Mobile health technology to promote physical activity in a workplace setting: A mixed methods pilot study in the police force

Submitted by Sarah Ann Buckingham, to the University of Exeter as a thesis for the degree of Doctor of Philosophy in Medical Studies, November 2019.

I certify that all material in this thesis which is not my own work has been identified and that any material that has previously been submitted and approved for the award of a degree by this or any other University has been acknowledged.

Signature:

Abstract

As work roles become more sedentary, new interventions to promote physical activity (PA) and reduce sedentary behaviour (SB) are required. Mobile health (mHealth) technology (including wearable activity monitors and smartphone apps) is increasingly popular, low cost and accessible. The project consisted of a systematic review of the existing evidence for mHealth interventions for PA and SB in workplace settings, followed by a real-world pre-post pilot study of a specific mHealth intervention in the police force (the Physical Activity Wearables in the Police Force, or PAW-Force study).

The review findings indicated that mHealth is a potentially effective, feasible and acceptable tool for promoting PA in a workplace context, at least in the short term. The longer-term impact and acceptability, and the impact on SB and wider outcomes (such as health, wellbeing and productivity) were less clear. The methodological quality of many existing studies was low and there were few mixed methods and qualitative studies.

The study aims were to address the evidence gaps identified in the review, including the use of mixed methods, a detailed exploration of feasibility, acceptability and engagement in both the short- and longer-term, in addition to the impact on PA, sedentary time, health and wellbeing, perceived stress, perceived productivity and sickness absence. This was the first known study of an mHealth intervention for PA and SB in the police force.

Police officers and staff (n = 180) within the Devon and Cornwall and Dorset Police received a 12-week intervention (a Fitbit® activity monitor and 'Bupa Boost' smartphone app) with 8 months follow-up. The results suggested a potential short- and long-term impact of the intervention on PA, particularly for less active officers and staff. Although the intervention was perceived as acceptable and feasible at various organisational levels, the findings highlighted the importance of contextual factors and opportunity, particularly for reducing SB. As in previous studies in other workplace settings, the impact on health, wellbeing, stress, productivity and sickness absence was unclear.

The fields of workplace wellness, mHealth and behaviour change are brought together in this PhD. The findings will inform future interventions in addition to policy and practice in the police force.

Table of Contents

List of Accompanying Material (Tables and Figures)

Figures

Tables

Acknowledgements

Firstly I acknowledge the support of the Devon and Cornwall Police and the University of Exeter Medical School, who jointly funded this project. I would like to thank my supervisors, Dr. Karyn Morrissey, Dr. Andrew James Williams, Dr. Lisa Price and Prof. John Harrison, who have provided invaluable support and encouragement throughout the past three years. I would also like to thank Dr. Cornelia Guell for the guidance on mixed methods and qualitative research.

The work would not have been possible without the continued support and interest of the Devon and Cornwall and Dorset Police. Special thanks are due to the following individuals who assisted in the setup and management of the pilot study: Inspector Andy Smith (Plymouth BCU), Chief Inspector Rob Chalkley, Sergeant James Dimmack and PCSO Matt Boon (North Dorset). Thanks are also due to the wider Health and Wellbeing Boards within the Devon and Cornwall and Dorset Police, who followed and supported the project from its inception to dissemination of the findings. I would also like to acknowledge the Communications, Admin and Performance and Analysis teams within the Devon and Cornwall and Dorset Police for their important inputs. I am grateful to Bupa and the Bupa Boost team for supplying the app that was evaluated in this study.

My colleagues within the European Centre for Environment and Human Health, and my family and friends have been very supportive and provided distractions when necessary. Finally, I am extremely grateful to all of the study participants who gave their valuable time to take part in this research.

Definitions and Abbreviations

95% CI;	95% confidence interval
BCT;	Behaviour change technique
BCU;	Basic Command Unit – the largest unit into which territorial British Police forces are divided. Each BCU has a number of specialist teams, including local investigation, local/neighbourhood and response policing.
BMI;	Body Mass Index
CALO-RE taxonomy;	'Coventry, Aberdeen and London – Refined' taxonomy, a standardised classification system of evidence-based behaviour change techniques (BCTs) for PA and healthy eating behaviours (Michie et al., 2011b)
COM-B;	Capability, Opportunity, Motivation and Behaviour model (Michie et al., 2011a)
CVD;	Cardiovascular disease
EPHPP;	Effective Public Health Practice Project quality assessment tool
HPQ;	Health and Work Performance Questionnaire
IPAQ;	International Physical Activity Questionnaire
MET;	Metabolic Equivalent of Task = a measure of intensity of physical activity based on oxygen consumption, compared to when at rest. One MET is the amount of oxygen consumed while sitting at rest (3.5 ml O_2 per kg body weight). A MET-minute refers to the intensity of activity in one minute. (Jette et al., 1990)
mHealth;	Mobile health
MVPA;	Moderate-to-vigorous physical activity
PA;	Physical activity
PCSO;	Police Community Support Officer
PICO;	Population, Intervention, Control (or Comparison) and Outcome
PRISMA;	Preferred Reporting Items for Systematic reviews and Meta-Analyses
PSS;	Perceived Stress Scale

