Date: 20060707

Files: 166-02-31991
166-02-31992

Citation: 2006 PSLRB 85

Public Service
Staff Relations Act

Before an adjudicator

BETWEEN

TERRANCE JAMES BARR AND
SHERRY ELIZABETH FLANNERY

Grievors

and

TREASURY BOARD
(Department of National Defence)

Employer

Indexed as
Barr and Flannery v. Treasury Board (Department of National Defence)

In the matter of grievances referred to adjudication pursuant to section 92 of the *Public Service Staff Relations Act*

REASONS FOR DECISION

Before: Guy Giguère, adjudicator

For the Grievors: Andrew Raven and Carolyn LeCheminant-Chandy, counsel

For the Employer: John Jaworski and Karl Chemsi, counsel

Heard at Ottawa, Ontario,
April 19 and 20, May 11 to 14 and 25 to 26, October 25 to 29,
November 22 to 25 and December 6 and 7, 2004, and April 25 to 27,
May 17 to 19 and September 29 and 30, 2005.

I. Grievances referred to adjudication

[1] The grievors, Terrance James Barr and Sherry Elizabeth Flannery, are firefighters with the Department of National Defence (DND). The DND requires its firefighters, as a condition for continued employment, to complete a fitness test within eight minutes. The grievors contend that a failure to meet this eight-minute standard can result in "career action", including recorded counselling, demotion and termination.

[2] The grievors allege that the eight-minute standard is discriminatory on the basis of age and gender. They submit that the DND has failed to provide any evidence that completing the fitness test within the eight-minute standard is necessary for the safe and efficient performance of their firefighting duties. The grievors are seeking an order that the eight-minute standard is discriminatory and that the DND be precluded from using this standard as a condition of employment.

II. Procedural background

[3] These grievances were filed in March and April 2002. Both grievors also filed complaints with the Canadian Human Rights Commission (CHRC) in late April 2002 pertaining to the same situation.

[4] In a memorandum of agreement dated April 18, 2002, the DND and the Union of National Defence Employees (UNDE), a component of the Public Service Alliance of Canada (PSAC), agreed that the present grievances and all similar grievances would be referred directly to the third level of the grievance process. The grievances would then be placed in abeyance until the CHRC decided to deal with the grievors' complaints or decline for the moment and refer the matter to the grievance process. If the CHRC referred the matter to the grievance process, the grievances of Mr. Barr and Ms. Flannery would be addressed at the third level. Other grievances would be held in abeyance pending a final decision on these two grievances.

[5] On December 19, 2002, the CHRC informed the grievors that it had decided not to deal with the complaints at that time pursuant to paragraph 41(1)(*b*) of the *Canadian Human Rights* Act, R.S.C., 1985, c. H-6 (the "*CHRA*") because the matter could more appropriately be dealt with through the grievance process and adjudication under the *Public Service Staff Relations Act*, R.S.C., 1985, ch. P-35 (the former "*Act*").

[6] On January 21, 2003, Rick Sullivan, Director General Employee Relations at the DND, informed the grievors that their grievances were denied at the third and final level of the grievance process. He explained that the firefighters' fitness test had been developed to reflect actual job demands and to ensure that firefighters are physically capable of carrying out their duties safely, efficiently and reliably. According to Mr. Sullivan, this was important not only for the health of the individual firefighters but also to their colleagues who fight fires alongside them and the clientele to whom they provide this service. He noted that as of December 31, 2003, firefighters who failed to successfully complete the test within the eight-minute standard would be given help to achieve that standard. However, if they were still unable to meet the eight-minute standard, every effort would be made to identify suitable alternative employment for those employees. He concluded that the eight-minute standard was not discriminatory and otherwise constitutes a *bona fide* occupational requirement (BFOR).

[7] On March 27, 2003, the grievances were referred to adjudication. On September 3, 2003, the grievors wrote to the Public Service Staff Relations Board (the Board) to request a pre-hearing conference to establish timelines for the exchange of expert witness reports. On November 26, 2003, the Chairperson of the Board met with the representatives of the parties to address the exchange of expert witness reports as well as setting dates for the hearing of these grievances.

[8] On April 1, 2005, the *Public Service Labour Relations Act*, enacted by section 2 of the *Public Service Modernization Act*, S.C. 2003, c. 22, was proclaimed in force. Pursuant to section 61 of the *Public Service Modernization Act*, these references to adjudication must be dealt with in accordance with the provisions of the former *Act*.

III. Summary of the evidence

A. Development of the eight-minute standard

[9] It is first important to explain that there are two groups of firefighters, one military and the other civilian. The Canadian Forces (CF) firefighters are military personnel providing firefighting and emergency services on CF bases outside of Canada or on CF vessels outside of Canadian harbours. DND firefighters are civilian employees providing firefighting services on CF bases in Canada, on CF vessels while in Canadian harbours and aircrafts on Canadian soil. They also provide these services to foreign military vessels in Canadian harbours. There are over 800 firefighters, half

are CF personnel and the other half are DND employees. DND firefighters are represented by the PSAC.

[10] The firefighters' fitness test used for CF and DND firefighters from the middle of the 1970s to the early 1990s was derived from what is known as the Cooper test. This Cooper-derived test consisted of running 1.5 miles, doing push-ups, sit-ups, pull-ups and carrying a 57-kilogram weight for a distance of 30 metres (Exhibit 58, tab 2, p. 3). There were different standards established depending on age and gender, as can be seen by the following table:

Annex A – Fitness Standard

Test Sequence	Age	Activity	Standard		
			Male	Female	
1	Under 30	Run -2400 M	10:45	13:30	
	30-34	Run -2400 M	11:15	14:00	
	35-39	Run -2400 M	11:45	14:30	
	40-44	Run -2400 M	12:15	15:00	
	45-49	Run -2400 M	12:45	15:30	
	50-54	Run -2400 M	13:15	16:00	
	55 and over	Run -2400 M	13:45	16:30	
		Walk-3200 M	28:00	34:15	

Push-Ups						
2	Age	17 to 19	Age 20 to 29	Age 30 to 39	Age 40 to 49	Age 50 and over
	Standard	27	23	19	16	13

Sit-Ups						
3	Age	17 to 19	Age 20 to 29	Age 30 to 39	Age 40 to 49	Age 50 and over
	Standard	39	35	29	25	20

4	**Pull-Ups or Chin-Ups**	5 consecutively
5		
6		

[11] In 1980, the CF Surgeon General cancelled the Cooper-derived test for the general military population because there had been cases of injuries and deaths of personnel doing the 1.5-mile run test. However, the Cooper-derived test was retained for specialty trades and continued to be administered to firefighters until the early 1990s.

[12] Dr. Wayne Lee, Director of Human Performance and Health Promotion, Canadian Forces Personnel Support Agency (CFPSA), explained that between 1983 and 1988 a contract was awarded to the Queen's University Ergonomics Research Group (ERG) to develop minimum physical standards for the general military population (Exhibit 56, p. 8). In 1989, the "Canadian Forces Minimum Physical Fitness Standards"

(CFMPFS), developed by the ERG, were approved and implemented. The CFMPFS consisted of four standards (Exhibit 58, tab 20, p. 12) depending on age and gender, namely: men less than 35 years old; women less than 35 years old; men 35 years old and older; and women 35 years old and older.

[13] On June 5, 1991, Lieutenant-Colonel Singleton, at the time the Canadian Forces Fire Marshal (CFFM), requested by memorandum (Exhibit 58, tab 16) that research be conducted to develop a new safe firefighter testing procedure to ensure that CF and DND firefighters were physically capable of carrying out their duties. He specified that the program should be "non-gender, non-biased and task-related." He explained that the Cooper-derived test was inadequate and that there were no built-in safety features in the test to stop the test if an individual had problems with his/her heart rate or high blood pressure.

[14] Dr. Lee approached the ERG, as it had assisted the DND in developing the CFMPFS, and on January 6, 1994 (Exhibit 53, tabs 20 and 21), a contract was awarded to the ERG to develop physical fitness standards for CF and DND firefighters.

[15] Dr. Janice Deakin, then professor at Queen's University School of Physical and Health Education, was named as the principal investigator for the ERG contract. There were several collaborative investigators from Queen's University: Dr. Pelot; Dr. Smith; Dr. Stevenson; and Dr. Wolfe. Dr. Lee was the scientific authority for the DND and any changes to work done for the contract were to be discussed with him. He was acting as a liaison between the DND and the ERG.

[16] Several graduate students also worked with the ERG, namely: S.A. Hughes; J.W. Dwyer; A.D. Hayes; and Ms. Jaenen. Ms. Jaenen was completing a master's degree at Queen's University and acting onsite as a DND coordinator for the ERG. At that time, she was employed at the CFPSA. She took care of the administrative and logistical details and got the ERG to different locations to conduct tests. As a graduate student, she assisted with data collection, reviewed scientific literature, and generally assisted the ERG. Captain Ken Hoffer, now Navy Fire Marshal, also participated as the Fire Service Representative.

[17] The ERG first conducted a review of the literature relating to the development and application of physical fitness standards for firefighters. The ERG then visited civilian and military training facilities for firefighters to familiarize themselves with

firefighting duties, firefighting equipment and operational requirements. The ERG went to the CF Base (CFB) at Borden and visited the Canadian Forces Fire Academy (CFFA).

[18] From the information that it gathered, the ERG established a list of the most common and demanding tasks for firefighters (Exhibit 26, Appendix A). They are:

i. perform tasks in a hazardous environment wearing self-contained breathing apparatus (SCBA);
ii. use and maintain fire department's ladders;
iii. perform forcible entry practices;
iv. participate in rescue operations during emergencies operating different equipment;
v. perform fire apparatus practices such as opening hydrants, connecting firefighting vehicles to other sources;
vi. perform search operations such as room search or search and rescue in a smoke maze while wearing a SCBA;
vii. conduct rescues from buildings by helping victims to walk or carrying them;
viii. perform rescues using breathing apparatus, cordage, ladders and rescue equipment;
ix. perform ventilation salvage and overhaul operations;
x. perform vehicle extrication;
xi. perform aircraft firefighting and rescue operations; and,
xii. fight structural fires by carrying hoses, carrying or using firefighting equipment, ladders, perform different duties during a fire.

[19] The ERG went back to CFB Borden to present this list of tasks to a panel of firefighters and instructors at the CFFA. The panel members were considered subject-matter experts, who were asked to review and verify this list (Exhibit 26, p. 22).

[20] The panel also discussed with the ERG the Cooper-derived fitness test in use at the time. The panel expressed dissatisfaction with this test. The panel was of the view that running a mile and a half was not related to the activities of firefighters. Ms. Jaenen testified that the panel also expressed dissatisfaction with the different standards for older firefighters and women. Dr. Deakin testified that the panel was unhappy with the "differential treatment". Dr. Deakin testified further that a member

of the panel made the following comment: "We don't have pink and blue hoses or different sizes of hoses for men, women or older firefighters." The implication was that a job is a job and everyone should meet the same standard.

