Profiling the New Zealand police physical appraisal test

Robin Marc Orr and Elisa Canetti
Tactical Research Unit, Bond University, Gold Coast, Australia

Jason Movshovich
Faculty of Health Sciences and Medicine, Bond University, Gold Coast, Australia

Robert Lockie
California State University Fullerton, Fullerton, California, USA

Jay Dawes
Oklahoma State University Stillwater, Stillwater, Oklahoma, USA, and

Ben Schram
Tactical Research Unit, Bond University, Gold Coast, Australia

Abstract

Purpose – The aims of this study were to evaluate fitness levels in a cohort of police trainees and compare these results to other police trainees and the general population.

Design/methodology/approach – Retrospective data for 274 male and 152 female police trainees were supplied. Measures included height, body mass and physical appraisal test (PAT; 2.4 km run, vertical jump, push-ups and grip strength) results, assessed twice, prior to commencement of training, separated by several months. Wilcoxon signed rank tests were used to analyze non-parametric initial and final PAT scores and Mann–Whiney U tests were used to determine variance between groups.

Findings – Male trainees were significantly quicker in the run (−12%, p < 0.001), completed more push-ups (+74%, p < 0.001) with greater grip strength (+52% left and +50% right, p < 0.001) when compared to female trainees. Following the second PAT assessment, the significant differences between male and female trainees remained (p < 0.001). Only female trainee 2.4 km run times improved significantly between initial and final PAT (−4%, p = 0.002).

Originality/value – When compared to the general population from which they were drawn and to other law enforcement trainees, the police trainees in this study were quicker, more powerful and stronger. While there was no loss of fitness between initial and final PAT performance, a conditioning program, spanning the periods between initial and final PAT may be of benefit to increase fitness prior to training commencement especially for female trainees who were generally less fit than, yet must complete the same training as, male trainees.

Keywords Law enforcement, Trainee, Cadet, Fitness assessment, Entry standards, Physical employment standards

Paper type Research paper

Introduction

Police officers have sworn duty to protect the general public, apprehend criminals, maintain order and uphold the law (Meena et al., 2018). As required of the profession, an officer may have to sprint short distances, pull/push objects, drag victims, lift and carry objects, jump and climb stairs (Anderson et al., 2001). These occupational demands highlight the
need for officers to be physically fit for duty, not only to maintain overall health, but to perform their duties optimally and reduce their risk of injury (Tomes et al., 2019; Orr et al., 2020).

Several studies have shown associations between low levels of fitness in police trainees and qualified officers, and their risk of injury and poorer task performance (Tomes et al., 2019; Orr et al., 2020; Nabeel et al., 2007; Violanti et al., 2017). In new police trainees, Orr et al. (2016a) found that Australian police trainees with the lowest vertical jump (VJ) height had a greater than three-fold relative risk of suffering an injury and/or illness than trainees with the highest VJ height. Levels of physical fitness have also been shown to be associated with the risk of injury in qualified police officers (Pinizzotto and Davis, 1999). As an example, Nabeel et al. (2007) reported that the higher an officer’s self-reported fitness, the less likely they were to experience back pain, chronic pain and sprains. When considering task performance and safety, research by Orr et al. (2017) revealed significant correlations between grip strength and measures of occupational performance, with poor grip strength associated with poorer defensive tactics performance. This finding is of note given that police training (which includes defensive tactics) was a leading cause of injury identified in the NZ Police, and dealing with a non-compliant offender and physical assaults are a leading cause of injury in police officers (Lyons et al., 2017).

