100%. This appears to be supported by both the combined population data as a whole and the independent gender subset data.

A limitation of this study was that not all PARE tests were maximal effort capacity defining assessments. Since a large number of these test, about 3,000 plus, were conducted by this author, clinical experience has shown us high fit offices can jog through at less than 4:00 minutes easily, while others go full out and fail at around 7:00 minutes. Although the best PARE times on file were taken for each individual officer, clinical safety issues required judgment and caution when testing an entire population. This is especially true for those over the age of 50 years, unused to maximal efforts and with multiple risk factors, even if optimally medically managed. Further, a widely diverse age range in incumbent officers (i.e., 20 to 60 years in the RCMP) does not afford a comfort in testing each officer every time with a maximal effort with maximum motivation. Some reasons for this were that supervising Certified Exercise Physiologists (CEPs) varied the motivation for some, slowing down older low fit officers or those displaying distress. In pretest screening and later test pacing advise, the Certified

Exercise Physiologist (CEPs), in the context of whole population testing, might vary instructions from “run hard, sprint now” to “slow down, walk, catch your breath” and “pause for a moment”. PARE testing CEP supervisors may actively limit full effort with advancing age of officers and other contra-indicators. Officers must be (a) medically cleared, (b) report being asymptomatic on the day of testing with same day health status screened by the industry standard *Physical Activity Readiness and You* questionnaire (PAR-Q+), (c) be symptomatic but optimally medically stabilized, and (d) be fully cleared for operational duty and PARE training by an occupational physician’s review in the last 90 days. Thus a limitation of this study was the mixing of real world objectives of PARE participation times of walking, jogging, and running hard with research goals of studying the full capacity or full performance ability profile for meaningful predictors of success. As reflection of real world conditions it is still valuable to draw conclusions supported by the data we have, wishing of course for the data we would like to have to make stronger conclusions, hence the following recommendations for other researcher.

A measure of perceived exertion would have helped delimit full effort from participatory efforts. If the purpose of using incumbent PARE data is to take a snapshot of officer’s capacity across all ages in real world conditions, then only maximal exertion tests should have been compared. Further, if the PARE standard were to be cross-validated by incumbent test scores, at +1 or +2 SD from the mean, as suggested by Shephard and Bonneau (2003), then exertion data of maximal efforts would be valuable. The newer supervisor rating of minimal cut-scores as standard reviewed by Rogers et al. (2014) could also be used for used for comparisons of full effort. Removing influential outliers did not equate with removing or keeping maximal efforts, as some 7:00 min:s efforts were maximal efforts; however, they were discarded and some 3:50 efforts were not maximal efforts for some fit officers capable of scoring a personal best of much faster times. Moreover, each test is more than likely an appropriate effort on the day of testing and does not represent the true full capacity of every officer every day and over time. Officers of all ages, present with a wide range of motivations, limited our ability to say the tests are not equal full efforts. This issue of motivation and encouragement make some of the testing results generalizations problematic. A subject for future research would be how to equate real world data in one population to another. A potential applicant or cadet population, with a job in mind, is differently motivated then a RCMP officer with a job. Knowing the percentage effort or rating of perceived exertion or physiological monitored values like METS in relative terms to an individual's maximal capacity after each PARE test might make equating full effort PAREs at different ages might yield additional valuable information. Physiological monitoring is suggested for field-testing by Reed (1982) and may help equate efforts amongst officer groups and age levels. Determinants of maximal aerobic or anaerobic power or other significant maximal work determinants, like anaerobic capacity (Rhodes & Farenholtz, 1992) would facilitate comparisons amongst police populations for fitness for duty.

In conclusion, the combined data supported nearly identical pursuit and body control predictive log odds. Men's and women's data also supported nearly equal pursuit and body control PARE segment predictive values within each sex group, with a slight difference between gender groups. Many subject matter experts and potential officers report expectations of the body control simulation as the limiter for females. This was not supported by a large data set, with thousands of women, with a very high model reported predictive accuracy rating in the high 90s. It is speculated that perception filters cover how we see individuals and potentially false generalization to a class of people. When we see a small female fail we might see female not low fitness performance. In reverse when we may see a larger male fail and we might see low fitness not male. In addition, when a larger female passes we may see fitness not female. Common expectations of gender success and failure on different simulations were not supported in this data analysis. Human factors of height, weight, BMI, and sex were not significant predictors whereas the square root of age was.

2.17 References

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., … Leon, A. S. (2000). Compendium of physical activities: An update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise*, 32(9), S498-S516. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Akerman, J. P., Heckman, G. A., & McKelvie, R. S. (2104). Exercise Capacity and Aging. *American Journal of Lifestyle Medicine.* Retrieved from http://ajl.sagepub.com/content/early/2014/07/10/1559827614542224 Doi: 10.1177/1559827614542224

Anderson, G. S., Plecas, D., & Segger, T. (2001). Police officer physical ability testing: Re-validating a selection criterion. *Policing: An International Journal of Police Strategies and Management*, *24*, 8-31. Retrieved from http://dx.doi.org/10.1108/13639510110382232

Anderson, G. S., & Plecas, D. (2007). *PARE Task Analysis.* Abbotsford, BC: RCMP Learning and Development. Retrieved from http://www.ufv.ca/media/assets/ccjr/reports-and- publications/PARE_Phase_2.pdf

Anderson, G. S., & Plecas, D. (2008). *Police abilities testing: Revalidation of the task analysis*. Ottawa, ON: RCMP Learning and Development Archives.

Andrews, A. & Risher, J. (2006). What does that have to do with being a cop? Employment Standards in law enforcement. *International Association of Chiefs of Police Conference,* Boston, MA. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=187913

Avery, R. D., & Faley, R. H. (1988). *Fairness in selecting employees.* (2nd edition). Reading, MA: Addison-Wesley.

Avery, R. D., Nutting, S. M., & Landon, T. E. (1992). Validation strategies in physical ability testing in police and fire settings. *Public Personnel Management*, *21*(3), 301-312. Retrieved from http://ppm.sagepub.com/

Ayoub, M. M., & McDaniel, J. W. (1974). Effect of operator stance and pushing and pulling tasks, *Transactions of American Institute of Industrial Engineers*, *6*, 185-195. Retrieved from http://www.iienet2.org/

Bar-Or, O. (1987). The Wingate anaerobic test: An update on methodology, reliability and validity. *Sports Medicine, 4*, 381-394. Retrieved from http://link.springer.com/journal/40279

Biddle, D. (2006). *Adverse impact and test validation: A practitioner's guide to valid and defensible employment testing. (*2nd Ed.). Folsom, CA: Biddle Consulting Group Inc.

Biddle, D., & Sill, N. S. (1999). Protective service physical ability test: Establishing pass or fail, ranking and banding procedures. *Public Personnel Management*, *28*(2), 217-225. Retrieved from http://ppm.sagepub.com/

Boocock, M. G., Haslam, R. A., Lemon, P., & Thorpe, S. (2006). Initial force and postural adaptations when pushing and pulling on floor surfaces with good land reduced resistance to slipping. *Ergonomics*, *49*(9), 801-882. Doi:10.1080/00140130600562876

Bonneau, J. (1998). *Submission to the Canadian Human Rights Commission: PARE. Briefing paper outlining PARE as a BFOR development and the 4 minutes standard.* Ottawa, ON: RCMP, Fitness & Lifestyle Archives.

Bonneau, J. (2001). *The development of PARE: An evolution*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J., & Brown, J. (1995). Physical ability, fitness and police work. *Journal of Clinical Forensic Medicine*, *2*, 157-164. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Brown, J. & Trottier, A. (1994). *Police health: A physician's guide for the assessment of police officers*. Ottawa, ON: Canada Communications Group.

Boyce, R. W., Jones, G. R., Schendt, K. E., Lloyd, C. L., & Boone, E. L. (2009). Longitudinal changes in strength of police officers with gender comparisons. *Journal of Strength and Conditioning Research, 23*(8), 2411-2418. doi: 10.1519/JSC.0b013e3181bac2ab

Campion, M. A. (1983). Personal selection for physical demanding jobs. *Personnel Psychology*, *36*, 527-550. doi: 10.1111/j.1744-6570.1983.tb02234.x

Chaffin, D. B., Anderson, G. B. J., & Martin, B. J. (2006). *Occupational Biomechanics* (4th Ed.). Hoboken, NJ: John Wiley and Sons.

Chaffin, D. B., Herrin, G. D., Lee, K. S., & Waikar, A. M. (1991). Effect of handle height on lower back loading in cart pushing and pulling, *Applied Ergonomics*, *22*(2), 117-123. doi: 10.1016/0003-6870(91)90310-E

Chaffin, D. B., & Resnick, M. L., (1995). An ergonomic evaluation of handle height and load on maximal and submaximal cart pushing, *Applied Ergonomics*, *26*(3), 173-178. doi: 10.1016/0003-6870(95)00014-4. Retrieved from http://www.journals.elsevier.com/applied-ergonomics/

Chapman, A. E. (2008). *Biomechanical analysis of fundamental human movements*. Champaign, Il: Human Kinetics.

Collingwood, T. R. (1988a). Implementing programs and standards for law enforcement physical fitness. *The Police Chief,* April *55*(4*),* 20-24. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1988b). Physical fitness leadership in law enforcement. *The Police Chief*, April *55*(4), 28-34. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1995). Physical fitness standards: Measuring job relatedness. *The Police Chief*, February *62,* 31-31. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (2004). Underlying physical fitness factors for performing police officer physical tasks. *The Police Chief,* March *71*(3), 1-7. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Deakin, J. M., Smith, J. T., Pelot, R. P., & Weber, C. L. (2000). Methodological considerations in the development of physical maintenance standards. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Deakin, J. M., R. P. Pelot, J. M. Smith, J. M. Stevenson, L. A. Wolfe, S. W. Lee, S. P. Jaenen, S. A. Hughes, J. W. Dwyer, and A. D. Hayes. (1996). *Development of a Bona Fide Physical Maintenance Standard for CF and DND Fire Fighters*. Final Report Submitted to Canadian Forces Personnel Support Agency. Kingston: Queen's University.

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). *The Sage Handbook of Qualitative Research* (3 rd ed.). Thousand Oaks, CA, Sage.

De Looze, M. P., van Greuningen, K., Rebel, J. Kingma, I., & Kuijer, P. P. F. M. (2000). Force direction and physical load in dynamic pushing and pulling. *Ergonomics*, *43*(3), 377-390. doi: 10.1080/001401300184477

DiDomizio, J., & Kier, P. J. (2010). Forearm posture and grip effects during push and pull tasks. *Ergonomics, 53*(3), 336-343. doi:10.1080/00140130903389076

Farenholtz, D. W., & Rhodes, E. C. (1986). *Police Officer Physical Abilities Study*. New Westminster BC: Justice Institute of British Columbia, Canadian Police College Library archives.

Fields, A. (2013). *Discovering statistics using IBM SPSS statistics*. London, UK: Sage.

Fleishman, E. A. (1964). *The Structure and Measurement of Physical Fitness*. Englewood Cliffs, New Jersey: Prentice Hall.

Gledhill, N., Bonneau, J., & Salmon, A. (2001). *Bona Fide Occupational Requirements: Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Gledhill, N. & Bonneau, J. (2001). Objectives, processes and consensus summary of the national forum on establishing Bona Fide Occupational Requirements. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Gledhill, N., & Shaw, C. (1996). *Report on the characterization, test construction and validation phase of the medical, physical, skills and abilities project.* Toronto, ON: Race Relations and Policing Unit of' the Ministry of the Solicitor General and Correctional Services.

Glitsch, U., Otterbach, H. J., Ellegast, R., Schaub, K., Franz, G., & Jager, M. (2007). Physical workload of flight attendants when pushing and pulling trolleys aboard aircraft. *International Journal of Industrial Ergonomics*, *37*, 845-854. doi:10.1016/j.ergon.2007.07.004

Hilde, J. M. (2009). *Logistic regression models: Texts in statistical science*. London, UK: Chapman and Hall.

Hoozemans, M. J. M., van der Beek, A., Frings-Dresen, M. H. W., van Duk, F. J .H., & van der Woulde, L. H. C. (1998). Pushing and pulling in relation to musculoskeletal disorders: A review of risk factors. *Ergonomics*, *41*(6), 757-781. Retrieved from http://www.tandfonline.com/toc/terg20/current

Hogan, J., & Quigley, A. M. (1986). Physical standards for employment and courts. *American Psychologist, 41*, 1193-1217. Retrieved from http://www.apa.org/pubs/journals/amp/

Hosmer, D.W. & Lemshow, S. (2000). *Applied logistic regression (*2nd edition).Toronto, ON: Wiley and Sons.

Innes, E. & Straker, L. (1999). Validity of work related assessments. *Work*, 13,125-152. Retrieved from http://www.iospress.nl/journal/work/

Jackson, A. S. (1994). Pre-employment physical evaluation. *Exercise and Sports Science Reviews, 22*(1), 53-90. Retrieved from http://journals.lww.com/acsm-essr/pages/default.aspx

Knapik, C.G. & Marras, W. S. (2009). Spine loading at different lumbar levels during pushing and pulling. *Ergonomics*, *52*(1), 60-70. doi:10.1080/00140130802480828

Kroemer, K. H. E. & Robinson, D. E., (1971). Horizontal static forces exerted by men standing in common working postures on surfaces of various tractions. *Applied Ergonomics, AMRL-TR 70-114.*

Kuijer, P. P. F. M., Hoozemans, M. J. M., & Frings-Dresen, M. H. W. (2007). A different approach for the ergonomic evaluation of pushing and pulling in practice. *International Journal of Industrial Ergonomics*, *37*, 855-862. doi:10.1016/j.ergon.2007.07.011

Kumar, S. (1995). Upper body push-pull strength of normal young adults in sagittal plane at three heights. *International Journal of Ergonomics*, *15*, 427-436. doi:10.1016/0169-8141(94)00062-8

Moss, P. A. (1998). Recovering a dialectical view of rationality, *Social Indicators Research*, *45*, 55-67. Retrieved from http://www.springer.com/social+sciences/journal/11205

Markus, K. (1998). Validity, facts, and values sans closure: Reply to Messick, Reckase, and Zimmerman. *Social Indicators Research, 45*, 73-82. Retrieved from http://www.springer.com/social+sciences/journal/11205

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts*. Thousand Oaks, CA: Sage.

Messick, S. (1998). Test validity: A matter of consequence. *Social Indicators Research, 45*, 34-45. Retrieved from http://www.springer.com/social+sciences/journal/11205

Pampel, F. C. (2000). *Logistic regression: A primer*. London, UK: Sage Publications.

Payne, W., & Harvey, J. (2010). A framework for the design and development of physical employment test and standards. *Ergonomics, 53*(7), 858-871. doi:10.1080/001/40139.2010.489964

Prenzler, T. (1996). Obstacles to the advancement of women in policing. *Australian Institute of Criminology Conference*. Retrieved from http://www.aic.gov.au/media_library/conferences/policewomen/prenzler.pdf

Quigley, A. (2008). Fit for duty: The need for physical fitness programs for law enforcement officers. *The Police Chief*, June *75*(6), 62-70. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Rackase, M. D. (1998). The interaction of values and validity: Does a test's level of validity depend on a researcher's values? *Social Indicators Research*, *45*, 45-54. Retrieved from http://www.springer.com/social+sciences/journal/11205.

Reed, A. (1982). Field tests. In *Physiological testing of the elite athlete*, J. D. MacDougal, H.A. Wenger, & H. J. Green (Eds.). Ithaca, NY: Movement Publications.

Rhodes, E. C. & Farenholtz, D. W. (1992). Police officer's physical abilities test compared to measures of physical fitness. *Canadian Journal of Sports Science*, *17*(3), 228-33.Retrieved from http://www.researchgate.net/journal/0833-1235_Canadian_journal_of_sport_sciences_Journal_canadien_des_sciences_du_sport

Schulze, C. (2012). The Masculine yardstick of physical competence: U.S. police academy fitness tests. *Women and Criminal Justice*, *22*, 89-107. doi:10.1080/08974454.2012.662117

Sharkey, B.J., & Davis, P.O. (2008). *Hard work: Defining physical work performances,* Windsor, ON: Human Kinetics.

Snook, S. H., & Ciriello, V. M. (1991). The design of manual handling tasks: revised tables of maximum acceptable weights and forces. *Ergonomics*, *34*(9), 1197-1213. doi:10.1080/00140139108964855

Stanish, H. I., (1993). Comparison of the Physical Ability Requirement Evaluation to measures of physical fitness. [Abstract]. *Canadian Journal of Applied Physiology, 18:*436 Retrieved from http://www.nrcresearchpress.com/journal/apnm

Stanish H., & Campagna P, (1993).Comparison of the Physical Ability Requirement Evaluation simulated fight component to measures of muscular strength, endurance and power. Unpublished manuscript, Ottawa, RCMP Fitness & Lifestyle Archives

Stanish, H. I., Wood, T. M., & Campagna, P. (1999). Prediction of performance on the RCMP Physical Abilities Requirement Evaluation. *Journal of Occupational and Environmental Medicine, 41*(8), 669-677. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Van der Beek, A. J., Hoozemans, M. J. M., Frings-Dresen, M. H. W., & Burdorf, A. (1999). *International Journal of Industrial Ergonomics*, *24*, 417-429. Retrieved from http://www.journals.elsevier.com/international-journal-of-industrial-ergonomics/

Wagner-Wisotzki, N. (2005). *Fitness and lifestyle management for law enforcement* (2nd Ed.). Toronto, ON: Edmund Montgomery Publishers.

Ward, A.W., Stoker, H. W., & Murray-Ward, M. (Eds.). (1996). *Educational Measurement: Origins, Theories and Explications.* Lanham, MD: University Press of America.

Wilmore J. H.j & Davis, J. A. (1979). Validation of a physical abilities test for the selection of state traffic officers. *Journal of Occupational Medicine, 21*(1), 33-39. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Zieky, M. J., & Livingston, A. (2004). *A Primer on setting cut scores on tests of educational achievement.* Citing material published elsewhere by Zieky & Perie (2004) in *Passing Scores: A Manual for Setting Standards on educational and occupational Tests*. Retrieved from https://www.ets.org/Media/Research/pdf/Cut_Scores_Primer.pdf

Zimmerman, D. W. (1998). Commentary on “Science, measurement, and validity: Is completion of Samual Messick’s synthesis possible?” *Social Indicators Research*, *45*, 69-72. Retrieved from http://www.springer.com/social+sciences/journal/11205

Zumbo, B. D. (1998). Opening remarks to the special issue on validity theory and methods used in validation: Perspectives from the social and behavioral sciences. *Social Indicators Research, 45*, 1-3. Retrieved from http://www.springer.com/social+sciences/journal/11205

Zumbo, B. D. (2000). Methodology and measurement matter in establishing a Bona Fide Occupational requirement for physically demanding occupations. In N. Gledhill, J.

Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Chapter 3

Pacing Effects in the PARE Pursuit Circuit

Policing requires physically demanding skills, some of which are rarely used, but must be available to the officer on demand in public critical incident callout (Anderson et al., 2001, Farenholtz & Rhodes, 1990). The short and long foot pursuit, of various distances, over various terrains and obstacles, is one of those uncommon but essential demands included in a police physical abilities test. Both the running pursuit construct and the body control pushing and pulling constructs are always included in police skill-related job simulations for police and corrections fitness for duty testing (Avery et al., 1992; Bonneau, 2001; Jackson, 1994; Jackson & Wilson, 2013; Jamnik et al., 2010; Osborn 1976; Strating et al., 2010; Trottier & Brown, 1994, Wilson & Bracci, 1982). The job simulation mirrors the critical incident callout flow of "getting to a problem, dealing with a problem, and removing a problem" (Anderson et al., 2001, p.8). The physical ability tests have been used for: applicant screening (Jackson, 1994), academy graduation testing, and incumbent officer formative learning during periodic recertification (Bonneau, 2001; Dunsmore & Hunter, 2001; Shell, 2003; Schulz, 2012). Research into the development, implementation, and ongoing evaluation of police bona fide occupational requirement (BFOR) tests (Gledhill, Bonneau, & Salmon, 2001) are dominated by unpublished government reports (Bonneau, 2001; Gledhill & Shaw, 1996; Lagasse, 1989) and some academic reports (Jamnik, Thomas, Gledhill & Shaw, 2010).

There is a limited amount of peer reviewed published academic research in occupational medicine that reports the results of large incumbent officer population testing for pass or fail performance groups in job simulation tests that include group factors of sex, administrative or operational police duties, and years of service (Gaul & Wenger, 1992; Jackson & Wilson, 2013; Strating, Bakker, Dijkstra, Lemmink & Groothoff, 2010). Gaul and Wenger (1992) studied incumbent RCMP officer's PARE testing over 2 years and reported sex differences for pass / fail. However, they did not analyze each of the two major sections of the PARE separately, namely the pursuit and the body control circuits, for gender differences. Strating et al. (2010) found regression factors for skill-related testing in a large numbers of Dutch police officers, but the test was not divided into separate pursuit and body control sections. Differences between men and women in the Dutch test were in order of 26 seconds. There were also differences between, older versus younger, and operational versus administrative groups of officers.

Stanish et al. (1999) studied PARE for predictors of success in a small student study sample. There was limited pass / fail dichotomy for men, but some pass / fail dichotomy for women, although women's average scores represented very low fitness levels. Missing in the academic literature is an examination of the pursuit circuit and any sex differences in pass and fail.

The common police and corrections fitness tests that include pursuit circuit or shuttle runs, to replicate the long demanding pursuit in a test, are all described below (Bonneau, 2001; Farenhotlz & Rhodes, 1986; Gledhill & Shaw, 1995; Jamnik et al, 2010). Additionally, the human bioenergetics energy system theory and task dependency theory informs this research and are briefly described in terms of fatigue and pacing effects.

3.1 Gender Issues

Human resources, police trainers, and human rights groups have easily accepted the third moment of skill-related job simulation testing (Deakin et al., 2001). The criterion reference times of young male inmates, as the most able and therefore the most difficult and demanding police arrest work has been critiqued as gender centric for a masculine metric (Lonsway, 2003; Schulze, 2012). Critics argue the cut-sore is not based on the intensity of the work sample, but the sex of the subjects used to set the cut-score. Thus, ability testing was perceived and still is perceived as a potential limiting factor for entry and advancement of women in policing in the RCMP Gender-Based Assessment (2013) and Gender and Respect Reports and amongst other peer reviewed articles (Lonsway, 2003; Prenzler, 1986, Schulze, 2012). Critics did not address the fact that the majority of police clients cited in critical incident reports were male (Anderson et al., 2001, Farenholtz & Rhodes, 1990). Methodologically speaking, too few women and too many low fit women were consistently used as subjects at the time of test development in the early and mid-1980s. Very few women were represented in the police demographic in validation of tests. Examples of this are the case of the PARE development validation study (Bonneau 1996, 2001) and PARE incumbent pilot study (Gaul & Wenger, 1992). Early validation studies consistently reported women as a group not reaching or passing the cut-score standard, with women's groups on average almost always reporting low fitness scores (Stanish, Wood, & Campagna, 1999) consequently, making the determination of validity for women officers problematic (Bonneau, 1996; Gaul & Wenger, 1992;). With a generation of PARE data available, this shortcoming in the published and grey literature (unpublished reports) can now be addressed by a larger population study reporting in published academic reports and in easily retrieved study databases.

The previous retrospective observational study of 13,709 officers found that Physical Abilities Requirement Evaluation (PARE) data did not support human factors of height, weight, BMI, or sex factors, used in moment one and two fitness testing, as significant PARE success prediction factors. Only the human factors of age, plus the performance factors of pursuit and body control times, were found to be statistically significant predictors, with the latter two being equal predictors of PARE success. The variables pursuit time, body control time, and square root of age had a predictive model 99 % accuracy rating in logistic regression analysis.

The PARE is used by the RCMP as both an applicant Physical Employment Standard (PES) and as a serving officer periodic recertification training, with other police skills training (Shell, 2003). PARE consists of a pursuit and a body control segment time, added for total time, with a pass at 4:00 min:s.

From the previous retrospective study for a generation of police officers the average time for the PARE for men (*n* = 11,582) was 217 s or 3:37 min:s for women (*n* = 2,127) it was 248 s or 4:08 min:s, and for the combined group (n=13,709) it was 222 s or 3:42 min:s. The pursuit course average times for the combined group was 156 s, (2:36), for the men's group it was 153 s, (2:33); and for women's group I was 171 s (2:51). The body control average times were for combined group was 64 s, (1:04 min:s); for the men's group was 63.3 s (1:03.3 min:s); and for women's group was 76.8 s (1:06.8 min:s). No data were available in the national database for progress through the six repeat laps of the PARE pursuit course for either sex. This leave unaddressed the issue of the long and demanding foot pursuit and fatigue during the pacing repeat laps times. The associated question is do low versus high fit women pace through the pursuit chase the same way as low versus high fit men?

A list of Canadian police and corrections BFOR tests is described below.

3.2 The Tests

Anderson et al. (2001) refers to the work of Osborn (1976), and Wilmore and Davies (1979) as the beginning of a dominant methodology of skill-related testing for researching critical tasks and demands, creating a skill-related test, and always implementing with a follow-up police officer validation check. This has been conducted for large numbers of municipal (Anderson et al., 2001; Farenholtz & Rhodes, 1990) and federal police like Canada and Australia (Anderson & Plecas, 2007; Bonneau, 1996, 2001). The job simulation tests that have an adequate pursuit component, according to officer follow-up responses for the initial test constructs are the PARE (Trottier & Brown 1994, Bonneau & Brown, 1995; Bonneau 1996, 2001), the Police Officer Physical Ability Test (POPAT), and the Correctional Officers Physical Abilities Test (COPAT) (Farenholtz, Rhodes, & Burrell, 1985). See Appendix A for additional detailed graphical layout of the footprint of the discrete items of each test. An original Canadian skill-related test was developed for municipal corrections in Ontario named the PAR

(Marchand, Reed, Thoden, & McNeely, 1995). Its performance standards cut-score was never developed; therefore, the test was never implemented. For federal corrections, the COPAT, see Appendix A for graphical layout, was developed but never implemented due to labour pushback, described by Green and Kreuter (2005) as common in a unionized environment for mandatory health promotion programs. With COPAT (Farenholtz & Rhodes, 1986), the incumbent officers who did not pass the standard, wanted to be accommodated in positions outside of jail duty. This is considered a simplistic early understanding of the duty to accommodate legislation (Stolz, Federal corrections personal communications, 2001). Recently the Fitness Test for Correctional Officers (FITCO) (Jamnik, Thomas, Shaw & Gledhill, 2010) was developed for municipal corrections in Ontario. When the above-published tests undergo ongoing evaluation, reports back from operational officers have a common critique for tests with shorter pursuit segments in the skill-related job scenario. Officers report the test was too easy or too short in total time to reflect a rare, but difficult, critical incident with a long foot pursuit. Officers gave feedback that a police job simulation test that does not raise effort to a maximal cardiovascular level in a pursuit circuit was deemed insufficient. Therefore, the tests with too short a run circuit, like the PREP, have add a health-related aerobic capacity construct field test, like the Leger-Boucher 20 m shuttle run. These tests are the FITCO and the Physical Readiness Evaluation for Police (PREP) (Ministry of the Solicitor General and Correctional Services, 1999). The latter two tests have compensated for the essential demanding component of the longer pursuit missing in their circuit test by adding the Leger-Boucher 20 m shuttle run. The reasons given by reviewing officers for the insistence of this element being included in a police skill-related test is that their pursuit circuits were too easy and or too short in time to completion versus the tougher longer pursuits some officers have experienced on the job (Gledhill & Shaw, 1996; Shaw personal communication, Ontario Police College). Other health-related tests for the cardiovascular fitness field test assessments that could be used are the outdoor 1.5-mile run, or the 12-minute Cooper test or prior to that, the 20-minute Balke run. The 1.5-mile run has been discontinued by most major police services as it has had deaths associated with it (Deakin et al., 1994). All the traditional field test methods for determining cardiovascular endurance or predicted maximal aerobic power, plus their score interpretation tables, have been practiced and published by the Cooper Institute in Texas for over a generation (Collingwood, 2004). These tests have always included test score interpretation by age and sex separate tables regardless of the fact that the most common arrest demographic in a critical incident callout was a young male suspect (Collingwood, 1995; Farenholtz & Rhodes, 1990; Hoffman & Collingwood, 2005; Shell, 2003).

3.3 Incumbent Testing

In service police training programs conduct periodic fitness testing certification of incumbent police officers concurrent with other police skills recertification (Shell, 2003). The original Canadian police skill-related abilities test that PARE, was based on the POPAT. Originally, POPAT had a three-foot vault rail to be cleared ten times as the last element of the run circuit task to solicit maximal cardiovascular intensity efforts. This was supposedly a low skill-related discrete item to ensure inclusion of a health-related construct – namely the maximal cardiovascular demand. This test adjustment predated by 15 years the clear distinction reported in the 2001 BFOR consensus conference for health and skill-relate testing constructs. Bonneau (2001) found the rail jump push to maximum heart rate redundant to the PARE pursuit circuit in the circuit already elicited a maximal heart rate, in some as early as lap three. Thus, he found it to be redundant to have the 10 repeat rail vaults at test end of POPAT (Bonneau, 2001; Reed, 2001; Sommerfeld, 1998). The PARE for federal police was created shortly after, as a modified POPAT. The PARE integrated the 10 three foot rail jumps of POPAT into both the long pursuit circuit, one per each of the six laps, and in the body control simulation.

The body control station included the four controlled falls between the six arcs pushing and pulling of 80 lbs. See appendix A for graphical representation of the PARE flow and stations. Incumbent testing has been reported by Dunsmore and Hunter (2001) for Ontario municipal and federal police as a well as Toronto area firefighters. The PARE was first applied to incumbent RCMP officers in the pilot fitness program reported by Gaul and Wenger (1992). A review of the scant academic published results for the RCMP PARE testing scores, for potential applicants to existing incumbent pools of officers, reveals very low female subject participation. Further the incumbent female subjects were low fit and barely able to pass the standard PES of 4:00 min:s (Anderson & Plecas, 2001, 2007, 2008; Bonneau, 1996, 2001; Gaul & Wenger, 1992; Strating et a;., 1999). However, fitness test scores for all male and female cadets in national police service training academies (i.e., FBI) are reported as high fit during police academy training (Knapik et al., 2011).

3.4 The PARE

Gaul and Wenger (1992) introduced the original PARE test, with various push pull arrest simulation machines, to incumbent populations in 1990-1992 in a two-year pilot project. The pilot project testing was conducted with four potential tests over two years, with a fitness training support program, in several provincial regions for both rural and urban detachment personnel. Only 14% of women officers passed initially, whereas over the two-year period, that percentage of passes increased to 48%, although there were less than 16 women tested in the 4 potential time periods over two years. Gaul and Wenger (1992) attributed most of the difficulty for women at the push pull station. The under 30 years of age group pass rate was 80% after the two years. The 31 to 40 years age group had similar results, with 75% passing. The authors provided limited age reporting for passing PARE at 4:00 min:s, with the 41 to 50 years age group. “Members between 41 – 50 years appeared to have more difficulty with PARE, however 75% of this group were able to complete the PARE in under 4:30 [min:s]” although 4:00 min:s was still the pass goal (Gaul & Wenger, 1992, p. 21). Insufficient data was reported as available for the over 50 years of age group.

The pilot project reported a combined group slight increase in pass rates over the 2 years with four possible retries. Initial test pass rates for all officers were 63% under 4:00

min:s, and 83% under 4:30 min:s. At project end, 72% passed under 4:00 min:s, 88% were under 4:30, and only 11% unable to achieve the 4:30 min:s mark (Gaul & Wenger, 1992). The mean improvement in PARE times, 12 s, was reported as statistically significant ($p < .001$) but no effect size was reported.

3.5 Energy Systems

The three major metabolic pathways of human bioenergetics for muscular energy production is reviewed by Wells et al. (2009) and many others (Brooks & Fahey, 1984; Fox & Mathews, 1976; Gore, 2000). The authors report that the use of the three pathways is intensity and duration dependent. The three system are described as: (a) the aerobic oxidative systems for "longer duration activities of low to moderate intensities"; (b) the aerobic glycolytic system for "short to moderate duration activities of higher intensity"; and, (c) the higher energy phosphate system for "short duration activities of high intensity" (Wells et al., 2009, p.83). Gore (2000) predicts energy system power drops will be strongest earliest, due to anaerobic power alactic capacity exhaustion of ATP re-synthesis first. Then power drops continue as anaerobic lactic capacity exhaustion ensues in the next 65 seconds, due to metabolic limitations of anaerobic glycolytic energy production pathways inhibition by metabolic acid accumulates. Finally, up to 90 s into the PARE test, lower ATP rates of production of energy sees a lower level aerobic power levels continue, at slightly diminished levels to exhaustion.

Total maximal power output in an all-out or sustained high level effort demonstrates sustained power drops resulting from the lag of these three dominant bioenergetics systems to provide peak muscular power in children (Wells, Selvadurai & Tein, 2009), and in athletes, and is well documented in anaerobic Wingate testing for sports and performance (Gore, 2000; MacDougal, Wenger & Green, 1991). The transition from one dominant muscular metabolic energy system pathway to another, as the dominant source of energy, predicts whole body muscular power drops for tests of near maximum efforts, if the test is over a time course of two to three minutes (Brooks & Fahey, 1986; Wells, 2009).

Central nervous system fatigue could be implicated in any of the "proper functioning of receptors, CNS [central nervous system]¸integrating centres, sensory cortex, spinal cord, α-motor neurons, γ loop, motor ends plates" (Brooks & Fahey, 1986, p. 709). Depletion of Ca^{2+} or other neural transmitters, may slow muscle function and power output.

3.6 Fatigue and Power Selection

In Weir, Beck, and Housh (2006) the authors review fatigue models like the central governor model or the task dependency model of fatigue. The authors cite the Bigland, Ritchie, and Woods (1984) and Gandevia's (2001) operational definition of fatigue as "an exercise-induced reduction in the ability to exert muscular force or power, regardless of whether or not the task can be sustained" (Wier, Beck, Cramer, & Housh, 2006, p. 574). Total maximal power output in an all-out or sustained high-level effort demonstrates that sustained irreversible short-term power drops. This is speculated to be the result of the lag of delivery of Adenosine Triphosphate (ATP) from various dominant bioenergetics pathways in providing peak muscular power to peak muscular ATP demand (Weir et al., 2006). The above power drops are well documented in laboratory testing of both aerobic and anaerobic testing modalities in sports and athletic groups by Gore (2000). The transition from one dominant muscular metabolic energy system pathway to another is speculated to be change in the dominant source of energy, resulting in muscular power output drops for tests of maximum, supra maximum and sustained near maximum efforts. This is especially true if the test is over a time course of two to three minutes long (Wells & Tein, 2009). Fatigue in the anaerobic system for a 2 to 5 minute event is not from glycogen substrate depletion, as moderate levels of exercise can be maintained for hours (Brooks and Fahey, 1986). Also discounted for fatigue are blood sugar depletion in short time bouts of exercise. In short high intensity bouts of exercise, with dissociated H^+ ion, from lactic acid and other metabolic accumulation from rapid glycolytic breakdown of muscle glycogen, can drop pH and inhibit the rate-limiting step of glycolysis by inhibiting Phosphofructokinase (PFK). H^+ may also displace Ca^{2+} from "troponin thereby interfering with muscle contraction" (Brook & Fahey, 1986, p. 707). Other effects of elevated $H^{+,}$ lowered pH is "inhibit the combination of O_2 with hemoglobin" (Brooks & Fahey, 1986, p. 707). Brooks and Fahey (1986) suggest NMR studies give weight to CP depletion over lactic acid accumulation as source of metabolic fatigue. Further, the sequestration of Ca^{2+} by the mitochondria may uncouple oxidative phosphorylation of ADP to ATP. Inadequate oxygen may cause both CP levels to drop and lactic acid and hydrogen ions to accumulate. Thus, energy systems theory informs our study, sets the expectations for sustained performance power drops results with the largest power drop in the first 25 seconds, as described below

3.7 Physiological Demands of PARE

In Bonneau's (2001) amalgam of unpublished government reports titled *The Evolution of PARE Report,* Reed (1992) reported direct oxygen consumption measures of the cardiovascular demands of the PARE job simulation test. When measured with a portable oxygen analyzer directly, PARE elicited maximum oxygen extraction per minute (VO_2), in l/min, for men and women reported as 2.88 L/min (± 0.12) and 1.88 L/min (± 0.18) or relative to body weight in ml/kg body weight for men and women, as 33.7 ml/kg/min (± 1.37) and 31.1 ml/kg/min (± 3.63). Data was collected during actual running of PARE while subjects wore the TEEM portable oxygen analyzer for some 20 males and 5 female subjects. Reported as equal were sex group differences peak relative oxygen consumption levels of $\dot{V}O_2$ s at 33.7 vs. 31.1 ml/kg/min, or about 10 METS, achieved for both sexes. The field testing solicited near age predicted maximal heart rates with 179 beats per minute (bpm) for males and 186 bpm for females. The relative intensity of oxygen demand in the simulation, in relation to maximal aerobic power (MAP), was documented as 66 % for the small subject pool of male and female officers. However, the difference in PARE scores for this small sample group of men and women were clinically very significant with 214 s (3:34 min:s; a good pass) versus 246 s (4:05 min:s; a slight fail) on average. The $\dot{V}O_2$ range of police work reviewed by Bonneau (2001), varies from Reed (1996), Gledhill et al. (1992), and Avery et al. (1992) were from 2.5 to 3.5 L/min.