[21] At the second visit to the CFFA, the ERG also compiled a list of firefighting equipment (Exhibit 26, Appendix B). The ERG then developed and validated a circuit that simulates representative tasks and some of the most demanding firefighting tasks.

[22] Dr. Deakin explained that the ERG did not know at the outset that it would develop a circuit of firefighting tasks. The comments received from the panel helped the ERG decide to develop a circuit. Dr. Deakin further explained that this was also supported by scientific literature.

[23] The most demanding and representative tasks were selected from the list that had been previously established (Exhibit 26, Appendix A). A circuit was developed over a period of performance trials at the CFFA and was presented to a second panel of subject-matter experts consisting of members of the CFFM Office, from both the CF and DND.

[24] After this presentation, the circuit was modified, as both panels recommended that the circuit be carried out in fire halls with the existing equipment and completed using just one tank of air of the SCBA (Exhibit 26, pp. 23 to 25). The test was to be conducted on a concrete slab floor in fire stations across the country. The circuit consisted of the following 10 tasks: one-arm hose carrying; ladder raise; charge hose drag; first ladder climb; high volume hose pull; forcible entry; victim drag; second ladder climb; ladder lower; and victim carry. Walking distances of either 15.24 or 30.48 metres were incorporated between the tasks to simulate tasks at the scene of an actual fire.

[25] A pilot study was done to investigate the physiological demands of the circuit and to determine its reliability before doing the main study. Twenty-five male firefighters between the ages of 21 and 42 participated in the pilot study.

[26] After being reviewed by the CFFM, the circuit was modified in the interest of safety. The victim-carry task was changed to spreader tool carry, as it represented a frequently occurring task performed by a single firefighter (Exhibit 26, p. 23).

[27] The ERG then proceeded to conduct the main study. This was the main data collection phase of its research for the purpose of assessing circuit performance across all of the CF and DND firefighter population. Data was collected on both genders, as well as across all ages of the CF and DND firefighting population. The ERG research team travelled across Canada to CFB Borden, CFB Chilliwack, CFB Comox, CFB Esquimault, CFB Halifax, CFB North Bay and CFB Petawawa. Firefighters from CFB Greenwood and CFB Shearwater also participated in this study.

[28] Between April and June 1994, data was collected from 202 male and seven female DND and CF firefighters who completed the circuit. The idea was to develop a single standard based on the performance of CF and DND firefighters who had completed the circuit (Exhibit 26, p. 44). However, to increase the number of women within the sample, additional testing had to be done outside the CF and DND population and 17 municipal female firefighters were tested in Toronto (October 1994) and in Winnipeg (December 1995).

[29] The average completion time for the men was 7 minutes and 30 seconds and 9 minutes and 57 seconds for the women. Based on its analysis of the data collected during the main study, the ERG recommended that a standard of eight minutes to complete the circuit be set for the firefighter fitness test (Exhibit 26, p. 95). The ERG explained that this was a performance objective that would provide a challenge for the young aerobically fit firefighters, while representing an attainable objective for older or less aerobically fit individuals.

[30] However, the ERG anticipated that between 29 and 42% of CF and DND firefighters could not complete the circuit within eight minutes. Therefore, the ERG recommended a gradual phase-in of the eight-minute standard over a three-year period. It noted that with physical fitness training and the opportunity to practice the circuit, more individuals would be able to meet the standard. The ERG recommended that, initially, there should be no job sanctions if a firefighter failed to complete the circuit in eight minutes. Rather, the firefighter would be required to participate in a mandatory physical fitness training program and be retested six months later.

[31] The ERG also recommended that it be emphasized that the standard for the circuit had been designed and developed to improve the fitness level of all CF and DND firefighters and not as a reason to terminate these individuals (Exhibit 26, p. 98).

[32] The ERG noted that the development of the eight-minute standard was consistent with the legal requirements for the establishment of a BFOR, as set out by the Government of Canada in 1988 (Exhibit 26, pp. 19 to 21 and p. 95).

B. Implementation of the eight-minute standard

[33] Lieutenant-Colonel Marc Desjardins, CFFM, explained at the hearing that the "Firefighter Physical Fitness Maintenance Program" (the Fitness Program) is not just a test but a complete wellness program. Firefighting is a risky operation and the employer has taken this measure so that employees are not at risk.

[34] The Fitness Program was instituted in 1998 (Exhibit 6). An implementation period starting in June 1998 allowed firefighters the opportunity to be evaluated without any time objective. Time was recorded for the individual's benefit only. From October 1, 1998, time was recorded for the benefit of the individual but a goal of 8 minutes and 30 seconds was set to complete the circuit.

[35] From October 1, 1999, all CF and DND firefighters were to be evaluated and the goal was to complete the circuit in eight minutes. The results were then to be reviewed by management. Personnel who failed to meet the eight-minute standard were to get assistance to meet the standard. Personnel who failed to meet the standard after a three-month retest could face administrative action.

[36] Several documents were prepared to explain the Fitness Program, specifically the circuit (Exhibits 21 and 22). As well, there was a video produced (Exhibit 48) entitled "Fire Fit". Captain Hoffer testified that he did several demonstrations of the circuit and his time to complete the circuit ranged between six and seven minutes. At the time of giving evidence, he was 52 years old. In November 2003, his time was 6:29 minutes.

[37] On May 1, 2000, Jim Judd, then Deputy Minister of the DND, announced that participation in the Fitness Program would become mandatory effective May 2000 for all CF and DND firefighters (Exhibit 7). At stage one, May 15, 2000, all personnel had to meet the eight-minute standard. However, no administrative action would be taken against anyone failing to reach the standard. Training and counselling would be offered to those failing to meet the eight-minute standard. Test results would be analyzed to determine the impact on age and gender groups.

[38] In stage two, scheduled for October 2000, the second annual testing of all personnel would have to be completed and administrative action would commence for those failing to meet the eight-minute standard.

[39] Implementation for CF firefighters was completed in 2000. However, implementation for civilian firefighters was difficult due to the UNDE's objection. Many civilian firefighters refused to be tested. In an effort to accommodate the UNDE's concerns, Mr. Judd announced, in February 2002, an additional phase-in period from the May 2000 initial deadline (Exhibit 9, p. 2). CF and DND firefighters had to be evaluated in 2002 (Exhibit 10) and had to meet the eight-minute standard by the end of 2003.

[40] In May 2002, a package was sent to CFB commanders to assist in the implementation of the eight-minute standard. The package included a series of letters for recommended disciplinary action for firefighters who refused to do the circuit.

[41] A directive of July 18, 2002 (Exhibit 11), stated that CF and DND firefighters were to be evaluated annually and that they were required to complete the circuit in eight minutes as a condition of employment. Failure to participate would be subject to administrative and disciplinary action (Exhibit 12). There have been a few cases of administrative action for firefighters refusing to participate in the test and in one case some disciplinary measure was imposed (Exhibits 17 and 18).

[42] In 2003, the Deputy Minister of the DND approved the Employment Continuity Program for firefighters who could not meet the standard, which included training for alternate positions, job placement counselling and potential deployment to other positions (Exhibit 16).

C. Ms. Flannery's evidence

[43] Ms. Flannery had been a volunteer firefighter for many years when she applied in 1987 to become a DND firefighter. She had to pass an initial fitness test. In each subsequent year, she was required to pass the Cooper-derived fitness test that was then in use. This test had different standards, depending on gender and age, and she successfully passed it (Exhibits 37 and 39).

[44] When she first started as a firefighter, Ms. Flannery was concerned that some tasks would be more difficult because women have a smaller physical stature than their male colleagues. At the firefighting school, she was shown techniques and alternate ways to handle the tasks. For example, she was shown a technique for the victim drag where, by using a rope, she could drag the victim, which required less strength. Another example concerned the new pumper truck. The connection for the hose was very high on the new truck and, given her height, it was difficult to reach. The truck was modified to enable shorter individuals to easily connect the hoses.

[45] Ms. Flannery explained that actual firefighting and rescue operations, as shown in her job description, correspond to only 4% of the duties. Other duties include: the inventory of the vehicle; fitness training; maintenance of equipment; fire hall maintenance; and fire prevention. She further explained the increasing demand for firefighters to respond to medical calls, which average two per day. Ms. Flannery responds to medical calls and she also gives first response instruction training.

[46] In 1994, Ms. Flannery was asked to participate in the ERG main study. She was 39 years old at the time and she was familiar with the tasks. On her first trial, it took her 11 minutes and 40 seconds to complete the circuit (Exhibit 41, p. 34); on her second trial, it took her 13 minutes and 42 seconds as she had a problem with the breathing apparatus.

[47] Since then, Ms. Flannery has been training regularly at home and at work and she participates in the Fitness Program. Ms. Flannery has not taken the test since 1994 as she feels that the test discriminates against female firefighters. She does not feel that the eight-minute standard is representative of the work of a firefighter. She testified that, for instance, the circuit does not test team work.

D. Mr. Barr's evidence

[48] When Mr. Barr testified on May 25, 2004, he had been employed as a firefighter for 23 years and was 56 years old. Until 1992, he had to pass an annual Cooper-derived fitness test. He has not taken a fitness test since 1992, but did participate in the ERG main study.

[49] Mr. Barr initially approved of the ERG circuit because it was all related to firefighters' tasks. Also, the information given to him led him to believe that there would be different standards for different age groups.

[50] Mr. Barr completed the circuit when he was 45 years old. He completed the 10 tasks of the initial trial in over nine minutes. At the second trial, he did not have to do the forcible entry task and it took him over seven minutes.

[51] Mr. Barr questioned the validity of the forcible entry task since, as a firefighter, he has never had to use a sledge hammer to move an object on a horizontal surface or for any other purpose. Mr. Barr has always done forcible entries with an axe or a prior bar specifically designed to open doors. As well, according to Mr. Barr, firefighters also use hydraulic equipment that is particularly useful when lifting large garage doors.

[52] Mr. Barr trains regularly. With the Fitness Program, all firefighters now have one hour per shift to train in the workplace. His level of fitness has never been questioned and he always received good performance evaluations by his supervisors (Exhibit 46).

[53] Mr. Barr's platoon consists of 14 firefighters, namely: one platoon chief; one captain; one lieutenant; and 11 firefighters. A minimum of 11 firefighters is required per platoon. He has been in the acting position of captain, also known as deputy platoon chief, for the last 11 and one-half years.

[54] More than half of the firefighters in the workplace are over the age of 50 and Mr. Barr is the oldest in this platoon. He explained that as a deputy platoon chief he would assign rescue tasks to the more agile or younger firefighters. He would utilize the younger firefighters in tasks where there is more heavy lifting required. As well, he would place the younger firefighters on the pumper truck. When he arrives at a scene, he has the authority to change assigned positions if he assesses that it is necessary. However, there is a rotation in the assignment of tasks and everybody gets to perform every task. The platoon works as a team. Team work is very important. Firefighters depend on and are very supportive of each other.