Rank (UK police officers);	The level or standing of police officers. In ascending order the ranks in the UK police force are: Constable, Sergeant, Inspector, Chief Inspector, Superintendent, Chief Officers.
SB;	Sedentary behaviour
SD or ±;	Standard deviation
WHO;	World Health Organisation

CHAPTER 1 INTRODUCTION

1.1 The importance of workplace wellness

The workplace is increasingly seen as an important setting for health promotion and preventive health activities, not only to prevent occupational injury, but also to improve people's overall health and wellbeing (Wyatt et al., 2015, Institute of Occupational Safety and Health, 2015). Workplace wellness programmes, or interventions designed for disease prevention and health promotion in employees, are increasing in prevalence (Batorsky et al., 2016). This is the result of the growing public health burden of lifestyle-related disease, and the high proportions of time spent in the workplace by the majority of adults. In addition to health and wellbeing benefits for employees, such programmes appeal to employers as there are potential economic benefits resulting from improved productivity (Guo et al., 2015), reduced healthcare costs (Baicker et al., 2010, Guo et al., 2015), lower absenteeism (Baicker et al., 2010) and lower presenteeism (Cancelliere et al., 2011). Workplace wellness programmes may be used to target a range of lifestyle behaviors, including the promotion of physical activity (PA) and reduction of sedentary behavior (SB).

1.2 Physical activity and sedentary behaviour

Physical activity (PA) is defined by the World Health Organisation (WHO) as "any bodily movement produced by skeletal muscles that requires energy expenditure" (World Health Organisation, 2019). This includes exercise and activities performed for leisure, work, active transport and household chores. Physical inactivity has been named as one of the biggest public health problems of the 21st century (Blair, 2009) and is the fourth leading cause of global mortality (World Health Organisation, 2019).

International guidelines recommend that adults should do at least 150 minutes of moderate-intensity PA (or 75 minutes of vigorous-intensity PA) per week, in addition to muscle-strengthening activity on at least two days per week (World Health Organisation, 2019). Failing to meet PA guidelines is associated with an increased risk of morbidity due to cardiovascular disease, metabolic syndrome and cancer (Wannamethee and Shaper, 2001, Metzger et al., 2010, Liu et al., 2016), and general mortality (Arem et al., 2015, Dwyer et al., 2015, Lear et al., 2017). Internationally, there is systematic evidence of a decline in occupation-related PA (Knuth and Hallal, 2009).

Sedentary behaviour (SB) is defined as “any waking behaviour characterised by an energy expenditure ≤1.5 METs (metabolic equivalents) while in a sitting or reclining posture” (Sedentary Behaviour Research Network, 2012). Although SB has been less extensively researched in comparison with PA, the existing evidence suggests that PA and SB are related but separate behaviours, with distinct determinants and causal pathways (Spence et al., 2016). More prospective and intervention studies are needed to investigate the complex relationships between PA, SB and health. However, there is now considerable evidence that SB is an independent predictor of morbidity and mortality, including cardiovascular disease (CVD) and cardiovascular and all-cause mortality (Owen et al., 2010, Dunstan et al., 2011, Wilmot et al., 2012). Systematic review evidence from prospective studies indicates that increased occupational sitting is a risk factor for diabetes mellitus and mortality (van Uffelen et al., 2010). It has been recommended that sedentary, desk-based workers should aim to accumulate four hours per day of standing and light activity during working hours (Buckley et al., 2015).

Historically, studies have taken a ‘dual-hinge’ approach to promoting PA and reducing SB, where a single intervention is used to substitute SB with PA. There is an increasing realisation that this might not be the most successful approach; a systematic review and meta-analysis showed that interventions specifically designed to target SB have a greater impact on SB than interventions that focus on PA alone or both PA and SB (Prince et al., 2014). Different behavioural systems are likely to be involved in PA and SB, for example SB may be based more on automatic motivation (e.g. habits and impulses) rather than reflective motivation (Spence et al., 2016).

1.3 The police force: Physical activity, sedentary behaviour, health and wellbeing

Policing is an increasingly sedentary occupation (Ramey et al., 2014). This is likely to be due in part to the changing nature of policing, with increasing rates of white collar and cybercrime relative to street crime (Caneppele and Aebi, 2017). Police officers have been shown to be more active on their off-duty days than when they are at work (Ramey et al., 2014). Policing is also a highly stressful occupation; police officers (and also staff) are exposed to a range of acute and chronic stressors as part of their role, in addition to organisational pressures (Ramey et al., 2016). Additional lifestyle issues associated with the policing occupation include shift work, poor sleep and unhealthy diets (Ruiz and Morrow, 2005).