3.8 Pacing and Power Drops

Abbiss and Laursen (2008) defined the “distribution of work, or pattern of energy expenditure” as “pacing or pacing strategy” (p. 240). The issue here is how participants “regulate their work output in order to optimize overall performance” (Abbiss & Laursen, 2008, p, 240). The term “pacing more accurately refers to performance times or velocities” (Abbiss & Laursen, 2008, p. 240) and in the case of police circuit test time per lap of a circuit. The “regulation of pacing is largely determined by the ability to resist fatigue” (Abbiss & Laursen, 2008, p. 240). The difference between all out short duration sprints of < 60 s, is contrasted by the authors with > 2 min paced endurance performance events. Strategies used include, negative, all out, positive and parabolic shaped and variable pacing strategies. Each of these strategies are described below from Abbiss and Laursen’s 2008 review article.

Negative pacing, usually called a negative split in track and field; sees the second half of a distance run faster than the first half. The negative split strategy sees an increase in motor unit recruitment near the end of performance using any stores of anaerobic energy supply left, finishing just before metabolic accumulations limit performance during the final surge or sprint. All-out pacing is common for short duration sprint events, of less than 30 – 60 s, with the optimal distance cited by Keller (1974) as 291 m. where the cost of breaking inertia is beast paid early and lesser submaximal effort is used to hold a slight slowing pace as long as possible (Abbiss & Laursen, 2008). Holding an all-out pacing effort for greater 45 seconds has been found in Wingate anaerobic power tests to be impossible (Gore, 2000) with everyone slowing down after 10 s. Positive pacing is defined as when a performer slowly and gradually declines throughout a race, in contrast to negative pacing. This strategy is reportedly responsible for demonstrating the highest fractional utilization of maximal aerobic power (VO_2 max.), the highest post-blood lactate concentrations, the highest respiratory gas exchange ratios, and the highest perceived exertions – compared with even split pacing. Self-selected pacing in distance events tends to be of this nature. Even pacing is self-evident, and most often observed in the longer hours plus events. Parabolic-shaped pacing efforts represent high initial paces, followed by slower intermediate or middle event pacing, finishing with an event-ending surge of effort, finishing faster as in a negative split. Performance profiling during distance events has seen athletes "progressively reduce speed during an endurance trial but tend to increase speed during the latter portion of the event" (Abbiss & Laursen, 2008p. 245). Commenting on the usual J shaped faster start and finish of parabolic shaped pacing, Abbiss and Laursen comment that "the choice of pacing strategy does not appear to be dictated by changes in any one physiological system, but instead may be influenced via a complex system of integrated feedback from a number of sources, including prior experience and anticipated duration" (p. 246). Variable pacing is most common with varying external environmental conditions, as changing resistance during an event, like varying wind speeds, and terrain incline changes over a route, where it is considered an optimal strategy to vary pacing accordingly. Physiological monitoring has shown this to produce more accumulated metabolic fatigue agents than constant pacing strategies.

Weir, Beck, and Housh (2006) states that power output as pace selection and the level of fatigue are responses solicited from a maximal effort that are driven by higher order cognitive functions in order to maintain homeostatic, when fatigue is explained by the central governor model. Running Canada's website (http://runningmagazine.ca/) call one of these strategies the teleoanticiaptory effects of looking to the end of the run and anticipating how much of a pace can be sustained without completely failing with fatigue. This is critiqued by Weir et al. (2006) in comparing equal power outputs in 30, 33 and 36 s Wingate tests were teleoanticipation would have seen dropped power outputs in the longer tests. The authors cite Bigland, Ritchie, and Woods' (1984) operational definition of fatigue as "an exercise-induced reduction in the ability to exert muscular force or power, regardless of whether or not the task can be sustained" (Wier et al., p. 574).

This bioenergetics energy systems theory sets the expectations for sustained performance time pacing decrements predictions and informs the current studies.

3.9 Purpose

A retrospective observation study should reveal whether a divisional cohort data set can support the contention that the PARE test in general, and the pursuit circuit particularly, is sex neutral for patterns of pacing effects. If not, the different time course of pacing per lap pattern for male and female officers should be documented to understand fully the sex differences in pacing. The purpose of this research was to assess the effects of pacing in the pursuit simulation of the RCMP PARE test segment and to see if this relationship changes across sex groups, while, controlling for human factors of height, weight, and age.

3.10 Ethical Approval

Permission to use divisional RCMP PARE data was granted in writing by Human Resources Officer in Charge (OIC) and front line supervisor in January 2013 with the proviso of confidentially of names and individual references of officers removed, through coding data and group reporting as outlined in Tri-Council policy on reporting research of human subjects.

3.11 Research Design

A retrospective observational study of PARE data for a divisional data set of convenience, downloaded from the RCMP national PARE database, was conducted to assess the effect of repeat lap times on pacing for sex and performance groups. The Divisional PARE test records represent all the testing conducted by this author in one federal enforcement non-contract division, from 2006 to 2013, where lap time records augmented the national records. The operational definition of pacing fatigue was a time deficit per lap from lap one, as a reduced work rate for a set distance or a power level drop per equal distance PARE lap. The research was conducted in two analysis steps. First a 2 x 2 x 6 complex mixed design repeated measures ANOVA was conducted on the divisional data of 620 cases with sex as male / female and performance, as pass / fail, as the independent between group factors, and six repeat lap times during PARE pursuit circuit as the within subject factor. The divisional data was to be assessed with covariates of interest like height, weight, BMI and age as covariates human factors entered into a ANCOVA but because of pre-existing group differences, between independent sex groups the use of these covariates was untenable as a valid statistical procedure for ANCOVAs (Fields, 2013; Wildt & Aholt, 2001). Thus, in a second step, a 2 x 2 x 6 complex mixed factorial ANOVA for the sub-set of 61 closely matched male and female officer data, was conducted on the equalized group for covariate data. Using the technique suggested by Zumbo (2001) in the BFOR consensus manual (Gledhill et al., 2001) equal groups of matched male and female officers (61) were created. This was achieved by closely matching (< 0.1 difference in years and BMI) all female records from the divisional database to an equal male case for BMI and age in years. Thus, equal group covariates of BMI and age were achieved through selection of the matched pairs. Then a contrast was possible for divisional data and the 61 matched pairs data sets using a 2 x 2 x6 complex mixed design ANOVA, with between subject groups of sex and performance assessed for within subject factors of six repeated lap times.

3.12 Data Treatment

PARE laps times were manually added to record lines from a download of the national Human Resources Management Information System (HRMIS) database for a division of convenience for each officer to complement height, weight, age, sex, PARE pursuit time, and body control times data. Sex was coded female = 1 and male = 0. Performance groups (pass) was coded pass or fail PARE at the PES 4:00 min:s level where fail = 1, and pass = 0. PARE data times in the both data sets, divisional records of 620 and 61 matched pairs records, were trimmed at +2 *SD* above the mean for total PARE time and repeat lap times in ANOVA procedures to remove influential outliers as suggested by Fields (2013). Pursuit and body control times were recorded to the nearest second by handheld timers. Physical activity was self-reported as days per week frequency and intensity out of a point scale of 0 = sedentary, 1 = light, 2 =moderate, and 3 = intense or vigorous. The original divisional data set had a significant between sex group difference for human factors of height, $F(1,614) = 219.4$, $p < .01$; weight, $F(1,614) = 252.0$, $p < .01$; BMI, $F(1,614) = 101.2$, $p < .001$; and age, $F(1,614) = 15.23$, $p < .01$. Therefore, a matched pairs data set was drawn from the divisional population with nearly identical values for, to closest 0.1, for BMI and age in order to control for historically important occupational assessment covariates. The records for both data sets span the time 2006 to April 2103. All divisional PARE was conducted and data gathered by this author.

3.13 Results

The subject characteristics for the divisional data of convenience of 85 female and 535 male officers are listed in Table 3.1 and Table 3.2. The subject characteristics for 61 matched pairs drawn from the above dataset, are described in Table 3.2. Also reported in the two tables are divisional and matched pairs' data for subject pre-test resting blood pressure and heart rate, as well as self-reported physical activity weekly frequency and intensity. Table 3.3 and Table 3.4 list total PARE times, pursuit times, body control times, for three groups, with the national PARE database (13,709) means inserted for comparisons with the divisional and matched pairs data. The standard deviation of the criterion reference young male group, was much more homogenous, 22 s (Bonneau, 2001) than the divisional data of RCMP officers which had a standard deviations of 30 seconds, and the national data from article one above at 32 s, see above. Also shown in Table 3.4 are lap times for the combined group, and male and female subsets groups for the divisional and matched pairs data. The national RCMP population total PARE time data, from previous studies inlcuded a mean ($\overline{X}$) of 222 s (SD ±32), a pursuit times

$\overline{\times}$ of 156 s (SD ±21.8), and body control time $\overline{\times}$ of 65 s (SD ±11.4), all close to the divisional data, and matched pairs data values, see Table 3.3 middle column. Thus the division and the matched pairs groups appear to be close samples to the national data values resulting low risk to validity for external generalization.

3.14 ANOVA Results

For both the divisional population data and the matched pair's data, drawn from the divisional population, a 2 x 2 x 6 complex mixed factorial design combined independent groups of sex and performance (pass fail), as between subject main factors, with six dependent repeat lap times as a within subject factor. The main effects results are described below.

3.15 Between Subjects Main Effects and Interactions

Muchley's test indicated that the assumption of sphericity for the ANOVA had been violated in both ANOVAs, with the divisional data set reporting a χ^2 (14) = 513, $p < .01$, and the matched pairs data set reporting a χ^2 (14) = 123, $p < .01$. Therefore, both data sets used the Greenhouse-Geisser corrections for the F-ratio and degrees of freedom. The divisional data used € = 0.73, and the matched pairs data set used € = 0.72.

3.16 Divisional Data

The divisional data set ANOVA summary tables reports all seven between and within subject factors, F ratios, corresponding significance levels and partial eta squared effect sizes for main effects, the 2 way, and 3 way interaction factors in Table 3.5. In the divisional data set sex group (male / female) and performance group (pass / fail) mean lap times were both statistically significantly different at $p < .05$. Figure 3.1 offers a bar graph summary comparison of the divisional and the matched pairs (a) showing sex and

Table 3.1

Subject Characteristics for One Division of RCMP Officers

	N	$\overline{X}$	SD	Var.	Skewness	Kurtosis
Combined						
Age (years)	601	40.6	7.8	61.3	0.1	-0.7
Height (cm)	616	177.8	8.1	65.2	- 0.3	0.7
Weight (kg)	609	86.5	13.9	192.9	0.1	0.5
BMI (kg/m^2)	608	27.2	3.3	11.1	0.5	1.7
HR at rest (bpm)	599	71.3	12.0			
BP S/D (mmHg)	608	127/81	12/8			
PA freq./wk	587	3.54	1.7			
PA intensity(0-3)	559	2.17	0.6			
Male						
Age (years)	519	41.0	7.9	62.1	0.1	-0.8
Height (cm)	531	179.5	7.0	49.4	0.3	1.3
Weight (kg)	524	89.5	11.9	141.8	0.6	0.7
BMI (kg/m^2)	523	27.7	3.1	9.5		
HR at rest (bpm)	515	71.3	12.0			
BP S/D (mmHg)	524	128/82	11/8			
PA freq./wk	485	3.48	0.7			
PA intensity(0-3)	485	2.15	0.3			
Female						
Age (years)	82	37.5	6.8	46.0	0.1	-0.8
Height (cm)	85	167.5	6.3	39.8	0.3	0.2
Weight (kg)	85	67.8	10.2	103.6	0.7	0.3
BMI (kg/m^2)	85	24.1	3.0	9.1	0.8	0.4
HR at rest (bpm)	84	71.6	11.9			
BP S/D (mmHg)	85	118/76	11/7			
PA freq. wk	84	3.9	1.7			
PA intensity(0-3)	79	2.3	0.5			

Note. $\overline{X}$ is means; *SD* is standard deviation; Var. is variance; BMI is body mass index, mmHg is millimeters of mercury; bpm is beats per minute; S/D is systolic over diastolic; BP is blood pressure; PA is physical activity; freq is frequency per week of physical activity, PA intensity is 0 for sedentary, 1 for light, walking 3.3. METS or about 60-70% effort, 2 for moderate 4.0 METS, jogging or 70-85% effort, 3 vigorous or > 8.0 METS.

Table 3.2

Subject Characteristics for 61 Male and Female Matched Pairs Data

	$\bar{x}$	*SD*	Variance.	Skewness	Kurtosis
Combined (*n*=122)					
Age (years)	37.8	6.9	48.0	0.1	-0.8
Height (cm)	173.4	8.1	66.5	0.3	-0.3
Weight (kg)	75.6	11.8	139.8	0.6	-0.1
BMI (kg/m^2)	25.1	3.0	8.7	0.7	0.1
HR at rest (bpm)	71.7	12.7			
BP S/D (mmHg)	123/78	12/8			
PA freq./wk	3.51	1.4			
PA intensity(0-3)	2.22	0.6			
Male (*n*=61)					
Age (years)	37.8	6.9	48.0	0.1	-0.8
Height (cm)	178.6	6.8	46.4	0.2	-0.6
Weight (kg)	80.3	11.8	138.8	0.5	-0.6
BMI (kg/m^2)	25.1	3.0	8.5		
HR at rest (bpm)	71	11.7	176		
BP S/D (mmHg)	126/79	12/8			
PA freq./wk	3.30	1.5			
*P*A intensity(0-3)	2.17	0.7			
Female (*n*=61)					
Age (years)	37.8	7.0	48.7	0.2	-0.8
Height (cm)	168.3	5.7	33.3	0.3	0.8
Weight (kg)	70.9	9.8	97.6	0.6	0.3
BMI (kg/m^2)	25.0	3.0	8.9	0.6	-.0.2
HR at rest (bpm)	72.0	12.1			
BP S/D (mmHg)	120/77	11/7			
PA freq. wk	3.7	1.3			
PA intensity(0-3)	2.3	.5			

Note. $\bar{x}$ is the mean; *SD* is the standard deviation; Var. is the variance; BMI is body mass index in kg/m^2, mmHg is millimeters of mercury; bpm is beats per minute; S/D is systolic over diastolic; BP is blood pressure; PA is physical activity; freq. is frequency per week of physical activity, PA intensity is 0 for sedentary, 1 for light, walking 3.3. METS or about 60-70% effort, 2 for moderate 4.0 METS, jogging or 70-85% effort, 3 vigorous or > 8.0 METS.

Table 3.3

Performance Scores for PARE Pursuit, Body Control, Lap Times for a Divisional

	Variable	$\bar{x}$(s)	National(13,709) $\bar{x}$ (±SD)	*SD*(s)	Variance (s)
	PARE	226	222 (±32)	30	883
	Pursuit	159	156 (±22)	20	405
	Control	66	66 (±11)	12	147
Combined	Pass/Fail	.78/.22			
	Lap 1	24.0		3.2	10
(*n*=620)	Lap 2	25.9		3.4	11
	Lap 3	26.2		3.6	13
	Lap 4	27.3		4.0	16
	Lap 5	27.5		4.2	17
	Lap 6	28.0		4.5	20
	PARE	222	217 (±30)	29	832
	Pursuit	157	153 (±21)	20	396
	Control	65	63 (±12)	11	126
	Pass/Fail	.82/.18			
Male	Lap 1	23.7		3.1	10
	Lap 2	25.7		3.4	11
(*n*=535)	Lap 3	25.8		3.5	12
	Lap 4	27.0		4.0	15
	Lap 5	27.1		4.1	17
	Lap 6	27.7		4.4	20
	PARE	248	249 (±31)	25	644
	Pursuit	171	171 (±20)	17	289
	Control	76	77 (±16)	14	186
Female	Pass Fail	.53/.47			
	Lap 1	25.7		2.8	8
(*n*=85)	Lap 2	27.7		3.0	9
	Lap 3	28.5		3.0	10
	Lap 4	29.3		3.2	13
	Lap 5	30.0		3.6	12
	Lab 6	30.0		4.2	18

Note. $\bar{x}$ is the mean; all values are in s, *SD* is the standard deviation; PARE is the Physical Abilities Requirement Evaluation; Pursuit is the police foot chase scenario of PARE; Control is the simulated body control push and pull station with four controlled falls to the ground.

Table 3.4

Performance Scores for PARE, Pursuit, Body Control, Lap Times and Percentage Pass/Fail for the Matched Pairs Data Set

	Variable	*M*	*SD*	Variance	Delta
	PARE	229.7	34.2	1169	
	Pursuit	160.5	22.3	495	
	Control	69.2	14.1	198	
Combined	Lap 1	24.2	3.3	10.9	--
(*n*=122)	Lap 2	26.1	3.6	12.6	1.9
	Lap 3	26.5	3.9	14.9	0.4
	Lap 4	27.4	4.4	19.2	0.9
	Lap 5	27.9	4.6	21.2	0.5
	Lap 6	28.5	4.8	23.0	0.6
	Pass/fail (%)	.71/.29		Total	4.3
	PARE	209.8	29.3	856	
	Pursuit	147.6	18.6	347	
	Control	62.8	12.1	347	
Male	Lap 1	22.6	3.0	9.3	--
(*n*=61)	Lab 2	24.2	2.9	8.5	1.6
	Lap 3	24.2	3.0	9.0	0.0
	Lap 4	25.1	3.6	12.6	0.9
	Lap 5	25.2	3.8	14.3	0.1
	Lap 6	26.3	4.1	16.5	1.1
	Pass/fail (%)	.90/.10		Total	3.7
	PARE	249.7	26.3	690	
	Pursuit	173.5	17.6	311	
	Control	75.7	13.0	168	
Female	Lap 1	25.7	2.9	8.1	--
(*n*=61)	Lab 2	27.9	3.1	9.8	2.2
	Lap 3	28.8	3.3	10.7	0.9
	Lap 4	29.8	3.8	14.7	1.0
	Lap 5	30.6	3.7	13.5	0.8
	Lap 6	30.7	4.5	19.9	0.1
	Pass/ fail (%)	.52/.48		Total	5.0

Note. *M* is the mean; SD is the standard deviation, pursuit is the police foot chase scenario of PARE, control is the simulated body control push and pull station for 61 male and female pairs of officers matched for BMI and age. Delta is difference from lap to next lap.

(b) performance means differences bars. Figure 3.1 (c) is a graphical summary of mean repeat lap times. Significant lap time differences for divisional data were found for contrasts for laps 1-2, laps 2-3, laps 3-4, laps 4-5, but not lap 5-6, see There was a significant two way sex*performance interaction, $F\ (1,118) = 9.0$, $p < 0.01$, partial eta squared (η^2) =.01, a weak effect. There is clearly an ordinal interaction of mean lap times by sex and performance levels in Figure 3.2 (a), as the separate groups sex lines are clearly trending towards crossing. Figure 3.4a expands the above mean values and depicts each lap time for sex and performance as a three-way interaction showing the male pass PARE times were lower on average than the female pass PARE times. However, the male fail PARE times were higher than the female fail PARE times. Therefor post hoc analysis of ANOVA main effects is precluded. However, Fields (2013) suggests reporting all main effects data as is without interpretation of post hoc effects. Table 3.6. In the within subject repeat effects there is a sex*laps interaction, as depicted in Figure 3.2 (a). For lap 3 and lap 6 we see an ordinal (non-parallel lines) interaction.

Figure 3.5 shows the mean lap times and the power drops for male and female officers. The largest pacing power drop effect size was the difference between lap 1 and lap 2, $F(1,616) = 335.1$, $p < .001$, 95% CI [-2.32, -1.87] with a strong effect size partial η^2 of 0.35 for lap 1 to lap 2 power drop. The next three repeat lap power drop contrasts were significant but with weak effect sizes, see Table 3.6. In the divisional data there was a two-way interaction of laps*performance, limiting the parsing of the main effects and any unique within subject factor interpretation. There was not a significant three way sex*performance*laps interaction as seen in Table 3.5, and graphically summarized in Figure 3.4 (a) for divisional data.

Table 3.5

Divisional Data Complex Mixed Factorial 2 x 2 x 6 ANOVA Reporting Summary Table

Variable		SS	df	MS	*F*	Sig.	Partial □²
Between Subjects	Sex	27.2	1	27.2	4.8	.03*	.01
	Performance	1633.4	1	1633.4	288.3	< .01*	.32
	Performance*Sex	50.7	1	50.7	9.0	< .01*	.01
	Error	20940	616	34.0	--	--	--
Within Subjects	Laps	3827.5	3.7	1027.2	195.1	< .01*	.24
	Performance*Laps	88.9	3.7	23.9	4.5	.02*	.01
	Sex*Laps	25.3	3.7	6.8	1.3	.27	.01
	Performance*Sex*Laps	44.9	3.7	12.1	2.3	.06	.01
	Error	12085	229	5.3	--	--	--

Note. Divisional retrospective observational study mixed factorial ANOVA for sex (male, female) and performance (pass, fail) data of 620 officers, 535 male, and 85 female. *=significance at $p < 0.05$ %. SS is the sum of squares; df is degrees of freedom, MS is the mean of summed squares.

Table 3.6

Divisional Data Repeat Lap Times Contrasts.

Repeat laps contrasts	*Lap – lap* Δ	*p* <	*F* value	*p* <	95% CI [upper, lower]	Partial η^{2} =
Lap 1 - lap 2	Δ= -2.09	< .001	*F* (1,616)= 35.1	< .001*	[-2.32, -1.87]	0.35
Lap 2 - lap 3	Δ= -0.59	< .001	*F* (1,616)= 25.1	< .001*	[-0.81, -0.36]	0.04
Lap 3 - lap 4	Δ= -1.12	< .001	*F* (1,616)= 52.2	< .001*	[-1.43, -0.82]	0.08
Lap 4 - lap 5	Δ= -0.45	< .008	*F* (1,616)= 7.0	= .008*	[-0.79, -0.12]	0.011
Lap 5 - lap 6	Δ= -0.25	= .21	*F* (1,616)= 1.6	= 0.33	[-0.64, 0.14]	0.003

*Significance at p < 0.05, CI is confidence interval

3.17 Matched Pairs

In the matched pairs data set sex group (male / female) and performance group (pass / fail) mean lap times were both statistically significantly different at $p < .05$ see Table 3.7. The between subjects two-way interaction for sex*pass, showed nearly parallel line with no interaction and was not statistically significantly different, see Figure 3.2 (b), allowing effect size and post hoc interpretation. Matched pairs data male versus female officers differences were statistically significantly different, F (1,118) = 13.5, $p < 0.01$ and the effect size, a partial η^2 of 0.10, was weak. Performance group passing PARE versus failing PARE were statistically significant, F (1,118) = 90.9, $p < 0.01$ and the effect size, a partial $\eta^2 = 0.44$, was strong. The repeat lap times differences were significantly different across all repeats, F (3.6, 420) = 52.9, $p < .01$, with a strong effect size partial $\eta^2 = 0.31$, see Figure 3.1 (c). Individual lap time differences were significant for contrasts lap1 vs lap 2, lap 3 vs lap 4 and laps 4 vs lap 5. Not significant were individual lap contrasts of lap 2 vs lap 3 and lap 5 vs lap 6. The strongest contrast effect was lap 1 vs 2 in both data sets. There were no two-way or three-way interaction in the matched pairs dataset. Post hoc analysis for matched pairs repeat lap effects collapsed for sex and performance follows next.

The within subject simple effects breaks down the average repeat lap times. Repeat lap differences were significant for lap 1- lap 2, lap 3 - lap 4, and lap 4 - lap 5. Not significant were differences for lap 2 – lap 3 and lap 5- lap 6 see Table 3.8. Table 3.9 depicts all the sex group pairwise comparisons, collapsed across performance groups, at each lap for male - female repeat mean lap time differences. All male versus female lap time differences, collapsed across performance groups, were significant, but with weak partial eta squared effect sizes. Table 3.10 depicts all performance group pairwise comparisons for repeat lap time differences at each lap, for pass versus / fail, collapsed across sex groups, were significant, with strong effects sizes.

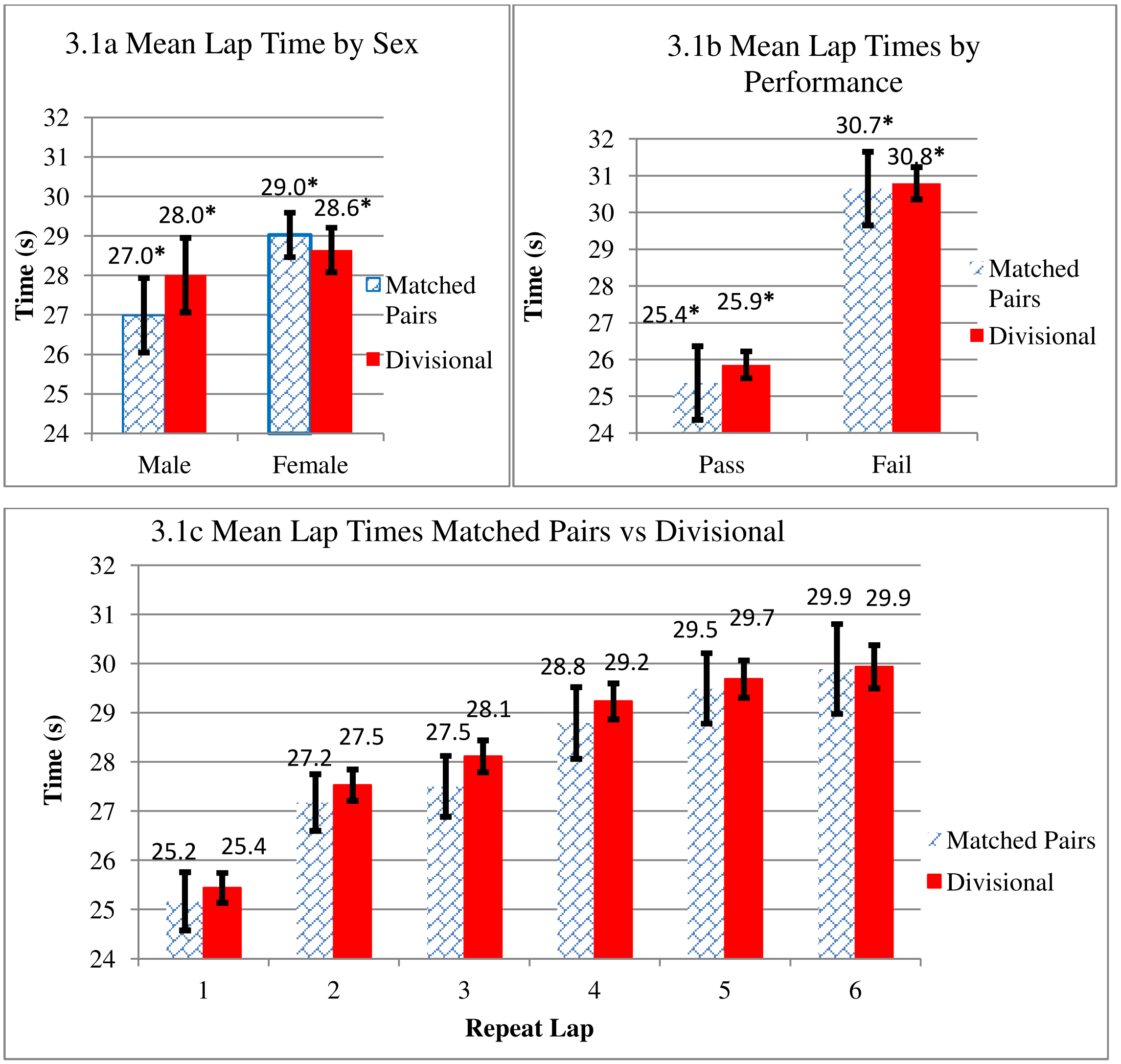
Table 3.7

Matched Pairs Data Mixed Fctorial 2 x 2 x 6 ANCOVA Summary Table

Factor	SS	df	MS	*F*	Sig.	Partial $\square^2$
Between Subjects						
Sex	66.2	1	66.2	13.5	.01*	·10
Performance	446.7	1	446.7	90.9	.01*	.44
Performance*Sex	7.3	1	7.3	1.48	.23	.01
Error	579	1	18	4.9		
Within Subjects						
Laps	982	3.6	276	52.9	.01*	.31
Sex*Laps	25.9	3.6	7.3	1.4	.23	.01
Laps*Performance	42.2	3.6	1.8	2.3	.07	.02
Laps*Performance*Sex	15.6	3.6	3.4	0.8	.49	.01
Error	2,190.5	420.1	5.2			

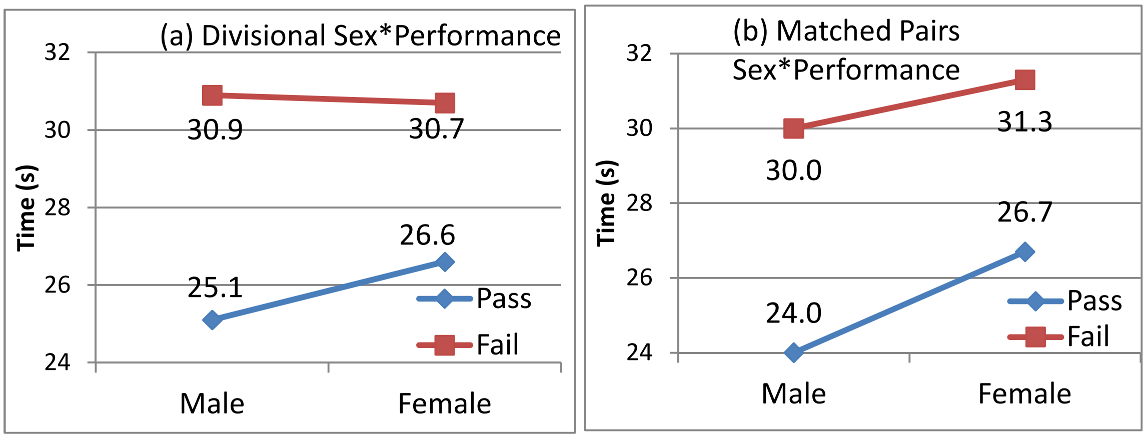
Note: 61 matched pairs data set results: SS is sum of squares; df is degrees of freedom; MS is for mean of summed squares; * represents an interaction term, *. Sig. represents significance levels, with * representing a statistically significant of effects at $p < 0.05$, and partial $\square^{\mathbf{2}}$ is the effect size.

Figure 3.1 Divisional and matched pairs data set main effects mean lap times for (a) sex groups, (b) performance groups (b), and by repeat lap times.



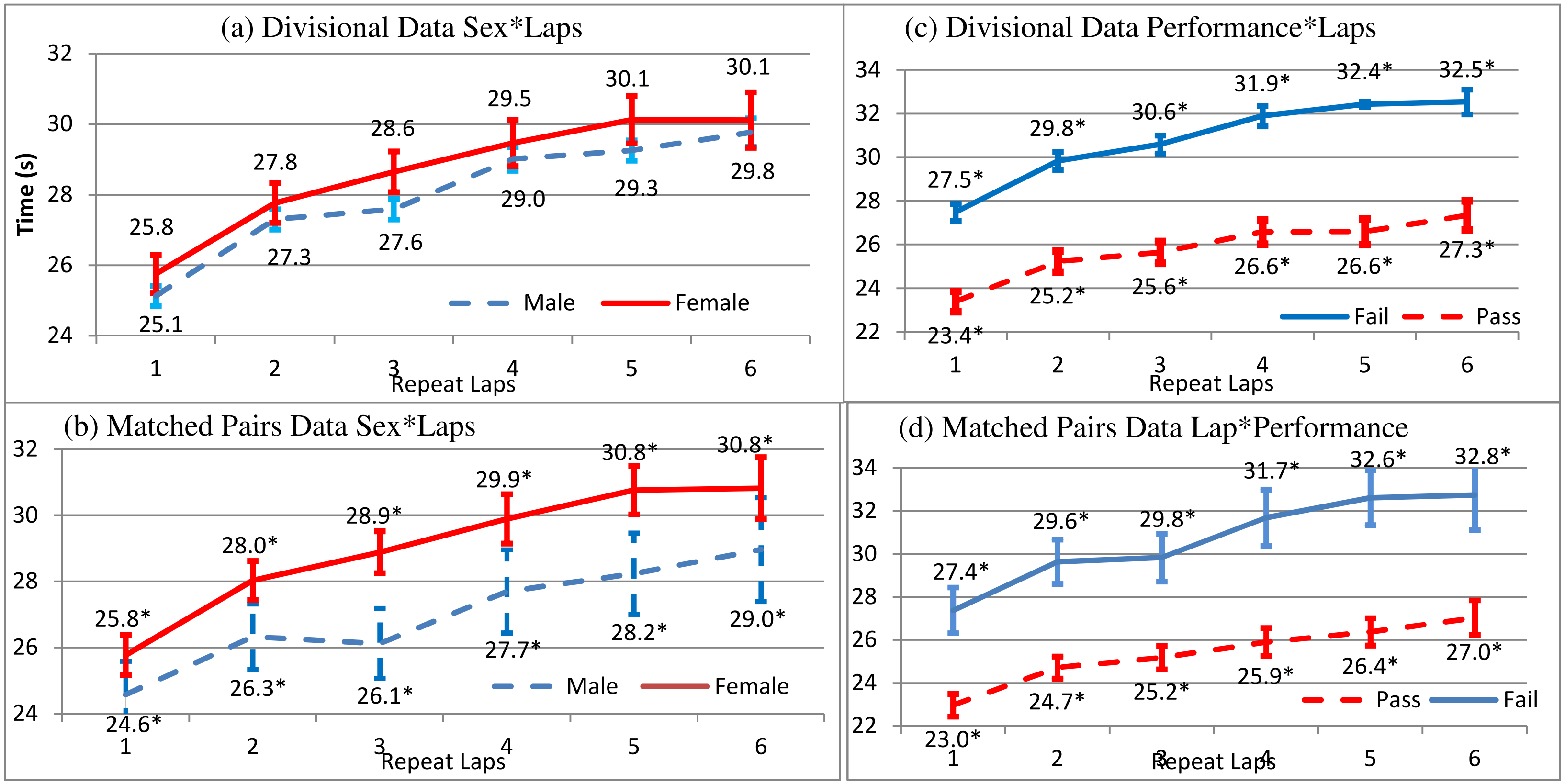
Note: Divisional and matched pairs main effects mean lap time for both (a) sex, (b) performance and (c) repeat lap times. Divisional data in reds bars and matched pairs data in patterned blue bars. Error bars are ± 95% confidence intervals. Divisional data within subject main effects for mean lap time differences are significant for laps: 1–2, 2-3, 3-4, and 4-5, but not 5-6. * Matched pairs data within subject main effects for mean lap time differences are significant for laps: 1-2, 3-4, 4-5, but not lap 2-3 and lap 5-6.