[55] Mr. Barr has not fought many fires because fire incidents in the DND environment are very controlled. All CF and DND personnel are trained in what action to take in case of a fire emergency. When incidents do occur, personnel take the appropriate measures. As such, when the firefighters arrive at the scene, the fire is often already extinguished.

E. Expert evidence

[56] Three witnesses, Stephen Brown, Dr. Deakin and Dr. Stuart Petersen, testified as expert witnesses. This evidence is grouped in this section along with more technical aspects found in the ERG report (Exhibit 26). However, some of Dr. Deakin's evidence was put in the preceding section, as it gave some context to the development of the eight-minute standard.

[57] For the reasons given in *Barr and Flannery v. Treasury Board (Department of National Defence)*, 2004 PSSRB 169, I did not recognize Ms. Jaenen as an expert. As an employee of the CFPSA, a non-public fund agency under the control of the employer, she is not independent from the employer. The employer had two other expert witnesses who could testify on its behalf on the same issues. Therefore, there was no necessity to have her testify for the employer, taking into consideration that she is not independent.

[58] Mr. Brown, a senior lecturer at Simon Fraser University's School of Kinesiology, was found to be qualified, without objection from the employer, in human performance and in the following subfields: exercise physiology; environmental physiology; ergonomics; determining the physical demands of a task or sport; exercise testing; exercise prescription; and active health.

[59] Mr. Brown prepared an expert witness report for the grievors (Exhibit 30), where he explained some key concepts of aerobic fitness, reviewed scientific literature on firefighters' fitness and analyzed the ERG project. He pointed out what he considered to be the weak links in the methodology used by the ERG to develop the eight-minute standard. He also analyzed the raw data compiled by the ERG during the main study.

[60] Dr. Deakin is the Director of the School of Physical and Health Education at Queen's University. Despite the grievors' objection, I found her to be qualified as an expert in kinesiology and human biodynamic. It was necessary for her to testify as an

expert witness to hear from her how the ERG came to establish the eight-minute standard. I found that the issue of the independence of Dr. Deakin could be addressed in the weight to be attributed to her evidence. Dr. Deakin did not produce an expert witness report, but testified on the ERG report (Exhibit 26).

[61] Dr. Petersen is a professor at the Faculty of Physical Education and Recreation at the University of Alberta. I recognized his qualification as an expert in exercise physiology. There was no objection to his expertise. However, the grievors submitted that his affiliation with the employer should be taken into account when assessing the weight to be given to his evidence.

[62] Dr. Petersen prepared an expert witness report dated June 14, 2004 (Exhibit 75) that, as agreed, was provided to the grievors prior to his testimony. However, as he testified eight months later, he had, in the meantime, reviewed his report. A revised version (Exhibit 76) was provided to the grievors only one week before he was to testify on February 14, 2005. As there were substantial differences between the two reports, both were admitted in evidence and the grievors were given the right to call reply evidence on the new report. Mr. Brown gave reply evidence for the grievors.

1. Aerobic fitness

[63] Aerobic fitness is required to perform tasks that use large muscle groups in a rhythmic sustained manner, such as firefighting. The best measure of aerobic fitness is maximal oxygen uptake (VO_2 max). The VO_2 max is the maximal rate at which the oxygen can be supplied to and taken up by cells (Exhibit 30, p. 5). The best test of aerobic fitness is direct measurement of the VO_2 max; the subject breaths in and out through an apparatus while exercising on a machine such as a treadmill. The most common method used to describe the oxygen uptake is relative to body mass (ml/kg/min.) (Exhibit 76, p. 9), which in this decision has been shortened to ml.

[64] Some tests, such as the Cooper test, estimate the VO_2 max and its variation. Other tests that are widely used are sub-maximal tests that attempt to predict the VO_2 max. In these tests, the subject is not pushed to the maximum performance and these tests are, therefore, considered safer. One of those tests is the step test, where the subject steps up and down a set of two 20 cm stairs to a special music CD. Every three minutes, the tempo increases and when the subject's heart rate reaches a certain

level, the test stops. However, there can be considerable error of plus or minus 10 to 20% in the prediction of the VO_2 max with the step test.

[65] On average, the VO_2 max declines with age, peaking at about age 30 and declining at the rate of about 10% per decade. The decline in the VO_2 max with age is related to decreased activity level, age-related diseases and the aging process. Scientific literature on firefighters shows the same age-related decrease in aerobic fitness as found in the general population (Exhibit 30, p. 10).

[66] Gender also affects aerobic fitness and it is reported in the scientific literature that after the age of 15, men have greater aerobic fitness than women. On average, VO_2 max is about 20% lower for women than for men of the same age and at the same level of activity. While elite female athletes have a VO_2 max that exceeds the VO_2 max of most men, their VO_2 max still falls 8 to 12% below that of elite male athletes (Exhibit 30, p. 13).

[67] The gender gap in aerobic fitness is due to biological differences between men and women. Men have larger hearts and can pump more blood with each beat. Men also have a greater percentage of body mass and muscles and a lower percentage of body fat than women. Mr. Brown and Dr. Deakin testified that women have approximately 50% of the upper body strength of men.

[68] Aerobic fitness can be improved with physical conditioning. However, it depends on a number of factors, one of which is the initial fitness level. An untrained person who starts a fitness program is likely to see rather sizable improvements over the first months of training and gains thereafter are more modest (Exhibit 30, p. 17).

2. Review of the scientific literature

[69] Firefighting is recognized as one of the most physically demanding and hazardous civilian occupations. It requires very high levels of muscular strength, endurance, aerobic and anaerobic power and motor abilities. Firefighters must have the physical capacity to complete firefighting tasks at irregular intervals. These tasks are performed under extreme conditions while wearing heavy protective equipment.

[70] The ERG reviewed scientific literature that measured the physical demands of firefighting. It found that the VO_2 max has been consistently identified in the literature as one of the most important physiological determinants of performance for firefighters.

[71] The ERG reported that four studies in particular had been undertaken to determine the minimum VO_2 max required to perform firefighting tasks: Lemon & Hermiston in 1977 recommended a VO_2 max of 39 ml; O'Connell in 1986, 39 ml; Davis & Datson in 1978, 42 ml; and Gledhill & Jamnik in 1992, 45 ml.

[72] The Gledhill & Jamnik study included climbing stairs of high-rise buildings as one of the most demanding tasks. In cross-examination, Dr. Deakin explained that research where municipal firefighters had to climb six flights of stairs would not be applicable to CF and DND firefighters as their job demands are lower; the buildings on CF bases are only three stories high.

[73] Dr. Petersen reviewed more recent scientific literature and found recommendations for a minimum VO_2 max to be in the range of 33.5 ml to 45 ml (Exhibit 76, p. 13). With the exception of the Bilzon study in 2001, only male subjects were used and the description of aerobic demands and the suggested standard are based consequently on male data. Dr. Petersen pointed out that "the aerobic demand on females performing firefighting tasks has not been adequately documented in the scientific literature".

[74] Mr. Brown reviewed 27 studies reported in literature that directly measured or predicted the VO_2 max of firefighters (Exhibit 30, pp. 21 and 22). The average VO_2 max reported ranged from about 26 ml to almost 56 ml. Mr. Brown concluded that the large range in reported VO_2 max between or within studies indicated that it was inappropriate to set a standard at 44 ml as the ERG has done. He also pointed out that in the studies where a minimum VO_2 max was recommended, the researchers' approach was to take the average VO_2 max of the subjects and set this as the minimum for firefighting. In Mr. Brown's opinion, this is an incorrect approach. Mr. Brown noted that in a study by *Romet and From* (Exhibit 30, p. 24), a different approach was recommended and rather than setting minimum standards for firefighters, the more physically fit individuals would be best suited for building entry and victim rescue.

[75] Mr. Brown also reviewed a number of studies that attempted to determine the oxygen cost (VO_2) of firefighting. The VO_2 is the oxygen that is used by the body to do the task, which is different from the VO_2 max, which is the maximum rate. Therefore, if the VO_2 of a firefighting task is 30 ml and the firefighter has a VO_2 max of 40 ml, then that firefighter's VO_2 for the task is 75% of the VO_2 max (Exhibit 30, p. 22). The VO_2 of firefighting tasks reported in the studies reviewed by Mr. Brown ranged from 23.4 ml to 44 ml (Exhibit 30, p. 25).

[76] Dr. Petersen also reviewed the VO_2 of firefighting as reported in four recent studies (Exhibit 76, p. 11). He averaged the VO_2 data in these studies and the mean VO_2 is about 32 ml. Dr. Petersen, however, cautioned that despite an apparently large number of studies, there was little consistency in the research design factors. Data should be judged on its own. However, these studies give some general insight on the VO_2 of firefighting.

3. ERG's pilot study

[77] The objectives for the ERG's pilot study (Exhibit 26, p. 27) were to compare the predicted VO_2 max, from tests used by the CF, with the VO_2 max measured directly in laboratory to better understand the relationship between the VO_2 max measured in laboratory and performance on the circuit; to examine the anaerobic contribution to circuit performance; to determine the reliability of time to complete the circuit; and to identify any test components that compromised the safety or integrity of the circuit.

[78] Dr. Deakin testified that, in view of these objectives, the ERG needed about 25 relatively fit firefighters located near Kingston to come to the laboratory for direct measurements of their VO_2 max. As Dr. Deakin explained, the ERG was testing the circuit and, therefore, it was not important at this stage that the subjects be representative of the workforce.

[79] The subjects of the ERG's pilot study were 23 male CF firefighters from the 7th Wing Ottawa who volunteered to participate. The subjects ranged in age from 21 to 42, with a mean age of 30.7 years. They were tested between October 18 and November 26, 1993. The subjects were first provided with a familiarization try-out, where they completed the circuit wearing physical training clothing. For the actual test, they wore full firefighter gear. The circuit performance times in the pilot study, with standard deviation (Std Dev), were the following (Exhibit 26, Table 6, p. 39):

Circuit performance times (min:sec) (pilot study)

Day	Mean	Std Dev	Minimum	Maximum
1	6:06	0:41	4:46	7:33
2	5:51	0:43	4:42	7:09

[80] Subsequently, all subjects attended Queen's University for laboratory testing. The tests included indirect and direct measurements of the subjects' VO_2 max. The mean predicted VO_2 from the step test was 44 ml. The mean predicted VO_2 max from the 1.5-mile run was 43 ml and the mean VO_2 max directly measured on a treadmill was 48.9 ml. The ERG found that, consistent with the literature, the predicted VO_2 max was less than actual.

[81] Mr. Brown pointed out that no female, older or DND firefighters participated in the ERG's pilot study. It was a relatively small sample that might not be representative of the larger population. They were all volunteers, which is not a random selection and creates a risk of selection bias. They were also more aerobically fit than subjects for the ERG's main study, as the pilot study subjects averaged a VO_2 max of 49 ml and took about six minutes to complete the circuit.