In addition to physically active duties, police officers go through long periods of sedentary tasks. This and increasing age are associated with declining fitness levels and performance in police officers (Orr et al., 2016b; Dawes et al., 2017). Police duties have been suggested to be 80–90% sedentary and, while the use of physical force is infrequent, an effective physical response may be vital (Bonneau and Brown, 1995). Furthermore, studies have reported that despite increases in age, the physical demands of officers remain constant, highlighting the need to retain fitness into older years (Dawes et al., 2017; Plat et al., 2011; Rhea, 2015). As such, while the job requirements may stay the same, the sedentary nature of the occupation combined with progressive age may mean that officers gradually become less physically fit. Lagestad et al. (2014) determined that the maximal strength and lean muscle mass of Finnish police officers’ declined after three years of service. In a follow-up study by the same authors, fitness scores decreased by 10–32% on bench press, pullups, 3000 m run, and a standing long jump after 15 years of police work (Lagestad et al., 2014). Further work by Orr et al. (2018b), who compared police trainees and officers of the same age, found that officers were less fit than the trainees, suggesting that a potential reason for this lower fitness was the nature of law enforcement duties and the environment (e.g. shift work). Furthermore, work by Meyers et al. (2018) has documented that even within general duties, police officers from the same country (the USA), differences existed in officer fitness levels. These results indicate that not only increasing age, but the nature of police work itself may lead to a decline in fitness.

The need for law enforcement officers is to not only possess sufficient fitness to perform their duties but to be at a level of fitness above that of the general population has been recommended (Collingwood et al., 1998). To date, minimal research has compared police officer’s fitness levels with that of the general population. Both general police (Dawes et al., 2017) and specialist police (Maupin et al., 2018) have been shown to have differing fitness levels when compared to the general population.

As such, if fitness assessment results are to have meaning in a law enforcement population, they are required to be considered in such context. Given the minimal information regarding fitness profiles in police officers, particularly how they compare to the general population, and noting potential differences between police services, the aims of this study were to evaluate the fitness levels in a cohort of New Zealand police officer trainees and compare the results to fitness profiles of other officers and the general population.
Methods
Data were provided by the New Zealand (NZ) Police Constabulary Recruitment database to inform this retrospective cohort study. The data included age, height, body mass and physical appraisal test (PAT) scores. Results of the PAT are used as selection criteria for enrollment in NZ Police College. Collection of height and weight data was performed by NZ Police College Nursing staff during college enrollment. PAT data were collected from trainees at two time points, (T1 — Initial) one year and (T2 — Final) eight to twelve weeks prior to commencement of police training. Prior to entry, prospective trainees had to be 18 years of age, have clear criminal record checks, and get medical clearance from a general practitioner for known conditions affected by exercise activity. In order to be included in the study, participants had (1) to be qualified to attend training at the NZ Police College through a satisfactory PAT score and (2) to attend Police College after passing a PAT score. Data were excluded if the results were incomplete, or clearly inaccurate (e.g. 130 cm VJ or 2:30 min 2.4 km run). Ethics was approved by the Bond University Human Research Ethics Committee (BUHREC, Research Protocol BS02086).

Procedures
The PAT evaluates performance on the following functional abilities: running, jumping, climbing, firmly grasping objects or people, using aerobic fitness for lengthy restraint of offenders and foot pursuit of fleeing suspects. There are four tests included in the PAT: the 2.4 km run, standing vertical jump (VJ), push-ups and left and right-handed grip strength via dynamometer. These are outlined in further detail below.

2.4 km run
Aerobic fitness assessments, such as the 2.4 km run, are commonly used in police populations (Lockie et al., 2019a, b; Dawes et al., 2016b; Bloodgood et al., 2019). Trainees completed six laps on a 400 m level track course to achieve a total of 2.4 km as quickly as possible and were given time to warm-up prior to completing the run test. Completed time was recorded in minutes and seconds (min: s) using a handheld stopwatch (as is common practice (Bloodgood et al., 2019)) and converted into points based on time, sex and the trainee's body mass index (BMI) score.

Vertical jump
The VJ is a measure of power commonly used in this population (Orr et al., 2016a; Andersen et al., 2018; Cocke et al., 2016). To perform the VJ, trainees stood with flattened feet, linked hands together to reach up as high as possible on a standardized height measurement tool. This mark was recorded and then the trainee jumped as high as possible from a standing position with a squatting movement allowed. The highest value reached by either hand was noted and subtracted from the measured maximal reach height with the resulting score recorded in cm. Three attempts were allowed, and the highest value was used in the final score.