Figure 3.2 Main effects interactions Sex*Performance for Divisional data (a), Matched Pairs data (b)



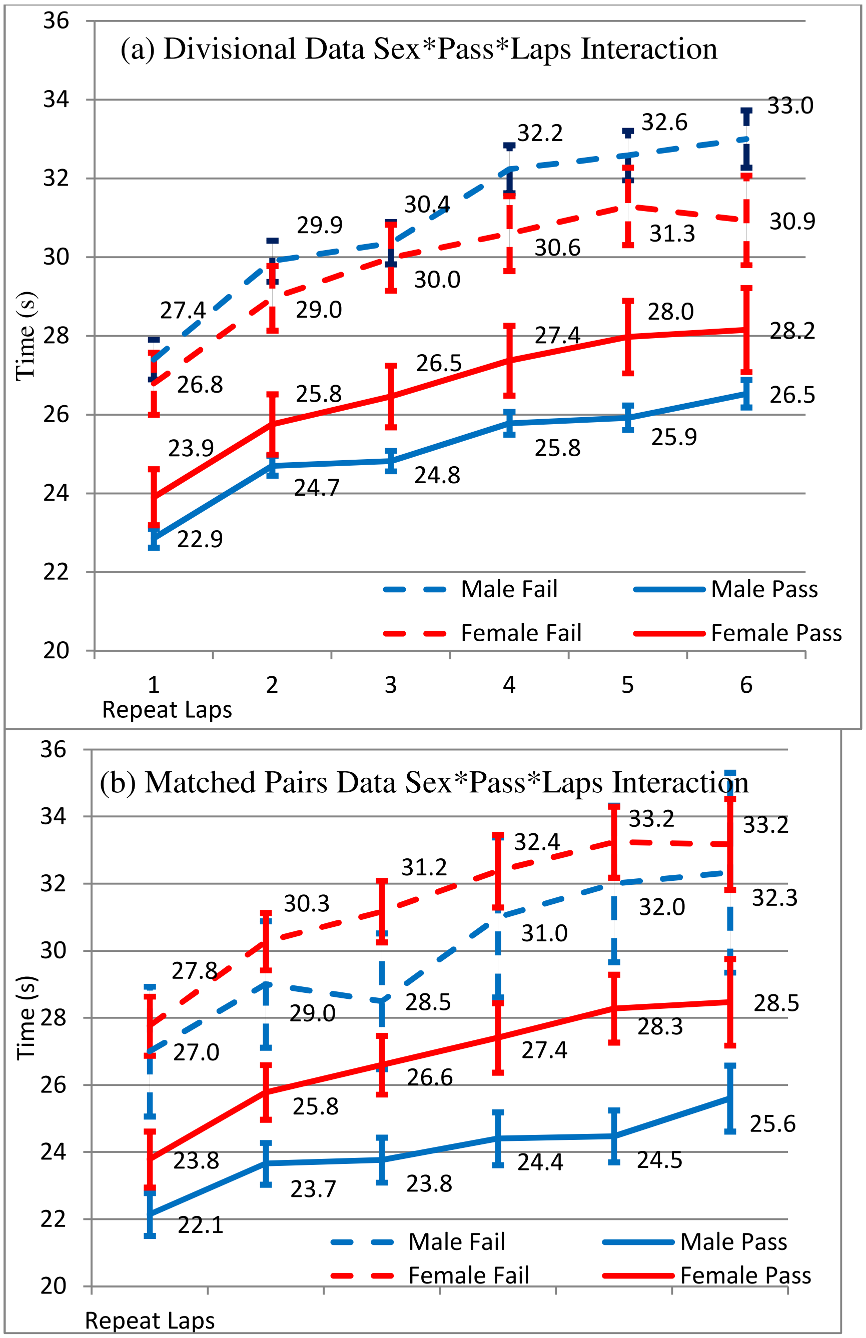
Note: (a) Divisional data between subject main effects 2 way interaction for sex*pass (performance) were statistically significant $F(1,616) = 9.0$, $p < 0.01$ showing a disordinal interaction trend of crossing lines whereas (b) Matched pairs data between subject main effects 2 way interaction for sex*pass (performance) were not statistically significant $F(1,118) = 1.48$, $p = 0.23$ showing a slight ordinal interaction trend of near parallel lines.

Figure 3.3 Divisional Sex*Laps (a) and Pass*Laps (c) and matched pairs sex*Laps (b) and Pass*Laps (d) two way interactions



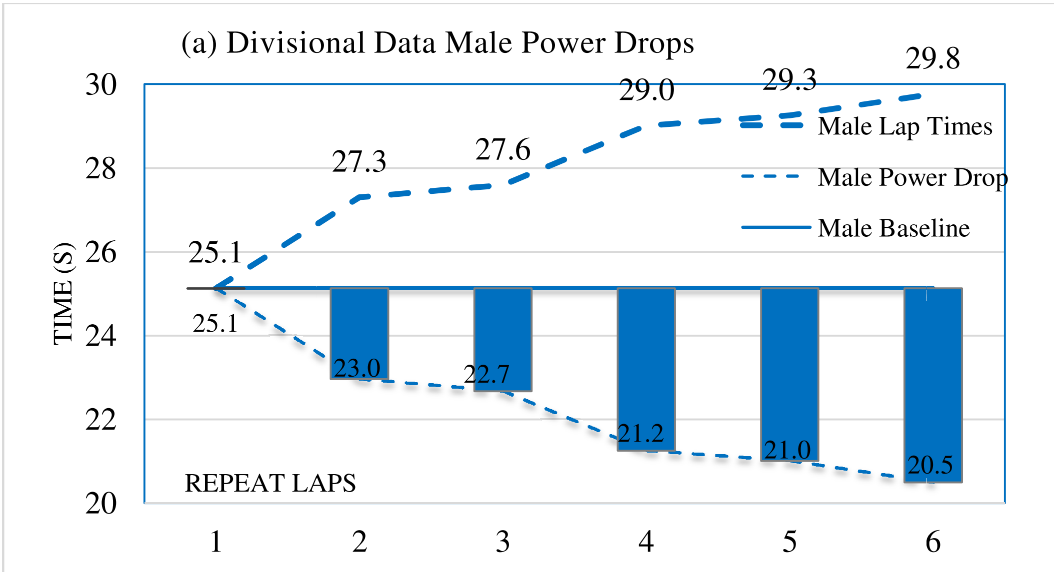
Note: Divional data is for 535 men and 85 women, and matched paris data is for 61 BMI and age matched male and female officers. Performance is coded pass / fail. * Significance (b) matched pairs mean lap time differnces by sex group collapsed for performance groups were for all laps 1 to 6.* Significane (d) for matched pais mean lap time differences by performacne group collapsed for sex groups were for all laps 1 to 6.

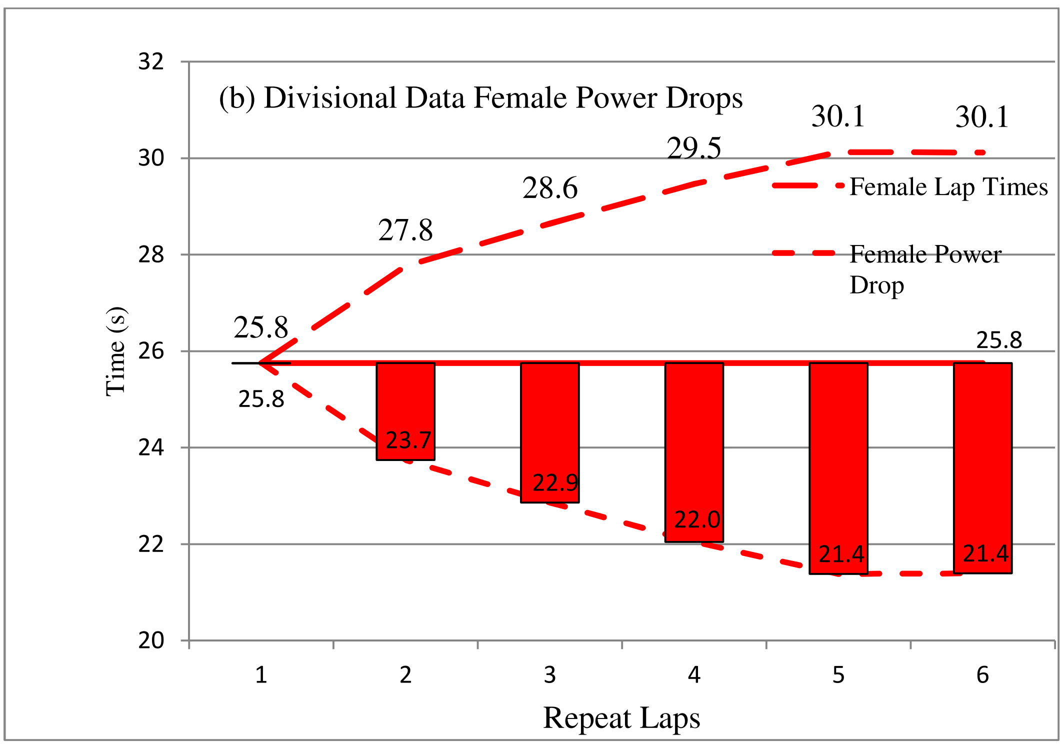
Figure 3.4 Divisional (a) and Matched Pairs (b) repeat lap times for sex by performance group



Note: Divisional and Matched Pairs data sets within subject factors 3 way interaction, which was not significant, for repeat lap times for both sex and performance factors.

Figure 3.5 Divisional data male (a) and female (b) lap times and power drops





Note: Divisional data mean lap times and power drops per sex group, (a) male, (b) female. A graphical summary of male and female officer lap times and power drops calculated as 2 times lap 1 minus the next lap time.

Table 3.8

Matched Pairs Data Set Mean Repeat Lap Time Contrasts.

Repeat laps contrasts	*Lap – lap* Δ	*p* <	*F* value	*p* <	95% CI [upper, lower]	Partial η^{2} =
Lap 1 - lap 2	Δ= -2.01	< .001	*F* (1,616)= 92.2	< .001*	[-2.42, -1.59]	0.44
Lap 2 - lap 3	Δ= -0.33	= .20	*F* (1,616)= 1.7	= .2	[-0.83, 0.17]	0.01
Lap 3 - lap 4	Δ= -1.29	< .001	*F* (1,616)= 19.6	< .001*	[-1.87, -0.71]	0.14
Lap 4 - lap 5	Δ= -0.70	= .24	*F* (1,616)= 5.2	= .024*	[-1.31, -0.01]	0.04
Lap 5 - lap 6	Δ= -0.40	= .33	*F* (1,616)= 1.0	= 0.33	[-1.19, 0.40]	0.01

*Significance at p < 0.05, CI is confidence interval

Table 3.9

Matched Pairs Data Repeat Lap Time Mean Contrasts for Male/Female Factor with Performance Groups Collapsed.

Repeat laps contrasts	*Lap – lap* Δ	*p* <	*F* value	*p*	95% CI [upper, lower]	Partial η^{2} =
Lap 1 ♂- ♀	Δ= -1.20	< .05	*F* (1,616)= 4.0	= .048*	[-2.38, -0.01]	0.03
Lap 2 ♂- ♀	Δ= -1.70	< .01	*F* (1,616)= 8.5	< .01*	[-2.86, -0.55]	0.07
Lap 3 ♂- ♀	Δ= -2.75	< .001	*F* (1,616)= 19.3	< .001*	[-4.00, -1.51]	0.14
Lap 4 ♂- ♀	Δ= -2.19	< .004	*F* (1,616)= 8.8	= .01*	[-3.66, -0.73]	0.07
Lap 5 ♂- ♀	Δ= -2.53	< .001	*F* (1,616)= 12.2	= .05*	[-3.96, -1.09]	0.09
Lap 6 ♂- ♀	Δ= -1.85	=.05	*F* (1,616)= 4.0	= .047*	[-3.68, -0.03]	0.03

*Significance at p < 0.05, Sex: ♂ is male, ♀ is female, CI is confidence interval

Table 3.10

Matched Pairs Data Set Repeat Lap Time Mean Contrasts for Pass / Fail with Sex Group Collapsed.

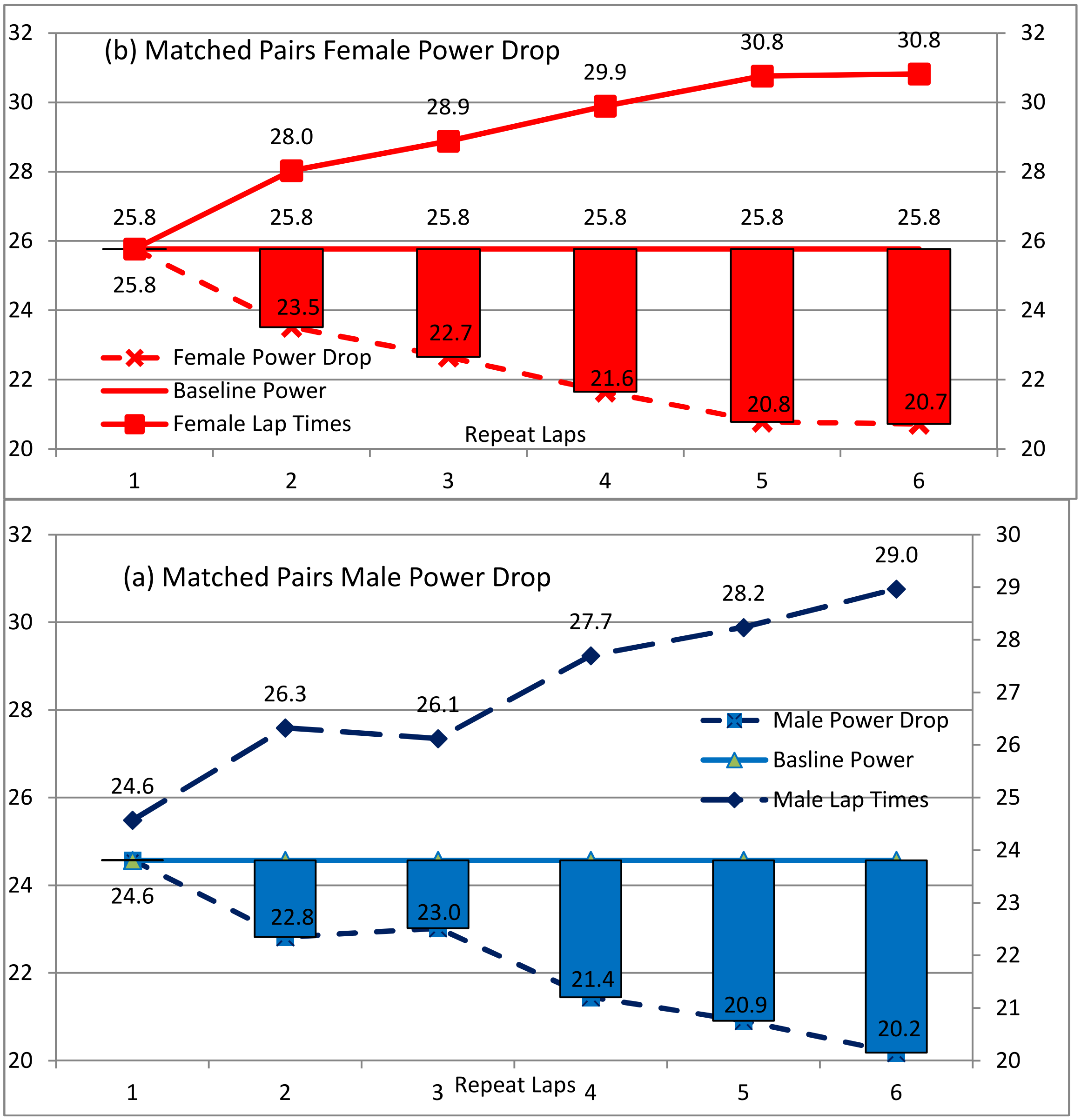
Repeat laps contrasts	*Lap – lap* Δ	*p* <	*F* value	*p* <	95% CI [upper, lower]	Partial $\eta^{2} =$
Lap 1 F-P	Δ= -4.42	< .001	*F* (1,616)=54.6	< .001*	[5.60, -3.23]	0.32
Lap 2 F-P	Δ= -4.92	= .001	*F* (1,616)=71.3	< .001*	[-0.83, -0.17]	0.38
Lap 3 F-P	Δ= -4.66	< .001	*F* (1,616)=55.3	< .001*	[-5.90, -3.42]	0.32
Lap 4 F-P	Δ= -5.79	< .001	*F* (1,616)=61.4	< .001*	[-7.25, -4.32]	0.34
Lap 5 F-P	Δ= -6.24	< .001	*F* (1,616)=74.7	< .001*	[-7.67, -4.81]	0.39
Lap 6 F-P	Δ= -5.72	< .001	*F* (1,616)=38.4	< .001*	[-7.55, 3.89]	0.25

*Significance at p < 0.05, Performance: P=pass, F=Fail, CI is confidence interval

3.18 Discussion

The purpose of this study was to examine the effect of pacing on success in the PARE and contrast differences across sex groups. This discussion starts with the descriptive statistical differences between sex groups for the main PARE scenario total time and pursuit and body control segment times followed by the inferential statisical interpretations of both data sets. Following that will be the discussion of pacing effects profiles for PARE success groups, with a definitve comparison of differnces across sex groups. Finally, limitations of the study will follow.

Figure 3.6 Matched pairs data male (a) and female (b) lap times and power drops



Note: Matched Pairs mean lap times and power drops per sex group, (a) male, (b) female. A graphical summary of male and female officer lap times and power drops calculated as 2 times lap 1 minus the next lap time.

3.19 Descriptive Statistics

The pass proportion for total PARE time in the divisional data group was .82 for males and .53 for females. This was higher than Gaul and Wenger's (1992) incumbent pilot project. The pilot project had four repeat PARE sessions over two years with very few women passing. The lower passing rate for females was a well anticipated finding from clinical experience, and in the peer reviewed literature in Shephard and Bonneau`s (2003) article on gender equity in occupational testing. That article suggested a 20% sex difference in performance ability. Also, Strating et al.'s (2010) large Dutch police sample of 7,000 for skill-related fitness testing had descriptive statistics of males as a group outscoring females as a group, on a skill related police test (means for women 3:43 vs men 3:17).

The pass proportion for total PARE time in the matched pairs data was .90 for males and .52 for females. The divisional data female –male mean difference for PARE scores were: total PARE 248 - 222 s = 26 s (100%); pursuit circuit 171-157 s = 14 s (14/26=54% of total); body control circuit 76 - 65 s = 11 s (11/26 = 46%). For the matched pairs data the female –male mean difference for the PARE scores were total PARE 250 - 210 s = 40 s (100%); pursuit circuit 174 - 148 s = 26 s (26/40 = 65%); body control 76 - 63 s = 13 s (13/40 = 33%). Previous research had not broken down skill-related tests into segments times and compared them across sex groups. Stanish (1993, 1999) examined both pursuit and body control segment separately in relation to health-related fitness testing component predictors in two studies.

Clinical importance in educating members about sex differences suggests informing clients that the PARE pursuit absolute time difference were slighter greater for both data sets: 14 s of 25 s (53% for combined group) in the divisional data, and 26 of 40 s 64% (combined group) in matched pairs data.

The mactched sample data were near to the population data for total PARE time, pursuit times and body control times as reported above. This suggest the sample is fairly representative of the divisional and national RCMP popultion profile of scores and bodes well for generalizing beyond the 61 matched officers, back to the divisional and national population. The current study times for all PARE elements were lower than the national data but higher than Gaul and Wenger (1992) pilot study. The female times were more in line with Gaul and Wenger (1992) incumbent pilot study PARE times for women.

3.20 Inferential Statistical Interpretations

The dramatically larger group size for male versus female officers in the Divisional population data plus the statistically significant unequal group differnces in height, weight and age placed constraints on entering covariates of height, weight, and age in the SPPS ANCOVA analysis. Closely matching males and females on BMI and age allowed covariates analysis from previous moments of fitness testing to be enetered into our analysis. On the later point, Fields (2013) is clear you cannot make groups that are different on a covariate become equal by simply inlcuding the factors in the covariate input box in SPSS analysis. However, using a sample of the divisional popultion data closely controlling for convariates (with $p < 0.5$ difference) resulted in the opputunity to compare between and within subject main effects and within subject main effects in post hoc analysis, as interpretation of where the varaince lies. This lead to the major findings of this study descibed below.

The two main findings of the mixed factorial ANOVAs were: first, that the performance pass or fail PARE group effect size was quiet strong compared to the male vesus female group effect size, which was weak and clinically insignificant. Second, the fastest PARE laps times group (males) had a marked two component power drop evident in both the divisional data Figure 3.5 (a) and matched pairs data Figure 3.6 (a). In the men's data the initial faster power drop of lap 1 to 2, is contrasted with a later lap 3 – lap 4 power drop. In the women data the intial power drop is also the greatest yet diminishes in a more constant pattern, as seen in Figure 3.5 and Figure 3.6. The divisioanla data sex*performance disordianl interaction of combined data Figure 3.2 (a) was diminsihed to an slight ordinal interaction in matched pairs data, see Figure 3.2 (b). The failing group started at a slower intial lap pace time and dropped less with lower pace time differences, equally for males and females, and not statitically significant. This difference between passing and failing groups pacing was amplified in the matched pairs group versus the divisional data group. Therefore the difference between men and women officers seems to be a slight pacing difference depicted as a slight ordinal interaction at lap 3.

In equating men and women's body mass and age the partial η^2 effect size for both performance and sex increased and the 2-way interaction was no longer significant. It is speculated that controlling for these factros as a potential confounder may have removed some but not all of sex component responsible for the sex*performance interaction seen in the divisional population. Since the sex main effects appear consistently small this may have tipped the scale to only the performance size effects being significant.

It is specualted that high levels of police ability performance outcomes are marked by a high level of physical literacy in police abilities, as high level of physical activity and or intense exercise, mediate high levels of health-reated fitness Bouchard's (2000). The 2000 Dose-Response conference reported physical activity and fitness levels affect health oucomes, as mediators to health outcomes. In simlilar fashion it is believed physical activity and physical literacy leads to skill-related outcomes, and intense exercise preparation acts as a mediator for high level apapcity in skill-related test scores. Although there was a trend for males to be finished lap 6 a split second or two before females as a group, and although this was statistically significant, it is not considered clinically significant, versus the many seconds that the performance groups pass group finsihed every lap 6, ahead of the fail group, which was statistically and clinically significant. The power drops for both genders followed a dominant energy systems effect with male and female officers fatiguing in similar patterns with sligth differences at lap 3 for the divisional group, a slight ordinal interaction that was not significant at lap 3 and 6 for the matched pair's sample, see Figure 3.2. This implies when gender, body mass and age are equated for male and female officers, in this case average 39.9 years, and 25.0 BMI, all officers complete the pursuit simultion in the same fatigue pattern with near equal fatigue power drops.

The second main finding is seen in contrasts between divisional data and the matched pair's data where the ANOVA results of within subject 2-way interaction of sex* performance and the within subject sex*laps interactions were no longer significant.

The similar fatigue patterning, of positive pacing power drops, albiet with a small time difference between male and female at lap 3 times, further supports PARE as a gender neutral test with performance supported as the major dicriminator, with the exisiting PARE using the PES 4:00 min:s cut-score. The strong performance effect size also implies fitness performance is the major construct for PARE success, as desired when body mass, and age are controlled for in both sexes taking the PARE. Sex was not as predictor of a diffenert pattern of pacing or fatigue for the long foot pursuit. Therefore we can summerize that people that pass versus fail within each sex group have a gereater difference in times versus the male versus female who either pass or fail.

3.21 Theory Predictions

The operational definition of fatigue and power drop used here are synonomous: a time per lap drop for equal PARE lap distance. There does seem to be "an exercise decline in the ability to use force and power" (Weir et al., 2013) depicting consistent power drops as expected from human bioenergetics energy system theory, seen in repeat lap times in Figures 3.5 and 3.6. The pattern of pacing seen in both Figure 3.5 an 3.6 for both men and women are termed "positive" pacing, where the client slows down through the race (Abbiss & Laursen, 2008, p. 242).

Early and strong first lap one and two time difference effects clearly suggest CP depletion in early alactic dominant energy system depletion from one to 45 s. There is qualitative evidence of no perception of speed differences between laps 1 and 2, when subjects are asked if they slowed down. Most do not notice the on average 2 second drop in time, from lap 1 to lap 2 for the first power drop. Early alactic fatigue appears to be asensory. The largest early power drops may be due to exhaustion of anaerobic alactic dominant energy provision period from 1 to 30 s range, but do not appear to be noticed. The perception of fatigue is however quiet another issue in the lactic range, of 50 seconds to 240 s (lap 2 to lap 6) where not only the officer running perceives the fatigue but no observer in the room can fail to notice the slowdown of latter laps pacing in most PARE tests. Further research should check if subjects can perceive speed changes early in the pursuit and is there a sex group difference to this perception. The latter type of fatigue is a very perceptible event with numerous kinesthetic, peripheral, and central cues associated with maximum cardiovascular efforts (Brooks & Fahey, 1986).

Finally, implications for clinical education can be immediately applied. Consistent power drops are evident for repeat lap times, as seen in Figure 3.5 and 3.6. The first power drop is common to both sexes and is pronounced. The fact that both sexes experience this pattern of fatigue in equal time progressions provides a teaching point for gender equity.

The implication for policing is that men and women officers on avearge complete skill-related performance pacing in identical postive fatiguing pacing patterns in the PARE pursuit. Meaning, men and women officers complete the pursuit in the same manner, with the same fatigue profile, although at a slower pace, for fit versus unfit officers. The fact that strong performance effects dwarfted the weaker sex group effects in pacing differences per lap, also suggests variance between genders is not as large as variance within gender performance groups. This also implies that pacing patterns in PARE are mostly not a gender issue, but a metabolic capacity training and/or perhaps a phsycial literacy issue. For BFOR testing, this means the reasonable accommodation for a pursuit course like PARE is to allow time for matabolic adaptation due to intense training and capacity enhancement during the recruiting process. Since high fit women complete the PARE in the same pattern as high fit men, the clinical conversation should be about metabolic enhancement. The confirmation of no pacing patterns sex differences, in matched pairs for BMI and age, in retrospective observational data documents that men and women officers do the job the same way, although not at the exactly the same speed.

3.22 Limitations of the Study

A limitation of this study was the population demographic of the federal police in the division of study was a 6 to 1, male to female ratio. This is slightly more disproportional than the national PARE database demographic of 5.3 to 1. A second limitation is the matching of pairs of officers to equate human factors for subjects. The higher BMIs for men over 30 could not be equally matched by a female officer from the divisional data of convenience. The lower BMIs for women, under about 20 could not be matched by men, from the divisional data.

A recommendation of the study would be to capture in the national HRMIS data baser individual repeat lap times.

A second limitation is the nature of mandatory skill-related testing in a free-living population. Although PARE is mandatory for participation, effort is self-determined unless the officer is a member of a unit with a mandatory PARE pass designation. These protective policing units with PARE standards at PES 4:00 min:s as policy, represent about 15% of a divisional population. Officers are tested annually or transferred out of protective units if they consistently fail the PES 4:00 min:s standard. Future testing should distinguish between a pass effort at PES 4:00 min:s for staffing reasons, and a full effort PARE over PES 4:00 min:s. The low score walking or jog participatory PARE, with a completion tick mark, should be distinguished from the maximal effort, regardless of pass or fail. This would enhance this research. Either motivation, metabolic field markers (Reed, 1998), or scales on a scale of perceived exertion scores (Borg, 1998), could enhance the distinction between capacity challenging best scores and average or sub average efforts. Eliminating influential outliers above 2 + *SD* only partly addressed this issue.

A final limitation of this study is that skill-related testing reported in academic journals is new and no other findings can be compared to for inferential effect size comparisons of the pursuit circuit for a large police sample or population study. Although Strating et al.'s (2010) descriptive statistics of a large Dutch police study indicated that older, administrative, and women officers score lower on skill-related tests, no pattern information was forthcoming to compare to. Predictive studies of PARE tests for small sample groups with low fit populations and limited sex representation makes any meaningful comparisons to previous research papers problematic. Further to this, the only previous PARE predictive studies used health-related field tests to predict success in PARE, at a different physical employment standard for PARE of 4:45, with low fit women and small sample sizes.

These PARE studies are the first skill-related fitness test that present evidence that show female officers complete the pursuit in an identical pattern of pacing to male officers, albeit at a slightly slower absolute pace, which has no clinical significance. With this in mind, other job simulation researchers should review and contribute their findings in a contradictory or confirming manner as a way of adding to this academic dialogue. The PREP, the COPAT, and the FITCO all have pursuit segment laps that could be analyzed in this manner.

Messick (1996) described construct validity as having social, political, and ethical considerations, beyond technical test validity values. Given women complete the pursuit simulation in the same pacing pattern as men, this implies that we should shift the discussion from accommodating a gender group to accommodating and enhancing small metabolic capacities for both genders equally in PARE training. This was Lonsway (2003) final viewpoint about inequitable test, which we do not perceive PARE to be, that we must however train men and women equally anyway in police training academies.

This is especially true as if the standard requires a period of reasonable accommodation for women to pass. When the pass cut score has been placed in reach of women, after reasonable accommodation fitness training, then this can be successfully reported. This was the case with the state of the art BFOR test by Jamnik et al. (2010), the new FITCO. Future research should look at the body control elements and continue sex differences research for that element of PARE.

In conclusion male and female groups of officers who pass PARE have faster and diffferent patterns in the PARE pursuit than those who pass or fail, where there is no sex differnces in pacing pattern. The divisional police populations that are dominated by taller, heavier men, can provide limited number of matched groups, for equal body mass and age, that have no statiscally significant difference in pacing PARE patterns in the PARE pursuit. When compared, male and female officers complete the PARE in identical pacing and fatigue patterns, with slight ordinal differnces at lap 3 and 6. Pacing and power drops in a sustained foot pursuit in the skill-related PARE test seem to be dominated by energy system human bioenergetics systems, rather than cognitive pacing skills.

3.23 References

Abbiss, C.R. & Laursen, P.B. (2008). Describing and understanding pacing strategies during athletic competition. *Sports Medicine*. 38(3): 239-252. dio: 0112-1642/08/003-0329

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M.L., Swartz, A. M., Strath, S. J., … Leon, A. S. (2000). Compendium of physical activities: An update of activity codes and MET intensities. *Medicine and Science in Sports and Exercise*, *32*(9), S498-S516.

Anderson, G. S., & Plecas, D. (2007). Physical Abilities Requirement Evaluation (PARE). Phase 1: Task Analysis. Retrieved from http://www.ufv.ca/media/assets/ccjr/reports-and-publications/PARE_Phase_1.pdf

Anderson, G. S., & Plecas, D. (2008). *Police abilities testing: Revalidation of the task analysis.* Ottawa, ON: RCMP Learning and Development Archives.

Anderson, G. S., Plecas, D., & Segger, T. (2001). Police officer physical ability testing: Re-validating a selection criterion. *Policing: An International Journal of Police Strategies and Management, 24(1)*, 8-31. doi:10.1108/13695110110382232

Andrews, A. & Risher, J. (2006). What does that have to do with being a cop? Employment Standards in law enforcement. *International Association of Chiefs of Police Conference,* Boston, MA. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=187913

Avery, R. D., Nutting, S. M., & Landon, T. E. (1992). Validation strategies in physical ability testing in police and fire settings. *Public Personnel Management*, *21*(3), 301-312. Retrieved from http://ppm.sagepub.com/

Bissett, D., Bissett, J., & Snell, C. (2012). Physical agility test and fitness standards: Perceptions of law enforcement officers. *Police Practice and Research: An International Journal, 13*(3), 208-223. doi.org/10.1080/15614263.2011.616142.

Boocock, M. G., Haslam, R.A., Lemon, P. & Thorpe, S. (2007). Initial force and postural adaptations when pushing and pulling on floor surfaces with good land reduced resistance to slipping. *Ergonomics*, *49*(9), 801-821. doi**:**10.1080/00140130600562876

Bonneau, J. (1996). *PARE: Standard summary presentation to the Canadian Human Rights Commission*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J. (1998). Submission to the Canadian Human Rights Commission: PARE. Briefing paper outlining PARE as a BFOR development and the 4 minute standard. Ottawa, ON; RCMP, Fitness & Lifestyle Archives.

Bonneau, J. (2001): *The development of PARE: An evolution*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J., & Brown, J. (1995). Physical ability, fitness and police work. *Journal of Clinical Forensic Medicine*, *2*, 157-164. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Bouchard, C. (2000). Physical activity and health: Introduction to dose-response symposium. *Medicine & Science in Sports & Exercise,* 33(6); S347-S350.

Brooks, G. A. & Fahey, T. D. (1986). *Exercise physiology: Human bioenergetics and its applications*. New York, NY: MacMillan

Brown, J. & Trottier, A. (1994). *Police health: A physician's guide for the assessment of police officers*. Ottawa, ON: Canada Communications Group.

Boyce, R. W., Jones, G. R., Schendt, K. E., Llyod, C. L. & Boone, E. L. (2009). Longitudinal changes in strength of police officers with gender comparisons. *Journal of Strength and Conditioning Research, 23*(8), 2411-2418. doi:10.1519/JSC.0b013e3181bac2ab

Casson, E. J. (2000). Establishing BFOR and validated standard for vision. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.), *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Canadian Society for Exercise Physiology. (2012). *The Canadian Physical Activity, Fitness & Lifestyle Appraisal: CSEP's Plan for healthy active living* (2nd Ed.). Ottawa, ON: Canadian Society for Exercise Physiology.

Collingwood, T. R. (1988a). Implementing programs and standards for law enforcement physical fitness. *The Police Chief,* April *55*(4*),* 20-24. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1988b). Physical fitness leadership in law enforcement. *The Police Chief*, April *55*(4), 28-34. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1995). Physical fitness standards: Measuring job relatedness. *The Police Chief*, February *62,* 31-31. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (2004,). Underlying physical fitness factors for performing police officer physical tasks. *The Police Chief.* March *71*(3), 1-7. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Cynrim, C., Mergner, T. & Maurer, C. (2009). Potential role of force cues in human stance control. *Experimental Brain Research*, *194*(3), 419-433. doi:10.1007/s00221-009-1715-7

Darcor and Ergoweb. (2013). The ergonomics of manual material handling: Pushing and pulling tasks. *Sutton Javelin Communications for Darcor Limited*. Retrieved from http:/www.darcor.com/technical/ergonomics of manual material handling/

Deakin, J. M., Smith, J. T., Pelot, R. P., & Weber, C. L. (2000). Methodological considerations in the development of physical maintenance standards. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.), *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

DeMaris, A. (1995). A tutorial in logistic regression. *Journal of Marriage and Family Life, 57*(4), 956-968. Retrieved from: http://www.jstor.org.proxy1.lib.uwo.ca/stable/353415

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2005). *The Sage handbook of qualitative research* (3rd Ed.). Thousand Oaks, CA: Sage.