4. <u>ERG's main study</u>

[82] Data on 226 firefighters consisting of 202 men and 24 women was compiled for the main study. More subjects participated but only data for those who completed all stages of the testing is shown on the following table (Exhibit 26, Table 9, p. 48):

Age distribution of subjects by gender and service

Age (years)	Male				Female					
	CF		DND		CF		DND		Professional	
	N	%	N	%	N	%	N	%	N	%
20-29	29	12.8	6	2.7	0	0	2	0.9	8	3.5
30-39	61	27.0	53	23.5	2	0.9	3	1.3	8	3.5
40-49	11	4.9	36	15.9	0	0	0	0	1	0.04
50-59	1	0.04	5	2.2	0	0	0	0	0	0

[83] The mean circuit performance time was 7 minutes and 46 seconds. The average time for completion for men was 7 minutes and 30 seconds and 9 minutes and 57 seconds for women. The breakdown by gender, service and age reads as follows (Exhibit 26, Table 12, p. 51):

Circuit performance times (min:sec) by gender, age and service

Gender	Service	Age	N	Mean	Std Dev	Minimum	Maximum
Male	CF	20-29	29	6:31	0:47	5:05	9:00
		30-39	61	7:10	1:22	5:02	12:25
		40-49	11	7:58	1:17	5:43	10:25
		50-59	1	12:10		12:10	12:10
	DND	20-29	6	7:04	1:02	6:05	8:46
		30-39	53	7:34	1:19	5:26	10:56
		40-49	36	8:13	1:35	5:46	12:17
		50-59	5	10:13	1:47	8:52	12:42
Female	CF	30-39	2	13:42	6:12	9:19	18:05
	DND	20-29	2	12:06	0:28	11:46	12:26
		30-39	3	11:32	2:08	9:26	13:42
	Professional	20-29	8	8:59	1:42	7:35	12:15
		30-39	8	9:05	1:57	6:32	11:30
		40-49	1	8:08		8:08	8:08
Grand			226	7:46	1:48	5:02	18:05

[84] The necessary equipment to test directly the VO_2 max of individual subjects for the main study was not available in locations outside of Kingston. Therefore, the ERG used the step test to estimate the VO_2 max of individuals for the main study.

[85] However, Dr. Deakin, in her testimony, recognized that the step test underestimates the aerobic fitness of women and fit people, while it overestimates that of overweight subjects.

[86] Mr. Brown explained that one reason why the eight-minute circuit completion time is not defendable as a minimum standard is that the test was not developed for or validated in the main study by a sample comprising enough women or older individuals. Of the 226 participants in the main study, only six males were between the ages of 50 and 59. There were too few subjects of that age category to make these results valid for a BFOR test for men of this age. As well, there were not enough older women, as only one was over 39 years old.

[87] Mr. Brown also pointed out that while men were evenly divided between CF and DND firefighters, most women were firefighters from municipal forces, which have different job requirements. Only two female CF and five female DND firefighters participated in the main study. In Mr. Brown's opinion, there were too few female CF and DND firefighters to make the results valid for a BFOR test for women in these service categories.

5. ERG's women's sub-study

[88] The ERG conducted a sub-study (Exhibit 26, p. 59) to evaluate the performance on the circuit of physically fit women and to examine the effects of practice sessions on their time to complete the circuit. Physically fit women were defined as having an above average VO_2 max for females.

[89] Dr. Deakin explained in her testimony that time to complete the circuit by the women in the main study was clearly longer than men in almost all age groups. The ERG knew that women would be disadvantaged in the circuit to a greater extent than men. The ERG wanted to understand if the performance of women resulted from the fact that they had a lower predicted VO_2 max. In cross-examination, Dr. Deakin was asked about a scientific paper that she co-wrote (Exhibit 96), where it indicates that the imposition of any physical standard necessarily establishes barriers to physically demanding occupations and results in an adverse impact on women. Dr. Deakin testified that any physical standard would be more difficult to attain for women and older firefighters.

[90] When the ERG analyzed the data from the main study, it found that female firefighters had a lower predicted VO_2 max than what was suggested in the literature as the minimum VO_2 max for firefighters. The fact that women in the main study had a predicted VO_2 max lower than the men could be the reason for their performance. Another explanation could be that something in the circuit made it more difficult for women. To find out, the ERG conducted a sub-study of women volunteers with a higher VO_2 max. Firefighters were not considered as an option for the study, because they were only 50 female firefighters in Canada.

[91] The physical characteristics of the women who completed the sub-study are shown in the following table (Exhibit 26, pp. 64 and 65):

Subject Characteristics (women's sub-study)

Characteristic	Mean	Std Dev	Minimum	Maximum
Age (years)	31.9	2.98	28	37
Height (cm)	164.6	5.66	154.6	173.1
Mass (kg)	61.1	7.30	53.8	73.8
Body Mass Index (BMI) (kg/m2)	22.7	2.83	18.1	27.0
Waist to Hip Ratio (WHR)	0.74	0.034	0.66	0.77
Sum of Skin Folds (SOS) (mm)	46.3	13.03	32.0	72.4
Sum of Trunk Skin Folds (SOTS) (mm)	19.6	7.14	13.0	36.0
Body Fat (%)	23.5	5.76	16.5	34.2
Step Test 1 (ml/kg/min)	39.9	2.27	35.1	43.0
Step Test 2 (ml/kg/min)	41.4	3.40	34.5	45.5
VO2 max (ml/kg/min)	45.2	4.31	38.3	52.0

Results of strength tests (women's sub-study)

Test	Mean	Std Dev	Minimum	Maximum
Bench Press (kg)	64.22	22.64	40.86	102.15
Leg Press (kg)	217.42	149.30	98.75	442.65
Pull Down (kg)	61.41	11.41	43.13	81.72

[92] The subjects in the women's sub-study had eight sessions where they completed the circuit. In the initial main session, the mean circuit time was 8 minutes and 52 seconds. This was reduced to 6 minutes and 56 seconds by the last session. The largest decline in mean total circuit time was observed between sessions one and two.

[93] There was no significant physiological improvement in the subjects' aerobic capacity as a result of repeated exposure to the circuit because of the participants' above average or excellent physical condition. Performance improvements by the subjects were attributed to an increased familiarity with the tasks and the adoption of pacing strategies (Exhibit 26, pp. 57 to 70). Experienced firefighters would likely improve their performance only with pacing strategies, as they are familiar with the tasks.

[94] Mr. Brown pointed out that the sample in the sub-study was small, with only nine participants. They had a high aerobic fitness level at 45 ml; the average VO_2 max for women is 35 ml. The sub-study group was not homogeneous, as some subjects were elite athletes. Thus, average or mean values may conceal important individual differences.

6. ERG's forcible entry sub-study

[95] As the ERG considered the forcible entry task in the circuit to be the least realistic simulation, a sub-study was done to validate that simulated task. The first objective of the ERG's forcible entry sub-study was to determine if hitting a tire as in the circuit was similar to hitting a structure in real life. The second objective was to establish the distance that the tire had to be moved and what hammer weight to use (Exhibit 26, pp. 72 and 73).

[96] Twenty male CF firefighters from 12th Wing Shearwater in Nova Scotia volunteered to participate in this sub-study and their age ranged from 29 to 46 years.

[97] The ERG concluded from the sub-study that hitting a tire was comparable to hitting a reinforced structure. The ERG also found that using a standard issued 4.54 kg sledge hammer was the proper weight to use. Finally, it was determined that the tire had to be moved a distance of 40 cm. The ERG noted that it was difficult to compare the results of the stub-study with previous research in the literature as the published studies did not report the mean, standard deviation and range for completion times and heart rates. However, the ERG believed that this sub-study demonstrated that the parameters for the simulated forced entry task made it a genuine test of a real life task for firefighters (Exhibit 26, p. 94).

[98] Mr. Brown noted that this sub-study was well-designed. However, the average time taken to complete these tasks (12.9 seconds) was much shorter than real life forcible entry (46 seconds), as measured by Gledhill & Jamnik (Exhibit 26, p. 13).

7. Setting the physical fitness standard for CF and DND firefighters

[99] The ERG analyzed the data from the main study to determine the percentage of subjects who would fail on different performance objectives. These three objectives were 8 minutes, 8 minutes and 30 seconds and 8 minutes and 15 seconds.

[100] The initial choice of an eight-minute performance objective was determined by choosing the circuit completion time corresponding to the VO_2 max that was cited in the literature as being adequate for performing firefighting duties (Exhibit 26, p. 52):

. . .

> *The range of minimum values reported was from 35 ml/kg/min to 45 ml/kg/min. The average circuit completion time associated with a VO_2 max of 44 ml/kg/min was 8 minutes. The average VO_2 max for the subjects not passing an 8 minute performance objective was 38 ml/kg/min.*

. . .

[101] In cross-examination, Dr. Deakin explained that the ERG report did not contain everything that was discussed and that the minutes of the ERG meeting would be more complete. It was pointed out to Dr. Deakin that the minutes of an ERG meeting of September 26, 1995 (Exhibit 103), indicated that it would be valuable to consider using the average time for each age group in setting a cut-off score. She stated that what was required of the ERG by the employer, as noted in the minutes of the ERG meeting of September 7, 1995 (Exhibit 103, p. 2), was "to develop a single cut-off score regardless of gender, age, service and fitness rating". She explained that the ERG looked at the data and the mean performance and worked from there. The average time to complete the circuit was 7 minutes 46 seconds, which the ERG rounded up to eight minutes.

[102] At eight minutes, the average VO_2 max of the failing group was 39 ml and the average of the passing group was 44 ml, which was within the recommended range of 39 ml to 45 ml of VO_2 max found in the scientific literature.

[103] Dr. Deakin testified that there was a typographical error in the ERG report and that the VO_2 max range was not 35 ml to 45 ml, but 39 ml to 45 ml. It was pointed out to her in cross-examination that the minutes of the ERG meeting of September 26, 1995 (Exhibit 103), state that the literature indicated that a VO_2 max of between 35 ml and 45 ml was required. She maintained that there was an error in the ERG report, which should have read 39 ml to 45 ml.

[104] The second performance objective was selected by choosing a completion time corresponding to a passing rate of approximately 75% of the subjects. This was established at 8 minutes and 30 seconds, which corresponded to a 73.5% passing rate.

The average VO_2 max for the passing group was 43.81 ml, while the failing group had an averaged VO_2 max of 37.9 ml.

[105] The third performance objective was set at the mid-point range of 8 minutes and 15 seconds. The average VO_2 max for the passing group was 44.1 ml, and 38.1 ml for the failing group.

[106] More analysis was then done on the relationship between these three performance objectives and the gender, age and service of the subjects. Moving from a performance objective of 8 minutes and 30 seconds to 8 minutes resulted in an increase in failure rate for men from 22 to 30% and for women from 59 to 78% (Exhibit 26, pp. 53 and 54).