Push-ups
Push-ups are used to measure upper body muscle endurance and are commonly used in this population (Orr et al., 2018a; Dawes et al., 2016a; Cocke et al., 2016; Cvorovic et al., 2018; Bloodgood et al., 2019). Push-ups were performed with each trainee lowering from a body straight position with only the hands and feet touching the floor until their upper arm was parallel to the floor and the elbow was flexed to 90°. For a repetition to count, the trainee had
to return to the initial fully raised position with the elbows fully extended. The score was based on the maximal number of completed repetitions completed without passive rest.

**Grip strength**
Grip strength is a commonly used measure in this population to measure isometric grip strength (Ört et al., 2013, 2017; Dawes et al., 2017; Lockie et al., 2019b). Measurements of grip strength were performed via a Jamar digital hand dynamometer (Sammons Preston, Boilingbrook, IL, United States) with trainees holding the dynamometer with an extended wrist and elbow, and the shoulder flexed to 90°. The trainees then squeezed the dynamometer as hard as possible in the hand being assessed. Three attempts were allowed, and the highest value (in kg) was used in the final score.

**Statistical analyses**
Data were entered into an Excel (Microsoft Corporation, Redmond WA, United States) spreadsheet, inspected by authors for data entry accuracy, entered into the SPSS Statistical Analysis software (IBM, Armonk NY, United States) and subsequently analyzed. As the data were not normally distributed, Wilcoxon Signed-rank tests were used to determine any change between the initial and final PAT scores within each sex-stratified group. Mann–Whitney U tests were used to determine differences between groups. As normative data from the New Zealand police and general population are scarce, comparisons to similar cohorts, such as against Australian and US populations, were made from available literature. Effect sizes (d) were also calculated for the between-group comparisons for the initial and final assessments, where the difference between the means was divided by the pooled SD (Cohen, 1988). Effect sizes for Wilcoxon signed-rank tests were performed converting z into r with 
\[ r = \frac{z}{\sqrt{N}} \] (Clark-Carter, 2009). In accordance with Cohen (1988), an effect size of \( r = 0.1 \) was considered small, \( r = 0.3 \) medium and \( r = 0.5 \) large.

**Results**
Data for 426 police trainees who completed an initial PAT met the inclusion criteria for this study. Of these, 341 trainees completed both the initial and final PAT. Anthropometric characteristics of trainees in each PAT are detailed in Table 1.

A large number of erroneous data for the VJ were considered unreliable in the initial measures (e.g. reaching height had not been subtracted from recorded height) and as such, initial VJ measures were excluded from analysis. While removing this data meant initial–final comparisons could not be made, the final results were still maintained in order to allow for some measure of comparison against normative data.

Male trainees’ initial PAT scores were significantly better than female trainees’ scores across all measures (\( p < 0.001 \)) and this difference remained at the final PAT assessment.

<table>
<thead>
<tr>
<th></th>
<th>Age (years)</th>
<th>Height (cm)</th>
<th>Mass (kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial PAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males (n = 274)</td>
<td>26.46 ± 6.55</td>
<td>179.64 ± 9.81</td>
<td>85.60 ± 15.73</td>
</tr>
<tr>
<td>Females (n = 152)</td>
<td>27.02 ± 7.09</td>
<td>167.71 ± 6.24</td>
<td>69.09 ± 10.13</td>
</tr>
<tr>
<td><strong>Final PAT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males (n = 225)</td>
<td>23.66 ± 2.52</td>
<td>179.77 ± 6.86</td>
<td>84.24 ± 15.26</td>
</tr>
<tr>
<td>Females (n = 115)</td>
<td>25.56 ± 3.86</td>
<td>166.50 ± 5.80</td>
<td>65.88 ± 8.29</td>
</tr>
</tbody>
</table>

Table 1. Anthropometric characteristics of participants
There was a slight improvement in male trainee times to complete the 2.4 km run (approximately 12 s), between their initial and final PAT; however, this was not statistically significant. Likewise, the number of push-up repetitions and grip strength increased slightly in male trainees between initial and final PAT assessments (Table 2). Again, however, these changes were not statistically significantly different.