Dunsmore, R. R., & Hunter, A. J. (2001). Maneuvering the obstacle course: Implementing incumbent physical fitness testing in the workplace. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.), *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Eid. E., & Geh, S. (2001). Challenges posed by the Supreme court of Canada in the Meiorin decision to employees in physically demanding occupations. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.), *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Equal Employment Opportunity Commission. (1978). *Uniform guideline on employee selection procedures*. Retrieved from http://www.uniformguidelines.com

Farenholtz, D. W, Rhodes, E. C. & Burrell, A. J. (1983). Development of a physical abilities test for correctional officers in British Columbia included in *Appendix II of*

Police Officers Physical Abilities Study. New Westminster BC: Justice Institute of British Columbia. Retrieved from Canadian Police College Library archives.

Farenholtz, D. W. & Rhodes, E.C. (1986). *Police Officer Physical Abilities Study*. New Westminster BC: Justice Institute of British Columbia. Retrieved from the Canadian Police College Library archives.

Farenholtz, D. W., Rhodes, E. C. & Bonneau, J. (1989). Validation of a police officer's physical ability test. *Canadian Association of Sports Sciences Poster Presentation,* Ottawa, ON. Retrieved from RCMP archives.

Farenholtz, D. W. & Rhodes, E. C. (1990). Recommended Canadian standards for police physical abilities. *Canadian Police College Journal, 14*(1), 37-49.

Frank, W. D., & Anderson, D. F. (1994). Relationship between physical activity and risk factors for cardiovascular disease among law enforcement officers. *Journal of Occupational Medicine, 36*(10), 1127-1132. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Gamst, G., Meyers, L. S. & Guarino, A .J. (2008). *Analysis of variance designs: Conceptual and computational approach with SPSS and SAS*. New York, NY: Cambridge University Press.

Gaul, W. & Wenger, H. (1992). *RCMP Physical Abilities Requirement Evaluation demonstration project: Final report.* Victoria, BC: University of Victoria, School of Physical Education. Retrieved from RCMP Archives.

Gillis, A. D., & Darby, B. (2001). To serve and protect, post Meiorin an argument for due diligence is an argument for undue hardship: The liability labyrinth faced by the public safety in recruitment and retention of competent employees. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.), *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Girard, G. (2012). *Five Year PARE compliance report.* Ottawa: ON: RCMP Learning and Development internal communications.

Gledhill, N., & Shaw, C. (1996). *Report on the characterization, test construction and validation phase of the medical, physical, skills and abilities project.* Toronto, ON: Race Relations and Policing Unit of' the Ministry of the Solicitor General and Correctional Services.

Gledhill, N. & Bonneau, J. (2000). Objectives, processes and consensus summary of the national forum on Bona Fide Occupational Requirements. In N. Gledhill, J.

Bonneau, & A. Salmon. (Eds.), *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Gledhill, N., Bonneau, J. & Salmon, A. (2000). *Proceedings of the national forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Gore, C.J. (Ed.). *Physiological tests for elite athletes*. Human Kinetics, Windsor ON.

Hoffman, B. (1996). How today's law enforcement officer ranks. *Joe Weider's Muscle and Fitness*, *76*, 232-235. Retrieved from http://www.muscleandfitness.com/news-and-features/athletes-and-celebrities/life-joe-weider

Howell, D. C. (1997). *Statistical methods for psychology*. Belmont, CA. Cengage Learning.

Jackson, A. S. (1994). Pre-employment physical evaluation. *Exercise Sports Science Review, 22*(1), 53-90. Retrieved from http://journals.lww.com/acsm-essr/pages/default.aspx

Jackson, C. A., & Wilson, D. (2013). The gender-neutral time obstacle course: A valid test of police officers? *Occupational Medicine*, *63*, 479-484. doi:10.1093/occmed/kqt102

Jamnik, V. K., Thomas, S .G., Shaw J. A., & Gledhill, N. (2010). Identification and characterization of the critical physically demanding tasks encountered by correctional officers. *Applied Physiology, Nutrition and Metabolism*, *35*, 45-58. doi:10.1139/H09-121

Jamnik, V. K., Thomas, S. G., Burr, J. F., & Gledhill, N. (2010). Construction, validation, and derivation of performance standards for a fitness test for correctional officer applicants, *Applied Physiology, Nutrition and Metabolism*, *35*, 59-70. doi:10.1139/H09-122

Jamnik, V.K., Thomas, S. G., & Gledhill, N. (2010). Applying the Meiorin decision requirements to the fitness test for correctional officer applicants: Examining adverse impact and accommodation. *Applied Physiology, Nutrition and Metabolism*, *35*, 71-81. doi:10.1139/H09-123

Keith, T.Z. (2006). *Multiple Regression and Beyond*. New York, NY: Pearson.

Kinnear, P.R. & Gray, C. D. (2010). *PASW 17: Statistics made simple*. New York, NY: Psychology Press.

Kleinbaum, D. G. & Klein, M. (2002). *Logistic regression: A self-learning text*. New York, NY, Springer.

Knapik, C. G. and Marras, W. S. (2009). Spine loading at different lumbar levels during pushing and pulling. *Ergonomics*, *52*(1), 60-70. doi: 10.1080/00140130802480828

Koller, M., Gledhill, N., Payne, N., Jamnik, R., Fisher, K., Gordon, C & Halligan, J. (1999). Evaluation of a new upper body strength meter for assessing upper body strength. *Medicine & Science in Sports and Exercise*, *31*, S114. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Kuruganti, U. & Rickards, J. (2004). The role of human factor engineering in establishing occupational fitness standards. *International Journal of Industrial Engineering*, *34*(6), 451-457. doi:1016/j.ergon.2004.05.004

Lagasse, P. P. (1989). *Assessment of Physical Standards for Admission in the Police Forces*. [Élaboration des normes physiques d'admission aux d'agents de la paix]. Unpublished manuscript, Faculty of Physical Education, Laval University, Quebec City, Canada.

Legault, J. P. (2000). Establishing a Bona Fide Occupational Requirement for occupational requiring a hearing abilities. In Gledhill, N., Bonneau, J. & Salmon, A. (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Lonsway, K.A. (2003). Tearing down the wall: Problems with consistency, validity, and adverse impact of physical agility test in police selection. *Police Quarterly*, 6(3): 237-276. doi:10.1177/1098611103254314

Maher, P. T. (1984). Police physical ability tests: can they ever be valid. *Public Personnel Management, 32*(2), 173-178. Retrieved from http://ppm.sagepub.com/

Marchand, R., Reed, A. Thoden, J.S., & McNeely, E. (1995). *Physical screening protocol for correctional officer applicants*. Toronto, ON. Ministry of the Solicitor General and Correctional Services. RCMP Fitness and Lifestyle-Archives.

McMurray, R. G., Ainsworth, B. E., Harrell, J. S., Griggs, T. R., & Williams, O. D. (1998). Is physical activity or aerobic power more influential on reducing cardiovascular disease risk factors? *Medicine & Science in Sports & Exercise, 30*(10), 1521-1529. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts*. Thousand Oaks, CA: Sage.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*. *119*(1), 40-48. doi: 10.103//0021-843.110.1.40

National Executive Institute Association. (2008). *Promoting health and wellness: Returning to full duty*. Proceedings of the Toronto 2008 Conference of the Major Cities Chiefs and the Federal Bureau of Investigation. Retrieved from www.neiassociates.org/PromotingHealth-fullduty.pdf

Nabeel, I., Baker, B. A., McGrail, M. P., & Flottemesch, T. J. (2007). Correlation between physical activity, fitness, and musculoskeletal injuries in police officers. *Minnesota Medicine, 90*(9), 40-43. Retrieved from http://www.minnesotamedicine.com/Past-Issues/Past-Issues-2007/September-2007/Clinical-Nableel-September-2007

Nutbeam, D. (Ed.). (1998). *Health Promotion Glossary*. World Health Organization. Geneva. Retrieved from http://www.who.int/hpr/NPH/docs/hp_glossary_end.pdf.

Ontario Ministry of Public Safety and Security. (2002). *The Ontario Constable Selection System: Technical guide for users of the Physical Readiness Evaluation for Police: the PREP*. Toronto, ON: Queens's Printers.

Osborn, G. D. (1976,). Physical agility testing: Validating physical agility tests, *The Police Chief*, January, 42-45. Retrieved from the Canadian Police College Library

Physical Education Canada. (2013). *What is physical literacy?* Ottawa ON: Retrieved from http:/phecanada.ca/programs/physical-literacy/what-physical-literacy.

Police Fitness Personnel of Ontario. (2012). *Certified Fitness Appraiser's Manual.* Aylmer, ON: Ontario Police College.

Robergs, R. A., & Landwehr, R. (2002). The surprising history of the "HRmax=220-age" equation. *Journal of Exercise Physiology, 5*(2), 1-10.

Royal Canadian Mounted Police. (2010). *Physical Abilities Requirement Evaluation (PARE) Course Training Standard (CTS).* Ottawa, ON: RCMP Learning and Development.

Royal Canadian Mounted Police. (2012). Gender–Based Assessment. *National Program Evaluation Services Royal Canadian Mounted Police*. Ottawa ON: Retrieved from http://www.rcmp-grc.gc.ca/aud-ver/reports-rapports/gba-eces/gba-eces-eng.pdf

Royal Canadian Mounted Police. (2013). *Gender and respect – The RCMP action plan*. Ottawa ON: Retrieved from http://www.rcmp-grc.gc.ca/gba-eces/action/index-eng.htm

Roitman, J. L. (Ed.) (2001). *ACSM's resource manual for guidelines for exercise testing and prescription* (4th edition). Baltimore: Lippincott, Williams & Wilkins.

Reed, A. (1998). The oxygen cost of running the Physical Ability Requirement Evaluation. In Bonneau (2001). *The development of PARE: An evolution*. Ottawa, ON: RCMP Learning and Development Archives.

Sassen, B., Cornelissen, V. A., Kiers, H., Wittink, H., Kok, G., & Vanhees, L. (2009). Physical fitness matters more than physical activity in controlling cardiovascular disease risk factors. *European Journal of Cardiovascular Prevention and Rehabilitation*, *16*, 677-673. doi:10.1097/HJR.0b013e3283312e94

Sassen, B., Kok, G., Schaalma, H., Kiers H., & Vanhees, L. (2010). Cardiovascular risk profile: Cross-sectional analysis of motivational determinants, physical fitness and physical activity. *BMC Public Health 10*: 59-69. Retrieved from http://www.biomedcentral/.com/1471/2458/10/592. doi:10.1186/1471-2458-10-592

Schulze, C. (2012). The Masculine yardstick of physical competence: U.S. police academy fitness tests. *Women and Criminal Justice, 22*, 89-107. doi:10.1080/08974454.2012.662117

Séguin, R.A. (2012). *Physical abilities requirement evaluation: Report on a 3 year PARE data cycle: 2008 to 2010*. London, ON: RCMP Training Unit.

Séguin, R. (2009). The seven step process. In *Certified Fitness Appraisers Manual*. Hamilton, ON. Police Fitness Personnel of Ontario.

Seo, N. J. and Armstrong, T. J. (2009). Biomechanical analysis for handle stability during maximum push and pull exertions. *Ergonomics*, *52*(12), 1568-1575. doi:10.1080/00140130903287999

Seo, N. J., Armstrong, T. J., and Young, G. J. (2010). Effects of handle orientation, gloves, handle friction and elbow posture on maximum horizontal pull and push forces. *Ergonomics*, *53*(1), 92-101. doi:10.1080/00140130903389035

Sharkey, B. J. & Davis, P. O. (2008). *Hard work: Defining physical work performances.* Windsor, ON: Human Kinetics.

Shell, D. E. (2003). Exercise science and law enforcement. *Strength and Conditioning Journal*, *25*(3), 52-57. Retrieved from Canadian Police College Library, Ottawa ON.

Shephard, R. & Bonneau, J. (2002). Assuring gender equity in recruitment standards for police officers *Canadian Journal of Applied Physiology*, *27*(3), 263-295.

Shephard, R. & Bonneau, J. (2003). Supervision of occupational fitness assessments. *Canadian Journal of Applied Physiology, 28*(2), 225-239.

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Research Methods and Statistics*, *19*, 1099-1104. doi:10.1111/j.1553-2712.2011.01185.x

Stanish, H. I., (1993). Comparison of the Physical Ability Requirement Evaluation to measures of physical fitness. [Abstract]. *Canadian Journal of Applied Physiology, 18:*436 Retrieved from http://www.nrcresearchpress.com/journal/apnm

Stanish H. & Campagna P, (1993).Comparison of the Physical Ability Requirement Evaluation simulated fight component to measures of muscular strength, endurance and power. Unpublished manuscript. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Stanish, H. I., Wood, T.M., & Campagna, P. (1999). Prediction of performance on the RCMP Physical Abilities Requirement Evaluation. *Journal of Occupational and Environmental Medicine, 41*(8), 669-677. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Strating, M., Bakker, R. H., Dijkstra, G. J., Lemmink, K. A. M .P., & Groothoff, J. W. (2010). A job related fitness test for the Dutch police. *Occupational Medicine, 60*, 255-260. doi:10.1093/occmed/kqq060

Thoden, H. C. (2002). *Testing for normality*. New York, NY: Marcel Decker Inc.

Thomas & Means. (2000). *Use of sex and age adjusted physical performance standards in the law enforcement profession*. PRIMA conference position paper. Aylmer ON: Retrieved from Ontario Police College Archives.

Thompson, W.R., Gordon, N. F. & Pascatello, L.S. (2009). *ACSM's: Guidelines for exercise testing and prescription, (Eighth Edition).* New York, NY: Lippincott Williams & Wilkins.

Trottier, A., & Brown, J. (1994). Occupational health in police work: a Canadian perspective. *Journal of Clinical Forensic Medicine*, *1*, 39-42. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Trottier, A., & Brown, J. (1994). Police health: *A physician's guide for the assessment of police officers.* Ottawa, ON: Canada Communications Group.

Wagner-Wisotzki, N. (2005). *Fitness and lifestyle management for law enforcement*, (2nd Ed.). Toronto, ON: Edmund Montgomery Publishers.

Welk, G. J. (2002). *Physical activity assessment in health-related research*. Windsor, ON: Human Kinetics.

Wenger, H. & Gaul, C. A. (1992). *RCMP Physical Abilities Requirement Evaluation demonstration project August 1990-August 1992, Final Report*. Ottawa, ON: RCMP Health Services Archives.

Weir, J. P., Beck, T. W., Cramer, J. T., & Housh, T. J. (2006). Is fatigue all in your head? A critical review of the central governing model. *British Journal of Sports Medicine, 40*(7), 573-586. doi:10.1136/bjsm.2005.023028

Wells, G. D., Selvaduria, H., & Tein, I. (2009). Bioenergetics provision of energy for muscular activity. *Paediatric Respiratory Reviews*, *10*, 83-90. doi:10.1016/j.prrv.2009.04.005

Whitehead, M. (2007). Physical Literacy: Philosophical considerations in relation to developing a sense of self, universality and propositional knowledge. *Sport, Ethics and Philosophy, 1*(3), 281-298. doi:10.1080/17511320701676916

Wilson, D., & Bracci, R. (1982). Police agility test: Another look at pre-employment screening. *Law and Order*, *August*, 36-42. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=85008

Wilmore J. H., & Davis, J. A. (1979). Validation of a physical abilities test for the selection of state traffic officers. *Journal of Occupational Medicine*, *21*(1): 33-39. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

World Health Organization. (1986). *Ottawa Charter for Health Promotion*. Retrieved from http://www.who.int/hpr/NPH/docs/ottawa_charter.hp.pdf

Zumbo, B. D. (2000). Methodology and measurement matter in establishing a Bona Fide Occupational requirement for physically demanding occupations. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.), *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Chapter 4

Is Self-reported Physical Activity Frequency and Intensity a Predictive Factor of PARE Success?

For over a generation now, the RCMP has identified somewhat fit applicants, through their physical employment standard (PES) Physical Abilities Requirement Evaluation (PARE) test. Those police applicants qualifying for federal police academy training were in part qualified based on an initial PES of PARE ≤ 4:45. Applicants accepted in the process were encouraged to continue to train and progress, until entering as cadets in the federal police academy. Then they graduate, based on passing the PES PARE graduation standard of ≤ 4:00 minutes. Since the early 1990s the RCMP has conducted tens of thousands of these applicant PARE tests, in yearly recruiting drives as a complement to the medical bring-to-the job screening (Bonneau, 2001; Trottier & Brown, 1994) amongst other for-hire competencies (Wagner-Wisotzski, 2005). Additionally, incumbent officers have experienced the same PARE with the same PES ≤ 4:00 min:s standard, since the 1992 pilot project (Gaul & Wenger, 1992). In 1997 serving police officer testing began sporadically with the hiring across Canada of Divisional Fitness and Lifestyle Advisors (DFLAs) who hold Certified Exercise Physiologists (CEP) credentials. CEPs were responsible for testing large police divisional populations (provincial groupings) of federal police officers. Population wide testing, using a potentially maximal effort test has inherent risk for those with chronic injuries and lifestyle diseases, including high or very high cardiovascular risks factors. Therefore, historically, and currently, on-the-job periodic occupational medicals act as and continue to act as gatekeepers for the operational skills competencies training that includes PARE at 4:00 min:s PES for serving officer fitness competencies recertification (Bonneau, 2001; Dunsmore & Hunter, 2001; Gaul & Wenger, 1992; Green & Kreuter, 2005; Sharkey& Davis, 2008; Shell, 2003; Trottier & Brown, 1994). The on-the-job officer PARE testing has become almost universal for the federal police since 2007. Pender, Murdaugh and Parsons (2011) saw physical activity as the first positive tension in health promotion, which is described below.

Health promotion activities have been classified as having either positive or negative tensions by Pender, Murdaugh and Parsons (2011). Health promotion practices that enhance health capacity, increase resiliency or optimize recovery have been described as having positive tensions of enhancement. Those activities that screen and prevent disease with primary and secondary prevention are seen as negative tensions that avoid or minimize the deleterious effects of acute and or chronic injuries, infectious or chronic diseases. Promoting being physically active has been classified as the first positive tension health promotion activity by Pender et al. (2011). However, getting someone ordered to attend a fitness test session, as a mandatory activity, and getting them to do the test with full effort, as a self-determination aspect of health promotion appears to be counter intuitive. The two seemingly separate but integrated activities are described next.

4.1 Self-Determinism to Mandatory Health Promotion

Historically, models like the Health Belief Model considered several factors as important in health protection behaviors. They were the individual beliefs of (a) the severity of the disease, (b) the personal perception of one`s susceptibility to the disease, and (c) the perception of individual self-efficacy in overcoming or avoiding the disease through inoculation behaviours (DiClementes, Crosby & Kegler, 2009). However, more and more health promotion activities marked by self-determination are now becoming mandatory health promotion activities in government legislated society wide settings and in a more limited fashion in work environment settings (Green & Kreuter, 1999). Now many health promotion practices are no longer exclusively self-determined but have become mandatory aspect of healthy public policy in federal policing as described in the 1986 Ottawa Charter for Health Promotion. Examples in the society are mandatory seat belts, non-smoking in public buildings and mandatory infectious disease protective inoculation. These health behavior activities have all shifted to become mandatory health promotion activities in schools, hospitals, now include first responders like police and ambulance as physical and medical fitness for duty (Green & Kreuter, 1999). In the work or school settings inoculation participation is mandatory and access to these sites can be denied when missing proof of inoculation. Positive organizational tensions to attend can be keeping your status as fully operational with no restrictions in order to benefit from full policing job functions, including overtime, travel, promotions by career updating of police skills when called on to recertified in a timely fashion (every three years general duty, yearly protective operations) (RCMP, 2010).

The promotion of physical activity with the goal of passing a PES fitness test for operational policing has become one of several dominant mandatory health cultural influences. Others organizational cultural influences are: (a) environmental health and safety under the Canada labor code Part II, and (b) medical surveillance status for occupational medicine review for safe job duty codes based on current medical status, and any duty to accommodate restrictions (Green & Kreuter, 1999; Plat, Frings-Dresen, & Sluiter, 2011; Trottier & Brown, 1994). The central question for officers to report, upon attending mandatory physical abilities testing is what is your current fitness and physical activity level and can you attempt full effort testing today? The research question that arises from the above question is does current self-reported physical activity (PA) frequency per week and intensity out of as 3-point scale add to the prediction model factors for PARE success? The context of the health promotion occupational abilities testing, hereafter called skill-related fitness testing, will be explored along with self-reported PA methods and the limited prediction studies for PARE test success.

4.2 Triadic Health Promotion Influences

An emerging health promotion ecological theory (DiClementes et al., 2007) titled "Triadic Influences" (Flay, Snyder & Pretaitis, 2007; Flay & Petratis, 1994) informs this investigation. This theory posits that the levels of causation and streams of influences for health promotion behaviors, for becoming physically active for example, can be informed from a continuum of influences, from distal to proximal to the individual, with 3 streams, namely: intra-personal, inter-personal and cultural-environmental. In light of the above Triadic Theory's ecological cultural influence stream the RCMP organizational work culture has three universal mandatory health promotion elements for officers: healthy medical status, functional physical fitness status, and safe and healthy work environments.

4.3 Mandatory Health Promotion

The first of three mandatory health promotion activities that influence federal policing is the periodic occupational medical examination, first described by Trottier and Brown (1994) in *Police Health*. Periodic medical examinations conducted by designated occupational physicians are transmitted to chief occupational physicians in each federal police division health services offices to act as gatekeepers. They create (a) police operational duty codes, (b) allowable training activities and intensities, or (c) any accommodated work restrictions for vision, geographical location for proximity to advanced health care, hearing, and operational duties codes (Trottier & Brown, 1994). Operational police duty and training specific activity clearances or restrictions for serving officers are described in a series of internal duty codes and clearance forms. (See Appendix C for the stand-alone PARE medical clearance form.)

The second of three mandatory health promotion activities is the occupational health and safety program with work site environment audits and work injury reporting by: professionally trained occupational health and officers, legislated building health and safety committees, and or union/ labour associations or representatives, supervisors and front line workers themselves (Green & Kreuter, 1999). Each area has prescribed roles and/or advocacy clearly outlined, like supervision for safe work in the Canada Labor Code, Part II.

The third of three powerful police cultural influences is the mandatory police operational skills maintenance (OSM) program conducted by certified police training officers and certified exercise physiologists (CEP). It is part of a larger public and officer safety program that considers safety to the public, to fellow officers and finally to themselves (Collingwood, 1988a, 1988b, 1995, 2004; Trottier & Brown, 1994). The latter police skills and physical activities training include mandatory participation in the occupational physical fitness test, the PARE (Bonneau, 2001; Bonneau & Brown, 1986; Trottier & Brown, 1994). The OSM recertification now includes concurrent and collocated PARE as a mandatory health promotion program activity, in most if not all federal police divisions across Canada. Although intensity of test effort is in part self-determined, actively participating with the goal of “striving for the standard” is a mandatory attendance event. The potential maximal effort of these tests comes from rich job reflections of the essential physical activities and tasks of policing that are well researched through labour intensive direct observation described below.

4.4 Police Tasks and Common Movement Demands

Well documented in policing is the sedentary or light demands of sitting, standing and or easy walking that predominant during a shift (Anderson et al., 2001). Also well documented are certain discrete tasks and potential maximal intensity activities (Bonneau, 2001; Farenholtz & Rhodes, 1990). These tasks were documented by officers themselves, research students, and subject matter experts with the results triangulated by researchers. The observing research students and officers used what Welk (2002) described as the gold standard of physical activity assessment: labor intensive direct observations to complement and triangulate officer critical incidence reports and recalls (Anderson et al., 2001; Farenholtz & Rhodes, 1990). Many services also used large-scale officer surveys as part of test development to confirm common, frequent and critical or essential physical tasks and demands of tasks analysis (Anderson & Plecas, 2008; Bonneau, 2001; Farenholtz & Rhodes, 1990; Gledhill and Shaw, 1995; Jamnik et al., 2010; Lagasse, 1989; Osborn, 1986; Wilmore & Davis, 1979). Essential, critical, and potentially demanding common police shift tasks that could occur at unpredictable times include movements of: running, jumping, leaping, bending, twisting, ascending and descending stairs, manipulating objects, lifting and carrying, pushing and pulling (Anderson et al., 2001; Bonneau, 2001; Osborn 1976; Wagner-Wisotzski, 2005).

4.5 RCMP Fitness Assessment Programming

Early FBI recruit academy training has shown cadets achieve a high level of health-related physical fitness scores (Knapik et al. 2001). Our federal police cadets here in Canada also have high initial skill-related fitness scores at the federal police academy (Skolney, 2001, personal communications). However, police populations historically have been less fit than age matched population groups and inmates for some time now (Bonneau & Brown, 1994; Pollock et al., 1977). In Canada, a recent Toronto police fitness conference by Chiefs of City Police and the FBI asserted that there is a need for incumbent police officers to plan ongoing training throughout their career, from hire to retire, to maintain a high level of physical fitness for operational duty (Major Cities Chiefs of Police, 2008). Periodic testing to PES, and ongoing supportive incumbent training programs (Major Cities Chiefs of Police, 2003), are both seen as elements to assist with this policing health promotion need, although only the incumbent testing with police academy graduation PES are more resisted by police unions groups and feminist officers, and not the ongoing training (Collingwood, 1998a, 1988b, 1995, 2005; Hoffman & Collingwood, 2005; Lonsway, 2003).

The RCMP divisional training program of periodic PARE recertification has tried to keep officers aware of their weekly physical activity guidelines and or exercise status needed throughout their police career, in light of the singular criterion reference of the 4:00 min:s PARE PES. However, as mentioned above, testing an entire population of police officers in a potentially maximal effort test requires cardiovascular risk levels identification and differential occupational medical clearances or restrictions for due diligences for low versus high coronary artery disease (CAD) risk levels.

4.6 Due Diligence in Universal Testing

There are three central concepts for risk stratification for medical clearance procedures for maximal effort police fitness testing of officers (a) age, (b) CAD risk factor levels for either current blood levels or lifetime plaque burden for arterial occlusion; and (c) recent PA levels (Fodor, Frohlich, Genest, & McPhereson, 2000). PA level recalls and passing mandatory medical exams have been conducted universally as pre-screening gatekeepers for PARE when testing large divisional populations of officers, from hire to retire, with officers having varying cardiovascular, pulmonary, and muscular skeletal health statuses.

4.7 Gate-Keeper Medicals and Pre-Test Screening

For safety purposes, PARE has been closely associated in time to the current occupational medical. Some officers with chronic health risks can be "optimally medically managed" with or without restrictions (Trottier & Brown, 1994). All PARE tests have physician signed and chief medical officer reviewed medical clearances apriori (for details see Periodic Health Assessment (PHA) questions published in *Police Health,* Trottier & Brown, 1984, and Appendix C PARE medical clearance form). Other health status pre-screening best practices aligned with the above are: (a) brief interview for self-reported PA weekly frequency and recent maximum intensity (last two months), and (b) the use of the industry standard Physical Activity Readiness Questionnaire and You (PAR-Q or now PAR-Q +) on the same day as a pre-screening for "apparent health", (c) industry standard same day blood pressure and heart rate cut offs determination, (d) same day informed consent, and (e) preliminary instructions contraindications for food, stimulants, heavy exercise, current infection, pregnancy etc.

The purpose of the federal police periodic occupational medical, as a gatekeeper, is to clear federal police officers for full operational police duties, through corresponding job codes and full effort police periodic training activities concurrently. The PARE as a potential maximal effort cardiovascular demand test requires risk stratification when pursuing universal officer population testing. High CAD risk or equivalent to high risk is determined based on the Canadian Society for Exercise Physiology's (CSEP) generic guidelines of risk factor counting. The threshold for risk is 2+ CAD risk factors. The risks include the following factors: (a) older (> 40 years male, > 50 years female), (b) 2 plus CAD risk factors, (c) attempting a maximum effort on a cardiovascular demanding test for those unaccustomed to maximal PA or exercise physical activity effort (CSEP, 1990). The RCMP uses the American College of Sports Medicine Guidelines (ACSM) specific clinical thresholds for seven potential CAD risk factors (Thompson, 2010). CEPs consider those individuals of any age used to regular maximal efforts with < 2 CAD risk as low risk. For individuals who are older, with 2 + high CAD risks or equivalent to high risk, the first maximal effort is experienced in the presence of the cardiologist's lab during 10 to 12 METs treadmill stress test. Those with comorbidities of a previous cardiac intervention, diabetes Type II or metabolic syndrome are also considered equivalent to high CAD risk by the National Institute of Health (NIH) publication *Adult Treatment Plan III*. This author, the chief Health Services Officer and the national Learning and Development policy centre, decided in the June 2005 policy meetings in Ottawa, jointly to accept and use the ACSM 7 modifiable risk factors threshold values, from ATP III, as the definition of two plus positive risk factors and as clinical guidelines for physicians screening older officers (Thompson, 2010). Rejected for screening CAD risk was the Framingham formulas as reviewed and critiqued by the Canadian Cholesterol Working Group as less then completely inclusive of all risk groups (Fodor et al., 2000). The older and 2+ CAD risk factors individuals referred to cardiologists, as a result of having their first maximal effort in the presence of a physician, would, if event free during that test, then have a PARE clearance for full effort. Further appropriate universal police population testing guidelines include using CSEP certified PARE supervisors (CEPs) (Shephard & Bonneau, 2003). Universal medical clearance with cardiovascular risk stratification and differential medical clearances practices for low versus high-risk officers, before potential maximal occupational fitness testing, is now considered a best practice for large police population testing (see Appendix C, PARE medical clearance form).

4.8 Self-Reported Physical Activity Screening

Self-reported regular PA levels up to and including a maximal cardiovascular demand effort, are integral to municipal and provincial health-related fitness testing screening by peer-to-peer fitness testers. It is also important to federal police skill-related fitness testing.

In RCMP PARE testing, the most recent level of PA is an essential component to test intensity supervision. A short-term physical activity history is conducted prior to PARE participation at OSM through a brief interview to counsel effort levels and maximize safety during PARE. Test administrators determine recent PA of participating officers through brief pretest screening for PA frequency per week and recent PA intensity levels. The frequency question is based on short term, last 2 months recall, for number of cardiovascular activities or workouts per week. The mandatory test intensity CEP recommendation should be congruent with: (a) no PA, no test, postpone or reschedule test; (b) walk PA walk test; (c) moderate jog or run PA history then moderate effort jog or run the test; (d) recent (last two months) hard or vigorous running PA, then potentially hard run through of the test. The PA intensity question is based on a “broader, pooled categories based on intensity levels” (Gabriel, Morrow, & Woolsey, 2012, p. S17). Levels of intensity are: (a) light activity demands up to 60% of a maximum cardiovascular capacity or walking; (b) moderate activity demands between 60 to 75% of a maximum cardiovascular capacity or jogging; or (c) 85% or greater of maximum cardiovascular capacity or hard / intense / vigorous training. This PA history determines a PARE participation recommendation from the supervising CEP that is very close to binding for the officer to follow as guidance during the testing. Following testing, attempts to counsel PA and fitness proceed with realistic and attainable PARE time goals. This applies especially for inactive and or older unfit officers who have trouble passing a

BFOR test PES with a singular pass standard of 4:00 min:s (Shephard & Bonneau, 2003). It was originally postulated in *Police Health* that PARE PES cadet graduation of 4:00 min:s should be pursued by operational constables or corporals, or equivalent public protection positions, and that longer serving older officers should pursue a reasonable goal of 4:30 min:s (Trottier & Brown, 1994). Post-test counselling advice is also based on the knowledge of how active the person was before their PARE experience, as ascertained in the PA history, and during PARE monitoring observations. It would enhance counselling aspects of CEP interventions and feedback to know the predictive power of pre-test PA levels on PARE success.

4.9 Post Test Counseling

All physical ability counselling starts with ascertaining the current PA level of the officer. This is determined via same day brief oral interview reported values for PA frequency and intensity, which also forms part of the basis of this research investigation. Now all federal officers who buy into the police fitness standard “strive for” the PES 4:00 min:s mark every three years for general duty officers (RCMP, 2010). This goal is mandated by staffing policy for protective operations units, as suggested by Bonneau (2001) and documented in limited PARE data records in the Human Resources Management Information System (HRMIS). Should officers fail the admittedly young male metric, counseling of a realistic and achievable goal is offered. A reasonable standard, like the 4:30 min:s standard for 30-year veterans is coupled with the knowledge that training status and PA mitigates the inevitable loss of aerobic capacity with age (Wilson & Tanaka, 2000). The CEP counselor also realizes the decline on average of 1 second per year for PARE scores is the normal for men based on above loss of aerobic power with aging. Thus, goal setting strives for an achievable and realistic PARE time score in light of the officer’s age, in short term training goals. What influences do current levels of self-reported frequency and intensity of personal PA have on predictions of PARE success?

4.10 Literature Review

The PARE is the current RCMP mandatory health promotion skill-related fitness test documented in the paper *The Evolution of PARE* by Bonneau (2001). The PARE, in objectives outlines in the *Uniform Guidelines,* is a test modified from another jurisdiction, POPAT from BC municipal police, and adapted to Canadian federal police, based on RCMP task analysis. Bonneau (2001) validated all the POPAT discrete item tasks against federal policing tasks / demands analysis and officer opinions, adjusting and reconciling it to federal police tasks, as suggested in the *Uniform Guidelines* process of test acquisition for use by another jurisdiction. The original change in POPAT to become PARE, as first published by Trottier & Brown (1984) in *Police Health*, was to integrate all 10 repeat 36 inch high rail vaults and controlled falls to the ground, all completed at the end of the POPAT circuit test, into the PARE. The stated purpose of the 10 vaults in POPAT was to drive heart rates to maximum by test end. In PARE the 10 vaults were integrated into the PARE pursuit and body control circuits, with one vault per lap followed by a controlled fall to the ground, in each of the six lap pursuit circuit, reducing the number of vaults to six and integrating them into the flow of getting to a problem (Bonneau, 2001). In addition, four controlled falls to the ground were introduced after the push and before the pull body control component. In PARE, maximal heart rate levels could be seen by some researchers by lap two or three (Sommerfeld, Dunlop & Neary, 1998) therefore the redundant stimulus to reach maximal efforts are not needed. Additional integration of the vaults into the pursuit circuit was also congruent with the over-arching construct of getting to a problem, potentially getting up and down off the ground during pursuit and or control activities.

The second minor change to POPAT, as PARE, was the shortening of the six-foot may jump to five feet. This occurred after an uncompleted human rights commission challenge that alleged women were adversely impacted (<80% of the men's pass rate by) by faulting with a 5 sec penalty for the six foot jump, and going over PES 4:00 as a result (Biddle, 2001; Jamnik et al, 2010; Shephard & Bonneau, 2003; Zumbo, 2001). PARE was challenged and the simplest response to the alleged adverse impact was to shorten the running leap to five feet, from six feet long, which was documented by Bonneau (2001) in the *PARE Evolution* paper and the human rights submission for the Genest case (RCMP Archives). Presently pass rates at RCMP federal police academy (Depot) show no adverse impact for graduating classes tested with PARE. Pass rates are 96 % for female cadets versus the 99% for male cadets (Skolney personal communications, 2009).

Notwithstanding the above slight changes to POPAT, the PARE test layout circuit follows almost identically the POPAT circuit; see Appendix A for a graphical comparison of all tests. The discrete physical tasks in the pursuit circuit are: running six laps of a circuit 20 feet wide by 80 feet long, with a three-foot rail vault, with a five-foot mat, with a five-step stair climb and descend, and with two18 inch running hurdles. The discrete physical task of the body control station is an 80 lb push and pull body control simulation, over 180 degrees with an untimed post-test 100-pound bag dead lift and carry over 25 feet and back.

Mat touch penalties and 18-inch jump hurdle penalties; have been kept at two seconds penalty for knocking a hurdle down, and at five seconds for touching inside the five foot running long jump distance. The flow of the pursuit and body control circuits still reflect the job simulation demands of an essential task during a critical callout for “getting to the problem, controlling the problem” and “removing the problem” (Anderson et al., 2001, p. 8; Bonneau, 2001).