[107] There were six subjects in the 50- to 59-year age category. They all failed to meet any of the three performance objectives. Moving from a performance objective of 8 minutes and 30 seconds to 8 minutes resulted in an increase in failure rate from 38 to 48% in the 40- to 49-year age category. Similarly, in the 30- to 39-year age category, the ERG found an increase in the failure rate from 22 to 32%. The failure rate rose slightly from 16 to 20% in the 20- to 29-year age category (Exhibit 26, p. 54).

[108] As for the impact of the performance objectives on the different services, the ERG noted in its report that the profile of the DND group reflected a larger proportion of both older firefighters and more women than the CF group (Exhibit 26, p. 55). These two groups, female subjects and subjects in the 50 to 59-year age category, had a higher failure rate with the eight-minute standard and would be the most adversely affected.

[109] The minutes of an ERG meeting of February 15, 1996 (Exhibit 102), note that because the step test does not correspond to the actual VO_2, its results could not be used to establish a standard. Instead, the standard would be based on the time required to complete the circuit. Dr. Deakin is quoted in these minutes as saying that the ERG could not make a test that was easier or perceived to be easier than the tests currently in use. Several participants at this ERG meeting noted that a performance objective of 550 seconds (9:10) would be too generous.

[110] Dr. Petersen testified that he would not set a performance objective based on the VO_2 max as it is not very accurate. He analyzed the VO_2 and VO_2 max of subjects in his study and concluded that factors other than the VO_2 were at play and contributed to performance. As many tasks in the circuit involve moving heavy objects, this suggests that strength and anaerobic power are also important. A review of the characteristics of the subjects suggests that body mass and stature may influence how different individuals accomplish the work (Exhibit 76, p. 39).

[111] Mr. Brown analyzed the raw data compiled by the ERG for the main study (Exhibit 50). He demonstrated with a graph (Exhibit 51, figure 19b) that a predicted VO_2 max of 44 ml was only 50% accurate in identifying which subjects would complete the circuit in eight minutes. The data for the 226 subjects was dispersed among four groups: true positives - 75 subjects who had a VO_2 max of 44 ml and over and completed the circuit in eight minutes; false positives - 72 subjects that completed the circuit in eight minutes but had a VO_2 max less than 44 ml; true negatives - 73 subjects who did not complete the circuit in eight minutes and whose VO_2 max is lower than 44 ml; and, finally, false negatives - six subjects who had VO_2 max of 44 ml or more but did not complete the circuit in eight minutes. He also analyzed the data by gender and age category and concluded that there should be separate standards (Exhibit 51, figures 14 to 18).

8. Dr. Petersen's research

[112] As part of Dr. Petersen's experiment, he included a summary of two studies that he conducted for the CFPSA. In the first study, he measured directly the VO_2 of completing the circuit in eight minutes. The subjects were 30 male and 23 female volunteers who were recruited by word of mouth and poster advertisements in fitness centers. The mean age and VO_2 max of the male subjects was 29 and 49.2 ml and of the female subjects it was 25.8 and 41.7 ml.

[113] Dr. Petersen found that the average VO_2 was the same for males and females for a performance of between 7 minutes and 30 seconds and 8 minutes and 30 seconds. However, the VO_2 required for achieving the eight-minute standard represented a higher fraction of the VO_2 max of women (Exhibit 76, page 43). Dr. Petersen also explained that a SCBA tank generally will last no longer than 10 minutes. The eight-minute standard stimulates the speed of intervention necessary in an actual fire emergency.

[114] In cross-examination, Dr. Petersen stated that it was risky to generalize beyond the sample of his study. However, Dr. Petersen testified that, in his opinion, older firefighters are capable of maintaining a good fitness level and meet the eight-minute standard. Mr. Brown testified on the same point and explained that the findings of Dr. Petersen could not be transferred to older subjects.

[115] Dr. Petersen's second study consisted of fire rescue scenarios where subjects participated as a team of either two men or two women. The subjects were 13 men and 12 women. The mean age and VO_2 max of the male subjects was 25.6 and 44.2 ml and those of the female subjects were 23.9 and 44.7 ml (Exhibit 76, p. 58). The mean body mass of the men was 86.2 kg and 67.7 kg for the women.

[116] After appropriate training, male and female pairs completed the scenarios while their VO_2 was measured directly by a portable respiratory measurement system. The subjects had to maintain a work rate consistent with the requirements of firefighting work, as determined by the subject-matter experts from the CF Fire Service. Men completed the scenarios in 10.87 minutes and women in 13.71 minutes (mean time). Even if the female subjects took on average 2:43 minutes more to complete the scenarios, the subject-matter experts rated their performance as satisfactory (Exhibit 76, pp. 60 and 72).

[117] In this study, the relative VO_2 max for men and women was the same and, therefore, women worked at the same fraction of VO_2 max as men. The peak VO_2 during the scenarios was similar for men (35.3 ml) and women (34.2 ml) at approximately 80% of their VO_2 max. Men, on average, completed the scenario faster than the women. This discrepancy is explained by differences in body mass and strength. The female subjects were shorter and lighter than the males. The added mass of the fire protection equipment represented about 28% of the body mass for men and 35% for women. Consequently, the relative burden of the fire protection equipment was substantially greater for the female subjects. The same could be said about the weight of the charged hose and rescue mannequins (Exhibit 76, p. 65).

[118] The grievors asked to be provided with the raw data of Dr. Petersen's fire rescue scenarios so that Mr. Brown could analyze it. Dr. Petersen testified in cross-examination that he lost half of the data as a result of a computer crash. He explained that, contrary to protocol, he did not make a hard copy or CD-Rom copy of

his data. Dr. Petersen did not supply the data that was saved, as he believed that having half of the data would not be useful.

IV. Reasons

A. Jurisdictional issue

[119] Are these grievances premature as the grievors have neither failed to achieve the eight-minute standard since its implementation, nor have they been subject to any disciplinary measures? If they are premature, is an adjudicator then without jurisdiction to hear these grievances?

1. Statutory framework and collective agreement

[120] Section 92 of the of the former *Act* reads, in part, as follows:

> ***92.*** *(1) Where an employee has presented a grievance, up to and including the final level in the grievance process, with respect to*
>
> *(a) the interpretation or application in respect of the employee of a provision of a collective agreement or an arbitral award,*
>
> *(b) in the case of an employee in a department or other portion of the public service of Canada specified in Part I of Schedule I or designated pursuant to subsection (4),*
>
> *(i) disciplinary action resulting in suspension or a financial penalty, or*
>
> *(ii) termination of employment or demotion pursuant to paragraph 11(2)*(f) *or* (g) *of the* Financial Administration Act*, or*
>
> *(c) in the case of an employee not described in paragraph (b), disciplinary action resulting in termination of employment, suspension or a financial penalty,*
>
> *and the grievance has not been dealt with to the satisfaction of the employee, the employee may, subject to subsection (2), refer the grievance to adjudication.*
>
> *(2) Where a grievance that may be presented by an employee to adjudication is a grievance described in paragraph (1)(a), the employee is not entitled to refer the grievance to adjudication unless the bargaining agent for the bargaining unit, to which the collective agreement or arbitral award referred to in that paragraph applies,*

> *signifies in the prescribed manner its approval of the reference of the grievance to adjudication and its willingness to represent the employee in the adjudication proceedings.*
>
> . . .

[121] Subclause 19.01 of the collective agreement states the following:

> ***ARTICLE 19***
> ***NO DISCRIMINATION***
>
> ***19.01*** *There shall be no discrimination, interference, restriction, coercion, harassment, intimidation, or any disciplinary action exercised or practised with respect to an employee by reason of age, race, creed, colour, national origin, religious affiliation, sex, sexual orientation, family status, mental or physical disability, membership or activity in the Alliance, marital status or a conviction for which a pardon has been granted.*

2. Summary of the arguments and analysis

[122] The employer submits that the grievances are prospective in nature and therefore premature. The grievors have not yet taken part in the fitness test. Moreover, they have not been subject to any administrative action for not taking the fitness test.

[123] The employer argues that what is being asked in these grievances is that the adjudicator issue a declaration on the validity of the eight-minute standard and to make a decision in a vacuum. The employer asserts that, unless the grievors have taken the firefighting fitness tests and have been adversely affected by doing so, they have no *prima facie* right to file a grievance.

[124] The grievors submit that an employee need not wait until discipline has been imposed before challenging a rule that violates the collective agreement. Furthermore, as the grievances involve human rights, employees do not need to suffer specific employment consequences before being permitted to challenge the policy. Employees are affected by a discriminatory test the moment that they are required to participate in the test. They do not need to wait until discipline has been imposed before challenging a rule that violates the collective agreement.

[125] The grievors also argue that the employer has waived its right to raise this argument. As a result of a memorandum of agreement between the employer and the UNDE in April 2002, 327 similar grievances have been put in abeyance pending a decision on the present grievances. The employer waited until a few weeks before this hearing to raise this argument for the first time.

[126] I have reviewed the jurisprudence submitted by the parties on the jurisdictional issued. Adjudicators will decline to hear a grievance that merely raises a hypothetical or moot issue. However, even if some aspects of a grievance are prospective or anticipatory, where a real difference between the parties has crystallized, and elements of certainty in implementation exist, the matter will not be considered premature.

[127] A difference will have crystallized when there are acts of the employer that suggest a breach of the collective agreement. This could be an active statement of intent to implement by the employer (*Loschiavo et al. v. Treasury Board (Fisheries and Oceans Canada)*, PSSRB File Nos. 166-02-15391, 15392 and 15389 (1986) (QL)) or a notification of change in shift schedule (*Leger et al. v. Treasury Board (Transport Canada)*, PSSRB File Nos. 166-02-18740 and 18616 (1989) (QL)). Under paragraph 92(1)(*a*) of the former *Act*, a grievance may be referred to adjudication if it involves the interpretation or the application of a provision of a collective agreement. It is not necessary for the employer to have applied its interpretation. Rather, it is sufficient if the employer has indicated that it intends to apply that interpretation to the employee.

[128] In the present case, there is obviously a real difference between the parties relative to the implementation of the eight-minute standard. Implementation started in 1998 with a phase-in period of two years. Initially, the standard was to be mandatory by May 2000, and firefighters failing to meet the standard by October 2000 would be subject to administrative action. The standard was implemented as planned in 2000 for CF firefighters. However, the UNDE objected to the implementation of the eight-minute standard and many of the DND firefighters refused to be tested.

[129] In February 2002, the employer announced an additional phase-in period where all DND firefighters had to do the circuit and be evaluated in 2002. A few weeks later, the grievors filed their grievances. A memorandum of agreement between the

employer and the UNDE was signed in April 2002. Three hundred and twenty-seven (327) similar grievances have been put in abeyance pending a decision on the present grievances. In May 2002, a package to assist in the implementation was sent out to CFB commanders. The package included a series of letters for recommended disciplinary action for firefighters who refused to do the circuit.