Female trainees completed the 2.4 km run approximately 31 s quicker on their final PAT, which was a significant improvement on their initial time (Table 2). Despite increasing the number of push-up repetitions by almost two repetitions, such improvement did not reach statistically significance ($p = 0.099$). Both left- and right-hand grip strength suffered a non-significant decline at the final PAT.

**Discussion**

The aims of this study were to evaluate the fitness levels in a cohort of New Zealand police officer trainees and compare the results to fitness profiles of other officers and the general population. As police officer trainees are at a stage between being a civilian and a licensed law enforcer, they are placed in a situation where they may be asked to perform fitness tasks required of an officer repeatedly as part of their training. As a result, premature exposure to a high volume of intensive fitness tasks may result in early onset of fatigue and a subsequent predisposition to trainee injury (Tomes et al., 2019).

The lack of any significant decrease in fitness between initial and final PAT assessments highlights the fact that the timing of this assessment may not be important prior to entry into police training when considering that trainees have met the standard. Considering this, a potential reason for this study finding that there was no significant loss of fitness between T1 and T2 may be that trainees were aware of a second fitness requirement. In this sense, a removal of the second fitness assessment requirement could lead to a decrease in fitness between time periods. Furthermore, while the timing is not crucial, the assessment itself is vital to ensure those with an elevated injury risk are identified early.

These results differ to those from a study with US law enforcement personnel undertaken by Lockie et al. in press who found that trainees improved their performance in push-up, 75 m pursuit run, arm ergometer revolutions and 2.4 km run and decreased their performance in sit-up between their initial “hire” assessments and the subsequent commencement of training. This same study by Lockie et al. in press identified successful candidates based on their PAT score. Push-up performance for those who passed training (initial = $40.32 + 14.25$ repetitions; final = $42.96 + 14.77$ repetitions) was notably higher than those presented in this study. It should be noted, however, that those who separated from academy training

---

### Table 2.

<table>
<thead>
<tr>
<th>Sex</th>
<th>Event</th>
<th>Initial PAT</th>
<th>Final PAT</th>
<th>$p$-value</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>2.4 km run (min: s)</td>
<td>10:44 ± 1:11*</td>
<td>10:32 ± 4:38*</td>
<td>0.507</td>
<td>−0.04</td>
</tr>
<tr>
<td></td>
<td>Push-ups (reps)</td>
<td>35.59 ± 7.32*</td>
<td>36.70 ± 5.80*</td>
<td>0.112</td>
<td>−0.11</td>
</tr>
<tr>
<td></td>
<td>Vertical jump (cm)</td>
<td>−</td>
<td>57.96 ± 6.25</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>(L) Grip strength (kg)</td>
<td>58.78 ± 8.70*</td>
<td>60.05 ± 6.12*</td>
<td>0.214</td>
<td>−0.08</td>
</tr>
<tr>
<td></td>
<td>(R) Grip strength (kg)</td>
<td>61.18 ± 8.33*</td>
<td>62.12 ± 8.50*</td>
<td>0.268</td>
<td>−0.07</td>
</tr>
<tr>
<td>Females</td>
<td>2.4 km run (min: s)</td>
<td>12:14 ± 1:14</td>
<td>11:43 ± 0:52†</td>
<td>0.002</td>
<td>−0.28</td>
</tr>
<tr>
<td></td>
<td>Push-ups (reps)</td>
<td>20.36 ± 6.26</td>
<td>22.15 ± 5.08</td>
<td>0.099</td>
<td>−0.15</td>
</tr>
<tr>
<td></td>
<td>Vertical jump (cm)</td>
<td>−</td>
<td>46.26 ± 7.71</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>(L) Grip strength (kg)</td>
<td>38.64 ± 5.14</td>
<td>38.11 ± 4.58</td>
<td>0.943</td>
<td>−0.01</td>
</tr>
<tr>
<td></td>
<td>(R) Grip strength (kg)</td>
<td>40.78 ± 5.57</td>
<td>39.47 ± 4.32</td>
<td>0.233</td>
<td>−0.11</td>
</tr>
</tbody>
</table>