The third change to POPAT was the pass / fail cut-score standard of 4:15 min:s was found to be equal to 3:57 min:s for the young male demographic used to validate the PARE, with the on average 18 s 10 rail jumps integrated in to the PARE pursuit circuit (Bonneau, 2001). The difference to 3:57 min:s was rounded to 4:00 min:s for simplicity (Bonneau, 2001). A large-scale member and subject matter opinion survey, stratified for RCMP Divisions across Canada, was conducted to validate the task analysis for PARE in federal policing in 2007. The same was conducted for PARE discrete test items in 2008 by Anderson and Plecas, ongoing test evaluation.

The resistance standards for the push and pull body control simulator machine has been maintained from the original published Canadian standards by Farenholtz and Rhodes (1990), with 80 lbs dynamic push and pull used as a graduation standard at RCMP police training academy (Depot) and 70 lb for applicants' recruiting processing entrance standard. The original POPAT researcher adjusted strength endurance push standards to what was achievable by women. In the validation study dynamic, side-to-side pushing activity resulted in a score that was lowered by one standard error of the mean, to be equal to 80.0 lbs. (Farenholtz, personal communication, 2010).

4.11 The PARE Pass Standard Criterion Reference

Skill-related police fitness tests are timed circuits interpreted as pass or fail based on total cumulative time. A pass or fail criterion reference is based on the average score of the most able bodied police clients, young male inmates, representing the hardest, most demanding work needed in an aggressive or resistive arrest scenario. The male metric criterion reference is controversial for critical feminist officers (Lonsway, 2003). The current PARE cut score was drawn from validation studies conducted over a generation ago by police researchers on young male inmates (Anderson et al., 2001; Bonneau, 1996; Bonneau, 2001; Farenholtz & Rhodes, 1990; Gledhill & Shaw, 1995; Lagasse, 1989).

4.12 Physical Activity Reporting: State of the Art

A series of papers reported from a scientific conference in 2012 covered the state of the art of PA self-reported methodology and measurement tools (Bowles, 2012). Published literature reviews of self-reported PA that lead to both active and sedentary behaviour measurement were written and compiled to address the "gap in understanding how to optimally assess physical activity" (Bowles, p. S1).

4.13 Definitions

Caspersen, Powel, and Christenson's (1988) definition of PA that has become standard in research is "any bodily movement produced by skeletal muscles that results in caloric expenditure" (Welk, 2002, p.4). Exercise is defined as any "physical activity that is planned, structured, and repetitive and result in improvements or maintenance of one or more facets of physical fitness" (Welk, 2002, p. 4). Physical fitness is also defined as "a set of outcomes or traits that relate to the ability to perform physical activity" (Welk, 2002, p. 4).

4.14 Self-Reported Physical Activity

Haskell (2012) reviewed the history of self-reporting methods from the early work of the British bus driver's studies that found walking ticket takers versus bus drivers occupational had divergent heart health levels. Haskell (2012) has reviewed the historical changes in self-reporting PA starting from occupation as a marker of PA, i.e. seated bus driver versus 8 hour walking ticket taker, to short medium and long time period recall questionnaires, to objective instruments that measure PA bout intensity and duration. The challenge described by Haskell (2012) has always remained as that of "how to obtain valid and reliable data on habitual physical activity in diverse free-living populations" (Haskell, 2012, p. S5) when the gold standard direct observation is not possible (Welk, 2002). Haskell (2012) suggests many of the current questionnaires that classify intensity into light, moderate, and or vigorous activities such as walking are very promising for historical developments of standardization of reporting. Welk (2002) reviewed objective data gathering instruments of pedometers, heart rate monitors, accelerometers, and other tools for data gathering PA levels, described further below.

Welk's text, *Physical Activity Assessment in Health-Related Research,* describes both subjective self-reporting and objective other reporting of PA. Dale, Welk, and Matthews' (2002) third chapter describe the methods of direct physical activity research measurement beyond subjective self-reporting using logs, diaries, interviews, and survey questionnaires. Objective measures included activity monitors, heart rate monitors, pedometers, direct observation, and indirect calorimetry using doubly labeled water methods. They report that objective measures do not provide specific type, context, or locations of PA. Haskell (2012) saw optimal reporting PA involving both some form of self-recall and objective data measures. Pre PARE oral interviews used only self-reported recall in the divisional cohort data gathering used in this retrospective study. The cognitive functions of autobiographical recall used in self-reporting, as well as the recommended framework of PA recall are covered next.

Mathews (2002) reviews self-reporting concepts by using “the basic cognitive model” (p.108) of generic or specific recalled data as seen in Baranowski and Daniel’s 1994 work (Matthews, 2002). By examining the “cognitive operations employed in recall” (p. 108) Gabriel, Morrow & Woolsey (2012) suggest a rigid framework for self-reporting physical activity.

4.15 Framework of Physical Activity Reporting

Gabriel, Morrow, & Woolsey (2012) examined the absence of a standardized conceptual framework for physical activity behavior that can be complex and multidimensional suggesting adopting a framework to advance standardized reporting the field. For Gabriel, Morrow & Woolsey (2012) the global concept of human movement includes complex and multidimensional PA behaviours. The movement concept also implies “a directional relationship…between the behavioral aspects of human movement, the characteristic of human movement, and the physiological result or consequence of movement” (p. S13). Gabriel et al. (2012) advocate a standardized framework of dimension and domain constructs that can be related to sedentary behavior, energy expenditure and fitness levels in other research. Gabriel et al.’s (2012) framework is identical to the American Heart Association Statement by Strath et al. (2013) in that it includes the four domains of PA, leisure, occupational, household, or domestic and transport. Additionally, characteristics of PA domain activity were solicited on recall cues as frequency, intensity, duration and mode or type of activity. Gabriel et al. (2012) defined dimensions of PA, in relation to health-related physical fitness as cardiorespiratory, musculoskeletal, flexibility, balance and coordination, and body composition. Strath et al. (2013) uses a slightly different descriptor, they add biomechanical to the list. PA frequency is usually reported as a number of sessions per time period (Gabriel et al., 2012; Strath et al., 2013). Duration of PA is usually defined as an absolute length of time for a “specified time period” such as a day (Haskell, 2012; Strath et al., 2013).

Gabriel, Morrow, and Woolsey (2012) described a conceptual framework for self-reporting physical activity starting with a departure from Caspersen et al.’s (1985) definitions of physical activity. PA was defined in Gabriel et al. (2012) as “the behavior that involves human movement, resulting in physiological attributes including increased energy expenditure and improved physical fitness” (p. S15). The global construct for their conceptual framework is human movement, with a subset for active and sedentary behavior. The upstream factors for either behavior are physiological, psychological, social, and environmental. Either behaviour is postulated to have a direct relationship between behaviours and characteristics of movement with the consequences of energy expenditure, physiological attributes, or health consequences as outcomes. Health enhancing component outcomes from physically active behaviours include physical fitness and energy expenditure. Physical fitness as an outcome is defined as a composite of balance, body composition, muscular fitness, cardiorespiratory fitness, and flexibility (Welk, 2002). Energy expenditure is a cumulative measure of basal metabolic rate, thermogenesis, and activity-related energy expenditure (Ainsworth et al., 2001). Sedentary behaviour of a discretionary or nondiscretionary type, not including sleep, is postulated to lead to health compromising outcomes. The various and large number of PA modes, with their corresponding energy demands, has been categorized to produce intensity values, by Ainsworth et al. (2001), which are used to calculate total energy expended.

Three other article on self-reporting reliability and validity emerged from the 2012 conference. Sternfeld and Goldman-Rosas (2012) reviewed questions to ask before choosing a self-reporting instrument from a database of appropriate measures of self-reported physical or sedentary behaviours. Hagstromer, Aimsworth, Kwak, and Bowles, (2012) reported a 21 item three sub-scale checklist for self-reported instruments quality validation. Masse and Niet (2012) discussed five sources of validity evidence and reviewed seven PA scales. Nusser, Beyler, Welk, Carriquiry, Fuller, and King (2012) developed a measurement error model for more accurate PA recall.

4.16 PA Intensity

PA Intensity is defined as either the rate of energy expenditure to count the metabolic demand or the total caloric costs (Gabriel et al., 2012; Strath et al. 2013; Welk, 2002). Two methods of quantifying PA intensity in units are, one the kilocalorie (kcal) costs of the oxygen uptake of the activity, and two the metabolic equivalent (MET) of the activity as a multiple of resting metabolic oxygen consumption cost of 3.5 ml/kg/min. Welk (2002) cites epidemiological studies that focus on absolute measures of PA, while chronic responses studies to exercise focus on the use of relative measures of PA. Howel's (2001) chart of reporting activity as a percentage of total capacity had four metabolic levels of: 12 METs, 8 METs, 6 METs, and 5 METs. Foundational to the kilocalorie metric is the equivalent of one liter of oxygen at five kcal of energy. Foundational to the METS reporting of PA intensity is the basal metabolic resting rate of oxygen consumption relative to body weight of sitting at 1 MET= 3.5 ml/kg/min, based on an average 70 kg person (Strath et al., 2013). Activity solicited intensity can then be reported in either multiplies of resting METs posited at any point in time or oxygen consumption needed to complete the entire PA. This represents the two standard energy costing scale metrics. Multiples of resting metabolism are assigned to either some physical activity descriptor, like walking = 3.5 ml/kg/min or general descriptors are equated to METS levels, like light = 4 METS and jogging or vigorous = 6 METS, vigorous or intense = 8 - 10 METS as global descriptors (Welk, 2002). The challenge to intensity reporting for populations of divergent people is that absolute and relative indicators of intensity are confounded by divergent fitness capacity levels of individuals for: aging loss of aerobic capacity and or enhanced capacity due to trained status. Welk (2002) describes Howley's adaptation of representing "the absolute intensity in terms of maximal aerobic capacity rather than age" (Welk, 2002, p. 5) as a way of comparing, using a common standard. Comparisons can then be made across age and fitness levels in a population. Durante and Ainsworth (1996) examined the limitations of only 50% of the variance of PA reporting being accurate which lead to a mode or activity classification system described below.

4.17 PA Mode or Activity

Ainsworth, Caspersen, Matthews, Masse, Baranowski and Zhu (2012) made recommendations to improve accuracy in self-reported PA and compiled a compendium of PA activities with energy expenditure levels for hundreds of physical activities. Ainsworth and colleagues (2000) expanded the definition of mode and settings with a comprehensive listing of some 21 major settings and 605 mode activities. The major heading were: bicycling, conditioning exercises, dancing, fishing and hunting, home activities, home repair, inactivity, lawn and garden and miscellaneous, music playing, occupation, running, self-care, sexual activity, spots, transportation, walking, water and winter activities, and religious and volunteer activities. Examples of constructs could be aerobic endurance or strength training activities or biomechanical aspects of lifting, carrying balancing, dragging, etc. (Strath et al., 2013).

To facilitate self-reporting human movement energy cost calculations Ainsworth et al. (2000) created a compendium of physical activity reporting for coding physical activity rates of energy expenditure. This could allow comparisons across reported activities and observational studies. The physical activities had a coding scheme of five digits for every activity with the first two digits representing major headings and the last two digits represent specific activity, with an associated MET intensity rating. Summary adjectives were light < 3 METs, moderate 3-6 METs, and vigorous > 6 METs. For sedentary behaviour, watching television would be coded as 07010 and 07020, inactive reclining as 07011, and sitting as 07021, thus further differentiating light sedentary behaviour that Haskell (2002) saw as the one of the challenge of self-reporting light or apparently sedentary behaviors. Ainsworth et al. (2000) report challenges and limitations to this system of absolute intensity reporting for PA for person's with varying body mass composition (i.e., lean or obese), varying age groups, and varying environmental conditions, for contrasting weight bearing and non-weight bearing PA. Correction factors for inter-individual differences are suggested.

4.18 Large Sample Testing

Testing of large sample police populations had shown dismal pass rates for females when using the progenitor test of PARE, the POPAT. POPAT, as originally created by Farenholtz and Rhodes (1985), was based on task analysis of British Columbian municipal police. The PES was set in part by inmate scores as a criterion reference group. Birzer and Craig (1996) reported that POPAT was the official test adopted by the Canadian Chiefs of Police in 1987. Birzer and Craig (1996) tested 743 male and 98 female officers with POPAT, with a pass rate of 92% for men and 28% for women, it was clearly adversely impacting women, 80% of 92% male pass rate is the critical no impact threshold level of (.8*92) 73.6%. Later the PARE, as a slightly modified POPAT, with the 10 rail jumps integrated into the circuit and the jump modified to five foot from six foot was used to test incumbent officers across three regions of Canada in a pilot test (Gaul & Wenger, 1992). The genesis of the jump modification was to accommodate the large percentage of females failing PARE with the six-foot jump, as outlined in the Genest challenge at the Human Rights Courts (Bonneau, 2002; Eid & Geh, 2001).

The RCMP PARE pilot project by Gaul and Wenger (1992) tested incumbents RCMP officers from several urban and rural detachments from 1990-1992. The pilot project found that the small number of incumbent women officers tested could not pass the PARE initially. Results reported that the pass rate for males changed from 71% to 78% over a 2-year period. Only 14% of women passed initially, but this figure increased in after repeated attempts over two years. The 31 to 40 years age group had similar results to men with 75% passing. The 41 to 50 years age group had only 75% under 4:30 min:s. Insufficient data was available for officers in the over 50 years of age group. However, the pilot work for incumbent testing did show that unlike the original POPAT, with the PARE fitness program development, women officers could be tested in larger numbers and could pass the standard with training. Although current cadet pass rates in the high 90S do not reflect well on the members being tested in the 1992 project, it does show officer loose fitness with years of service.

The major change in PARE is the more valid and reliable, current day highly standardized push pull machine. It has consistent resistance levels, complementary non-slippery floor surfaces protocols, and variable grip heights for optimal peak power for smaller statured persons. It also has other clinically important biomechanical posture education cues in current 2104 protocol for optimal pushing or pulling, allowing short statured persons and female a higher pass rate.

4.19 Prediction of Success Studies

There are no published research studies on the predictors of PARE score success that have included a large number of women, the only study used small sample sizes (Stanish, Wood & Campagna, 1999). Only one Dutch police study (Strating et al., 2011) reported descriptive statistical information on passing for a large sample of federal Dutch police. Operational officers', male officers', and younger officers' mean scores were faster on a Dutch police skill-related job simulation test than for the female, older, or administrative officer groups.

Wilson and Jackson (2013) reported the 3-lap skill-related police fitness, termed the gender-neutral timed obstacle course for UK police had a high male pass rate and a low female pass rate. They concluded the test was adversely impacting, although no mention was made of reasonable accommodation training time. Like in Jamnik et al., 2010or test familiarization.

Stanish, Wood, and Campagna (1999) contrasted health-related fitness test scores to PARE scores. Fitness test constructs of aerobic power, anaerobic power, muscular strength and endurance, and body composition were contrasted with PARE scores. A three-variable model was found in multiple regression analysis to be predictive of PARE score success, with 70 lb bench press, standing long jump, and agility “explaining 79% of the variance” for men (Stanish et al. 1999, p. 2). The authors reported a small percentage of male scores were failures (9%) making the “classification” of pass or fail problematic for their proposed logistic regression analysis which requires sufficient dichotomous variables data to fit the sigmoid curve model (Fields, 2013; Menard, 2010). However, female data had sufficient dichotomous data (passes and failures) to conduct a logistic regression of success predictors. The authors reported two variables as predictive of PARE score success in the women’s group; namely, the 1.5 mile aerobic power endurance run and the agility test, with “93% overall classification accuracy” (Stanish et al., 1999, p. 2).

4.20 Research Purpose

The purpose of this study is to examine the predictive effects of self-reported PA frequency and intensity on PARE pass / fail success outcomes along with other potential covariates of height, weight, age, and sex from previous fitness testing historical moments.

4.21 Methods

All PAREs conducted by this author was retrieved from the Human Resources Management Information System (HRMIS) from every incumbent police officers in a divisional of convenience for records spanning 2006 to April 2013. These records were downloaded to a spreadsheet to facilitate selection of one unique best PARE test score per officer, representing the fastest performance time on PARE per officer, with duplicate records sorted and automatically removed using excel duplicates elimination function. Frequency of PA was defined as a scale of 0 to 7 days a week. Intensity of PA was defined as: inactive scored as 0; light or 40 - 70% effort or slow walking up to brisk walking, scored as a 1; moderate or 70 - 85% effort or brisk walking to jogging, scored as a 2; 85%+, intense effort or running, scored as a 3.The dependent or outcome variable pass or fail PARE was dummy coded 1 for fail and 0 for pass based on the 4:00 min:s PES. Sex coding was 0 for male and 1 for female. PARE times were recorded to the nearest second with hand-held stopwatches. Excel data records were uploaded to SPSS (version 21) for logistic regression analyses.

Logistic regression quality control procedural steps included (Fields 2013, Stoltzfuz, 2010): (a) influential outlier removal; (b) model building through the generation of a short list of significant predictor variables using a cumulative data entry method; (c) verification of linearity of the logit (natural log of the odds, (P/1-P)) for significant independent predictor variables; and, (d) mulitcollinearity diagnostics checks. Final model logistic regression was run on only significant predictors from earlier significant model entries at $p < .05$. Every significant predictor factor was checked for linearity of the logit with the predictor variable interaction with its natural log, as suggested by Fields (2013). Any significant interaction had the variable transformed until linearity was achieved with the transformed variable (Fields, 2103; Hilde, 2009). Finally, odds ratios and Chi square statistics for significant predictors of each model building steps for maximum likelihood technique were assessed at an alpha of .05, as suggested by Menard

(2010) and Fields (2013). Any penalty scores were removed because residuals for mat scores were extreme at near 15 times acceptable levels.

4.22 Results

The descriptive statistics of the retrospective observational study for divisional data set of 626 officers' PARE records including age, height, weight, and body mass index (BMI), resting heart rate, and blood pressure pre-screening values are displayed in Table 4.1. The total PARE time, pursuit, and body control job simulation times for combined, as well as male and female officers, are in Table 4.2. The log liner assumptions for dependent variables and their natural log interactions were all met. Multicollinearity testing diagnostics run through linear regression with identical output and independent variables showed no tolerance ranges for residuals < 0.1 and no residuals termed VIF factors > 10 by Fields (2103). Tolerances for all variables ranged from .8 to .5 while VIF values ranged from 1.1 to 2.0.

The generation of federal police population data in a previous studies (N = 13,709), had a younger average age, were slightly faster in the combined group PARE scores (3:41 vs. 3:47 min:s), slightly faster for male PARE scores (3:37 vs. 3:43 min:s), and slightly slower for female PARE scores (4:08 vs. 4:07 min:s). The national population PARE pursuit run times were faster for the combined group scores (2:36 vs. 2:40 min:s), faster for male scores (2:33 vs. 2:38 min:s), and slightly faster for females scores (2:50 vs. 2:51 min:s). The national population PARE body control times were slightly faster for combined groups (1:05 vs. 1:06 min:s), and for male groups (1:03 vs. 1:05 min:s), and slightly slower for female groups (1:17 vs. 1:14 min:s). Women self-reported being more frequently active on average 3.78 (±1.4 *SD*) than men 3.42 (±1.6 *SD*), and reported participating in PA at a higher intensity level on average 2.29 (±.54 *SD*) than men 2.15 (±.62 *SD*).

The initial predictor model reported a baseline prediction of 77.8% for 528 PARE tests for the divisional data, with 411 passing and 117 failing PARE. Of the predictor factors entered (i.e., height, weight, sex, age, run time, control time, PA frequency, and PA intensity), only PARE pursuit and body control times showed significance in the final reduced model of significant predictors. The log odds (LO) or the PARE pursuit run was 0.66 and for the body control simulation was 0.69 – which were nearly equal, as was shown for the generation of national police population data in previous studies.

4.23 Discussion

The main purpose of the study was to assess if adding PA self-reported frequency and intensity data supported further significant predictive factors for PARE success beyond those seen in a previous studies, showing body control, pursuit, and square root of age as significant predictors. The divisional data did support the pursuit time and the body control time as significant predictors, but not additional PA factors. Although both scores comprise the total PARE score, the pass fail dichotomous score, of 0 / 1, relies on all inputs to calculate significant predictors. A review of some of the aspects of test safety limitations on officer effort and some aspects of PA self-reporting may indicate some issues, which may help future researchers, assess this question in a broader context. Reasons for not finding significant effects from PA self-reporting are explored first.

4.24 Physical Activity Frequency

The average level of PA frequency of 3.47 times/week for the combined male and female officer population, of 3.42 times/week for the male officers, and of 3.78 times/week for the female officers may not be a large enough frequency effect to have clinical or statistical significance. The findings of PA levels frequency, if accurate, may be a true reflection of real world officer PA, and would lead us to conclude average self-reporting habitual PA is not frequent enough to effect success in a skill-related maximum effort

Table 4.1

Subject Characteristics for the Divisional Sample of Convenience

	N	Mean	SD	Variance	Skewness	Kurtosis
Combined						
Age (years)	607	40.6	7.8	61.8	0.1	-0.7
Height (cm)	623	177.8	8.7	76.2	0.2	-0.3
Weight (kg)	615	88.7	13.9	193.1	0.2	0.5
BMI (kg/m^2)	613	27.3	3.3	11.1	0.5	1.6
HR (bpm)	605	71.3	11.9			
BP S/D (mmHg)	615	127/80	12/8			
PA freq./wk	594	3.5	1.5			
PA intensity	563	2.2	0.6			
Men						
Age (years)	526	41.1	7.9	62.6	0.1	-0.8
Height (cm)	539	179.3	8.0	64.6	-2.4	28.8
Weight (kg)	531	88.6	12.1	146.0	0.6	1.0
BMI (kg/m^2)	529	27.8	3.1	9.7	0.7	2.4
HR @ rest(bpm)	522	71.2	11.9			
BP S/D (mmHg)	531	128/81	11/8			
PA freq. /wk	511	3.4	1.5			
PA intensity	485	2.2	0.6			
Women						
Age (years)	84	37.5	6.8	46.0	0.1	-0.8
Height (cm)	84	168.0	6.3	40.1	0.4	0.3
Weight (kg)	84	68.3	10.0	97.9	0.7	0.4
BMI (kg/m^2)	84	24.2	3.0	9.0	0.8	0.4
HR @ rest(bpm)	83	71.6	12.0			
BP S/D (mmHG)	84	118/76	11/7			
PA freq. wk	83	3.8	1.4			
PA intensity	78	2.3	0.5			

Note. M is mean; *SD* is standard deviation; BMI is body mass index; with weight in kg over height in m squared; S/D is systolic over diastolic; BP is blood pressure; HR bpm is heart rate in beats per minute; PA freq. wk. is physical activity frequency per week.

Table 4.2

Performance Scores for PARE, Self-reported PA Frequency and Intensity Scores for a Division Data set of Convenience

Group	Factor	N	Mean	SD	Variance	Skewness	Kurtosis
Combined							
	PARE	626	227	31	972	.61	.24
	Pursuit	626	160	21	452	.60	.51
	Control	626	66	12	134	.84	.80
	PA Frequency	594	3.47	1.5	2.34	.03	.07
	PA intensity	563	2.17	1.5	0.39	-.12	.10
Men							
	PARE	542	223	31	944	.74	.60
	Pursuit	535	158	21	453	.74	.84
	Control	535	65	11	124	.72	-.01
	PA Frequency	511	3.42	1.6	2.40	.08	.06
	PA intensity	485	2.15	0.6	0.39	-.11	-0.5
Women							
	PARE	84	247	26	659	.49	-.64
	Pursuit	85	171	17	287	.30	-.38
	Control	85	74	11	129	.61	.86
	PA Frequency	83	3.78	1.4	1.90	-.30	.45
	PA intensity	78	2.29	0.5	0.30	.12	-.55

Note. SD is the standard deviation; pursuit is the police foot chase 6 lap scenario of PARE; Control is the simulated body control pushing and pulling station with four controlled falls between

Table 4.3

The Logistic Regression Predictor Statistics for All Variables

Variables	*β*	Wald	df	*p*	LO	95% CI [lower, upper]
Sex	1.4	.9	1	.34	4.1	[.22, 76.54]
Age (years)	-.1	3.4	1	.06	.91	[.81, 1.0]
Height (cm)	-.3	1.5	1	.21	.73	[.45, 1.2]
Weight (kg)	.3	1.1	1	.28	1.3	[.80, 2.1]
BMI (m/kg^2)	-.75	1.1	1	.30	.47	[.12 – 1.9]
Pursuit (s)	-.44	30.8	1	< .001*	.64	[.55 - .75]
Control (s)	-.40	32.0	1	< .001*	.67	[.59 - .77]
PA week		5.77	6	.45		
Breakdown						
PAweek1	2.6	.01	1	.93		
PAweek2	5.7	.04	1	.84		
PAweek3	3.0	.02	1	.84		
PAweek4	5.4	.19	1	.66		
PAweek5	1.3	.06	1	.82		
PAweek6	.59	.1	1	.75		
PA intensity		1.98	2	.37		
Breakdown						
PAint1	1.0	.9	1	.34		
PAint2	-.84	1.67	1	.2		

police fitness test. The self-reported PA frequency is well below the recommended five times, 30 minute, moderate intensity level workouts within the public health PA guidelines (Strath et al., 2013).

4.25 Physical Activity Intensity

The average level of PA intensity of 2.17 for the combined male and female officer group, of 2.15 for the male officers, and of 2.29 for the female officers, also seems insufficient for success in a potentially maximum level 3.0 intensity test.

Table 4.4 Final Logistic Regression Predictive Model.

Reduced Model

Variable	β	Wald	df	*p*<	Log Odds	95% CI [lower, upper]
Pursuit	-.43	38.75	1	.001*	.66	[0.57, 0.75]
Control	-.38	41.37	1	.001*	.69	[0.61, 0.77]

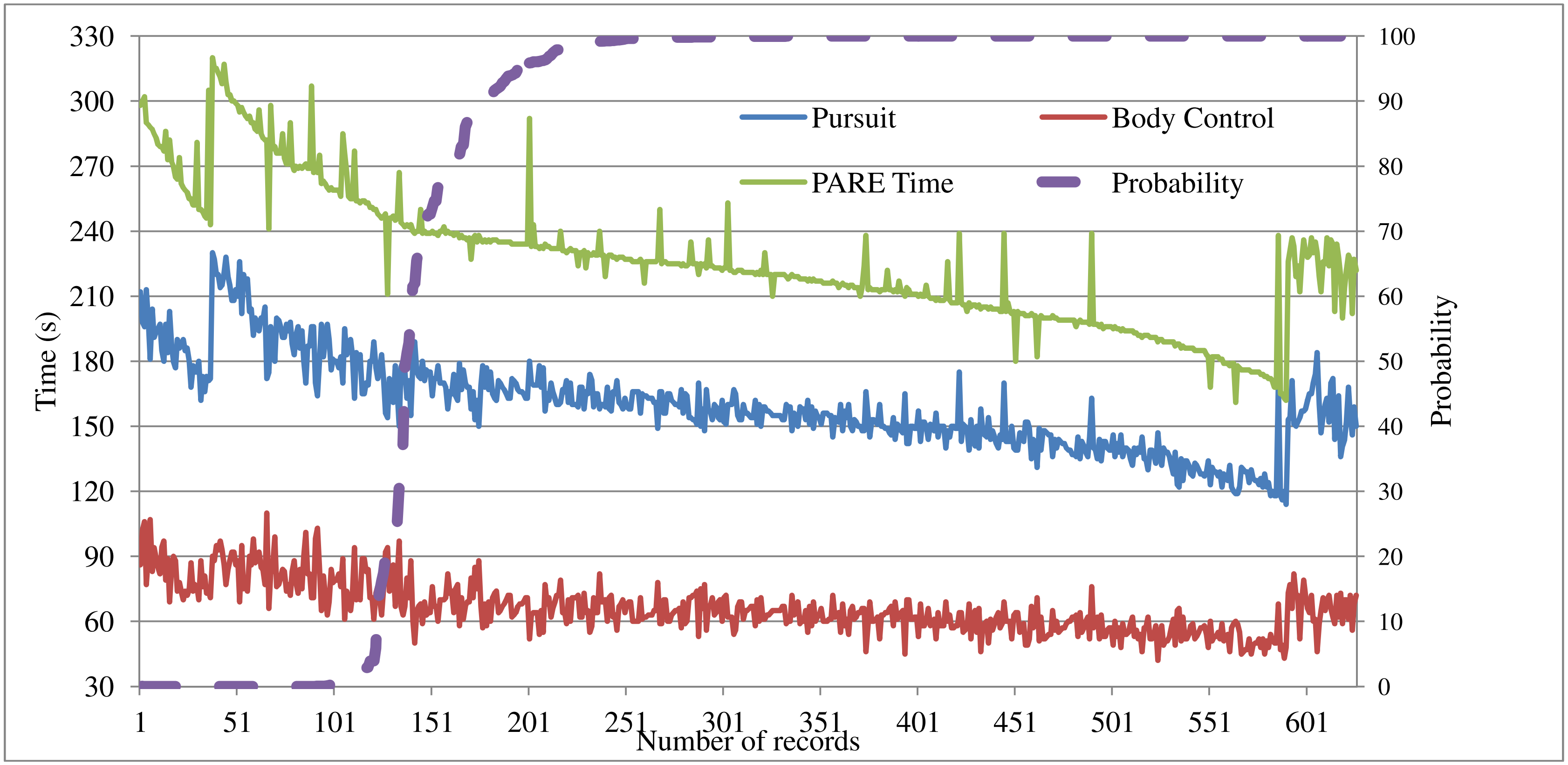
Note. Original factors reported and final reduced model only with statistical significance (*, $p < .05$) for the logistic regression with 98 missing cases for all PA frequency and PA intensity, therefore 524 cases. $P(\mathrm{Y}) = 0.43(\text{pursuit}) + -.38(\text{control})$.

A level two PA intensity is considered a moderate effort, at an equivalent metabolic demand of 4.0 METs (Ainsworth et al., 2001). A level three is considered hard or vigorous activity at an equivalent metabolic demand of 8.0 + METS. The PARE was documented as having a demand of 10 to 12 METs (Reed, in Bonneau, 2001). A higher level of PA, nearer the maximum or vigorous test intensity may have been the threshold we wished for to influence PARE success. Future researcher will have to find a population of much higher weekly PA freq. and PA intensity to validate the concept of sufficient training intensity to effect PARE success.

4.26 Test Effort Mediator

For those highly motivated and or with recent maximal effort familiarity and training, it is expected that they would be able to affect their PARE results and motivation is postulated as a mediating variable for achieving PES standard. For those few with a

Figure 4. 1 Probability of passing PARE changes from 0 to 1 for Divisional data set



Probability of passing PARE, dark purple sigmoid shaped line, indicates at which records (between 101 and 151) in the progressive faster and faster PARE times for both the pursuit and body control -- changes pass probability from the near 0.0 asymptote to transition quickly to nearly 1.0 asymptote. Note: Total PARE scores green, body control times red, pursuit circuit times blue, and probability change sigmoid curve purple.

maximal effort in their recent PA history, maximum efforts are not only allowed, but are highly supported and motivated. Unfortunately, the intensity levels were insufficient to achieve significance in the data set to support a predictor factor for intensity.

A categorical variable of usual or average recollected intensity and frequency may not have the fidelity to predict pass or fail effort or be sophisticated enough to show any effect (Gabriel et al., 2002). Without knowing the population of federal police PA levels in previous studies from the federal HRMIS database, it is problematic to say if this division was close in habitual PA to the population or other divisional samples. Objective PA data recorded for trial periods before PARE would help triangulate with self-reporting PA frequency and intensity.

4.27 Physical Literacy and PA Mode

The mode specificity of pre PARE activity might be the missing recall data that could support or yield a PARE success prediction factor. In particular, two areas are included in the reported activity patterns. The limited self-reported recall questions do not include physical literacy questions of the mode of movement against gravity, in weight bearing exercises, and around items in space and with resistance (Whitehead, 2007). They also do not solicit functional training modes of movement in the recall questions of how hard and how often an officer has participated in PA before presenting for the PARE in the last two months. Pre-test PARE questions ask only for weekly frequency and intensity levels as habitual last two months recall questions. As preparation for PARE, the functional movement skills specificity of the push and pull 80 lb in a controlled arc, vaulting, climbing, and descending stairs, and multiple controlled falls to the ground demands of the test should probably be solicited as pretest specific modes of PA activity preparation. Perhaps functional training that includes the same mode of activities as in the PARE scenarios would yield significant predictive factors at lesser than 3.0 or hard intensity (8 METs) training levels.

The role of PA behaviours on skill-related abilities has only recently been addressed by the Whitehead's (2007) physical literacy definitions and Mandigo, Francis, Lodewyk, and Lopez's (2009) position statement for Physical and Health Education Canada. Specifically, physical literate individuals "move with competence in a wide variety of physical activities that benefit the development of the whole person" (p. 6). Further, they "consistently develop the motivation and ability to understand, communicate, apply, and analyze different forms of movement. They are able to demonstrate a variety of movements confidently, competently, creatively, and strategically across a wide range of health-related physical activities. These skills enable individuals to make healthy, active choices throughout their life span that are both beneficial to and respectful of themselves, others, and their environment" (Mandigo et al., 2009, p. 7). So it is speculated that police officer of a small detachment with a small limited gym of only machine weights might add balance mobility and other functional training movements, reflective of balance and agility, to their routine, and supposedly have a better chance of passing a skill-related abilities test marked by agility, balance, coordination and reaction time.

The limited recall data remains problematic and a subject for future research interventions. Just as habitual PA effects for varying groups of age, physiological capacity, and BMI affect health-related fitness test levels, (Kesaniemi et al., 2001; Welk, 2002) so might physical literacy activities of "the development of agility, balance, coordination and skill across a wide range of activities" (Mandigo et al., 2009, p. 5) affect skill-related fitness test levels. The sensitivity of the two pre-PARE screening PA questions might be enhanced if mode or physical literacy-type motions were included in skill-related testing questions for mode activities of balance, agility, mobility, speed, and coordination.

Skill-related fitness testing research literature will only be comparable if future intervention researchers value skill-related testing; program appropriate movements for training and preferably use similar or identical highly standardized skill-related police protocols like PARE, FITCO, and PREP to assess results.

4.28 Limitations

One of the limitations of this study is it represents a federal police division subset of data that are older, at 40.6 years of age, than the national data average age of 13,709 PARE records on file, with 37.2 years of age. A second limitation is that these results cannot be generalized to a single year divisional population profile, as this studies data was amassed over a period of six years, with at least two, and some times, and in some cases six repeat PARE opportunities for those in annual protective police (15%) recertification.

4.29 Conclusion

Habitual PARE pre-test interview results do not support additional significant predictive values for PARE beyond what has been shown for a generation of police data. These factors join other non-significant predictors of human factors like sex, height and weight or BMI. Only test items of long foot pursuit and body control performance times were significant predictors, in line with population findings for federal police previously conducted.

Future research should examine adding perceived exertion or physiological monitoring of cardiovascular variables like heart rate and VO_2 during testing or in pre-test physical activity assessments as a way of offering further fidelity to the data predictors, especially for delineating a full maximal capacity effort from a participatory sub-maximal effort. Heart rate or step counting data downloaded could also help with either equivalent oxygen costs or preparatory total work volume for self-reported PA. This could be addressed in future research designs. Groups of light, moderate, or intense average self-reported PA preparation with in test effort reported -- may be more revealing in terms of significant predictors.