[130] There is a general rule in labour law known as "obey now and grieve later" that aims to avoid insubordination in the workplace. The trade-off for this rule is that employees can grieve a policy of the employer if its breach will likely lead to discipline. The eight-minute standard was implemented when the grievances were filed and refusing to participate in the circuit could subject an employee to a range of disciplinary measures. If discipline was imposed, it could be grieved and could be referred to adjudication under paragraph 92(1)(*b*) of the former *Act*. The same issues would be discussed and the only difference would be the insubordination of the employee and that insubordination could be held against the employee. Accordingly, it is in the interest of harmonious labour relations that these grievances be adjudicated now.

[131] As J.A. Tarnopolsky of the Ontario Court of Appeal has held in *Metropolitan Toronto (Municipality) v. C.U.P.E.* (1990), 74 O.R. (2d) 239 (C.A.), at pp. 254-55:

. . .

> *. . . However, on the strength of the foregoing, it seems clear that under an "obey now, grieve later" rule, an arbitrator is practically required to take jurisdiction to hear a grievance against a directive, at least in a case where a breach is likely to constitute insubordination and subject the employee to disciplinary action. In my respectful opinion the Board, in taking jurisdiction, acted in accordance with both the letter and spirit of the collective agreement; its actions were neither patently unreasonable nor (using the more interventionist test) wrong in law. To decide otherwise would be to invite anarchy in the workplace.*

. . .

[132] It seems that it would be in the interest of the employer as well to have these grievances adjudicated so that it can move on to testing the fitness of DND firefighters or to find alternatives if the standard is found to be discriminatory and not a BFOR.

[133] The issue is not hypothetical for the grievors. Both of them participated in the ERG's main study in 1994 and twice they were unable to achieve the standard. Some of the equipment now used in the circuit is lighter, but the grievors are 12 years older and they are no more likely to achieve the eight-minute standard today. The odds that the grievors would fail again are higher as all women and CF and DND firefighters 50 years and older failed to complete the circuit in eight minutes during the main study. There is also certainty that, as a consequence of not meeting the eight-minute standard after a few attempts, the firefighter loses his or her job. Given the importance of human rights legislation and its *quasi*-constitutional status in Canada, there are even more compelling reasons for not subjecting the grievors to what could potentially be found to be a discriminatory standard prohibited under the law.

[134] These are the reasons for the decision that I rendered in *Barr and Flannery v. Treasury Board (Department of National Defence)*, 2004 PSSRB 24. I find that the grievances before me are not premature and that an adjudicator appointed under section 93 of the former *Act* does have jurisdiction to entertain them.

[135] Finally, the Federal Court of Appeal decided in *Canada (Attorney General) v. Boutillier*, [2000] 3 F.C. 27 (C.A.), that, for the purposes of subsection 91(1) of the former *Act*, the *CHRA* complaint process is ". . . an administrative procedure for redress. . . ." In this case, the CHRC has informed the grievors that their complaints of discrimination could be more appropriately dealt with according to the grievance process and adjudication under the former *Act*. In these circumstances, I find that an adjudicator appointed under section 93 of the former *Act* has jurisdiction to entertain these grievances alleging discrimination.

B. Main issue

[136] Does the requirement to perform the circuit in eight minutes constitute *prima facie* discrimination against women and older firefighters? If yes, is the standard a BFOR under the *CHRA*?

1. Statutory framework

[137] Under section 3 of the *CHRA*, age and sex are prohibited grounds of discrimination. Sections 7 and 10 of the *CHRA* prohibit discriminatory practices in employment and read as follows:

. . .

> ***7.*** *It is a discriminatory practice, directly or indirectly,*
>
> *(a) to refuse to employ or continue to employ any individual, or*
>
> *(b) in the course of employment, to differentiate adversely in relation to an employee, on a prohibited ground of discrimination.*
>
> . . .
>
> ***10.*** *It is a discriminatory practice for an employer, employee organization or employer organization*
>
> *(a) to establish or pursue a policy or practice, or*
>
> *(b) to enter into an agreement affecting recruitment, referral, hiring, promotion, training, apprenticeship, transfer or any other matter relating to employment or prospective employment,*
>
> *that deprives or tends to deprive an individual or class of individuals of any employment opportunities on a prohibited ground of discrimination.*
>
> . . .

[138] However, paragraph 15(1)(*a*) of the *CHRA* states that a BFOR is not a discriminatory practice. To be considered a BFOR, it must be established that accommodating the needs of an individual or class of individuals affected would impose undue hardship on the employer, considering health, safety and cost. Paragraph 15(1)(*a*) and subsection 15(2) of the *CHRA* read as follows:

. . .

> ***15****(1) It is not a discriminatory practice if*
>
> *(a) any refusal, exclusion, expulsion, suspension, limitation, specification or preference in relation to any employment is established by an employer to be based on a bona fide occupational requirement;*
>
> . . .
>
> *(2) For any practice mentioned in paragraph (1)(a) to be considered to be based on a bona fide occupational requirement and for any practice mentioned in paragraph (1)(g) to be considered to have a bona fide justification, it must be established that accommodation of the needs of an*

> *individual or a class of individuals affected would impose undue hardship on the person who would have to accommodate those needs, considering health, safety and cost.*
>
> . . .

2. Summary of the arguments and analysis

a) Is the eight-minute standard *prima facie* discrimination against women and older firefighters?

[139] The grievors submit that all expert witnesses in this case have agreed that the eight-minute standard had an adverse effect on women and older firefighters. Expert witnesses agreed that aerobic fitness is lower, on average, in women than in men and that aerobic fitness generally declines with age, peaking at around age 30 and declining at a rate of about 10% per decade thereafter. Both Mr. Brown and Dr. Deakin testified that women have approximately 50% of the upper body strength of men. Furthermore, the adverse impact of the eight-minute standard upon women and older firefighters is documented in the ERG's report itself (Exhibit 26, page 56). With a standard of eight minutes, 78% of women would fail, as would 100% of men over the age of 50. Dr. Petersen's research (Exhibit 88(b), page 37) highlights that women have to work at a significantly higher relative intensity to complete the circuit in eight minutes.

[140] The employer submits that the grievors have not demonstrated a *prima facie* case of discrimination. The firefighter fitness test draws a distinction between persons who are fit and those who are not. The evaluation disqualifies unfit men as much as it disqualifies unfit women and older persons. The job is physical, the tasks are physical, and there has to be a certain fitness level to apply to firefighters to do this job. The evidence shows that older individuals and women are capable of passing the ERG test in eight minutes. The difference between those passing and those not passing the test is their fitness level; thus, the comparator is not age or gender but rather fitness level, which is not a ground for discrimination.

[141] The employer argues that the question is whether female CF and DND firefighters are fit. Exhibit 50 (page 6) shows that few female CF and DND firefighters completed the circuit in eight minutes or less. However, a significant number of women in the ERG's sub-study were able to complete the circuit in eight minutes. There is also evidence that older firefighters can complete the circuit in eight minutes

or less. Captain Hoffer was 52 years old when he testified and had completed the circuit in less than eight minutes, a few months earlier. In Dr. Peterson's study, two men aged 50 completed the circuit in under eight minutes. As well, women and older firefighters can train and perform at better times.

[142] Dr. Deakin testified that any physical standard would be more difficult to attain for women and for older firefighters. In a scientific paper that she co-wrote (Exhibit 96), she pointed out that the imposition of any physical standard necessarily establishes barriers to physically demanding occupations and results in an adverse impact on women.

[143] Mr. Brown explained at the hearing and in his expert report (Exhibit 30, pages 13 to 17) that women have an average VO_2 max of about 20% less than that of men of the same age and level of activity. This gap in aerobic fitness is due to biological differences; men have larger hearts that can pump more blood. A greater percentage of body mass in men is muscle and they have a lower percentage of body fat than women.

[144] In his report, Dr. Petersen noted (Exhibit 88(b), pp. 86 and 87) that, on average, women do not perform as well as men on task-simulation or firefighting related physical fitness tests. He explained that most women are shorter and lighter and have a lower level of strength and a lower level of aerobic fitness than most male firefighters. Dr. Petersen's research showed that even with equal aerobic fitness, fit women will take more time to complete the circuit than men because of their smaller body mass and lower strength. Women also have a lower percentage of body mass that is muscle and a higher percentage of body fat than men. It is therefore clear that, as a group, female firefighters are disadvantaged in having to complete the circuit at the same required standard as for men in the same age group.

[145] All expert witnesses agreed that aerobic fitness declines with age, peaking at about age 30 and declining at the rate of 10% per decade thereafter. The decline with age is similar for firefighters to the decline observed in the normal population (Exhibit 31, tab 14, p. 339). While some of the decline in the VO_2 max with age is related to decreased activity levels, the decline is also due to age-related diseases and the aging process *per se* (Exhibit 30, p. 10). Also, with aging, there will be increased body fat versus a lower percentage of muscle in body mass. While training will increase fitness and counterbalance some of the effects of aging, as a fit person ages

aerobic fitness decreases. Age-related diseases and the aging process will also decrease aerobic fitness as an individual gets older. It is, therefore, clear that older firefighters are adversely affected in completing the eight-minute standard.

[146] As a fitness test, completing the circuit in eight minutes appears neutral on its face. However, women and older firefighters are adversely affected by the standard and it has a discriminatory effect on them.

[147] I therefore conclude that the grievors have met the burden and that the requirement that firefighters complete the circuit test in eight minutes is a *prima facie* discriminatory standard on the basis of age and sex. As in *British Columbia (Public Service Employees Relations Commission) v. BCGSEU* ("*Meiorin*"), [1999] 3 S.C.R. 3, the eight-minute standard has a disproportionate negative effect on women as a group. Likewise, this requirement has a disproportioned effect on older firefighters.

b) Is the eight-minute standard a BFOR?

[148] In 1999, the Supreme Court of Canada issued its landmark decision known as *Meiorin.* In this decision, the Court instead of distinguishing between direct or indirect discrimination adopted a new unified approach that could be applied to all cases. The court ruled that once it is established that a workplace standard is *prima facie* discrimination, then the onus shifts to the employer to establish on a balance of probabilities that the discriminating standard is a BFOR. The court developed the following three-step test in this regard:

. . .

> *(1) that the employer adopted the standard for a purpose rationally connected to the performance of the job;*
>
> *(2) that the employer adopted the particular standard in an honest and good faith belief that it was necessary to the fulfillment of that legitimate work-related purpose; and*
>
> *(3) that the standard is reasonably necessary to the accomplishment of that legitimate work-related purpose. To show that the standard is reasonably necessary, it must be demonstrated that it is impossible to accommodate individual employees sharing the characteristics of the claimant without imposing undue hardship upon the employer.*

[149] The Supreme Court of Canada in *British Columbia (Superintendent of Motor Vehicles) v. British Columbia (Council of Human Rights)*, [1999] 3 S.C.R. 868, extended the *Meiorin* test to all claims of discrimination under human rights legislation.