**Note(s):** *Significant difference to female scores ($p < 0.001$); †Significant difference to initial scores ($p < 0.05$)
presented similar performance to the trainees in this study (initial $= 33.24 \pm 11.88$ repetitions; final $= 35.36 \pm 13.43$ repetitions). Conversely, the 2.4 km times reported by Lockie et al. in press were notably slower for both successful and unsuccessful candidates (Pass initial $= 12.49 \pm 1.32$ min, Pass final $= 11.55 \pm 1.25$ min, Separated initial $= 13.44 \pm 1.27$ min: Separated final $= 13.17 \pm 1.12$ min) when compared to those reported in this study. Conversely, a study of Australian Army Officers found that the officers lost fitness between their initial officer selection board and the commencement of training prior to the implementation of a self-directed conditioning program, which then mitigated this loss (Orr, 2010).

Due to a lack of normative data available for the New Zealand and Australian general population, data for the 2.4 km run were compared to the North American general population using data obtained from the American College of Sports Medicine. The mean score of the males in their initial PAT 2.4 km run placed them within the 70th percentile for their age group when compared to normative data (Baechle and Earle, 2008). Despite improving their mean time in the final PAT, the male trainees' time remained within the 70th percentile for their age group when compared to normative data. Female trainees were initially placed within the 70th percentile when compared to normative data of their age group (Baechle and Earle, 2008).

The significant improvement in run times in the final PAT resulted in a run time within the 90th percentile when compared to normative data. This may be explained by the initially greater number of older female candidates that were lost to attrition and were not available for the final PAT, resulting in a higher mean score. In relation to other tactical populations, the male trainees in this study presented higher levels of fitness than law enforcement officers reported by Dawes et al. (2016b) (mean 1.5 mile run time of 12:45 ± 1.10 min). The mean age of officers in that study was 39.42 ± 8.41 years which may have influenced the results, but were still faster than those reported in a study by Orr et al. (2018b) (male cadets times = 11.01 ± 1.17 min; female cadet times = 12.35 ± 0.82 min) and Lockie et al. (2019c (accepted)) (male cadets times = 11:38 ± 1:24 min; female cadet times = 12:48 ± 1:11 min). In a US police cadet cohort (males and females) assessing 2.4 km run times and failure to complete training, cadets who passed training run times averaged 11.55 ± 1.25 min, while cadets who failed training run times averaged 13:17 ± 1:12 min (Lockie et al., in press). Thus, this cohort ran 2.4 km faster than the general population and other police trainees.

Despite not presenting a comparative analysis between initial and final VJ scores due to unreliable measurement of VJ scores in T1, data obtained in the final VJ assessment are reliable and comparable to normative data obtained from the Canadian general population, which describes for the female age group of 20–29 years an average VJ of 30 ± 1 cm ($n = 83$) and for the male age group of 20–29 years an average VJ of 50 ± 1 cm ($n = 73$) (Payne et al., 2000). As such, the trainees in this study were notably superior in lower leg power when compared to the general population. Similarly, a cohort of male law enforcement officers from the USA recorded VJ scores at 58.47 ± 8.79 cm ($n = 82$, age = 20–29 years) (Dawes et al., 2017). Beck et al. (2015) reported jump heights of 51.4 ± 10.2 cm from a group of 15 police officers with a mean age of 33.1 ± 8.7 years, which was older than the mean age of officers in this study. Female data have been reported in military personnel 39.37 ± 6.6 cm (mean age = 27.0 ± 6.5 years) (Grier et al., 2017), suggesting that the female trainees in this study had greater jump heights.

The push-up results of both the male and female officers in this study did not significantly change between the initial and final PAT assessment. The males result would be classified as “very good” according to normative values from the US general population, while the females, initially being classified as “good”, progressed to be classified as “very good” after their final PAT assessment (Baechle and Earle, 2008). Normative general population fitness data cites male and female values for the 20–29 years age group as 22 ± 1 repetitions ($n = 83$) and 31 ± 2 repetitions ($n = 73$) respectively although females were
reported to have performed the push-ups from a modified position using the knees as a pivot
point (Payne et al., 2000).