4.30 References

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M.L., Swartz, A. M., Strath, S. J., … Leon, A.S. (2000). Compendium of physical activities: An update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise*, *32*(9), S498-S516. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Ainsworth, B. E., Casperson, C. J., Matthews, C. E., Massse, L. C. Baransowski, T., & Zhu, W. (2012). Recommendations to improve the accuracy of estimates of physical activity derived from self-report. *Journal of Physical Activity and Health*, *9*(Suppl.), S76-S84. Retrieved from http://journals.humankinetics.com/jpah

Anderson, G. S., & Plecas, D. (2007). PARE Task Analysis. *RCMP Learning and Development.* Ottawa, ON*:* Retrieved from RCMP Internal Learning and Development intranet web page archives.

Anderson, G. S., & Plecas, D. (2008). *Police abilities testing: Revalidation of the task analysis*. Ottawa, ON: RCMP Learning and Development Archives.

Anderson, G. S., Plecas, D., & Segger, T. (2001). Police officer physical ability testing: Re-validating a selection criterion. *Policing: An International Journal of Police Strategies and Management, 24*, 8-31. doi:10.1108/13695110110382232

Andrews, A., & Risher, J. (2006). What does that have to do with being a cop? Employment Standards in law enforcement. *International Association of Chiefs of Police Conference,* Boston, MA. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=187913

Avery, R. D., Nutting, S. M., & Landon, T.E. (1992). Validation strategies in physical ability testing in police and fire settings. *Public Personnel Management*, *2*(3), 301-312. Retrieved from http://ppm.sagepub.com/

Biddle, D. (2006). *Adverse impact and test validation: A practitioner's guide to valid and defensible employment testing*, (2 nd Ed.), Folsom, CA: Biddle Consulting Group Inc.

Birzer, M. L., & Craig, D. E. (1996). Gender differences in police physical ability test performance. *American Journal of Police*, *15(2*): 93-108. doi:10.1108/07358549610122494

Boocock, M. G., Haslam, R. A., Lemon, P., & Thorpe, S. (2006). Initial force and postural adaptations when pushing and pulling on floor surfaces with good land reduced resistance to slipping. *Ergonomics*, *49*(9), 801-882. doi: 10.1080/00140130600562876

Bonneau, J. (1996). *PARE: Standard Summary Presentation to the Canadian Human Rights Commission*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J. (1998*). Submission to the Canadian Human Rights Commission: PARE. Briefing paper outlining PARE as a BFOR development and the 4 minutes standard*. Ottawa, ON; RCMP, Fitness & Lifestyle Archives.

Bonneau, J. (2001): *The development of PARE: An evolution*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J., & Brown, J. (1995). Physical ability, fitness and police work. *Journal of Clinical Forensic Medicine*, *2*, 157-164. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Brown, J. & Trottier, A. (1994). *Police health: A physician's guide for the assessment of police officers*. Ottawa, ON: Canada Communications Group.

Boyce, R. W., Jones, G.R., Schendt, K.E., Llyod, C.L., & Boone, E.L. (2009). Longitudinal changes in strength of police officers with sex comparisons. *Journal of Strength and Conditioning Research, 23*(8), 2411-2418. doi: 10.1519/JSC.0b013e3181bac2ab

Casson, E. J. (2000). Establishing BFOR and validated standard for vision. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Canadian Society for Exercise Physiology. (2012). *The Canadian physical activity, fitness & lifestyle appraisal: CSEP's plan for healthy active living* (2nd Ed.). Ottawa, ON: Canadian Society for Exercise Physiology.

Collingwood, T. R. (1988a). Implementing programs and standards for law enforcement physical fitness. *The Police Chief,* April *55*(4*),* 20-24. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1988b). Physical fitness leadership in law enforcement. *The Police Chief*, April *55*(4), 28-34. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1995). Physical fitness standards: Measuring job relatedness. *The Police Chief*, February *62,* 31-31. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (2004). Underlying physical fitness factors for performing police officer physical tasks. *The Police Chief,* March *71*(3), 1-7. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Cynrim, C., Mergner, T., & Maurer, C. (2009). Potential role of force cues in human stance control. *Experimental Brain Research*, *194*(3), 410-433. doi:10.1007/s00221-009-1715-7

Darcor and Ergoweb. (2013). The ergonomics of manual material handling: Pushing and pulling tasks. *Sutton Javelin Communications for Darcor Limited*. Retrieved from http:/www.darcor.com/technical/ergonomics of manual material handling/

Deakin, J. M., Smith, J. T., Pelot, R. P., & Weber, C. L. (2000). Methodological considerations in the development of physical maintenance standards. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

DeMaris, A. (1995). A tutorial in logistic regression. *Journal of Marriage and Family Life, 57*(4), 956-968. Retrieved from: http://www.jstor.org.proxy1.lib.uwo.ca/stable/353415

Denzin, N.K., & Lincoln, Y. S. (Eds.). (2005). *The Sage Handbook of Qualitative Research* (3 rd Ed.). Sage, Thousand Oaks, CA.

DiClemente, R. J., Crosby, R, A., & Kegler, M. C. (2009). *Emerging Theories in Health Promotion Practice and Research*. San Francisco, CA: Josey-Bass.

Dunsmore, R. R., & Hunter, A. J. (2001). Maneuvering the obstacle course: Implementing incumbent physical fitness testing in the workplace. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Eid. E., & Geh, S. (2000). Challenges posed by the Supreme court of Canada in the Meiorin decision to employees in physically demanding occupations. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Equal Employment Opportunity Commission. (1978). *Uniform guideline on employee selection procedures*. Retrieved from http://www.uniformguidelines.com.

Farenholtz, D. W, Rhodes, E. C., & Burrell, A. J. (1983). Development of a physical abilities test for correctional officers in British Columbia, included in *Appendix II of Police Officers Physical Abilities Study.* New Westminster BC: Justice Institute of British Columbia. Retrieved from the Canadian Police College Library archives.

Farenholtz, D. W., & Rhodes, E. C. (1986). *Police Officer Physical Abilities Study*. New Westminster, BC, Justice Institute of British Columbia, November 1986 Retrieved from the Canadian Police College Library archives.

Farenholtz, D. W., Rhodes, E. C., & Bonneau, J. (1989). Validation of a police officer's physical ability test. *Canadian Association of Sports Sciences Poster Presentation paper*. Ottawa ON: Retrieved from RCMP archives.

Farenholtz, D. W., & Rhodes, E. C. (1990). Recommended Canadian standards for police physical abilities. *Canadian Police Journal, 14*(1), 37-49. Ottawa, ON: Retrieved from http://www.questia.com/library/p90673/canadian-journal-of-police-and-security-services

Fodor, J. G., Frohlich, J. J., Genest, J. J. G., & McPherson, P. R. (2000). Recommendations for the management and treatment of dyslipidemia: Report on the working group on hypercholesterolemia and other Dyslipidemias. *Canadian Medical Association Journal, 162*(10), 1441-1447.

Frank, W. D., & Anderson, D. F. (1994). Relationship between physical activity and risk factors for cardiovascular disease among law enforcement officers. *Journal of Occupational Medicine, 36*(10), 1127-1132. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Gabriel, K. K. P., Morrow, J., & Woolsey, A-L. (2012). Framework for Physical Activity as a complex and multidimensional behavior. *Journal of Physical Activity and Health*, *9*(Suppl), S11-S18. Retrieved from http://journals.humankinetics.com/jpah

Gamst, G., Meyers, L. S., & Guarino, A. J. (2008). *Analysis of variance designs: Conceptual and computational approach with SPSS and SAS*. New York, NY: Cambridge University Press.

Gaul, W., & Wenger, H. (1992). *RCMP Physical Abilities Requirement Evaluation Demonstration Project: Final Report.* Victoria, BC: University of Victoria, School of Physical Education. Retrieved from RCMP Learning and Development Archives, Ottawa.

Gillis, A. D., & Darby, B. (2000). To serve and protect, post Meiorin an argument for due diligence is an argument for undue hardship: The liability labyrinth faced by the public safety in recruitment and retention of competent employees. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Girard, G. (2012). *Five Year PARE Compliance Report*. Ottawa: ON: *RCMP Learning and Development* internal communications.

Gledhill, N., & Shaw, C. (1996). *Report on the characterization, test construction and validation phase of the medical, physical, skills and abilities project.* Toronto, ON: Race Relations and Policing Unit of' the Ministry of the Solicitor General and Correctional Services.

Gledhill, N. & Bonneau, J. (2001). Objectives, processes and consensus summary of the national forum on establishing Bona Fide Occupational Requirements. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Gledhill, N., Bonneau, J., & Salmon, A. (2000). Proceedings of the consensus forum on establishing BONA FIDE requirements for physically demanding occupations. Toronto, ON: York University Press.

Green, L. W., & Kreuter, M. W. (2005). *Health promotion planning: An ecological and educational approach*. New York, McGraw Hill.

Haskell, W. L. (2012). Physical activity by self-report: A brief history and future issues. *Journal of Physical Activity and Health*, *9*(Suppl.), S5-S10. Retrieved from http://journals.humankinetics.com/jpah

Hagströmer, M., Ainsworth, B. E., Kwak, L., & Bowles, H.R. (2012). A checklist for evaluating the methodological quality of validation studies on self-reported instruments for physical activity and sedentary behavior. *Journal of Physical Activity and Health*, *9*(1), S29-S36.

Hogan, J., & Quigley, A. M. (1986). Physical standards for employment and courts. *American Psychologist, 41*, 1193-1217. Retrieved from http://www.apa.org/pubs/journals/amp/

Hooker, S. P., Fulton, J. & Mudd, L. M. (2012). Practice-based evidence: A novel approach for self-report physical activity measurement. *Journal of Physical Activity and Health*, *9*(Suppl.), S85-S87. Retrieved from http://journals.humankinetics.com/jpah

Howell, D. C. (1997). *Statistical methods for psychology*. (8th Ed.). Wadsworth, CA: Cengage Learning.

Jackson, A. S. (1994). Pre-employment physical evaluation. *Exercise and Sports Science Reviews, 22*(1), 53-90. Retrieved from http://journals.lww.com/acsm-essr/pages/default.aspx

Jackson, C. A., & Wilson, D. (2013). The gender-neutral time obstacle course: A valid test of police officers? *Occupational Medicine, 63*, 479-484. doi:10.1093/occmed/kqt102

Jamnik, V. K., Thomas, S. G., Shaw J. A., & Gledhill, N. (2010). Identification and characterization of the critical physically demanding tasks encountered by correctional officers. *Applied Physiology, Nutrition and Metabolism*, *35*, 45-58. doi:10.1139/H09-121

Jamnik, V. K., Thomas, S. G., Burr, J. F., & Gledhill, N. (2010). Construction, validation, and derivation of performance standards for a fitness test for correctional officer applicants. *Applied Physiology, Nutrition and Metabolism*, *35*, 59-70. doi:0.1139/H09-122

Jamnik, V. K., Thomas, S. G. & Gledhill, N. (2010). Applying the Meiorin decision requirements to the fitness test for correctional officer applicants: Examining adverse impact and accommodation. *Applied Physiology, Nutrition and Metabolism*, *35*, 71-81. doi:10.1139/H09-123

Keith, T. Z. (2006). *Multiple Regression and Beyond*. New York, NY: Pearson.

Kesaniemi, Y. A., Danforth, E., Jensen, M. D., Kopelman, P. G., Lefebvre, P. & Reeder, B. A. (2001). Consensus statement: Dose response issues concerning physical activity and health: an evidence-based symposium. *Medicine & Science in Sports & Exercise, 33*(6), S351-S358. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Kinnear, P. R., & Gray, C. D. (2010). *PASW 17: Statistics made simple*. New York, NY: Psychology Press.

Kleinbaum, D. G., & Klein, M. (2002). *Logistic regression: A self-learning text*. New York, NY: Springer.

Knapik, C. G., and Marras, W. S. (2009). Spine loading during different levels during push and pull. *Ergonomics*, *52*(1), 60-70. doi:10.1080/00140130802480828

Koller, M., Gledhill, N., Payne, N., Jamnik, R., Fisher, K., Gordon, C., & Halligan, J. (1999). Evaluation of a new upper body strength meter for assessing upper body strength. *Medicine & Science in Sports and Exercise*, *31*, S114. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Kuruganti, U. & Rickards, J. (2004). The role of human factor engineering in establishing occupational fitness standards. *International Journal of Industrial Engineering*, *34*(6), 451-457. doi:1016/j.ergon.2004.05.004

Lagasse, P. P. (1989). *Assessment of Physical Standards for Admission in the Police Forces*. [Élaboration des normes physiques d'admission aux d'agents de la paix]. Unpublished manuscript, Faculty of Physical Education], Laval University, Quebec City, Canada.

Legault J.P. (2000). Establishing a Bona Fide Occupational Requirement for occupational requiring a hearing abilities. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Lonsway, K. A. (2003). Tearing down the wall: Problems with consistency, validity, and adverse impact of physical agility test in police selection. *Police Quarterly*, *6*, 237-276. doi: 10.1177/1098611103254314/

Maher, P. T. (1984). Police physical ability tests: Can they ever be valid. *Public Personnel Management, 32*(2), 173-178. Retrieved from http://ppm.sagepub.com/

Mandigo, J., Francis, N., Lodewyk, K., & Lopez, R. (2009). Position paper: Physical literacy for educators. *Physical and Health Education Canada*. Ottawa ON: Retrieved from: http://www.phecanada.ca/sites/default/files/pl_position_paper.pdf

Marchand, R., Reed, A. Thoden, J.S., & McNeely, E. (1995). *Physical screening protocol for correctional officer applicants*. Toronto, ON: Ministry of the Solicitor General and Correctional Services.

Masse, L. & de Niet, J. E. (2012). Sources of validity evidence needed with self-report measures of physical activity. *Journal of Physical Activity and Health*, *9*(Suppl.), S44-S55. Retrieved from http://journals.humankinetics.com/jpah

McMurray, R. G., Ainsworth, B. E., Harrell, J. S., Griggs, T. R., & Williams, O.D. (1998). Is physical activity or aerobic power more influential on reducing cardiovascular disease risk factors? *Medicine & Science in Sports & Exercise, 30*(10), 1521-1529. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts*. Thousand Oaks, CA: Sage.

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology, 119*(1), 40-48. doi:10,1037//0021-843x.110.1.40

National Executive Institute Association. (2008). *Promoting health and wellness: Returning to full duty.* Proceedings of the Toronto 2008 Conference of the Major Cities Chiefs and the Federal Bureau of Investigation. Retrieved from www.neiassociates.org/PromotingHealth-fullduty.pdf

Nabeel, I., Baker, B. A., McGrail, M. P., & Flottemesch, T. J. (2007). Correlation between physical activity, fitness, and musculoskeletal injuries in police officers. *Minnesota Medicine, 90*(9), 40-43. Retrieved from http://www.minnesotamedicine.com/Past-Issues/Past-Issues-2007/September-2007/Clinical-Nableel-September-2007

Nusser, S. M., Beyler, N. K., Welk, G. J., Carriquiry, A. L., Fuller, W. A., & King, M. N. (2012). Modeling errors in physical activity recall data. *Journal of Physical Activity and Health*, *9*(Suppl), S56-S67. Retrieved from http://journals.humankinetics.com/jpah

Nutbeam, D. (Ed.). (1998). *Health promotion glossary*. World Health Organization. Geneva. Retrieved from http://www.who.int/hpr/NPH/docs/hp_glossary_end.pdf.

Ontario Ministry of Public Safety and Security. (2002). *The Ontario constable selection system: Technical guide for users of the physical readiness evaluation for police the PREP*. Toronto, ON: Queens's Printers.

Osborn, G. D. (1976). Physical agility testing: Validating physical agility tests. *The Police Chief*, *January*, 45-50. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives.

Plat, M. J., Frings-Dresen, M. H. W. & Sluiter, J. K. (2011). A systematic review of job-specific workers' health surveillance activities for fire-fighting, ambulance, police and military personnel. *International Archives of Occupational Environmental Health*, *84*(8), 839-857. doi:10.1007/s00420-011-0614-y

Pender, N., Murdaugh, C. L., & Parsons, M. A. (2010). *Health promotion in nursing practice*, 6th Ed. Boston, Mass: Pearson Publishing.

Police Fitness Personnel of Ontario. (2012). *Certified Fitness Appraiser's Manual.* Aylmer, ON: Ontario Police College.

Royal Canadian Mounted Police. (2010). *Physical Abilities Requirement Evaluation (PARE) Course Training Standard (CTS).* Ottawa, ON: RCMP Learning and Development, Ottawa.

Royal Canadian Mounted Police. (2012). Gender based assessment. *National Program Evaluation Services Royal Canadian Mounted Police*, Ottawa, Ontario. Retrieved from RCMP Intranet under National Research and Evaluation Unit

Police Fitness Personnel of Ontario. (2012). *Certified police fitness appraiser manual*. Hamilton, ON: Hamilton Regional Police.

Roitman, J. L. (Ed.) (2001). *ACSM's resource manual for guidelines for exercise testing and prescription* (4th Ed.). Baltimore, MD: Lippincott, Williams & Wilkins.

Reed, A. (1998). The oxygen cost of running the Physical Ability Requirement Evaluation. In J. Bonneau (2001). *The development of PARE: An evolution*. Ottawa, ON: RCMP Learning and Development Archives.

Royal Canadian Mounted Police. (2012). Gender–Based Assessment. *National Program Evaluation Services Royal Canadian Mounted Police*. Retrieved from http://www.rcmp-grc.gc.ca/aud-ver/reports-rapports/gba-eces/gba-eces-eng.pdf

Sassen, B., Cornelissen, V. A., Kiers, H., Wittink, H., Kok, G., & Vanhees, L. (2009). Physical fitness matters more than physical activity in controlling cardiovascular disease risk factors. *European Journal of Cardiovascular Prevention and Rehabilitation*, *16*, 677-673. doi:10.1097/HJR.0b013e3283312e94

Sassen, B., Kok, G., Schaalma, H., Kiers H., & Vanhees, L. (2010). Cardiovascular risk profile: Cross-sectional analysis of motivational determinants, physical fitness and physical activity. *BMC Public Health, 10*, 592-601. doi:10.1186/1471-2458-10-592.

Seo, N.J., & Armstrong, T.J. (2009). Biomechanical analysis for handle stability during maximum push and pull exertions. *Ergonomics*, *52*(12), 1568-1575. doi:10.1080/00140130903287999

Seo, N.J., Armstrong, T.J., & Young, G.J. (2010). Effects of handle orientation, gloves, handle friction and elbow posture on maximum horizontal pull and push forces. *Ergonomics*, *53*(1), 92-101. doi:10.1080/00140130903389035

Sharkey, B.J. & Davis, P.O. (2008). *Hard work: Defining physical work performances.* Windsor, ON: Human Kinetics.

Shaw, & Gledhill, N. (1995, March). *Job review phase: Medical, physical, skills and abilities recommendations, tests and standards for the constable selection report*. Ottawa, ON: Ministry of the Solicitor General and Correctional Services. Retrieved from Ontario Police College Archives.

Shields, M., Tremblay, M. S., Laviolette, M., Craig, C. L., & Gorber, S. C. (2010). Fitness of Canadian adults: Results from the 2007-2009 Canadian health measures survey. *Statistics Canada Health Reports, 21*(1), 1-15. Catalogue no 82-003.

Shell, D.E. (2003). Exercise science and law enforcement. *Strength and Conditioning Journal*, *25*(3), 52-57. Ottawa ON: Retrieved from Canadian Police College Library.

Shephard, R. & Bonneau, J. (2002). Assuring gender equity in recruitment standards for police officers. *Canadian Journal of Applied Physiology*, *27*(3), 263-295.

Shephard, R. & Bonneau, J. (2003). Supervision of occupational fitness assessments. *Canadian Journal of Applied Physiology*, *28*(2), 225-239.

Sommerfeld, H., Dunlop, B., & Neary, J. P. (1998). Heart rate response during the RCMP PARE test [Abstract]. *Canadian Journal of Applied Physiology*, *23*(5), 509.

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine, 18(*10*),* 1099-1104. doi:10.1111/j.1553-2712.2011.01185.x.

Stanish, H. I., Campagna, P. D., & MacKinnon, S. C. (1993). Comparison of the Physical Ability Requirement Evaluation to measures of physical fitness. *Applied Physiology, Nutrition, and Metabolism, 41*(8), August 1999, 669-677. Retrieved from http://www.nrcresearchpress.com/journal/apnm

Stanish, H. I. (1993). A comparison of the physical ability requirement evaluation simulated fight component to measures of muscular strength, endurance and power. Unpublished paper Ottawa, ON: Fitness & Lifestyle Program.

Stanish, H. I., Wood, T. M., & Campagna, P. (1999). Predictors of performance on the RCMP physical abilities requirement evaluation. *Journal of Occupational and Environmental Medicine, 4*(8), 669-677. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Strating, M., Bakker, R. H., Dijkstra, G. J., Lemmink, K. A. P. M., & Groothoff, J. W. (2010). A job related fitness test for the Dutch police. *Occupational Medicine, 60*, 255-260. doi: *10.1093/occmed/kqt102*

Strath, S. J., Kaminsky, L. A., Ainsworth, B. E., Ekelund, U., Freedson, P.S., Gary, R. A., Richardson, C.R., Smith. D.T., & Swartz, A.M. (2013). Guide to the assessment of physical activity: Clinical and research applications. A scientific statement of the American Heart Association. *Circulation*, *128*, 2259-2279. doi:10.1161/01.cir.0000435708.67487.da

Sternfeld, B., & Goldman-Rosas, L. (2012). A systematic approach to selecting an appropriate measure of self-reported physical activity or sedentary behavior. *Journal of Physical Activity and Health*, *9*(Suppl.), S19-S28. Retrieved from http://journals.humankinetics.com/jpah

Thoden, H. C. (2002). *Testing for normality*. New York, NY: Marcel Decker Inc.

Thomas & Means. (2000). *Use of sex and age adjusted physical performance standards in the law enforcement profession*. PRIMA conference position paper. Alymer ON: Retrieved from Ontario Police College Archives.

Thompson, W. R., Gordon, N. F. & Pescatello, L. S. (2009). *ACSM's: Guidelines for exercise testing and prescription, (8th Edition).* New York, NY: Lippincott Williams & Wilkins.

Trottier, A., & Brown, J. (1994). Occupational health in police work: A Canadian perspective. *Journal of Clinical Forensic Medicine*, *1*, 39-42. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Wagner-Wisotzki, N. (2005). *Fitness and Lifestyle Management for Law Enforcement*, (2nd Ed.). Toronto, ON: Edmund Montgomery Publishers.

Welk, G. J. (2002). *Physical activity assessment in health-related research.* Windsor, ON: Human Kinetics.

Wenger, H. A. & Gaul, C. A. (1992). *RCMP physical abilities requirement evaluation demonstration project: August 1990-August 1992, Final Report.* Ottawa, ON: RCMP Health Services Archives.

Weir, J. P., Beck, T. W., Cramer, J. T. & Housh, T. J. (2006). Is fatigue all in your head? A critical review of the central governing model. *British Journal of Sports Medicine, 40*(7), 573-586. doi:10.1136/bjsm.2005.023028

Wells, G. D., Selvaduria, H., & Tein, I. (2009). Bioenergetics provision of energy for muscular activity. *Paediatric Respiratory Reviews*, *10*, 83-90. doi:10.1016/j.prrv.2009.04.005

Whitehead, M. (2007). Physical Literacy: Philosophical considerations in relation to developing a sense of self, universality and propositional knowledge. *Sport, Ethics and Philosophy, 1*(3), 281-298. doi:10.1080/17511320701676916

Wilson, D. & Bracci, R. (1982). Police agility test: Another look at pre-employment screening. *Law and Order*, *30*(8), 36-38, 41-42. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=85008

Wilson, T.M. & Tanaka, H. (2000). Meta-analysis of the age-associated decline in maximal aerobic capacity in men: relation to training status. *American Journal of Physiology and Heart Circulation, 278*: H829-834.

Wilmore J.H. & Davis, J.A. (1979). Validation of a physical abilities test for the selection of state traffic officers. *Journal of Occupational Medicine*, *21*(1), 33-39. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

World Health Organization. (1986). *Ottawa charter for health promotion*. Retrieved from http://www.who.int/hpr/NPH/docs/ottawa_charter.hp.pdf

Zumbo, B.D. (2000). Methodology and measurement matter in establishing a Bona Fide Occupational Requirement for physically demanding occupations. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Chapter 5

Implications in PARE Predictor Studies

The three retrospective observational studies were completed to assess predictors of PARE success. The purpose, major findings, and implications of these three integrated studies are discussed below.

5.1 The Relative Predictive Power of the Pursuit and Body Control Simulations and their Relationship Across Sex Groups

The purpose of the first study was to examine the relative predictive power of the pursuit and body control segments of PARE and assess if this relationship differs across gender groups. The main findings of the study were that the data supported an equal prediction model composed for pursuit and body control scenario times log odd (*LO)* = 2.7 and 2.7. A significant predictor but weak predictor effect was the square root of age, with a *LO* = .48 in the combined data of 13,709 officers. The relative predictive power of these two elements was nearly identical, notwithstanding their 70-30 % descriptive statistical differences for time portions of the total PARE test score. The same almost equal log odds results was true for female officers' population data (*n* = 2,127) for pursuit *LO* = 2.40 and body control *LO* = 2.45 and male officers' population data (*n* = 11,582) for pursuit *LO* = 3.00 and body control scenario *LO* = 3.0. However, the data did not support first and second moment fitness testing covariate factors of height, weight, and sex as predictive of PARE success.

5.2 Discussion

The unexpected results of equality of log odds for both skill-related test scenarios are unique in occupational skill-related fitness testing literature. That sex was not a significant PARE predictor at PES of 4:00 min:s, is equally unknown in the literature and surprising to most exercise physiology clinicians. More specifically, the fact that sex group differences were not found to predict success is a major finding for a purportedly gender neutral skill-related testing. Sex group descriptive statistics have an average PARE times difference between men and women as 32 seconds apart, with mean PARE times of 3:37 min:s for men versus 4:09 min:s for women. This difference in mean PARE time for each sex group did not translate into an unequal predictive log odds for pursuit and body control scenarios. Similar to Gaul and Wenger (1992), incumbent testing incumbent male officers outperformed women officers as a group; however, this federal population data set is much larger than the 1992 pilot study.

Many of the issues highlighted in Gaul and Wenger's (1992) pilot incumbent PARE testing study a generation ago have been removed through standardization of equipment that minimizes what Messick (1996) termed irrelevant task difficulty. Additionally, biomechanically informed teaching cues and supportive equipment changes have provided instruction that negates the historical influences of height / weight and sex and made equipment accessible for all smaller statured applicant and incumbent officers. Examples of equipment and teaching changes include (a) standard non-slip rubber or carpet flooring for greater than 50% co-efficient friction; (b) accessible 18 in vertical push handles versus older 12 in push handles; (c) a removal of orthogonal plane of motion of the handles of older models; (d) and removal of the neck level supportive pad in the older model push pull machine in 2007 that blocked power transfer in shorter men and women. Force output results are now speculated to be more indicative of sufficient metabolism and functional capacity and the true construct of the test's whole body strength endurance push and pull scenario, and not upper body arm maximal strength (Armstrong & Young, 2010; Ayoub & McDaniel, 1974; Boocock, Haslam, Lemon, & Thorpe, 2006; Chaffin, Herrin, Lee, & Waikar, 1991; Chaffin & Resnick, 1995; Ciriello & Snook, 1991; DARCOR and ERGO web, 2103; Knapik & Marras, 2009; Kroemer, 1971; Seo & Armstrong, 2009; Seo et al., 2010). This supports the status of PARE as a BFORs defensible test that is not subjectively based on prohibited discriminatory movements or demands and not objectively sexist in cut score standards (Eid & Geh, 2001; Zumbo, 2000). However, the near equality in log odds predictive weight has clinical implications for training programming, namely equally weighted instruction is important for cardiorespiratory and upper body – core strength training instruction.

5.3 Implications

The implications of the first study is that the data supports sex being removed as a predictive factor. Clinical implications support education that continues to have an emphasis on biomechanics of equipment set up and proper coaching of pushing and pulling technique that contributes to measuring the maximal push construct alone. Balance or posture constructs like upright pushing with on foot forward and back (offset parallel stance) should be avoided. Another implication of study 1 is that the square root of age was a statistically significant predictive factor, although only slightly above unity. In pilot work before the removal of 2 + SD outliers of PARE times, the square root of age's log odds was 1.0 or a 50-50 probability, and was significant in the national data set. However, the statistically significant value of the age predictor did not have any clinical value as it represented unity or 1.0. A log odds of 1.0 is a 50-50% prediction chance and not much of a predictive factor (Pampel, 2000). Only when influential outliers were removed from the data set, with the data representing higher levels of effort and performance with lower PARE time scores, did the square root of age become more statistically and clinically important as a PARE predictor variable. Age then became statistically significant just above the 50% prediction or 1.0 log odds level. It is speculated that further outliers removal to below the current + 2 SD cut off above the mean, would have seen the trend for age towards greater statistical significance occur.

An implication for future research is that if a performance marker is found, like motivation or physiological effort as a percentage of maximum, then perhaps not only outlier removal of raw time scores for walkers would produce sex and age trending towards significance. Perhaps groups of maximal effort versus participatory effort could enhance analysis with another group above pass / fail performance data groupings.

The overarching implication for covariates is that PARE is gender, height, and weight neutral. Only age from previous moments of fitness testing covariates of height, weight, age, and sex predicts success.

5.4 Study Two

The purpose of the second study was to determine the effect of pacing in the pursuit circuit on performance success in the PARE and any differences for sex groups. Although the divisional data set could not address main effect size post hoc analysis when the interaction was present, matching pairs data allowed for main effect size post hoc analysis and comparison.

There were three major findings in this retrospective observational study's divisional data and 61 matched officer pair's data cohort study. Pass / fail PARE groupings were statistically significantly different for repeat lap times with strong effect sizes in both data groups, whereas the male versus female grouping were statistically significant, but with weak effects sizes and no clinical significance in both data groups. Repeat laps times were signfiicant in both data groups, with the strongest power drop for lap 1 to lap 2. Divisional data had low fit men a lower scores than low fit women, but hi fit men were faster than high fit women, showing an interaction of sex and performance. There was no interaction in matched apirs data, and no disordinal sex group interaction for within subject repeated lap times and only a slight ordinal effect at repeat lap number 3 and 6.

The first implication is sex is the minor influence that, although statisically significant, has little clinical significance. Performance is the main strong effect size. This leads to a conclusion that perfomance is a major influencer and that fitness variance within a gender far exceeds fitness variance between genders. In the larger divisional data analysis, men outnumbered women 535 to 85 or roughly 6.3 to 1 and sex groups were uneqaul for age, height, and weight values, and therefore their these human factors use as covariates was untenable in an ANCOVA (Fields, 2012; Gamst, Meyers & Guarino, 2008). Further real world data samples that are unbalanced cannot be used for sex differences analysis unless moderators of sex differences for available data like height and weight and age are taken into consideration. Zumbo (2000) suggests comparing groups on an underlying construct of interest to better analyze for disparate group functioning. The controlling of groups for equal covariates values for closely matched pairs ($\pm$ 0.1) was the data we had that we could equate, after the fact for 61 men and women officers. This allows some groups with levels of covariates to be analyszed. Future researcher will have to look to prospective studies that can use or enhance equating techniques to further study sex group differnces.

5.5 **Implications**

The implications for policing are that male and female officers fatigue in almost indentical patterns. So the only consideration in a scenario is the absolute performance score of the individual for the task. It is clear that men were faster and had more pronounced 2 power drops between lap 1 to 2 and lap 3 to 4 during the pursuit scenario. However, no sex specific instructions are needed as the fatigue patterns are very close. Instead energy systems instructions are implicated equally for both male and female officers, for both alactic and lactic peak power and energy capacity out of those systems (Gore, 2000). The implication from the above failing group data findings is also that rehabilitation of policing skills in a long foot pursuit is gender neutral when failing times are considered.

This study's results cannot be automatically generalized to the complete federal police population with its 4 to 1 male to female ratio. At the divisional level, the 6.3 to 1 ratio also limits generalizing the matched pairs' data findings to other divisional populations. Causation statements would require random sampling for findings to be generalized or other samples of matched pairs. However, future research can use the matched pairs' data that support an identical pacing pattern for high men and women, to construct ever-larger more representative samples and to test the null hypothesis of disparate fatigue patterns across sex groups.

The second study’s limitation was its inability in real world data sets to represent enough low BMI men and high BMI women, in matched pairs research, as no equal matched pair could be found at the upper and lower ends of BMI values. Future research would have to purposely solicit this group, with sufficient > 30 BMI, < 20 BMI officers, and not count on real world data gathering to provide it. The implication is that minimal sex difference may only exist in the moderate BMI range, as no representative hi and low group members were present in this data. Only the middle portion of the BMI range was a match from the limited divisional data of about 800 cases. Thus, the findings could not be extended to the whole range of police officers’ anthropometrics, which may provide unique health promotion fitness challenges for policing fitness for duty.

Increasing the data capture available, from some 800 divisional records spanning 6 years captured by this author, with laps times hand recorded, to all the 14,000 unique PAREs of the national population captured by all 12 DFLAs, with lap times captured systematically in HRMIS, would expand the range of research and its ultimate findings. Should the national database expand (instead of its recent contracting of its data capture) to include individual repeat lap times, then larger matched samples could be run. The implication of enhanced data capture would be broadened research potential available to the divisions and the national policy centre.

Although PARE was mandated, participatory in-test effort was described as self-determined, after CEP test supervision instruction. It still remained evident that fast PARE scores require metabolic capacity and motivation. As suggested by Reed (1998), having knowledge of what constitutes a capacity challenging effort, by using a physiological marker in field tests, can help create performance sub-groups. Motivation markers can also do the same using perceived exertion scales like the original 6 to 20 perceived exertion scale or the newer modified 1 to 10 Borg scale of perceived exertion. Eliminating influential outliers does not address motivation and metabolic capacity issues, although age, a third important inter-individual difference, was accounted for.

No implications could be drawn from comparing this large data record retrospective study predictive findings with other PARE prediction studies. Other studies to date were dominated by small sample sizes and the difference between studies is so disparate, with so few women in the older studies, and so few total subjects, and so few fit women with mean PARE scores far from the PES of 4:00 min:s, that findings from this study appear to be unique.

5.6 Study Three

The purpose of the third study was to examine additional predictive factors of PA for the PARE success predictor model. Specifically the purpose was to examine the predictive effects of self-reported PA frequency and intensity on PARE pass / fail success outcomes along with other potential covariates of height, weight, age, and sex from previous fitness testing historical moments.

5.7 Discussion

PA weekly frequency and PA intensity out of 3, variables were not additional significant predictors of PARE success. PA self-report at PARE may be invalid due to the brevity of the brief oral interviews, only asking for two of many PA factors of varying descriptions, including dimensions, energy costs, and or modes of PA activity. However, if the reported values were real and valid, indicators, at the levels reported, they may never influence success due to their insufficiency.

The divisional data did not support additional PA factors of frequency and intensity (three levels: light, moderate, or intense) as significant predictors. Problematic with self-reporting PA were the following issues from the literature. One, the lack of a comprehensive conceptual framework for reporting PA as a complex multidimensional behaviour, with sedentary behaviour (Gabriel et al., 2012). Two, the limited pre-test interview lacked some missing domains of PA, like occupational, domestic, transportation/utilitarian, and leisure time cues (Strath et al., 2013) to stimulate the more accurate recall episodic events versus lesser accurate generic PA memory recall (Matthews, 2002). Additionally, missing was any mode specific dimension recall of activity related to the skill-related testing of police movement types or biomechanical descriptors of movement challenges (Strath et al., 2012). Missing from the limited pre-test interview was the energy expenditure cost standardization of METs minutes per week from a list of PA to facilitate comparisons of work across individuals (Ainsworth et al., 2000). Problematic to reporting intensity is also the inter-individual differences that come with age and varying metabolic capacity, which usually necessitates relative versus absolute values intensity reporting (Welk, 2002).