[150] Both parties agreed that the *Meiorin* test is to be applied in the present case. Having established that completing the test in eight minutes is a *prima facie* discriminatory standard against women and older firefighters, the burden shifts to the employer to establish, on a balance of probabilities, that the standard is a BFOR. The employer must meet the *Meiorin* three-step test.

Step 1: Did the employer adopt the eight-minute standard for a purpose rationally connected to the performance of the firefighting job?

[151] The focus of *Meiorin's* first step is not on the validity of completing the circuit in eight minutes but rather on the validity of the standard's more general purpose. Is there a rational connection between firefighting and completion of the circuit in eight minutes?

[152] The employer explained that the purpose of the eight-minute standard is to attain operational efficiency and its objective is to ensure that firefighters have the requisite fitness to perform their job. Operational efficiency is the ability to carry out the work efficiently and safely. It is to protect the firefighters, the public and property. The evidence is uncontradicted that firefighting is one of the most physically demanding occupations and that firefighters have to be fit to do their job. Without the requisite fitness, the functions and duties of firefighters cannot be carried out and therefore operational efficiency is compromised. This was not challenged by the grievors. The purpose of achieving operational efficiency is rationally connected to the performance of the firefighting job.

[153] Accordingly, I find that the first part of the test has been met. There is a rational connection between the eight-minute standard and the job to be performed.

Step 2: Did the employer adopt the eight-minute standard in an honest and good faith belief that it was necessary to the fulfilment of that legitimate work-related purpose?

[154] The evidence supports the employer's position that it adopted the eight-minute standard in an honest and good faith belief that it was necessary to the fulfilment of

operational efficiency. This was never questioned by the grievors. Therefore, I find that the second part of the test has been met.

Step 3: Is the eight-minute standard reasonably necessary to accomplish this legitimate work-related purpose?

[155] In the third step, the focus shifts to the standard itself. To establish that completing the circuit in eight minutes is reasonably necessary to the fulfilment of operational efficiency, the employer must demonstrate that it is impossible to accommodate female firefighters or older firefighters without imposing undue hardship on the employer. As in *Meiorin*, the contentious issue is whether the employer has demonstrated that completing the circuit in eight minutes is reasonably necessary ". . . to identify those persons who are able to perform the tasks of a . . . firefighter safely and efficiently."

[156] The employer argues that the fact that a firefighter attends work does not mean that he or she is physically fit to perform the essential duties of firefighting. While a firefighter can work as part of a team on a regular basis, in an actual fire the firefighter must be able to perform all the tasks of the job and be able to work with all the equipment. Therefore, one minimum physical fitness standard regardless of age or gender is justified. The employer retained experts to devise a non-discriminatory test.

[157] The employer submits that it cannot make further accommodations because of the risk to others. Further accommodations would place the individual firefighter, his or her co-workers, and members of the general public at risk. The eight-minute standard simulates the level of fitness required of DND firefighters to carry out their actual firefighting duties safely. The circuit allows an individual to rely on other abilities, such as techniques, strength and agility, to compensate for his or her physical condition.

[158] The employer argues that the methodology followed by the ERG is sound. The participation rate of CF and DND female firefighters was 87.5% in the ERG's main study. In addition, 17 female municipal firefighters were recruited for the main study. Women represented 10.6% of the total group tested.

[159] The employer explained that the performance objective is not a 44 ml VO_2 max but eight minutes, which allows for performance improvement through training,

acquiring better techniques and understanding the circuit. Another factor that came into play in setting the eight-minute standard was that the circuit must be completed with one bottle of air. As Dr. Petersen testified, a SCBA tank generally will last no longer than 10 minutes. The eight-minute standard simulates the speed of intervention necessary in an actual fire emergency where firefighters have to take important steps quickly.

[160] At the request of the employer, Dr. Petersen conducted a study to determine whether the VO_2 associated with the eight-minute standard was different for men and women. The results revealed that the average VO_2 is the same for men and women in meeting the eight-minute standard. Dr. Petersen also developed a rescue training scenario that demonstrated that the aerobic demands of firefighting are similar for men and women. Dr. Petersen testified that older firefighters are capable of maintaining a good level of physical fitness that enables them to perform the work and to meet the eight-minute standard.

[161] The grievors argue that there were significant problems with the way in which the research was conducted, namely: there were not enough women and older subjects in the sample study; it was inappropriate to combine data from different services; the ERG's pilot study used only male subjects; there was no analysis done of the connection between subjects who did not complete the circuit and gender or age; volunteers were used for the main study, which created a selection bias because the less fit are less likely to volunteer; the circuit is not representative of the fact that work is done in teams; and the use of a predicted VO_2 max for the main study underestimated the aerobic capacity of women and fit people while it overestimated that of overweight people.

[162] The grievors submit that the rationale in setting the eight-minute standard was flawed and, therefore, chosen arbitrarily. The ERG's report explained that the eight-minute standard was determined by choosing the circuit completion time corresponding to the VO_2 max identified in the scientific literature as being adequate for performing firefighting duties. The minimum VO_2 max reported were from 35 ml to 45 ml. However, the eight-minute standard corresponds to a 44 ml VO_2 max. Dr. Deakin testified that there was a typographical error and that the report should have read "39 ml to 45 ml". The grievors question the credibility of Dr. Deakin on this.

[163] The grievors explain that, to set the eight-minute standard, the ERG also looked at how many individuals would fail at 8 minutes, 8 minutes and 15 seconds and 8 minutes and 30 seconds. The grievors submit such norm-referencing is contrary to *Meiorin*. The ERG simply lumped together the results of men and women in the main study to establish a standard. The men's average circuit completion time was 7 minutes and 30 seconds. However, they represented 90% of the subjects. The women's average circuit completion time was 9 minutes and 57 seconds. However, they represented only 10% of the subjects.

[164] The grievors stress that no evidence was presented to support the view that the average completion time of the circuit is an appropriate minimum requirement to perform the tasks of a firefighter safely and efficiently.

[165] The grievors also submit that Dr. Petersen's new research does not validate the eight-minute standard. The purpose of Dr. Petersen's study was to measure the VO_2 of performing the circuit in eight minutes and not to validate or establish a new standard. Dr. Petersen testified that it is not appropriate to set a standard based on VO_2 max values. He did not provide the grievors with the VO_2 and heart rate data that they requested for the ERG circuit. This would have permitted a comparison between male and female subjects, as was done by Mr. Brown for the ERG's report. Dr. Petersen also lost half of the data of the subjects who completed his fire rescue scenarios and he did not provide any raw data on the VO_2 for these subjects.

[166] Dr. Petersen's scenarios were longer in duration and lower in intensity than the ERG's circuit. These scenarios were completed in times of, on average, 10.87 minutes for men and 13.71 minutes for women. Therefore, while women required more time, the subject-matter experts were completely satisfied that the work was done at an acceptable rate, taking both safety and performance into consideration. The grievors submit that Dr. Petersen's evidence does not support a finding that the eight-minute standard is a BFOR.

[167] To establish whether the eight-minute standard is a BFOR, I reviewed the evidence in light of some of the important questions that *Meiorin* suggested may be asked.

(a) Has the employer investigated alternative approaches that do not have a discriminatory effect, such as individual testing against a more individually sensitive standard?

[168] The employer did not investigate alternative approaches such as individual testing against a more individually sensitive standard. Only one approach was investigated and it was to develop a single standard. The employer explained, in its written submissions, at paragraph 137, that employees had expressed dissatisfaction with the Cooper-derived test, which had different standards depending on age and gender. As well, Dr. Deakin pointed out that when she met with an expert panel of firefighter instructors she was told that the job requirements were the same for all and that there should not be different standards depending on age and gender. The minutes of a meeting of the ERG research team of September 7, 1995 (Exhibit 103, p. 2) note that what was required of the ERG group by the employer "is to develop a single cut-off score regardless of gender, age, service and fitness rating".

(b) Is it necessary to have all employees meet the single standard for the employer to accomplish its legitimate purpose or could standards reflective of group or individual differences and capabilities be established?

[169] The Cooper-derived test used by the employer had different fitness standards for age and gender. This test was abandoned because there were concerns that the 1.5-mile run was not safe, but there was no evidence that having different standards had caused any problems to the employer in accomplishing its legitimate work-related purpose of attaining operational safety and efficiency.

[170] At a meeting of the ERG on September 26, 1995 (Exhibit 103), it was discussed that in setting a cut-off score, it would be valuable to consider using the average time for each age group. Unfortunately, it was not in the ERG's mandate to look at having different standards reflective of differences in age groups and gender. As Dr. Deakin testified, the expert panel told the ERG that there should be only one standard and the requirement was to develop a single cut-off score.

[171] The fact that some firefighters were unhappy that there were different standards in the Cooper-derived test is not relevant and does not justify the discrimination against women and older firefighters. In *Meiorin*, at paragraph 81, the Supreme Court of Canada clearly explained that having different standards for different groups is not reverse discrimination against men. In Dr. Petersen's fire rescue scenarios, it took women, on average, 2.43 minutes longer than men. However,

this performance was considered satisfactory by experts in firefighting. These women were maintaining a work rate consistent with firefighting work. This indicates that different aerobic standards could be developed to ensure that women and older firefighters perform the different firefighter tasks safely and efficiently. In any event, the employer has not established, on a balance of probabilities, that the development of different standards would impose an undue hardship on the employer, considering safety and cost.

(c) Is there a way to do the job that is less discriminatory while still accomplishing the employer's legitimate purpose?

[172] One of the big challenges of the job for people with a smaller stature and body mass is the weight of the protective equipment and the tools used in firefighting. There has been some improvement in the protective clothing, which is now made with nylon and weighs less. There was some change to the pumper truck as the height of the equipment on the truck was lowered to permit shorter firefighters to operate it. On the evidence presented, the employer has not shown that it has investigated and ruled out further improvements to the protective equipment and tools used in firefighting to diminish the physical requirements to do the job.

[173] Mr. Barr testified that, in assigning duties, he would likely assign rescue tasks to the more agile or younger firefighters in his platoon. He indicated that there is flexibility in the assignment of tasks when arriving at a fire scene. It would seem, therefore, that having diversity in a platoon is attainable without compromising operational efficiency and safety.

[174] For training purposes, it might be preferable to have a rotation in the assignments of tasks. However, a platoon chief has the flexibility to reassign tasks during an actual emergency, depending on the nature of the emergency and the individual strengths of the members of the platoon.

[175] An approach that might be considered would be to assign primary responsibilities for tasks on a permanent basis instead of having one standard or having different standards for each age, group or gender. Tasks requiring more strength and aerobic fitness would be assigned to the stronger and more fit firefighters.