The average of 36 repetitions performed by male officers in this study is similar to an
older cohort (33.1 ± 8.7 years) reported by Beck et al. (2015), with 34.8 ± 12.6 repetitions
recorded in a group of US campus law enforcement officers. The average male push-up
scores in this study were similar to the overall cohort scores reported by Lockie et al. in press,
which included female data (which would have lowered the overall average), in trainees who
separated from training in the US but lower than the cohort average (both male and female)
of those who passed. Furthermore, the results of this population are lower than what has
been reported by a similar aged group of US police officers who averaged 47.70 ± 14.74
repetitions (Dawes et al., 2017). Similarly, Meyers et al. (2018), who reported on two different
US law enforcement populations, found that average number of pushups ranged from
37.84 ± 7.72 repetitions in one agency (male n = 315: female n = 4) to 39.31 ± 7.98 repetitions
in another agency (male n = 72: female n = 7). These scores, however, were obtained from
trained officers who are known as a population to be generally less fit than trainees
(Orr et al., 2018b).

Higher scores were reported by Crawley et al. (2016) in males (n = 61, push-ups = 47 ± 13
repetitions) but not in females (n = 7, push-ups = 18 ± 10 repetitions). The female results in
this study were lower than those reported by Dawes et al. (2017) (30.50 ± 9.95 repetitions) for
a group of female officers of a similar age. As such, the push-up results reported in this police
population, while above that of the general population, were generally lower than that of
other police services.

Grip strength results for both the male and female trainees in this study are classified
as “excellent” when compared to normative values (Ratamess, 2012). Both male
(range 59–62 kg) and female (range 38–41 kg) grip strengths reported in this study are
similar to values reported by Dawes et al. (2017) for law enforcement officers
(males = 54.67 ± 7.47 kg, females = 37.67 ± 5.57 kg). The data for the trainees from this
study were above those reported by Crawley et al. (2016) in a US law enforcement agency
(males = 52–55 kg, females = 31–34 kg) and by Orr et al. (2017) in Australian agency (right
grip strength = 42.15 ± 8.29 kg, left grip strength = 41.91 ± 8.53 kg). The grip strength
results reported in this police population were well above that of the general population and
similar if not higher than that of other police services.

One major limitation of the study is the unreliable collection of vertical jump data in which
initial participant reach height may not have been subtracted from the final score. As the
initial reach data was not available, final jump height could not be accurately determined for
analysis. Another limitation is the lack of information about how standardized data were
collected in other studies. There is a large body of evidence which suggests that cognitive and
physical task performance are significantly impacted by the time of day and motivation of the
participants (Sonnentag and Frese, 2002). This information, however, is mostly unavailable in
previous literature. Also, volunteers for fitness normative data may not have been under the
same pressure as police candidates whose fitness performance may determine whether they
successfully pass all tests to progress in their careers. Additionally, tactical populations of
differing age groups were compared to the current cohort due to limited availability of studies
including homogenous groups, limiting the reliability of such comparison. Future studies
should aim to compare similar age groups among the general population and various tactical
populations who identically perform fitness tests. This should ensure stronger validity
among comparisons.

Conclusion
The results from this study suggest that, while male trainees were generally fitter than female
trainees, both sexes were generally fitter than the population for which normative data were
available. When compared to police cohorts from other countries, trainees in this cohort run 2.4 km faster, and were similar, if not superior, in VJ height and grip strength measures, but similar, if not below, in push-up repetitions. Given the general lack of differences between an initial and final PAT assessment, the agency should not be concerned about loss of fitness between timing of the PAT prior to trainee training. Trainee candidates may be aware of the follow-up assessment, and thus may have motivation to at least maintain their level of fitness. However, female trainees did show a moderate improvement in run times between their initial and final PAT and the initial PAT results may have led them to increase training in order to achieve higher fitness and scores prior to entering training. Given the slightly lower level of push-up performance when compared to other police trainee populations, a conditioning program spanning the period between initial and final PAT may be of benefit.

References


Corresponding author
Robin Marc Orr can be contacted at: rorr@bond.edu.au

For instructions on how to order reprints of this article, please visit our website: www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: permissions@emeraldinsight.com