Finally, it may be that an average PA intensity of 2.15 of 3, for men and 2.29 of 3 for women, in the divisional data, may not be sufficient activity to affect a maximal effort test. If moderate PA were defined as 70 to 85 % of a maximum effort, 2.15 is rather close to a 70% effort, or what Howell (2001) would describe as a 12-13 on a 20 point rating of perceived exertion scale, which could be a brisk walk. Success on a max running police test that shows consistent power drops throughout a six lap test, seen in the study two, which represents the same divisional data, clearly implies fatigue through two energy systems to a third system in a few minutes (Wells, Selvaduria, & Tein, 2009' Gore, 2000). Subject pre-test preparation of a steady modest PA effort of slow steady jogging (i.e., 12-13 on the Borg Scale) for a fatiguing maximal effort run test may not be sufficient to be a significant predictor of PARE pass / fail success without more PA recall information.

5.8 Implications

Subjects exhibited insufficient PA activity in the federal police divisional data of convenience to influence a maximal effort test. Even though average reported PA intensity and frequency are below what we would expect to be influential levels, lack of mode specific type of police movements also needs remediation. Validity of self-recall PA research efficacy is also suspect with such a short PA recall list of variables. Greater sophistication in integrating PA recall into skill-related testing, especially mode specific movements of running, jumping, pushing and pulling, lifting and carrying would be suggested. In like manner, maximal effort and maximal motivational recording could group participants separate from performance not just on the PES 4:00 min:s but on full personal motivation of perceived exertion.

As a group, these studies support additional insight into what appears to be a gender-neutral skill-related occupational fitness test. Further, they provide such large sample data groups as to approximate federal police populations means and standard deviations for performance descriptions. This data is more representative of large police population's parameters than small sample groups and small much older validation study groups, from a previous generation of police officers. These studies also include the most significant number of women (2,127) police officer subjects in the research literature representing real world habitual testing data.

Future research implications are also evident in looking at the actual discrete test elements for disparate mat and stick penalty times that caused them to be excluded from logistic regression analysis in study one. This should be examined for future test impact. Additionally if the residual was so large in the regression analysis: Is the five-second mat penalty too severe? Is the stick two second penalty necessary? Finally, can another way of grouping performance pass / fail, other than time, or other than the 4:00 min:s PES develop different insights into skill-related PARE testing amongst RCMP incumbent officers?

5.9 References

Ainsworth, B. E., Haskell, W. L., Whitt, M. C., Irwin, M. L., Swartz, A. M., Strath, S. J., … Leon, A.S. (2000). Compendium of physical activities: An update of activity codes and MET intensities. *Medicine & Science in Sports & Exercise*, *32*(9), S498-S516. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Ainsworth, B. E., Casperson, C. J., Matthews, C. E., Massse, L. C. Baransowski, T., & Zhu, W. (2012). Recommendations to improve the accuracy of estimates of physical activity derived from self-report. *Journal of Physical Activity and Health*, *9*(Suppl.), S76-S84. Retrieved from http://journals.humankinetics.com/jpah

Anderson, G. S., & Plecas, D. (2007). PARE Task Analysis. *RCMP Learning and Development.* Ottawa ON: Retrieved from RCMP Internal Learning and Development intranet web page archives.

Anderson, G. S., & Plecas, D. (2008). *Police abilities testing: Revalidation of the task analysis*. Ottawa, ON: RCMP Learning and Development Archives.

Anderson, G. S., Plecas, D., & Segger, T. (2001). Police officer physical ability testing: Re-validating a selection criterion. *Policing: An International Journal of Police Strategies and Management, 24*, 8-31. doi: 10.1108/1369510110382232

Andrews, A., & Risher, J. (2006). What does that have to do with being a cop? Employment Standards in law enforcement. *International Association of Chiefs of Police Conference,* Boston, MA. Retrieved from https://www.ncjrs.gov/App/publications/abstract.aspx?ID=187913

Avery, R. D., Nutting, S. M., & Landon, T. E. (1992). Validation strategies in physical ability testing in police and fire settings. *Public Personnel Management*, *2*(3), 301-312. Retrieved from http://ppm.sagepub.com/

Bar-Or, O. (1987). The Wingate anaerobic test: An update on methodology, reliability and validity. *Sports Medicine, 4*, 381-394. Retrieved from http://link.springer.com/journal/40279

Birzer, M. L., & Craig, D. E. (1996). Sex differences in police physical ability test performance. *American Journal of Police, 15(*2), 93-108. doi:10.1108/07358549610122494

Boocock, M. G., Haslam, R. A. Lemon, P., & Thorpe, S. (2006). Initial force and postural adaptations when pushing and pulling on floor surfaces with good land reduced resistance to slipping. *Ergonomics*, *49*(9), 801-882. Retrieved from http://www.tandfonline.com/toc/terg20/current

Bonneau, J. (1996). *PARE: Standard summary presentation to the Canadian Human Rights Commission*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J. (1998). *Submission to the Canadian Human Rights Commission*: PARE. Briefing paper outlining PARE as a BFOR development and the 4 minutes standard. Ottawa, ON; RCMP, Fitness & Lifestyle Archives.

Bonneau, J. (2001): *The development of PARE: An evolution*. Ottawa, ON: RCMP Fitness & Lifestyle Archives.

Bonneau, J., & Brown, J. (1995). Physical ability, fitness and police work. *Journal of Clinical Forensic Medicine*, *2*, 157-164. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Brown, J., & Trottier, A. (1994). *Police health: A physician's guide for the assessment of police officers*. Ottawa, ON: Canada Communications Group.

Boyce, R. W., Jones, G. R., Schendt, K. E., Llyod, C. L., & Boone, E. L. (2009). Longitudinal changes in strength of police officers with sex comparisons. *Journal of Strength and Conditioning Research, 23*(8), 2411-2418. doi:10.1519/JSC.0b013e3181bac2ab

Casson, E. J. (2000). Establishing BFOR and validated standard for vision. In N. Gledhill, J. Bonneau, A. & Salmon (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Canadian Society for Exercise Physiology. (2012). *The Canadian physical activity, fitness & lifestyle appraisal: CSEP's plan for healthy active living* (2nd Ed.). Ottawa, ON: Canadian Society for Exercise Physiology.

Collingwood, T. R. (1988a). Implementing programs and standards for law enforcement physical fitness. *The Police Chief,* April *55*(4*),* 20-24. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1988b). Physical fitness leadership in law enforcement. *The Police Chief*, April *55*(4), 28-34. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (1995). Physical fitness standards: Measuring job relatedness. *The Police Chief*, February *62,* 31-31. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Collingwood, T. R. (2004,). Underlying physical fitness factors for performing police officer physical tasks. *The Police Chief.* March *71*(3), 1-7. Retrieved from www.policechiefmagazine.org/magazine/index.cfm?fuseaction=archives

Chaffin, D. B., Herrin, G. D., Lee, K. S., & Waikar, A. M. (1991). Effect of handle height on lower back loading in cart pushing and pulling, *Applied Ergonomics*, *22*(2), 117-123. doi:10.1016/0003-6870(91)90310-E.

Cynrim, C., Mergner, T., & Maurer, C. (2009). Potential role of force cues in human stance control. *Experimental Brain Research*, *194*(3), 410-433. doi:10.1007/s00221-009-1715-7

Darcor & Ergoweb. (2013). *The ergonomics of manual material handling: Pushing and pulling tasks*. Sutton Javelin Communications for Darcor Limited. Retrieved from http:/www.darcor.com/technical/ergonomics of manual material handling/

Deakin, J. M., Smith, J. T., Pelot, R. P., & Weber, C. L. (2000). Methodological considerations in the development of physical maintenance standards. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

DeMaris, A. (1995). A tutorial in logistic regression. *Journal of Marriage and the Family, 57*(4), 956-968. Retrieved from: http://www.jstor.org.proxy1.lib.uwo.ca/stable/353415

Dunsmore, R. R., & Hunter, A. J. (2001). Maneuvering the obstacle course: Implementing incumbent physical fitness testing in the workplace. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Eid, E., & Geh, S. (2001). Challenges posed by the Supreme Court of Canada in the Meiorin decision to employees in physically demanding occupations. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, ON: York University Press.

Equal Employment Opportunity Commission. (1978). *Uniform guideline on employee selection procedures*. Retrieved from www.uniformguidelines.com

Farenholtz, D. W., Rhodes, E. C., & Burrell, A. J. (1983). Development of a physical abilities test for correctional officers in British Columbia, included in *Appendix II of Police Officers Physical Abilities Study (November 1986).* New Westminster, BC. Justice Institute of British Columbia. Retrieved from the Ontario Police College Archives.

Farenholtz, D. W., & Rhodes, E. C. (1986). *Police officer physical abilities study*. Justice Institute of British Columbia, New Westminster BC. Retrieved from the Canadian Police College Library archives.

Farenholtz, D. W., Rhodes, E. C., & Bonneau, J. (1989). Validation of a police officer's physical ability test. *Canadian Association of Sports Sciences Poster Presentation paper*. Retrieved from RCMP archives.

Farenholtz, D. W., & Rhodes, E. C. (1990). Recommended Canadian standards for police physical abilities. *Canadian Police Journal, 14*(1), 37-49. Retrieved from http://www.questia.com/library/p90673/canadian-journal-of-police-and-security-services

Frank, W. D., & Anderson, D. F. (1994). Relationship between physical activity and risk factors for cardiovascular disease among law enforcement officers. *Journal of Occupational Medicine, 36*(10), 1127-1132. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Gabriel, K. P., Sidney, S., Jacobs, D. R., Quesensberry, C. P., Reis, J. P., Jiang, S., & Sternfeld, B. (2014). Convergent validity of a brief self-reported physical activity questionnaire. *Medicine & Science in Sports & Exercise, 46*(8), 1570-1577. doi: 10.1249/MSS.0000000000000278

Gamst, G., Meyers, L. S., & Guarino, A. J. (2008). *Analysis of variance designs: Conceptual and computational approach with SPSS and SAS*. New York, NY: Cambridge University Press.

Gaul, W., & Wenger, H. (1992). *RCMP Physical Abilities Requirement Evaluation demonstration project: Final report.* Victoria, BC: University of Victoria, School of Physical Education. Retrieved from RCMP Archives.

Gillis, A. D., & Darby, B. (2001). To serve and protect, post Meiorin an argument for due diligence is an argument for undue hardship: The liability labyrinth faced by the public safety in recruitment and retention of competent employees. In N. Gledhill, J. Bonneau & A. Salmon. (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto, 2000. 85-112. York University Press.

Girard, G. (2012). *Five year PARE compliance report*. Ottawa: ON: RCMP Learning and Development internal communications.

Gore, C.J. (Ed.). *Physiological tests for elite athletes*. Human Kinetics, Windsor ON.

Gledhill, N., & Shaw, C. (1995, March). Medical, physical, skills and abilities recommendations, tests and standards. *Constable Selection System: Report on Job Review Phase*, 1. Aylmer ON: Retrieved from the Ontario Police College Archives.

Gledhill, N., & Bonneau, J. (2001). Objectives, processes and consensus summary of the national forum on establishing Bona Fide Occupational Requirements. In N. Gledhill, J. Bonneau, & A. Salmon. (Eds.). *Proceedings of the National Forum on Bona Fide Occupational Requirements.* Toronto*,* ON*:* York University Press.

Gledhill, N., Bonneau, J., & Salmon, A. (2000). Proceedings of the consensus forum on establishing BONA FIDE requirements for physically demanding occupations. Toronto, ON. York University Press.

Haskell, W. L. (2012). Physical activity by self-report: A brief history and future issues. *Journal of Physical Activity and Health*, *9*(Suppl.), S5-S10. Retrieved from http://journals.humankinetics.com/jpah

Haslam, R. A., Boocock, M., & Lemon, S. (2000). *Proceedings of the Human Factors and Ergonomics society Annual Meeting July 2000*, 44(28); 479-481. Windsor ON: Human Kinetics. doi: 10.1177/154193120004402817

Hagströmer, M., Ainsworth, B. E., Kwak, L., & Bowles, H.R. (2012). A checklist for evaluating the methodological quality of validation studies on self-reported instruments for physical activity and sedentary behavior. *Journal of Physical Activity and Health*, *9*(1), S29-S36. Retrieved from http://journals.humankinetics.com/jpah

Hoffman, B. (1996). How today's law enforcement officer ranks. *Joe Weider's Muscle and Fitness*, *76*, 232-235. Retrieved from http://www.muscleandfitness.com/news-and-features/athletes-and-celebrities/life-joe-weider

Hooker, S.P., Fulton, J., & Mudd, L.M. (2012). Practice-based evidence: A novel concept for self-report physical activity measurement. *Journal of Physical Activity and Health*, *9*(Suppl. 1), S85-S87. Retrieved from http://journals.humankinetics.com/jpah

Howell, D.C. (1997). *Statistical methods for psychology*. (8th Ed.). Wadsworth CA, Cengage Learning.

Jackson, A.S. (1994). Pre-employment physical evaluation. *Exercise and Sports Science Reviews, 22*(1), 53-90. Retrieved from http://journals.lww.com/acsm-essr/pages/default.aspx

Jamnik, V.K., Thomas, S.G., Shaw J.A., & Gledhill, N. (2010). Identification and characterization of the critical physically demanding tasks encountered by correctional officers. *Applied Physiology, Nutrition and Metabolism*, *35*, 45-58. doi: 10.1139/H09-121

Jamnik, V. K., Thomas, S. G., Burr, J. F., & Gledhill, N. (2010). Construction, validation, and derivation of performance standards for a fitness test for correctional officer applicants, *Applied Physiology, Nutrition and Metabolism*, *35*, 59-70. doi:10.1139/H09-122

Jamnik, V.K., Thomas, S.G. & Gledhill, N. (2010). Applying the Meiorin decision requirements to the fitness test for correctional officer applicants: Examining adverse impact and accommodation. *Applied Physiology, Nutrition and Metabolism*, *35*, 71-81. doi: 0.1139/H09-123

Keith, T. Z. (2006). *Multiple regression and beyond*. New York, NY: Pearson.

Kesaniemi, Y. A., Danforth, E., Jensen, M. D., Kopelman, P. G., Lefebvre, P., & Reeder, B. A. (2001). Consensus statement: Dose response issues concerning physical activity and health: An evidence-based symposium. *Medicine & Science in Sports & Exercise, (Suppl.) 33*(6), S351-S358. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Kinnear, P. R. & Gray, C. D. (2010). *PASW 17: Statistics made simple*. New York, NY: Psychology Press.

Kleinbaum, D. G., & Klein, M. (2002). *Logistic regression: A self-learning text*. New York, NY, Springer.

Knapik, C. G., & Marras, W. S. (2009). Spine loading at different lumbar levels during pushing and pulling. *Ergonomics*, *52*(1), 60-70. doi:10.1080/00140130802480828

Koller, M., Gledhill, N., Payne, N., Jamnik, R., Fisher, K., Gordon, C., & Halligan, J. (1999). Evaluation of a new upper body strength meter for assessing upper body strength. *Medicine & Science in Sports and Exercise*, *31*, S114. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Kuruganti, U., & Rickards, J. (2004). The role of human factor engineering in establishing occupational fitness standards. *International Journal of Industrial Engineering*, *34*(6), 451-457. doi:1016/j.ergon.2004.05.004

Lagasse, P. P. (1989). *Assessment of Physical Standards for Admission in the Police Forces*. [Élaboration des normes physiques d'admission aux d'agents de la paix]. Unpublished manuscript, Faculty of Physical Education, Laval University, Quebec City, Canada.

Legault, J. P. (2000). Establishing a bona fide occupational requirement for occupational requiring a hearing abilities. In N. Gledhill, J. Bonneau, & A. Salmon (Eds.). *Proceedings of the consensus forum on establishing BONA FIDE requirements for physically demanding occupations*. Toronto, ON: York University.

Lonsway, K.A. (2003). Tearing down the wall: Problems with consistency, validity, and adverse impact of physical agility test in police selection. *Police Quarterly*, *6*, 237-276. doi:10.1177/1098611103254314/

Maher, P. T. (1984). Police physical ability tests: Can they ever be valid. *Public Personnel Management, 32*(2), 173-178. Retrieved from http://ppm.sagepub.com/

Marchand, R., Reed, A., Thoden, J. S., & McNeely, E. (1995). *Physical screening protocol for correctional officer applicants.* Toronto, ON: Ministry of the Solicitor General and Correctional Services. RCMP Fitness and Lifestyle-Archives.

Masse, L., & de Niet, J. E. (2012). Sources of validity evidence needed with self-report measures of physical activity. *Journal of Physical Activity and Health*, *9*(Suppl.), S44-S55. http://journals.humankinetics.com/jpah

McMurray, R. G., Ainsworth, B. E., Harrell, J. S., Griggs, T. R., & Williams, O. D. (1998). Is physical activity or aerobic power more influential on reducing cardiovascular disease risk factors? *Medicine & Science in Sports & Exercise, 30*(10), 1521-1529. Retrieved from http://journals.lww.com/acsm-msse/pages/default.aspx

Menard, S. (2010). *Logistic regression: From introductory to advanced concepts*. Thousand Oaks, CA, Sage.

Messick, S. (1998). Test validity: A matter of consequence. *Social Indicators Research, 45*, 34-45. Retrieved from http://www.springer.com/social+sciences/journal/11205

Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, *119*(1), 40-48. doi: 10.1037//0021-843x.110.1.40

National Executive Institute Association. (2008). *Promoting health and wellness: Returning to full duty by the Major Cities Chiefs and the Federal Bureau of Investigation*. Proceedings of the 2008 Toronto Conference. Quantico, Virginia. Retrieved from www.neiassociates.org/PromotingHealth-fullduty.pdf

Nabeel, I., Baker, B.A., McGrail, M.P., & Flottemesch, T.J. (2007). Correlation between physical activity, fitness, and musculoskeletal injuries in police officers. *Minnesota Medicine, 90*(9), 40-43. Retrieved from http://www.minnesotamedicine.com/Past-Issues/Past-Issues-2007/September-2007/Clinical-Nableel-September-2007

Nusser, S. M., Beyler, N. K., Welk, G. J., Carriquiry, A. L., Fuller, W. A., & King, M. N. (2012). Modeling errors in physical activity recall data. *Journal of Physical Activity and Health*, *9*(Suppl. 1). S56-S67. Retrieved from http://journals.humankinetics.com/jpah

Nutbeam, D. (1998). *Health promotion glossary*. World Health Organization. Retrieved from http://www.who.int/hpr/NPH/docs/hp_glossary_end.pdf.

Ontario Ministry of Public Safety and Security. (2002). *The Ontario Constable Selection System: Technical guide for users of the physical readiness evaluation for police the PREP*. Toronto, ON: Queens's Printers.

Osborn, G. D. (1976,). Physical agility testing: Validating physical agility tests. *The Police Chief, January*, 45-50. Retrieved from Canadian Police College Library, Ottawa

Physical Education Canada. (2013). *What is physical literacy*? Retrieved from http:/phecanada.ca/programs/physical-literacy/what-physical-literacy.

Police Fitness Personnel of Ontario. (2012). *Certified Fitness Appraiser's Manual.* Aylmer, ON: Ontario Police College.

Royal Canadian Mounted Police. (2010). *Physical abilities requirement evaluation (PARE) course training standard (CTS).* Ottawa, ON: RCMP Learning and Development.

Royal Canadian Mounted Police. (2012). Gender–based assessment. *National Program Evaluation Services Royal Canadian Mounted Police*. Retrieved from http://www.rcmp-grc.gc.ca/aud-ver/reports-rapports/gba-eces/gba-eces-eng.pdf

Police Fitness Personnel of Ontario. (2012). *Certified police fitness appraiser manual*. Hamilton, ON: Hamilton Regional Police.

Roitman, J. L. (2001). *ACSM's resource manual for guidelines for exercise testing and prescription* (4th edition). Baltimore, MD: Lippincott, Williams & Wilkins.

Reed, A. (1998). The oxygen cost of running the Physical Ability Requirement Evaluation. In J. Bonneau (Ed.). *The development of PARE: An evolution*. Ottawa, ON: RCMP Learning and Development Archives.

Sassen, B., Cornelissen, V. A., Kiers, H., Wittink, H., Kok, G., & Vanhees, L. (2009). Physical fitness matters more than physical activity in controlling cardiovascular disease risk factors. *European Journal of Cardiovascular Prevention and Rehabilitation*, *16*, 677-673. doi:10.1097/HJR.0b013e3283312e94

Sassen, B., Kok, G., Schaalma, H., Kiers H., & Vanhees, L. (2010). Cardiovascular risk profile: Cross-sectional analysis of motivational determinants, physical fitness and physical activity. *BMC Public Health, 10,* 592-602. Retrieved from http://www.biomedcentral/.com/1471/2458/10/592.

Seo, N. J., & Armstrong, T. J. (2009). Biomechanical analysis for handle stability during maximum push and pull exertions. *Ergonomics*, *52*(12), 1568-1575. doi: 10.1080/00140130903287999

Seo, N.J., Armstrong, T.J., and Young, G.J. (2010). Effects of handle orientation, gloves, handle friction and elbow posture on maximum horizontal pull and push forces. *Ergonomics*, *53*(1), 92-101. doi:10.1080/00140130903389035

Sharkey, B.J., & Davis, P.O. (2008). *Hard work: Defining physical work performances.* Windsor, ON: Human Kinetics.

Shell, D. E. (2003). Exercise science and law enforcement. *Strength and Conditioning Journal*, *25*(3), 52-57. Retrieved from Canadian Police College Library.

Shephard, R., & Bonneau, J. (2002). Assuring gender equity in recruitment standards for police officers. *Canadian Journal of Applied Physiology*, *27*(3), 263-295. Retrieved from http://journals.hil.unb.ca/index.php/CJAL/

Shephard, R., & Bonneau, J. (2003). Supervision of occupational fitness assessments. *Canadian Journal of Applied Physiology, 28*(2), 225-239. Retrieved from http://journals.hil.unb.ca/index.php/CJAL/

Stanish, H.I., Campagna, P.D., & MacKinnon, S.C. (1993). Comparison of the physical ability requirement evaluation to standard measures of physical fitness. *Canadian Journal of Applied Physiology, Nutrition, and Metabolism. (Abstract). Annual*

General MeetingProceedings 1993. Retrieved from http://www.nrcresearchpress.com/journal/apnm

Stanish, H.I., Wood, T.M., & Campagna, P. (1999). Predictors of performance on the RCMP physical abilities requirement evaluation. *Journal of Occupational and Environmental Medicine, 41*(8), 669-677. Retrieved from http://journals.lww.com/joem/Pages/default.aspx

Stanish, H. I. (1993). A comparison of the physical ability requirement evaluation simulated fight component to measures of muscular strength, endurance and power. Ottawa: ON. An unpublished partial paper in the RCMP Fitness & Lifestyle Program archives.

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. *Academic Emergency Medicine 18(10): 1099-1104*. Des Plaines, Ill. Society for Academic and Emergency Medicine. doi: 10.1111/j.1553-2712.2011.01185.x.

Strating, M., Bakker, R. H., Dijkstra, G. J., Lemmink, K. A. P.M., & Groothoff, J. W. (2010). A job related fitness test for the Dutch police. *Occupational Medicine, 60*, 255-260. doi:10.1093/occmed/kqq060

Sternfeld, B. & Goldman-Rosas, L. (2012). A systematic approach to selecting an appropriate measure of self-reported physical activity or sedentary behavior. *Journal of Physical Activity and Health*, *9*(Suppl.), S19-S28. Retrieved from http://journals.humankinetics.com/jpah

Thoden, H. C. (2002). *Testing for normality*. New York, NY: Marcel Decker Inc.

Thomas & Means. (2000). *Use of sex and age adjusted physical performance standards in the law enforcement profession*. PRIMA conference position paper. Retrieved from Ontario Police College Archives.

Thompson, W. R., Gordon, N. F., & Pescatello, L. S. (Eds.). (2009). *ACSM's: Guidelines for Exercise Testing and Prescription, (8 th Edition).* New York, NY: Lippincott Williams & Wilkins.

Trottier, A., & Brown, J. (1994). Occupational health in police work: A Canadian perspective. *Journal of Clinical Forensic Medicine*, *1*, 39-42. Retrieved from http://www.sciencedirect.com/science/journal/13531131

Trottier, A., & Brown, J. (1994). *Police health: A physician's guide for the assessment of police officers.* Ottawa, ON: Canada Communications Group.

Wagner-Wisotzki, N. (2005). *Fitness and Lifestyle Management for Law Enforcement*, (2nd Ed.). Toronto, ON: Edmund Montgomery Publishers.

Wells, G. D., Selvaduria, H., & Tein, I. (2009). Bioenergetics provision of energy for muscular activity. *Paediatric Respiratory Reviews*, *10*, 83-90. doi:10.1016/j.prrv.2009.04.005

Welk, G. J. (2002). *Physical activity assessment in health-related research*. Windsor, ON: Human Kinetics.

Appendices

Appendix A

The graphical layout of the PARE, PREP, COPAT, FITCO

Figure 3. The PARE flow and layout

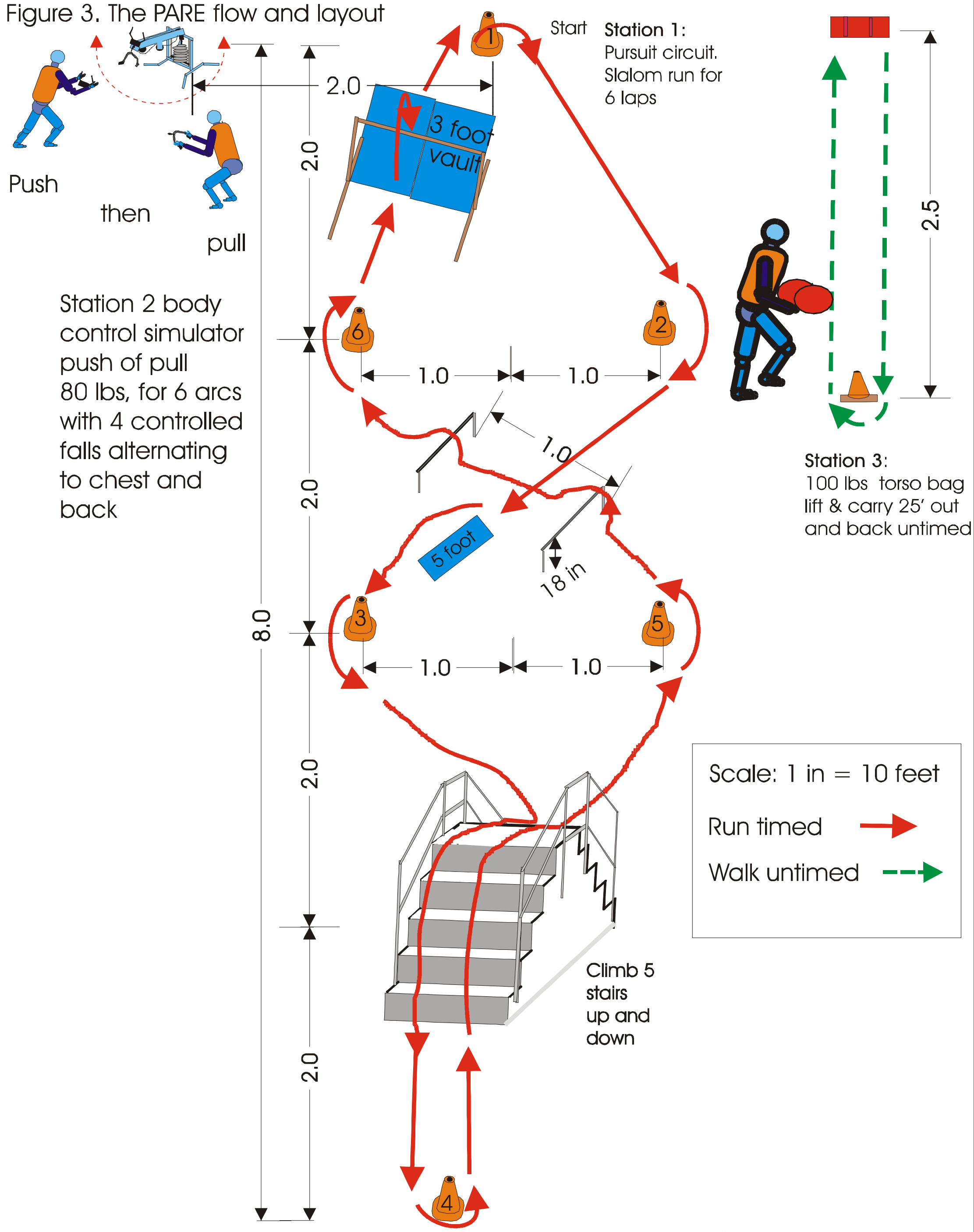


Figure 3. Pursuit circuit 6 laps running, with 5 ' jump, stair climb & descend, two 18 " hurdle jumps, 3 ' vault and controlled fall. Body control simulation, 6 arcs pushing & pulling 80 lbs with 4 controlled falls between. Rescue: 50 ' 100 lbs bag lift and carry.

Figure 4: Physical Readiness Evaluation for Police (PREP) Pursuit Restraint Circuit.

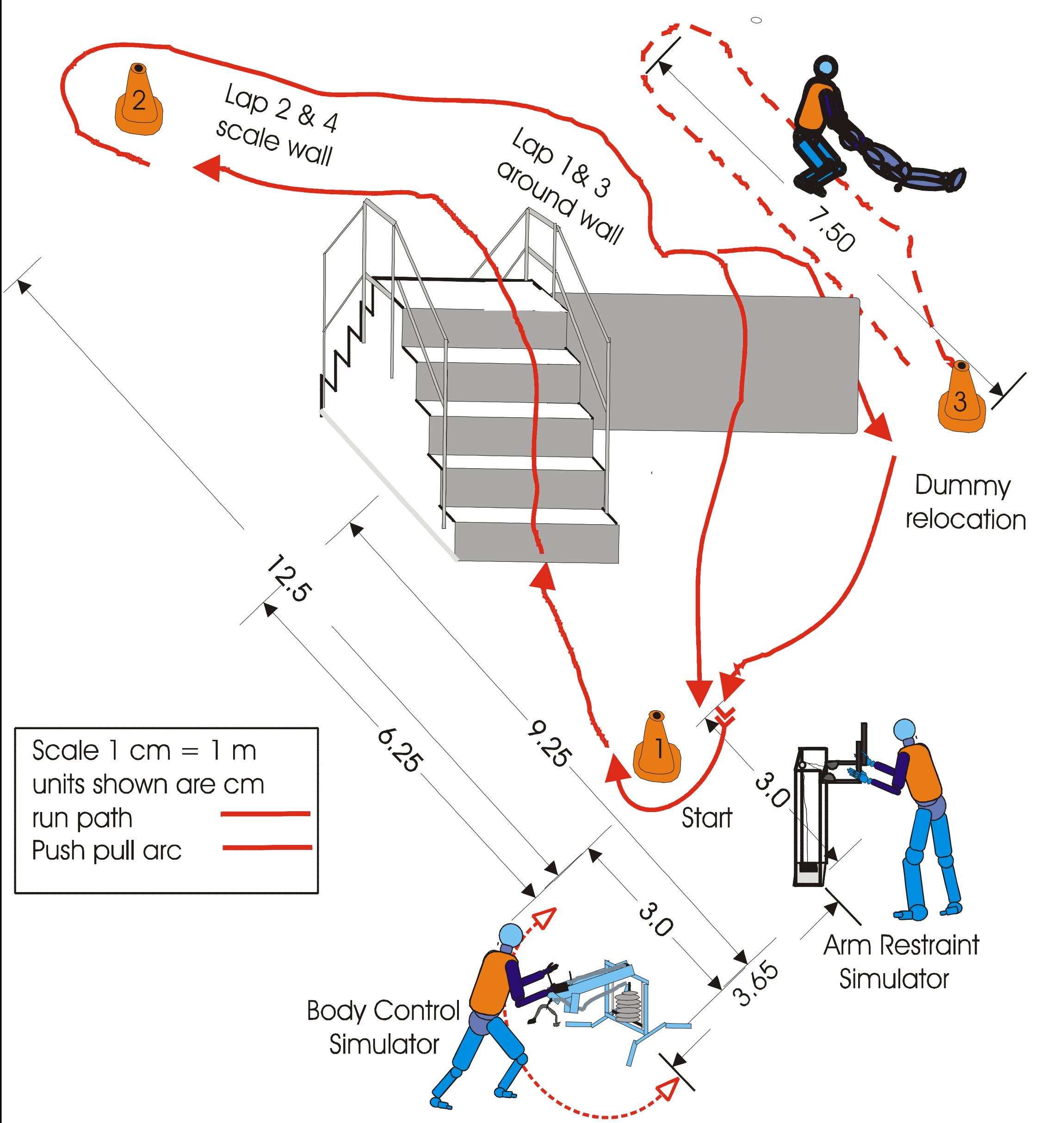


Figure 4. The PREP layout. Timed stations. Station 1, a 4 lap circuit run, laps 1 & 3 scale the 4 foot wall, lap 2 & 4 go around. Station 2, the body control simulation, lift 70 lbs by \ pushing, moving laterally through six 180 degree arcs. Station 3, the arm apprehension simulation, power grip the resistance handles with > the 34 lb resistance, bring arms and handles together (35 lb resistance) to touch, then in the same one motion return the handles apart. Return to station 2 complete 6 arcs pulling. Station 4, rescue simulation drag the 150 lb mannequin 7.5 m and back. Within 15 min of circuit testing complete stage 6.5 of the Ledger Boucher shuttle run. Redrawn from the *Technical Guide for PREP Users, 1998,* Ministry of the Solicitor General & Correctional Services.

Figure 5. COPAT layout and action.