Studies in various countries have indicated that the police force has a higher than average risk of certain health conditions compared with the general population, with the aforementioned factors deemed to contribute to this risk. High body mass index (BMI) and obesity have been reported at high levels in the police force (Soroka and Sawicki, 2014, Can and Hendy, 2014). There is evidence for a high prevalence of metabolic syndrome (a cluster of risk factors including central obesity, hypertension, dyslipidaemia and glucose intolerance) in police officers (Leischik et al., 2015, Anderson et al., 2016). Police officers have also been shown to have an elevated risk of CVD (Franke et al., 2002, Hartley et al., 2011, Zimmerman, 2012).

Meeting healthy PA guidelines is important for members of the police force. A physically active lifestyle will potentially mitigate the health risks associated with the policing occupation (Can and Hendy, 2014, Anderson et al., 2016). In addition to enhancing the wellbeing of individuals, there may be potential organisational benefits such as improved productivity and reduced absenteeism (Guo et al., 2015, Baicker et al., 2010). Furthermore, police officers need to meet professional standards of fitness in order to pass the annual fitness test and deal with situations that necessitate fitness, endurance or the use of physical force.

Reported barriers to PA in the police force include lack of time, work pressures and lack of access to exercise facilities (Soroka and Sawicki, 2014, Lagestad and Van Den Tillaar, 2014). However, there have been few in-depth studies exploring the context of PA and SB, including barriers and facilitators, in this population. The need for novel interventions to promote PA and reduce SB in the police force has been emphasised (Ramey et al., 2014, Lagestad and Van

Den Tillaar, 2014, Anderson et al., 2016). One potential intervention is the use of mobile health technology.

1.4 The rise of mobile health and gaps in evidence

Mobile health (mHealth) is defined by the WHO as “medical and public health practice supported by mobile devices” (World Health Organisation, 2011). mHealth technology includes wearable PA monitors or trackers and smartphone applications (apps). In recent years, there has been a rapid increase in the use and popularity of mobile technology, and an associated increase in the use of mHealth in the general population, which allows people to monitor and manage their own health and wellbeing. In 2018, 172.2 million wearable devices were purchased worldwide (International Data Corporation, 2019). At the beginning of 2019, there were over 45,000 health and fitness apps available for download from the Apple App Store (Appfigures, 2019). The estimated value of the global mHealth market in 2019 is around $37 billion (Roland Berger, 2016).

mHealth technology has significant potential value in health promotion, due to its widespread appeal, accessibility, scalability and cost-effectiveness (Sullivan and Lachman, 2016). It also offers the potential to be tailored to meet the needs of individuals or specific groups (Huang et al., 2019). The technology may be used to minimise barriers to improving health and has the potential to incorporate evidence-based techniques to initiate and sustain behaviour change (Teyhen et al., 2014).

However, despite its widespread use and perception that it may be used for health promotion, there are many gaps in our understanding of how mHealth technology may be utilised to change behaviour and its potential impact on health outcomes. While there is some evidence that mHealth may be an effective way to promote healthy behaviours including increasing PA and reducing SB (Schoeppe et al., 2016, Direito et al., 2017), reviews have found that many mHealth interventions are neither theory-based nor evidence-informed (Fanning et al., 2012, Bort-Roig et al., 2014, Buijink et al., 2013, Knight et al., 2015). As a result, it is unclear which components (or app features, for example) may be most impactful for changing behaviour. Also, many studies of mHealth interventions have been short in duration (Fanning et al., 2012, Bort-Roig et al., 2014, Afshin et al., 2016), and consequently little is known about their longer-term potential to sustain behaviour change. Furthermore, many existing systematic reviews in this field are of poor quality. Bardus and colleagues (2016) conducted a comprehensive 'review of reviews' of mHealth and Web 2.0 technologies for weight management, PA and diet-related behaviours and found a number of limitations. These included failing to report whether reviews were based on a protocol, not including grey literature, and omitting full lists of included and excluded studies. As a result, only 16% of reviews were rated as high methodological quality. Heterogeneity of studies and reviews was high and the authors identified a need for higher quality reviews to improve confidence in the findings (Bardus et al., 2016).

mHealth interventions are complex in nature, as they are composed of multiple interacting components. Further complexity is added through factors such as difficulty of implementation (Craig et al., 2008) and contextual factors including setting, population and timing of the intervention (Moore et al., 2014). There is an increasing recognition of the importance of understanding not only what works, but how a complex intervention works, for whom, and in which circumstances (Moore et al., 2014). Aspects such as feasibility, acceptability and engagement are of vital importance, and are emphasised by the UK National Institute for Health and Care Excellence (NICE) and Medical Research Council (MRC) (Moore et al., 2014, National Institute for Health and Care Excellence, 2007), but remain understudied in mHealth research (McCallum et al., 2018). The use of qualitative and mixed methods studies that explore these factors has been recommended by the MRC (Moore et al., 2014). There is a need to improve the process of planning, conducting and reporting digital health and mHealth studies and evaluating interventions according to MRC guidelines (Huang et al., 2019).The need to consider context and setting (including social, environmental and economic factors) in behaviour change interventions has been recognised (National Institute for Health and Care Excellence, 2007, Moore et al., 2014). Different interventions are likely to be impactful and acceptable for different groups or populations. To date, there are no known studies of the use of mHealth technology for the promotion of PA in the police force.