[176] A similar approach was quoted in Mr. Brown's expert report (Exhibit 30, p. 24). In a study by *Romet and From*, the authors concluded that, instead of setting minimum standards for firefighters, the more physically fit individuals should be given tasks involving building entry and victim rescue. The employer has not established that such an approach would impose an undue hardship, considering safety and cost.

(d) Is the standard properly designed to ensure that the desired qualification is met without placing an undue burden on those to whom the standard applies?

[177] I reviewed the evidence and found significant problems with the methodology and findings of the ERG's main study. Most importantly, it has not been shown that completing the circuit in eight minutes is reasonably necessary for the safe and efficient performance of the work of a firefighter.

[178] The ERG developed a single standard based on the performance of the CF and DND firefighters on the circuit (Exhibit 26, p. 44). Data on firefighters who completed the circuit was compiled for the main study. All five female firefighters from the DND participated, as well as two female firefighters from the CF. This represented only 3% of the total of the sample used to establish the standard. The ERG recognized this as problematic and supplemented the sample with female municipal firefighters. This approach assumed that there would be no difference in fitness or job demands between CF and DND firefighters and municipal female firefighters. This assumption was not validated. On the contrary, the expert evidence was that the job demands for CF and DND firefighters are not as high, as the buildings that they service are a maximum of three stories high. Municipal firefighters may have to climb stairs of high-rise buildings.

[179] None of the CF and DND female firefighters completed the circuit in eight minutes or less (Exhibit 26, p. 51). Even when including the municipal female firefighters, 78% of the female firefighters failed to complete the circuit in eight minutes.

[180] Data on only six firefighters between the ages of 50 and 59 was compiled for the main study. However, as of the dates of the hearing, 30% of DND firefighters were over the age of 50. All six male firefighters in the age range of 50 to 59 failed to complete the circuit in eight minutes. About half of the firefighters between the ages

of 40 and 49 (Exhibit 25, Table 12, p. 51) failed to complete the circuit in eight minutes.

[181] Following *Meiorin*, workplace standards must be designed to reflect all members of society insofar as is reasonably possible. Standards must reflect the differences between individuals and differences that characterize groups of individuals. Otherwise, such standards are discriminatory under the *CHRA* and must not be used. The eight-minute standard failed all CF and DND female and older firefighters in the age group of 50 to 59 years. As noted in *Meiorin*, the burden is on the employer to demonstrate that, in the course of accomplishing its legitimate purpose, it cannot accommodate individual or group differences without experiencing undue hardship. The employer has not established that it would experience undue hardship if different standards were used to accommodate these two groups.

[182] The ERG considered that the fitness level of CF, DND and municipal female firefighters who participated in the main study was poor because they had a lower VO_2 max than what was reported in the literature to be the minimum required. However, as Dr. Petersen noted, only male subjects were used in the research referenced in the ERG's report. In light of this, a women's sub-study was conducted; no sub-study was conducted for older male firefighters. The purpose of the women's sub-study was to determine if physically fit women could complete the circuit in eight minutes. The women in the sub-study were all able to complete the circuit within that time. However, as Mr. Brown emphasized in his testimony, some were elite athletes and the range of their VO_2 max was between 38.3 ml and 52 ml. These women were very strong and could bench press between 40 and 102 kilos. One woman was able to leg press 442.65 kilos.

[183] The ERG had established early on that the standard would be based on the performance of subject CF and DND firefighters. However, this was not the case for female CF and DND firefighters, as they all failed to meet the eight-minute standard. Then it was decided to add to the main study female municipal firefighters and they also had to meet the standard. The women in the women's sub-study were not firefighters and I do not find that their results validate the eight-minute standard for CF and DND female firefighters. This sub-study, in any event, does not assess the issue of accommodation for female CF and DND firefighters. Similarly, no assessment of the issue of accommodation was done for the older firefighters.

[184] I was not presented with evidence establishing that completing the circuit in eight minutes represents the minimum fitness level to perform the job of a firefighter safely and efficiently. The ERG established the standard based on a review of the scientific literature that indicated that a VO_2 max of 35 ml to 45 ml was the minimum to do the job. The average circuit completion time of eight minutes was associated with a VO_2 max of 44 ml, and Dr. Deakin explained that there was a typographical error in the ERG report and that the minimum VO_2 max recommended was actually between 39 ml and 45 ml.

[185] In his report (Exhibit 88(b), p. 13), Dr. Petersen pointed out that new studies conducted since the ERG circuit was developed put the range for recommended a minimum VO_2 max for firefighters at between 33.5 ml and 45 ml. Mr. Brown reviewed the scientific literature and found that the VO_2 max of firefighters reported in the literature ranged from about 26 ml to almost 56 ml. In many of the studies, the VO_2 max was predicted and not measured directly, which adds considerable variability. Mr. Brown pointed out that, in some studies, the average VO_2 max of firefighters was deemed to represent the minimum VO_2 max to do the job, which is not correct.

[186] A VO_2 max of 44 ml associated with completing the circuit in eight minutes is at the far end of the range reported by the ERG in its report. Furthermore, the wider range in VO_2 max reported by Mr. Brown and Dr. Petersen is evidence that a VO_2 max of 44 ml does not constitute the minimum aerobic capacity to perform the job of a firefighter safely and efficiently.

[187] The ERG report does not capture all that was discussed in establishing the eight-minute standard. Dr. Deakin explained that the ERG looked at the average time to complete the circuit and the characteristics of the participants and rounded the time to eight minutes. In the minutes of the ERG meeting held on February 15 and 16, 1996 (Exhibit 102), Dr. Deakin stated that the ERG could not make a test that is easier or perceived to be easier than the test that was currently in use (i.e. the 1.5-mile run). The minutes of the ERG meeting then go on to state that at 550 seconds (9:10 minutes) it would be too generous as most of the 30 to 39-year-old men would pass and half of the 40 to 49-year-old men would not. At 8 minutes and 30 seconds, between 25 and 35 percent of the 40 to 49-year-old men would fail.

[188] Therefore, the eight-minute standard did not represent the minimum aerobic capacity to perform the job safely and efficiently and other considerations were at play. What the ERG did was lump together all the data, took the mean completion time of the main study at 7 minutes and 46 seconds, where younger male firefighters were over-represented, and rounded the figure to eight minutes. The problem here does not lie with female or older firefighters but with the use of a younger male norm in place of a fair-minded gender and age neutral job analysis.

[189] The eight-minute standard does not represent the minimum VO_2 max or fitness level to do the job. Instead, it is a cut-off score selected by the ERG so that it would be challenging and not perceived to be too easy to pass. The ERG could have determined different cut-off scores for males and females by age groups. The different standards could be challenging for each gender and age group and still ensure a fitness level necessary to do the job safely.

[190] The accommodation for female or older firefighters in the standard itself was not explored. It was argued that the employer could not make any accommodations because of the risk to the safety of others. However, in Dr. Petersen's fire rescue scenario, it was found to be acceptable by CF subject-matter experts that it took the female subjects almost three minutes more than the male subjects to complete the test. The previous firefighters' fitness test used by the employer had different standards and there was no evidence that this had resulted in a higher risk to the safety of firefighters or the public. The employer has failed to establish that it would face any undue hardship if it accommodated women or older firefighters. I therefore conclude that the eight-minute standard is not a BFOR.

3. Future search for accommodation in assessing fitness of firefighters

[191] As I have found that the eight-minute standard is not a BFOR, the parties have to consider what lies ahead and how can the fitness of firefighters be assessed while accommodating differences. As noted in *Central Okanagan School District No. 23 v. Renaud*, [1992] 2 S.C.R. 970, the task of determining how to accommodate individual differences also falls on the bargaining agent.

[192] There was no evidence that the bargaining agent was consulted in the development of the eight-minute standard or in the search for a possible accommodation in the standard itself. However, there were discussions with the

UNDE that resulted in the employer delaying implementation of the eight-minute standard. Nevertheless, these measures did not result in any greater acceptance of the standard by employees who are members of the bargaining unit. To this day, many employees refuse to take the test. Joint development of a standard and accommodation measures should be attempted, as it could lead to an acceptable standard.

[193] I understand that the employer is concerned that incumbent firefighters might not have the minimum physical fitness level to perform firefighting duties. Without testing, this could go undetected, thereby posing a risk to the safety of their colleagues and to the public. I recognize the bargaining agent's concern for the job security of its members, but it also has an interest in the safety of its members and the general public. I believe that it would be in everybody's interest to discuss how to assess the physical fitness of firefighters in a non-discriminatory way. It would be important during these discussions for the bargaining agent and the employer to explore options to address the cases where a firefighter is no longer fit to perform his or her duties.

[194] The firefighters' reaction to the circuit was positive when it was first introduced. It was seen to be representative of actual firefighters' tasks. I believe that if proper accommodations were provided for women and older firefighters, the circuit could be used for testing physical ability to do the job. Moreover, regular use of the circuit could promote greater awareness by firefighters of their own level of physical fitness. Dr. Petersen's fire rescue scenarios could also be useful as they involve team work and do not set a single standard. Instead, an acceptable work rate is the criteria, which even allows more time to complete the scenario for women and possibly older firefighters.

[195] The employer has developed some options, such as career reassignment or early retirement, for those who cannot meet the standard. These options should be further discussed with the bargaining agent, as they could promote some self-identification by firefighters who feel that they are no longer fit to do the work.

[196] The grievance process and adjudication have been lengthy and do not lead to creative solutions to improve the safety of firefighters or the public. It is not an easy task, but it is in everybody's interest that the bargaining agent and the employer sit down and address the issues together. Mediation services, if required, are available at

the Public Service Labour Relations Board to help the parties in finding these solutions.

4. Scope of the review

[197] In their submissions, the grievors ask that the order, if granted, extend to all firefighters who have faced career action or disciplinary action as a result of failing to take the test or to complete the ERG circuit in eight minutes.

[198] The present decision should have an impact on grievances that the parties have agreed to hold in abeyance. However, there may be other factors to consider in some of those grievances or career actions. The employer should review those grievances or career actions in light of this decision to avoid unnecessary litigation.

[199] For all of the above reasons, I make the following order:

(The Order appears on the next page.)

V. Order

[200] I declare that an adjudicator appointed under section 93 of the former *Act* has jurisdiction to entertain these grievances alleging discrimination.

[201] I declare that the grievors have met their burden of persuasion and that the requirement that firefighters complete the Fitness Program circuit in eight minutes is a *prima facie* discriminatory standard on the basis of age and sex.

[202] I declare that the employer adopted the eight-minute standard for a purpose rationally connected to the performance of firefighting.

[203] I declare that the employer adopted the eight-minute standard in an honest and good faith belief that it was necessary to the fulfilment of operational efficiency.

[204] I declare that the eight-minute standard is not a BFOR.

[205] The grievances are granted.

[206] The employer is to immediately cease using the eight-minute standard as a condition of employment for DND firefighters.

July 7, 2006.

Guy Giguère,
adjudicator