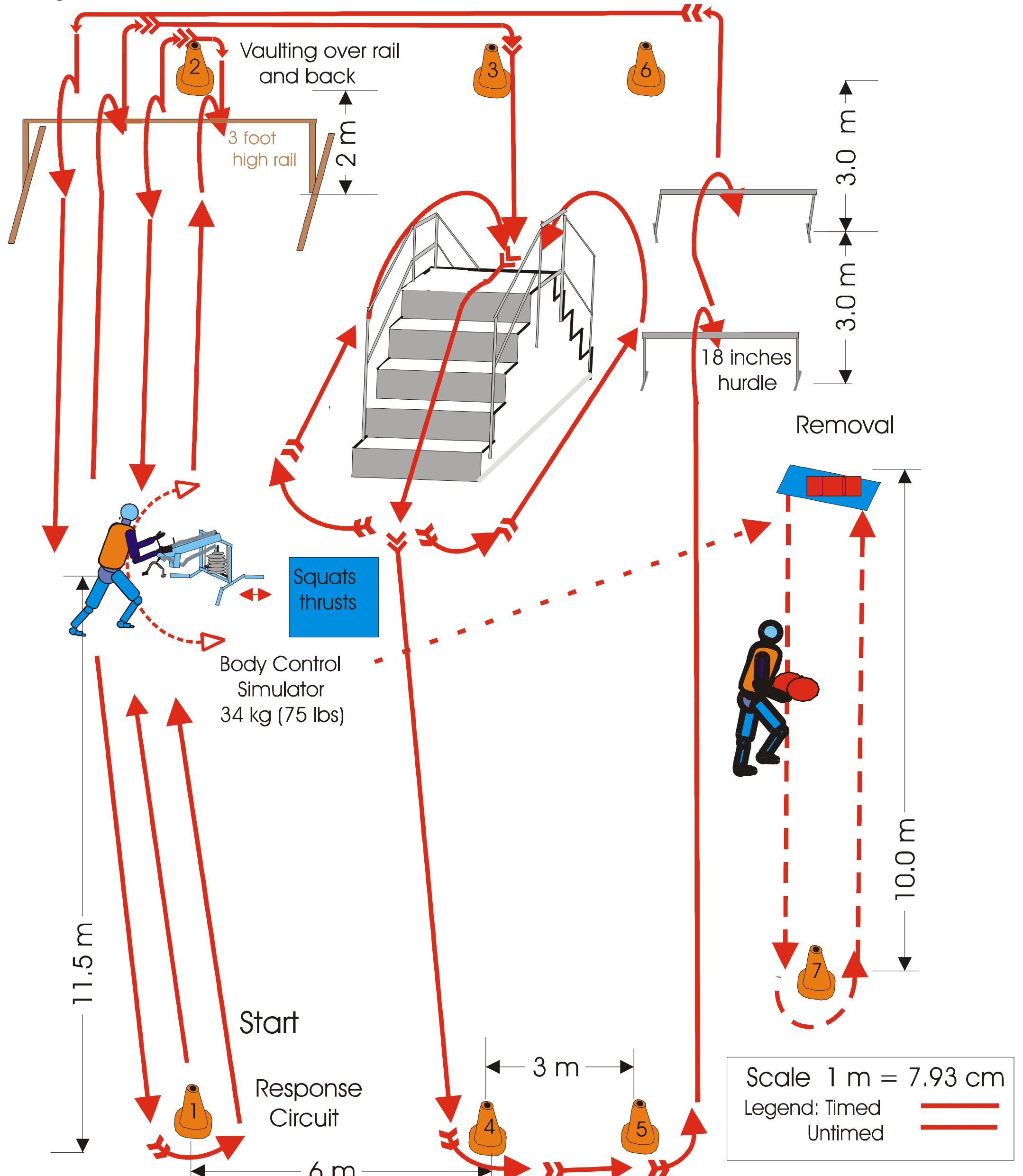
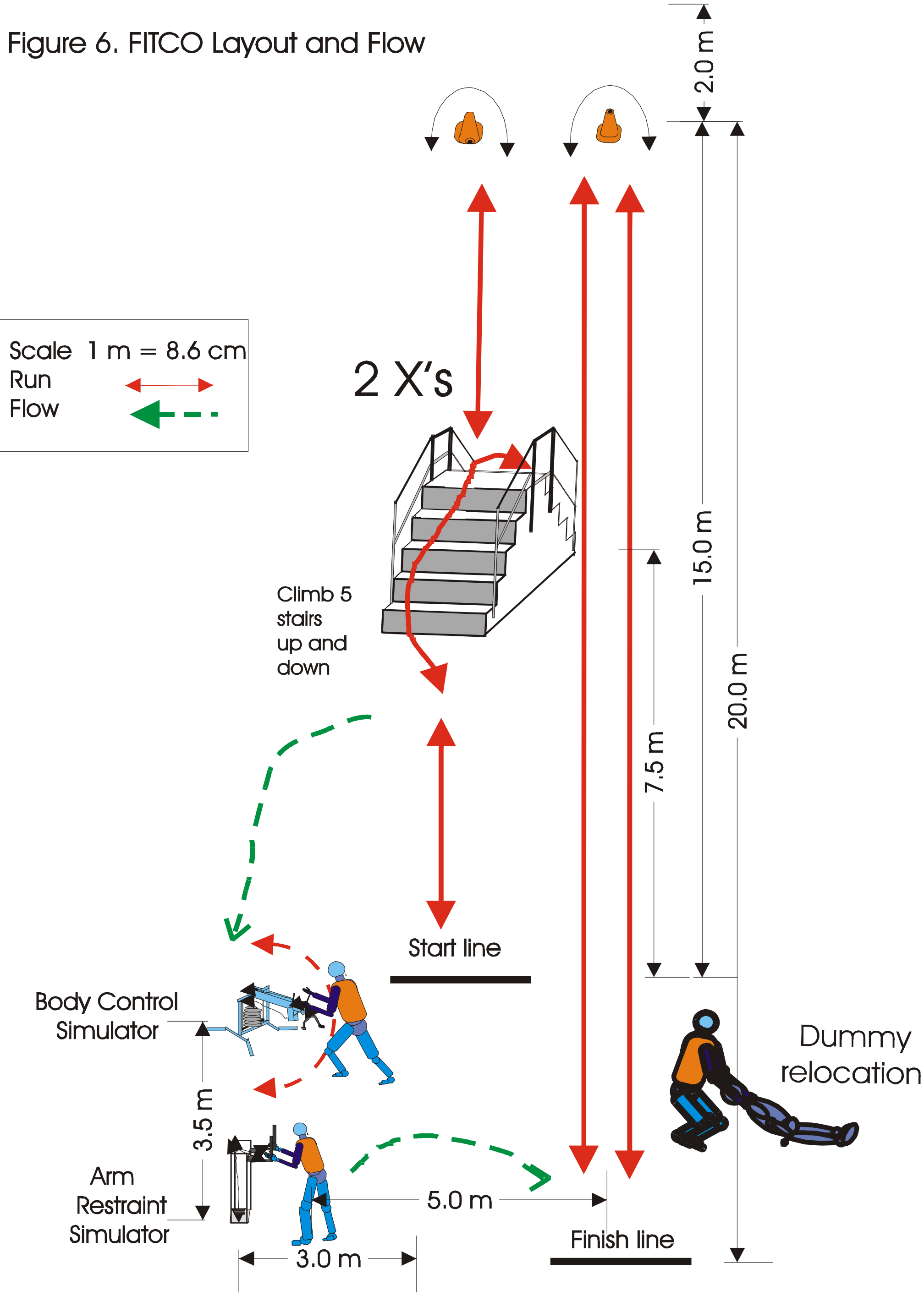


Figure 5. COPAT circuit: Starting at cone 1, move past door simulation, then vault the rail. Repeat 3 Xs. On to cone 3, then climb 5 stairs, 3 Xs. To cone 4,5 & 6 jumping 2 hurdles. Return to Body Control Simulator. Lifting 75 lbs, complete six 180 ° arcs. Then 6 squat thrusts, 6 arcs pulling, 4 squat thrusts. Rescue: Cone 7, lift and carry 34 kg (75 lbs) 10 m out & back, untimed. Redrawn from Wagner-Wisotzki, (2003)

Figure 6. FITCO Layout and Flow



Appendix B

Media graphical representative of the PARE

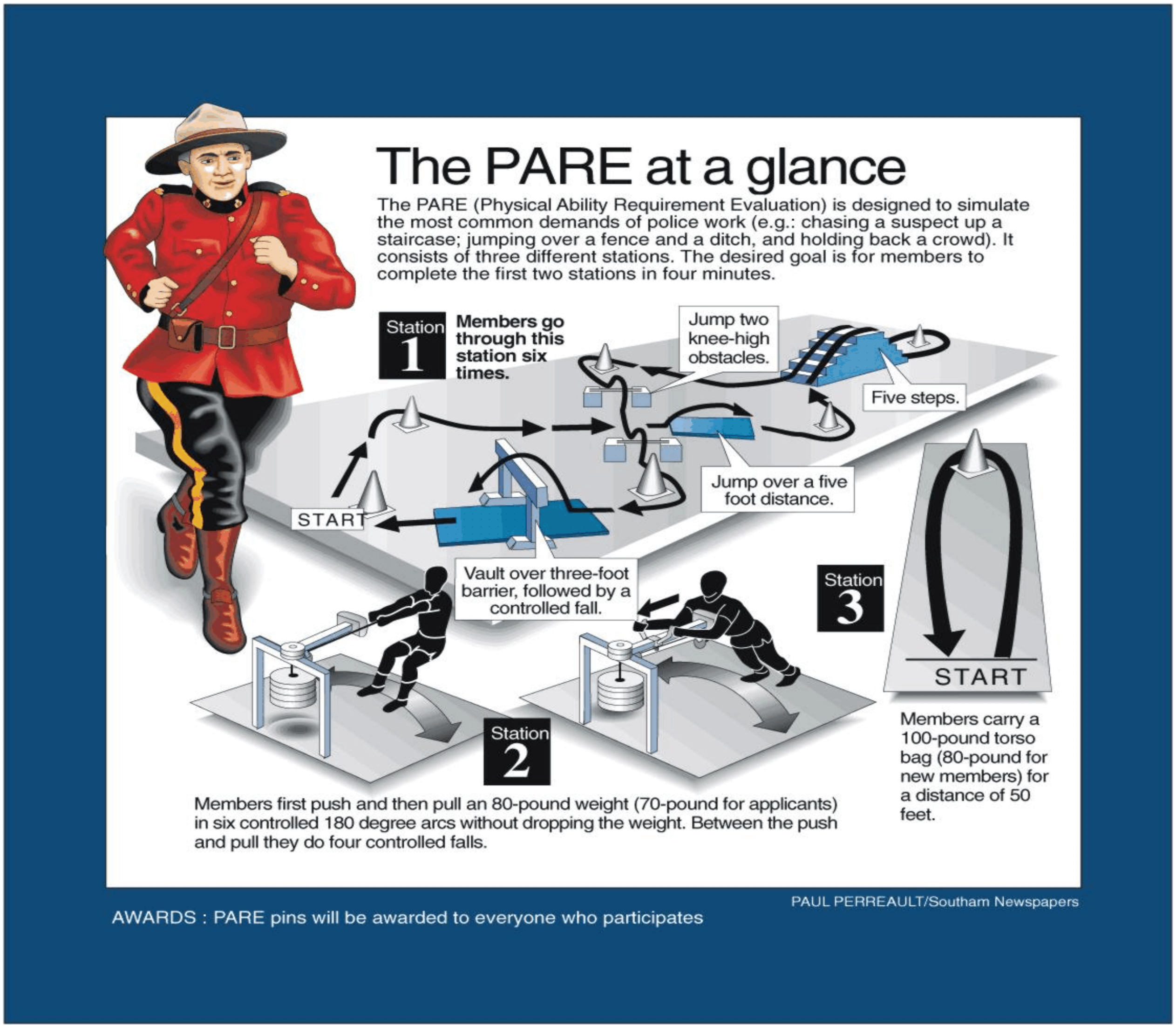


Figure 2. The Graphical representation of the PARE used in promotion activities showing the 3 stations. Stations 1, the pursuit course slalom run with 5 step stair climb, two 18 in jumps, a 5 foot running broad jump, a 3 foot rail for vaulting per lap of 6 laps. Pacing each lap to 25 sec is a crucial skill to avoid premature fatigue. Station 2, the body control simulator, requires 6 arcs of lifting a 70 / 80 lbs net resistive stack and carriage and moving laterally left and right through 180 ^{o}, first pushing, followed by 4 controlled falls to the ground completed alternately to the chest and back, then pulling. The timed portion of the PARE stops at station 2. An untimed 80 / 100 lbs bag lift and carry over 50 feet, has no time limit, with 3 unsuccessful attempts as a failure. Station 1 has a pass time standard of 4:45 for recruits entry / retention in police academy training at 70 lb push, 80 lb lift/carry, and graduation is 4:00 push, 80 lb push pull and 100 lb lift and carry

Appendix C

PARE Medical Clearance Form

Go to Part 2

Print Part 1

Royal Canadian Mounted Police Gendarmerie royale du Canada

Protected B when completed

PIB CMP/P-PE-808
PIB CMP PPU 070

PARE Medical Clearance - Part 1

Det. / Unit - For Member Use Only
HRMIS / Applicant ID

Patient Information

Surname	Given Names	Gender	Date of Birth (yyyy-mm-dd)
		M F	

Height (cm)	Weight (kg)	Resting Heart Rate	Resting Blood Pressure

Risk Factors

Note to Physician

The Physical Abilities Requirement Evaluation (PARE) is a maximal physical exertion test equivalent to an Exercise Stress Test at the 12 Metabolic Equivalent for Task (MET) level. The following are risk factors to consider when assessing suitability for PARE.

Section A - For All Individuals - Pulmonary And Musculoskeletal Restrictions

If yes to **any one** risk factors in Section A, patient should not undertake PARE.

Pulmonary obstruction / restriction that would prevent maximal testing.	Yes	No
Needs to use a short acting inhaler immediately prior to participate in maximal testing. (Short acting inhalers can only be used after the test if needed. Long acting or combined inhalers are allowed.)	Yes	No
Musculoskeletal restrictions that could interfere with strenuous activities or maximal testing.	Yes	No

Section B - For All Individuals - High Or Very High Cardiovascular Risk Factors (ATP III [1] & CMAJ [2])

If yes to **one or more** risk factors in Section B, it is recommended to send patient to an Exercise Stress Test before clearing for PARE.

Previous CVA, MI, vascular surgery or any clinical evidence of atherosclerosis	Yes	No
Diabetes [3]	Yes	No
Metabolic Syndrome	Yes	No

Section C - For Men > Age 40 And Women > Age 50 - Coronary Artery Disease Risk Factors (ACSM [4] & CSEP [5])

If yes to **two or more** risk factors in Section C, it is recommended to send patient for an Exercise Stress Test before clearing for PARE.

Family history of premature cardiovascular disease	Yes	No
Cigarette smoking	Yes	No
Hypertension [6]	Yes	No
Dyslipidemia	Yes	No
Abnormal fasting glucose level	Yes	No
Obesity [7]	Yes	No
Physical inactivity	Yes	No

Section D - Exercise Stress Test (when Required)

Clinically positive for ischemia	Yes	No
Electrically positive for ischemia	Yes	No

Number of MET reached (12 MET are required prior to undertaking PARE)

Additional tests (if needed, specify):

Medical References

1) Detection, Evaluation and Treatment of High Blood Cholesterol in Adults. (Adult Treatment Panel III). National Institute of Health. National Heart, Lung and Blood Institute.
2) Recommendations for the management of dyslipidemia and the prevention of cardiovascular disease: 2003 update. CMAJ appendix 2003; 168 (9) 921-924.
3) Report of the Expert committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2003; 26 (supp 1):S5-S20.
4) American College of Sports Medicine, Cited in ACSM Guidelines for Exercise Testing and Prescription, Seventh Edition. 2006.
5) Canadian Society of Exercise Physiology. Professional Fitness & Lifestyle Consultant. Resource Manual. 2004.
6) Canadian recommendations for the management of hypertension (2005)
7) Canadian Guidelines for Body Weight Classification in Adults (2003)

Return to Part 1

Print Part 2

Protected B
when completed

PIB CMP/P-PE-808
PIB CMP PPU 070

Royal Canadian Mounted Police Gendarmerie royale du Canada

PARE Medical Clearance - Part 2

Det. / Unit - For Member Use Only
HRMIS / Applicant ID

To be presented to the PARE Administrator

Patient Information

Surname		Given Names	Gender	Date of Birth (yyyy-mm-dd)
			M F	
Height (cm)	Weight (kg)	Resting Heart Rate	Resting Blood Pressure	

Physician's Recommendations

After reviewing Part 1 of the PARE Medical Clearance and evaluating the following risk factors:

- Pulmonary Obstruction / Restriction
- Musculoskeletal Restrictions
- High or Very High Cardiovascular Risk Factors
- Coronary Artery Disease Risk Factors
- Exercise Stress Test to 12 MET, if applicable

It is my professional opinion that the above named patient is:

- medically fit to undertake the Physical Abilities Requirement Evaluation.
- not medically fit to undertake the Physical Abilities Requirement Evaluation.

Comments

Physician's signature	Date (yyyy-mm-dd)	Physician's stamp

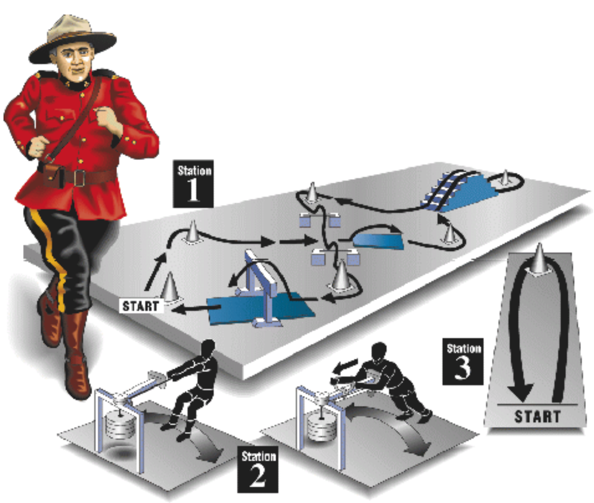
Appendix D

June 2014 Physical Abilities Requirement Evaluation Protocol

The Royal Canadian Mounted Police

General Duty

Physical Ability Requirement Evaluation (PARE)



General Duty PARE Protocol

Updated June 2014

PARE Description and Standards

PARE is divided into three sections which represent a situation where a police officer must engage in a foot chase (obstacle course), physically control a suspect (push/pull activity) and carry a person or an object away from the scene (weight carry).

1. **Obstacle Course:** involves completing approximately a ¼ mile (400 m) run which includes 6 mat jumps, 120 stairs, 12 hurdles, 6 vaults and 3 each of back/front falls.

2. **Push/pull Activity:** involves controlling a 70 or 80 lbs (32 or 36 kg) weight and completing six 180□ arcs while pushing, 4 controlled falls and six 180□ arcs while pulling.

3. **Weight Carry:** involves lifting and carrying an 80 or 100 lbs (36 or 45.5 kg) bag over a distance of 50 ft (15 m).

PARE Standards and Requirements

The first two sections, obstacle course and push/pull activity, are timed. The third section, the weight carry, is a pass or fail activity and is not timed. The participant is allowed a maximum of three trials to complete the task properly. Failure to complete this section constitutes a failure of the entire PARE test.

General Duty PARE as a Requirement

- Applicant standard is PARE in ≤4:45 (70 lbs (32 kg) push/pull), 80 lb. (36 kg) bag
- If required the Auxiliary constable standard is PARE in ≤4:45 (70 lbs (32 kg) push/pull), 80 lbs (36 kg) bag
- Lateral Applicant (a Police Officer from another agency) standard is PARE in ≤4:00, 100 lbs (45.5 kg) bag
- Reservist must complete PARE in ≤4:00 (80 lbs (36 kg) push/pull), 100 lbs (45.5 kg) bag
- An RM should strive to complete regular PARE within ≤4:00 (80 lbs (36 kg) push/pull), 100 lbs (45.5 kg) bag

GD PARE Standard for Specialized Units

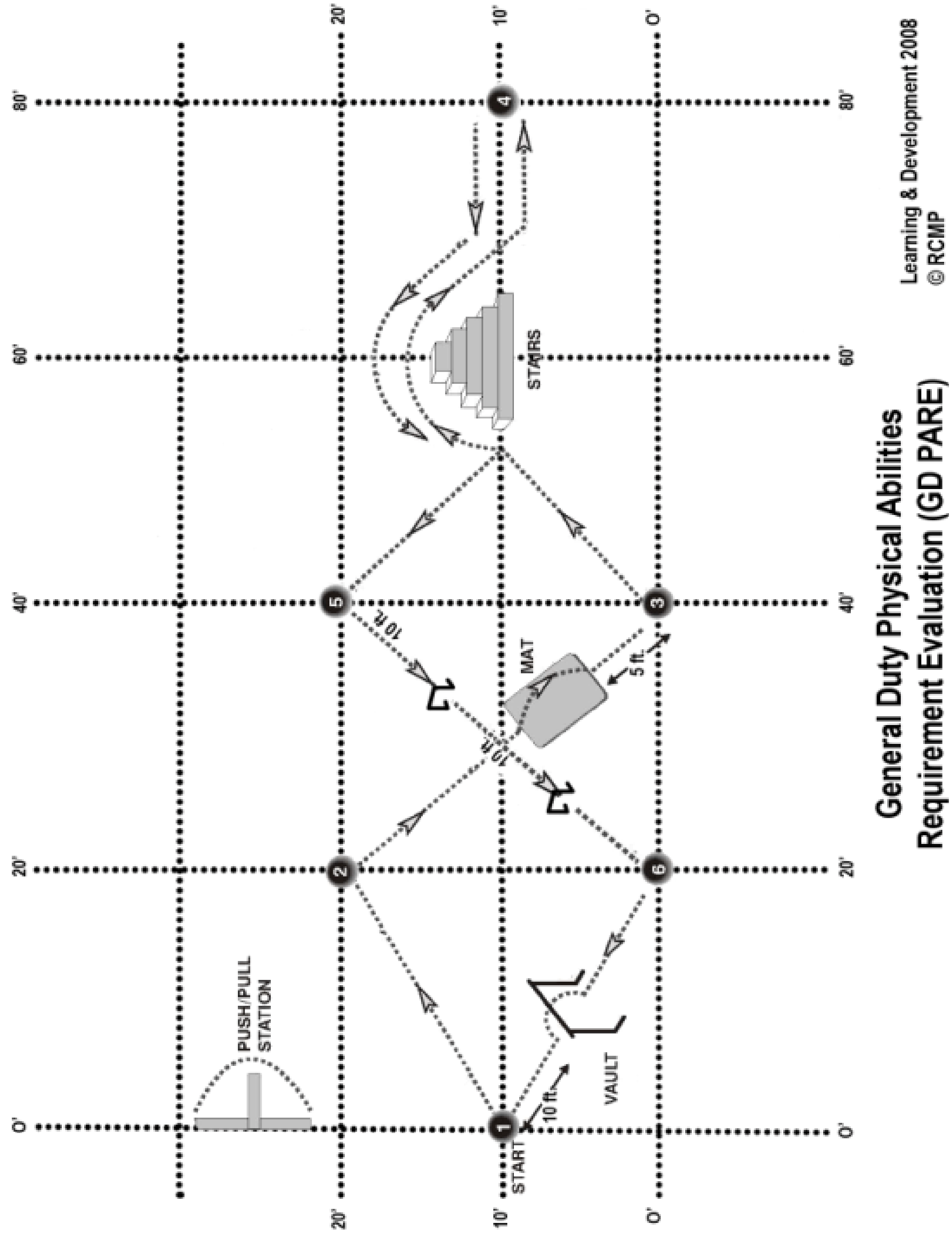
The regular PARE in ≤4:00 is a standard for consideration of, and continuation in, the following specialized units. RMs in these units must meet the regular PARE ≤4:00 standard each calendar year:

Protective Policing (bodyguard duties):

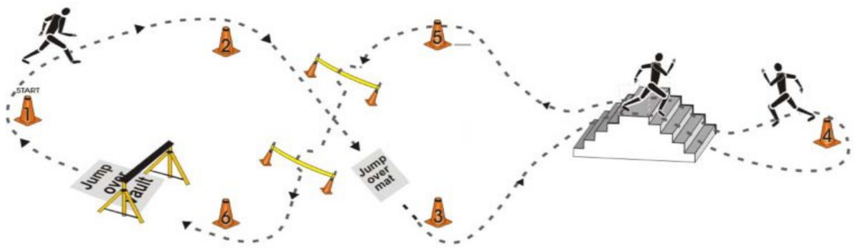
- Tactical Troops;
- Aircraft Protection;

- Underwater Recovery;
- Musical Ride;
- Explosive Disposal Unit;
- Crime Scene Methods and Procedures Analyst;

PARE Obstacle Course Graphic



1. **Obstacle Course:**



The first section of PARE consists of a six lap obstacle course measuring 1116 ft (340 m) resulting in approximately a ¼ mile (400 m) run.

This activity typically takes 2:30 – 3:00 (averaging 25 – 30 sec/lap). The course is laid out in the following manner:

a. From the start marker #1, the participant runs left towards marker #2.

b. Going around the left side of marker #2, the participant turns right toward marker #3. Before reaching marker #3, the participant must jump over and clear a 5 ft (1.5 m) mat. The far edge of the mat is placed 5 ft (1.5 m) from marker #3. After landing, the participant turns left around marker #3 and proceeds towards the stairs.

c. The stairs are placed in such a manner that the centre of the top platform is exactly 60 ft (18 m) from marker #1, and 20 ft (6 m) from marker #4. It must be directly in line with the centre of the course. The participant must run up and down the stairs towards marker #4 touching at least one step on the way up, the top platform and at least one step on the way down. They may touch as many steps as they like. Participants may use the handrails as a guide up and down the stairs.

d. Marker #4 is set on the centre line exactly 80 ft (24 m) from marker #1. The participant runs around this marker, from either the right or the left side, and runs back over the stairs then turns right towards marker #5.

e. The participant runs around the right side of marker #5, then turns left heading towards marker

#6. Before reaching marker #6, the participant must leap over the 2 hurdles raised 18 in. (45 cm) from the floor. The first hurdle is 10 ft. (3 m) from marker #5. The second hurdle is located 10 ft. (3 m) from the first hurdle. The hurdles are lined up parallel with each other between markers 5 and #6.

f. Reaching marker #6, the participant runs around the left side of the marker and turns right, heading towards marker #1. Before reaching marker #1, the participant must get over a 3 ft (0.9

m) high vault situated approximately halfway between markers #6 and #1. The participant must land in control, on both feet, on the opposite side of the vault and then perform a controlled fall on their front or their back, alternating each lap. The front fall requires that the participant's chest, abdomen and hips be in contact with the floor simultaneously. During the back fall, the participant must touch both shoulder blades simultaneously on the floor. The participant must get up and proceed around marker #1 to complete the lap.

Six laps must be completed before proceeding to the push/pull unit which must be located within 20 ft (6 m) of marker #1.The following faults and assigned penalties may occur during the Obstacle Course:

a. **Markers:** A participant failing to go around the outside of a marker must come back and go around the marker.

b. **Mat:** A participant not clearing the full 5 ft (1.5 m) length of the mat will be penalized 5 seconds. A participant can only incur one mat penalty per lap.

c. **Stairs:** For safety reasons, the participant is instructed to run up and down the stairs, stepping on at least one step on the way up, the top platform, and at least one step on the way down. Failure to follow instructions may lead to the termination of the test.

d. **Hurdles:** A participant knocking a stick off the hurdle will be penalized 2 seconds per hurdle. If the first stick knocks off the second stick, the participant will be penalized 4 seconds. The PARE Administrator will replace the stick(s).

e. **Hurdles:** A participant jumping over a hurdle with either leg outside the cone, even if a stick is not knocked off, will be penalized 2 seconds.

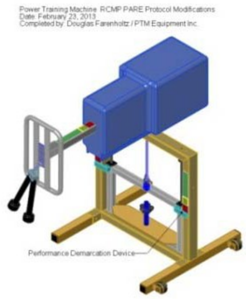
f. **Vault:** If the participant cannot get over the vault, the PARE is terminated. The manner in which the participant gets over the vault is not specified. They may touch the vault with any part of their body (i.e., foot up, straddle, etc.) but must remain in control at all times.

g. **Vault Landing:** A participant landing in an uncontrolled manner after going over the vault is required to go back around and over the vault again, landing in control on both feet.

h. **Controlled falls:** A participant using the mat or vault to raise or lower themselves must repeat the controlled fall.

i. **Controlled falls:** A participant failing to touch their chest, abdomen and hips on the floor simultaneously during the front fall, or their shoulder blades to the floor simultaneously during the back fall, must repeat the controlled fall.

2. **Push/pull Activity**



Upon finishing the obstacle course, the participant moves immediately to the push/pull unit which provides a resistance of 70 lbs (32 kg) for applicants or 80 lbs (36 kg) for RMs. The push/pull unit must be located within 20 ft (6 m) of the obstacle course. A floor mat measuring approximately 8'x 24' (2.5 m x 7 m) may be placed under the push/pull unit for traction if the floor surface is slippery.

The participant may perform this activity in the order they choose, push first and then pull, or vice versa. Proper body mechanics must be maintained throughout the push/pull activity in order to demonstrate adequate muscular ability.

The PARE administrator stops the clock at the end of the 6th arc once the participant's body and the lever arm are directly in line with the frame of the push/pull unit, prior to the participant lowering the plate carriage. The participant is encouraged to let the weight down slowly as this will prevent damage to the machine and will not affect the overall time.

The push/pull activity typically takes 55 – 80 seconds.

Push Activity

Upon reaching the push/pull unit, the participant grasps any portion of the handle and pushes the plate carriage off the base of the unit and then proceeds to complete six controlled 180 degree arcs while keeping the weight suspended. In order to complete an arc, the participant's body and the lever arm of the push/pull unit must be directly in line with the frame of the push/pull unit at the start and end of each arc.

During the push the participant must demonstrate muscular control throughout the arcs by keeping both elbows bent at all times and refraining from bracing their elbows against their body. Placing the chest or shoulders on the handle is not accepted, however, incidental contact may occur which will not provide a mechanical advantage.

For coaching purposes, participants can be encouraged to grip the vertical bars at optimal push height (just above their hip bones while standing upright). With both feet back and the core stabilized, they can then use the legs as primary movers. The core should be fully activated and the chest kept low in a full forward leaning posture to avoid hyper extension or arching of the back while pushing. The participant pushes the handles through slightly bent, stabilized arms in a controlled, motion, until the weight carriage lifts off the base of the unit. S/he then proceeds to complete six controlled, 180° arcs while keeping the weight suspended off the frame while moving laterally.

This activity typically takes 20 - 30 seconds. Controlled falls

After six arcs are completed, the weight must be lowered with control and the participant must complete four controlled falls; two to the front and two to the back, in alternating sequence. During the front falls, the participant must ensure their chest, abdomen and hips touch the floor simultaneously. During the back falls, both shoulder blades must touch the floor simultaneously. After each fall, the participant must come to a standing "ready" position, and tap the handle with both hands. Participants are not allowed to use the wall or any part of the push/pull unit to raise or lower themselves.

This activity typically takes 15 - 20 seconds.

Pull Activity

Once the sequence of controlled falls is complete, the participant grasps any portion of the rope with both hands and pulls, lifting the plate carriage off the base of the push/pull unit. The participant must then complete six controlled 180 degree arcs while keeping the weight suspended. In order to complete an arc, the participant's body and the lever arm of the push/pull unit must be directly in line with the frame of the push/pull unit at the start and end of each arc.

During the pull, participants must demonstrate muscular control by facing the general direction of the machine and maintaining an observable bend in their elbows, hips and knees at all times.

This activity typically takes 20 - 30 seconds.

The following faults and assigned penalties may occur during the Push/pull Activity:

Arcs

If any of the following faults occur, the participant must perform an additional arc, or completed the activity properly.

Dropping the weight: This occurs when a participant fails to maintain the plate carriage off the base of the push/pull unit during a controlled arc. Testers may use the coloured tape on the push/pull unit to cue the participant on the position of the plate carriage. The carriage does not have to be up all the way, but it must not touch the base during the arc, if it touches, the participant must redo that arc. Dropping the weight at either end of the arc is acceptable.

Incomplete arc: This occurs when a participant fails to complete an entire 180 degree arc. The participant's body and the lever arm of the push/pull unit must be directly in line with the frame of the machine at the start and end of each arc.

Pushing off the wall: This occurs when a participant intentionally hits the wall or pushes off the wall with any part of their body. Participants will be allowed one warning. Any subsequent infraction will require another arc.

Locking the elbows: This occurs when a participant fails to maintain a bend in the elbows during the push activity. Participants will be allowed one warning then must redo any arc that was completed without an observable bend in the elbows.

Bracing the elbows during the push activity: This occurs when a participant braces their elbows against the body. Participants will be allowed one warning then must redo any arc that was completed while bracing their elbows.

Improper pull position: This occurs during the pull activity when a participant turns away from the machine or fails to keep an observable bend in their elbows, hips and knees while completing an arc. Participants will be allowed one warning then must redo any arc that was completed with improper body position.

Controlled Falls

If any of the following faults occur, the participant must perform an additional arc, or completed the activity properly.

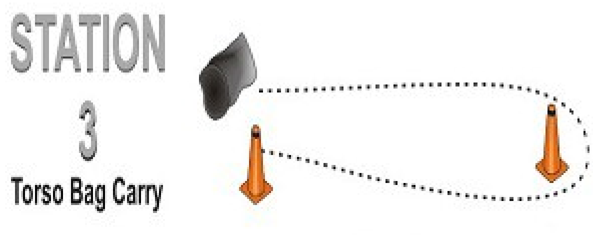
Improper controlled fall technique: This occurs when a participant fails to touch their chest, abdomen and hips simultaneously on the floor during the front fall, or fails to touch both shoulder blades simultaneously on the floor during the back fall. The participant must finish the task correctly before moving on. The participant will be allowed one warning before continuing. Subsequent infractions will require the participant to repeat the controlled fall.

Improper positioning: This occurs when a participant fails to come to a controlled, standing "ready" position after each fall. The participant must finish the task properly before continuing. The participant will be allowed one warning before continuing. Subsequent infractions will require the participant to come to a controlled, standing "ready" position.

Not tapping the handle: This occurs when a participant fails to tap the handle with <u>both</u> hands after rising from each fall. The participant must redo the tap prior to continuing. The participant will be allowed one warning before continuing. Subsequent infractions will require the participant to tap the handle with both hands.

Using assistance to get up or down: This occurs when participants use any part of the push/pull unit, or the wall, to raise or lower themselves during the controlled falls. The participant will be allowed one warning before continuing. Subsequent infractions will require the participant to get up without aid.

3. **Weight Carry**



Within two minutes of completing the push/pull activity, the participant must begin the weight carry. The participant must be able to pick up a bag, weighing 80 lbs (36 kg) for applicants or 100 lbs (45.5 kg) for RMs. The participant is allowed to wrap their arms around the bag or grasp the excess material to lift the bag. Participants are not allowed to use the Velcro straps to grip the bag. The bag must be carried in a controlled manner in front of the participant, NOT over their shoulders and NOT resting on their thighs while walking. To avoid injuries, participants must use proper lifting technique which includes bending the knees, keeping the back straight and lifting with the legs. The participant must walk forward 25 ft (7.5 m) and turn around the marker and return to the start. When they reach the start they must lower the bag in a controlled manner using proper lowering technique, using the legs, not the back, and set it down gently.

A participant will have a maximum of three attempts to perform the task completely. Three unsuccessful trials will result in a participant failing the entire PARE test. This section is not timed; it is a pass or fail activity.

The following faults and assigned penalties may occur during the Weight Carry:

Unable to lift the bag: A participant unable to lift the bag from the floor fails PARE.

Using the Velcro straps: A participant picking up the bag using the Velcro straps must repeat the lift without using the straps.

Dropping the bag: A participant picking up the bag but dropping it before completing the 50 ft (15 m) must repeat the task.

Improper carry: A participant carrying the bag over their shoulder or walking with the bag resting on their thighs must repeat the task.

Lowering the bag: A participant failing to lower the bag in a controlled manner as described above must repeat the task.

Curriculum Vitae
Robert A. Séguin

EDUCATION

BACHELOR IN PHYSICAL EDUCATION 1983-87
University of Ontario *Ottawa, Ontario*

BACHELOR IN EDUCATION 1987-88
University of Ontario *Ottawa, Ontario*

MASTER IN SCIENCE KINANTHROPOLOGY (KINESIOLOGY) 1989-1992
University of Ottawa *Ottawa, Ontario*
Master thesis: *A Comparison of the Effects of Three Selected Exercise Intensities on Lactate Recovery Ability of Elite Distance Runners*.

PHD HEALTH PROMOTION 2009-2015
University of Western Ontario *London, Ontario*
PhD dissertation: Factors Effecting Success in PARE Among RCMP Officers

RELATED EMPLOYMENT

DIVISION FITNESS & LIFESTYLE ADVISOR 1997-PRESENT
Royal Canadian Mounted Police *London, Ontario*

INSTRUCTOR HEALTH PROMOTION 2005 TO PRESENT
Fanshawe College *London, Ontario*

BFOR **CONSULTANT FOR BRUCE POWER EMERGENCY SERVICES** 2002
Bruce Power Nuclear Facility *Port Elgin, Ontario*

PERSONAL TRAINING- REHABILITATION KINESIOLOGIST 1996-1997
Ottawa Athletic Club *Ottawa, Ontario*

ABILITIES ASSESSOR 1993
Performance Plus Rehabilitative Care *Ottawa, Ontario*

STAFF SCIENTIST 1992-1993
Science North *Sudbury, Ontario*

RESEARCH / TEACHING ASSISTANT 1989-1992
University of Ottawa *Ottawa, Ontario*