1.5 A case study in workplace wellness: Devon and Cornwall and Dorset Police and the PAW-Force intervention

Health and wellbeing is a key priority for the Devon and Cornwall Police. The organisation prides itself on setting a leading example by promoting a culture of health, wellbeing and resilience amongst its officers and staff (Devon and Cornwall Police, 2016). A three-year wellness programme, 'ActivAte 2020' (also involving Dorset Police) began in 2017. The programme included a range of individual and local initiatives within an organisational framework, designed to target three health behaviours:

- Promotion of PA
- Improving diet and nutrition
- Improving sleep quality

The PhD focused on the first component of the programme, the promotion of PA. This involved a study of an mHealth intervention consisting of a smartphone app (Bupa Boost) and a wearable activity monitor (Fitbit® Charge 2). Bupa Boost is an app run by the private health insurer Bupa, and is available on both the Apple and Android platforms. Bupa describe the app as:

> *"a health and wellbeing tool designed to inspire you to set and achieve your own health goals. Focusing on four key areas: nutrition, fitness, mindfulness and relaxation, you'll be able to compete with colleagues and have fun as you improve your health"* (Google Play, 2019).

Bupa had an agreement with the Devon and Cornwall and Dorset Police forces to provide the app for free to officers and staff and to support research around the app. The Fitbit® Charge 2 activity monitor was selected by the Devon and Cornwall Police, and was also supplied free of charge to study participants. The Fitbit® was able to synchronise (sync) with the Bupa Boost app. Bupa Boost also has the ability to connect with other health and fitness apps and wearable devices to provide a single platform through which data can be combined and viewed.

Together the Bupa Boost app and Fitbit® are known as the 'Physical Activity Wearables in the Police Force (or PAW-Force) intervention'. Discussions between stakeholders from the University of Exeter and Devon and Cornwall Police resulted in the decision to undertake a study to assess feasibility and acceptability of the intervention for officers and staff, including an exploratory analysis of outcomes. The study was labelled as a pilot study in accordance with recent definitions. Pilot studies are considered a subset of feasibility studies; while feasibility studies are used to explore whether something can be done, should we proceed with it, and how, pilot studies additionally involve implementing an intervention and study processes on a smaller scale in preparation for a future study (Eldridge et al., 2016). It was agreed that the results of the study may potentially be used to inform a future larger scale effectiveness trial, depending on funding and resources.

Recognising the limitations of the 'dual-hinge' approach to PA and SB (see **section 1.2**), the primary aim of the PAW-Force study was to promote PA amongst officers and staff. However, as the intervention included some components that could be used to target SB (such as prompts and cues), the reduction of SB was considered a secondary aim, with sedentary time included as an outcome. A full description of the intervention and study methods (including rationale for the study design and outcomes) is given in **Chapter 3**.

A systematic review of mHealth interventions for promoting PA and reducing SB in workplace settings (see **Chapter 2**) was conducted in parallel to the PAW-Force study. While the published review may be read as a standalone study, a key purpose of the review in this thesis was to provide a background for the PAW-Force study, including helping to inform the study design and specific research questions.

1.6 Theoretical basis

The PAW-Force intervention was provided 'ready-made' by the Devon and Cornwall Police Force and was therefore not based on any overarching theory. Despite this, concepts from a number of theories were used for various purposes throughout the thesis. These included describing the intervention, designing the pilot study, and evaluating the findings.

Specifically, the components of the PAW-Force intervention were described and classified according to the **'Coventry, Aberdeen and London – Refined' (CALO-RE) taxonomy**, a standardised classification system of evidence-based behaviour change techniques (BCTs) for PA and healthy eating behaviours

(Michie et al., 2011b). The **Socio-Ecological Model** (Bronfenbrenner, 1989, Stokols, 1996, Robinson, 2008) was used to explore the context of PA in the police force, including barriers and facilitators at different 'levels', i.e. individual, interpersonal, organisational, and community/environmental. In addition, this model together with **social influence theories** such as Bandura's social cognitive theory (Bandura, 1997), provided the basis for the 'individual' and 'social' phases of the study. The **Technology Acceptance Model** (Davis, 1989, Holden and Karsh, 2010) was used to illustrate the findings in relation to the importance of perceived usability and usefulness in engagement and acceptability of the intervention. A number of behaviour change models and theories were used in the interpretation of quantitative and qualitative findings, and to explain how the intervention worked to change behaviour. The **Capability, Opportunity, Motivation and Behaviour (COM-B) Model** (Michie et al., 2011a) was the main model used to explain the mechanisms of change in the short and long term. The **Transtheoretical Model** (Prochaska and DiClemente, 1983) was used to conceptualise readiness to change. Concepts of intrinsic and extrinsic motivation, a component of **Self-Determination Theory** (Deci and Ryan, 2000), were also drawn on. A description of each of the models and theories is given in the chapters in which they are used.

1.7 <u>Research paradigm and methodological approach</u>

A mixed methods approach is taken throughout the thesis. This involves the use of quantitative and qualitative methods in the same study and integration of data (Tashakkori and Creswell, 2007). Mixed methods can overcome some of the problems associated with the use of quantitative and qualitative methods alone, and can produce a whole greater than the sum of its parts (Barbour, 1999). A mixed methods approach is recommended for the evaluation of complex interventions (Moore et al., 2014). Further detail on how mixed methods were used in the PAW-Force study is given in **Chapter 3**.

A pragmatist paradigm is adopted for this research; this is the approach most commonly taken by mixed methods researchers (Tashakkori and Teddlie, 2010). Pragmatism emphasises the importance of methodology rather than metaphysics, and methodology is placed at the centre of epistemology (i.e. theory of knowledge) and methods. The three main characteristics of pragmatism are abduction, intersubjectivity and transferability (Morgan, 2007). The pragmatic approach allows abductive inference, i.e. a combination of inductive (data-driven) and deductive (theory-driven) reasoning. The concept of intersubjectivity recognises that neither complete objectivity nor complete subjectivity is possible. Transferability replaces the traditional qualitative/quantitative dichotomy of context-specific and generalisable findings (Morgan, 2007).

1.8 Objectives of the PhD and PAW-Force study

The overarching aims of the PhD were to:

a) Conduct a systematic review to establish the evidence base surrounding the feasibility, acceptability and effectiveness of mHealth for the promotion of PA and reduction of SB in workplace settings
b) Explore the context of PA and SB in the police force (including prevalence, opportunities, barriers and facilitators for these behaviours) and develop recommendations for future organisational policies to promote PA and reduce SB
c) Design and conduct a mixed methods pilot study (with an embedded process evaluation) of a specific mHealth intervention for PA in the police force to determine feasibility, acceptability and impact. This work was informed by the findings of the systematic review (a).
d) Describe and evaluate the intervention according to:
    - Identification of key BCTs (using the CALO-RE taxonomy)
    - Relevant theories (e.g. theories of social support and influence, Socio-Ecological Model, COM-B Model, Technology Acceptance Model)
    - UK Medical Research Council (MRC) guidance for evaluation of complex interventions

Specific research questions to be addressed by the PAW-Force study were:

1. What is the context (prevalence, opportunities, barriers, facilitators) of PA and SB in the police force?
2. Is mHealth and fitness technology (the Fitbit® and Bupa Boost app) a feasible and acceptable intervention in the police force? Including:
    - Participant expectations, experiences and engagement
    - Longer-term engagement and acceptability
    - Adverse events and unintended consequences
    - Implementation and delivery
    - Feasibility and acceptability at all levels of the organisation (i.e. as perceived by managers, commissioners and occupational health staff)
3. Does the intervention assist police officers and staff in increasing PA and reducing sedentary time? If so, who is likely to benefit most (e.g. baseline activity level, age, gender, occupation)?
4. Which intervention components are preferred and potentially most likely to result in behaviour change?
5. Are there any wider benefits in terms of improved health and wellbeing, reduced stress, improved productivity and reduced sickness absence?
6. Will a future larger scale effectiveness trial be feasible and acceptable in this context? Including a reflection on:
    - Reach and recruitment
    - Adherence and attrition
    - Comparison of study completers and non-completers
    - Feasibility and acceptability of data collection procedures and outcomes
    - Acceptability of study participation

1.9 Overview of thesis

The thesis is organised into eight chapters. In **Table 1**, each chapter is related to the overall aims of the PhD and the specific research questions that were stated in **section 1.8**.

Table 1 Chapter titles with associated aims and research questions

Chapter	PhD aims	PAW-Force study research questions
Chapter 2 Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: A systematic review	a	N/A
Chapter 3 The PAW-Force study: Intervention and methods	c, d	N/A
Chapter 4 Participants, key baseline findings and context	b, c, d	1, 6
Chapter 5 Impact of the PAW-Force intervention	c, d	3, 4, 5, 6
Chapter 6 Acceptability of the intervention for police officers and staff	c, d	2, 3, 4, 5, 6
Chapter 7 Feasibility for the wider workforce and reflection on study methods	c, d	2, 6
Chapter 8 Integration of findings, discussion and conclusions	a, b, c, d	1, 2, 3, 4, 5, 6

Chapter 2 includes the published systematic review of mHealth interventions to promote PA and reduce SB in the workplace. This is a systematic evaluation of the existing evidence base relating to the effectiveness, feasibility and acceptability of mHealth in various workplace settings. Twenty-five studies are included. The review is framed within the thesis as a whole.

In **Chapter 3**, the PAW-Force intervention and study methods are described. The rationale for the study design and specific methods is given. The chapter includes a full description of the intervention and its components including coding of BCTs according to the CALO-RE taxonomy. Quantitative and qualitative outcomes, data sources and analysis methods are outlined.

In **Chapters 4 to 7**, the findings of the PAW-Force study are presented. These chapters include detailed discussions of findings in relation to the current literature.

In **Chapter 4**, the baseline study data (both quantitative and qualitative) are used to explore the context of PA and SB within the police force. This includes an exploration of participant characteristics, baseline activity levels, perceived barriers and facilitators, and recommendations to improve PA and reduce sedentary time.

In **Chapter 5**, the impact of the intervention is assessed, using quantitative data only. This includes an exploratory analysis of the outcomes of PA, sedentary time and wider outcomes (perceived health, wellbeing, stress, productivity and sickness absence).

In **Chapter 6**, the quantitative and qualitative data are used to build a picture of the acceptability of the intervention, including aspects such as engagement, perceived usability and usefulness, and perceived impact.

Chapter 7 includes an exploration of feasibility for the wider workforce and a reflection on study methods. Feasibility for the wider workforce is considered according to the results of a survey with managers, commissioners and occupational health staff within the Devon and Cornwall and Dorset Police forces. Implementation and delivery of the intervention is discussed in this chapter. Quantitative and qualitative data are used together to reflect on the feasibility of study methods.

Chapter 8 is the concluding chapter of the thesis. The findings are discussed in relation to the overall aims of the PhD, and summaries of findings from the systematic review and the PAW-Force study are presented. For each of the six study research questions, the main quantitative and qualitative findings are presented separately and overall contributions to knowledge are discussed. This chapter includes a discussion of strengths and limitations of the study, implications and impact for the police force, and implications for academia including recommendations for future research.

1.10 Outputs from the thesis

The following publications and presentations have arisen from the research presented in this thesis:

Publications

Buckingham, S.A., Williams, A.J., Morrissey, K., Price, L. & Harrison, J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: A systematic review. *Digital Health* 2019; 5: 1-50.

Buckingham, S.A., Morrissey, K., Williams, A.J., Price, L., & Harrison, J. OP67 The Physical Activity Wearables in the Police Force (PAW-Force) trial: Feasibility, acceptability and impact. *Journal of Epidemiology and Community Health* 2019; 73: A32-A33. (conference abstract)

A paper based on the results of the PAW-Force study is in preparation at the time of writing the thesis.

Presentations

Buckingham, S.A., Williams, A.J., Morrissey, K., Price, L. & Harrison, J. (2019). The Physical Activity Wearables in the Police Force (PAW-Force) trial: Feasibility, acceptability and impact (oral presentation). Society for Social Medicine & Population Health and European Congress of Epidemiology, Cork, Ireland, 4th - 6th September 2019.

Buckingham, S.A. (2019). Physical Activity Wearables in the Police Force: The PAW-Force Trial Summary of Findings (oral presentation). Plymouth BCU Health and Wellbeing Board Meeting, Crownhill, Plymouth, 10th June 2019.

Buckingham, S.A. (2019). Physical Activity Wearables in the Police Force (PAW-Force) trial: Summary of findings (oral presentation). Peninsula Forum for Environment and Human Health, Truro, 23rd May 2019.

Buckingham, S.A., Williams, A.J., Morrissey, K., Price, L. & Harrison, J. (2019). The Physical Activity Wearables in the Police Force (PAW-Force) trial: Quantitative findings (oral presentation). UCL Centre for Behaviour Change Conference, London, 9th - 10th April 2019.

Buckingham, S.A. (2019). The Physical Activity Wearables in the Police Force (PAW-Force) trial: Summary of findings (oral presentation). Devon & Cornwall Police and Dorset Police Joint Health and Wellbeing Strategic Coordination Group meeting, Middlemoor, Exeter, 19th February 2019.

Buckingham, S.A., Williams, A.J., Morrissey, K., Price, L. & Harrison, J. (2018). Physical Activity Wearables in the Police Force, the ‘PAW-Force Trial’: Protocol and baseline results (poster presentation). UCL Centre for Behaviour Change Conference, London, 21st - 22nd February 2018.

Buckingham, S.A. (2017). The PAW-Force (Physical Activity Wearables in the Police Force) Trial Protocol (oral presentation). Devon & Cornwall Police and Dorset Police ActivAte2020 Launch Event, Middlemoor, Exeter, 13th January 2017.

Although many of the above papers and presentations have joint authorship, the work contained in them is based on the research presented in this thesis.

CHAPTER 2 MOBILE HEALTH INTERVENTIONS TO PROMOTE PHYSICAL ACTIVITY AND REDUCE SEDENTARY BEHAVIOUR IN THE WORKPLACE: A SYSTEMATIC REVIEW

2.1 Introduction

A systematic review was undertaken to establish the existing evidence base relating to the use of mobile health (mHealth) interventions to promote physical activity (PA) and reduce sedentary behaviour (SB) in a workplace setting. The review was accepted for publication in February 2019. This chapter includes a brief overview of the rationale and development of the review (**section 2.2**), followed by the published manuscript (**section 2.3**) and concludes with a discussion of the findings within the wider context of the project (**section 2.4**). The review protocol including details of the search strategy, data extraction and quality assessment tools, is included in **Appendix 1**.

2.2 Rationale and development

It was necessary to conduct a systematic review for several reasons:

1. To summarise the existing evidence base and gain a comprehensive understanding of the field;

2. To identify gaps in the current literature and use these to inform the specific research questions for the PAW-Force study;

3. To examine the quality of existing studies and use any identified methodological strengths or limitations to guide the design of the PAW-Force study.

Initial scoping searches of existing systematic reviews in the field of mHealth interventions for PA and SB were performed. It was apparent that there had been no prior reviews of such interventions specifically in a workplace setting, and that previous reviews had various limitations and inconsistent findings (see Introduction of published manuscript, **section 2.3**). The content and format of these existing reviews were used as a guide for the development of the review protocol, and helped to identify relevant search terms, databases and sources of grey literature. Development of the review protocol and subsequent conduct of the review was also informed by the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009), with further guidance from the Cochrane handbook for systematic reviews of interventions (Higgins and Green, 2011) and the NHS Centre for Reviews and Dissemination (NHS Centre for Reviews and Dissemination, 2009). The review question was developed around the 'PICO' framework, as recommended by Cochrane (Higgins and Green, 2011). The '**Population**' was defined as working adults (recruited in the workplace and/or the intervention is delivered in the workplace). The '**Intervention**' was defined as any programme with an mHealth tool as a major component, including smartphone apps and wearable activity monitors or trackers. Experimental studies needed a '**Comparison**' or '**Control**' group (e.g. wait-list control) to be included in the review, but quasi-experimental and observational studies were also included if they had pre- and post-exposure data. This was to maximise the number of studies for inclusion, increasing the breadth of the review. '**Outcomes**' of objective or self-reported (quantitative) measures of PA or SB were necessary for studies to be included, and secondary outcomes (e.g. health, wellbeing, productivity) were also examined, in addition to quantitative and qualitative measures of feasibility and acceptability. Aspects such as feasibility and acceptability had been recognised as particularly lacking in mHealth studies (McCallum et al., 2018). The overarching research question was therefore: "Are mHealth interventions effective, feasible and acceptable for promoting PA and reducing SB in workplace settings?" Secondary questions were:

- What are the most commonly used behaviour change theories and techniques?
- Is there any evidence for subgroup differences (e.g. age, gender, shift workers, different occupations) in the effectiveness (or acceptability) of workplace mHealth interventions for PA and SB?
- Are any other related outcomes (e.g. health, wellbeing, productivity) improved after mHealth interventions for PA promotion?
- Where are the gaps in current knowledge surrounding the use of mHealth technology in PA promotion and SB reduction, and what are the implications for future research?

The search strategy was developed from the 'PICO' framework and existing reviews. A comprehensive list of search terms (including free text and controlled vocabulary) was produced, piloted, and developed further based on the comprehensiveness and relevance of the search results, before being applied to all of the electronic databases (and sources of grey literature) defined in the protocol. Details of data sources are provided in the Methods section of the published manuscript (see **section 2.3**). The searches were updated twice (in February 2018 and December 2018) to identify new evidence; updated searches were recommended by peer reviewers and guidance (NHS Centre for Reviews and Dissemination, 2009, Higgins and Green, 2011). The full search and screening log is presented in **Appendix 1c**.

A data extraction form was designed to systematically and rigorously capture data from each of the studies, including the 'PICO' and the main quantitative and qualitative findings. The form was piloted on a small number of studies (and checked by a second reviewer) before progressing with the complete data extraction process. While various tools for assessment of study quality were available (such as the Cochrane Collaboration's tool for assessing risk of bias (Higgins and Green, 2011), the Effective Public Health Practice Project (EPHPP) quality assessment tool for quantitative studies (Thomas et al., 2004) was felt to be most appropriate. The decision to use this tool was taken due to its applicability to a wide range of study designs (i.e. not only randomised controlled trials, RCTs), its established validity and reliability, and suitability for health promotion interventions. The templates for data extraction and quality assessment are included in **Appendix 1d**.

The review protocol included the specification of methods for meta-analysis. While meta-analysis was attempted, it was found to be infeasible due to the heterogeneity of study designs and outcomes in the included studies. This was further compounded by inconsistent and incomplete reporting of outcome data. The authors of the review agreed that a narrative synthesis approach was a suitable alternative; provision for this had been made in the protocol.

The review protocol (**Appendix 1a**) was registered with the University of York Centre for Reviews and Dissemination PROSPERO database (CRD42017058856).

2.3 Published manuscript

Title: Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: A systematic review

Authors: Sarah Ann Buckingham, Andrew James Williams, Karyn Morrissey, Lisa Price, John Harrison

Citation: Buckingham, S.A., Williams, A.J., Morrissey, K., Price, L. & Harrison, J. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: A systematic review. *Digital Health* 2019; 5: 1-50.

ABSTRACT

Objective: This systematic review aimed to assess the effectiveness, feasibility and acceptability of mobile health (mHealth) technology (including wearable activity monitors and smartphone applications) for promoting physical activity (PA) and reducing sedentary behaviour (SB) in workplace settings.

Methods: Systematic searches were conducted in seven electronic databases (MEDLINE, SPORTDiscus, Scopus, EMBASE, PsycINFO, Web of Science and the Cochrane library). Studies were included if mHealth was a major intervention component, PA/SB was a primary outcome, and participants were recruited and/or the intervention was delivered in the workplace. Study quality was assessed using the Effective Public Health Practice Project (EPHPP) tool. Interventions were coded for behaviour change techniques (BCTs) using the Coventry, Aberdeen and London – Refined (CALO-RE) taxonomy.

Results: 25 experimental and quasi-experimental studies were included. Studies were highly heterogeneous and only one was rated as ‘strong’ methodological quality. Common BCTs included self-monitoring, feedback, goal-setting and social comparison. 14/25 (56%) studies reported a significant increase in PA, and 4/10 (40%) reported a significant reduction in sedentary time. 11/16 (69%) studies reported a significant impact on secondary outcomes including reductions in weight, systolic blood pressure and total cholesterol. While overall acceptability was high, a large decline in technology use and engagement was observed over time.

Conclusions: While methodological quality was generally weak, there is reasonable evidence for mHealth in a workplace context as a feasible, acceptable and effective tool to promote PA. The impact in the longer term and on SB is less clear. Higher quality, mixed methods studies are needed to explore the reasons for decline in engagement with time and the longer term potential of mHealth in workplace interventions.

Protocol registration: The review protocol was registered with PROSPERO: CRD42017058856

Key words: Systematic review, mobile health, physical activity, sedentary behaviour, workplace, occupational health, behaviour change.

INTRODUCTION

Physical inactivity is considered one of the biggest public health problems of the 21st century (Blair, 2009). Failing to meet the recommended guidelines is associated with an increased risk of morbidity due to cardiovascular disease, cancer and metabolic syndrome (Wannamethee and Shaper, 2001, Liu et al., 2016, Metzger et al., 2010) and general mortality (Arem et al., 2015, Dwyer et al., 2015). There is now also substantial evidence that sedentary behaviour is an independent predictor of poor health and mortality (Owen et al., 2010, Dunstan et al., 2011, Wilmot et al., 2012).

Interventions to increase physical activity (PA) levels and reduce sedentary behaviour (SB) are clearly vital. The workplace is viewed as an important setting for health promotion and disease prevention (Institute of Occupational Safety and Health, 2015). Around half of weekday sitting time is work-related (Miller and Brown, 2004, Kazi et al., 2014) and up to 71% of working hours in office workers are spent sedentary (Clemes et al., 2014). Occupational sedentary time is predicted to further increase in future with rises in automation and information technology use (Hendriksen et al., 2016a). Promotion of PA in the workplace has many potential benefits, including improved health and wellbeing of employees and economic benefits for employers (Hendriksen et al., 2016b).

Mobile health (mHealth) technology has rapidly gained popularity in the general population. mHealth technology includes wearable PA monitors or trackers and smartphone applications (apps) designed to help people to manage their own health and wellbeing. The potential value of mHealth in health promotion lies in its widespread appeal, accessibility and ability to reach large populations at a low cost (Sullivan and Lachman, 2016). It also offers the potential for tailoring of interventions to the needs of individuals or specific groups.

Studies have investigated the use of mHealth to promote PA in a range of settings, including the workplace (Sullivan and Lachman, 2016). Whilst the results of clinical and general population studies suggest that mHealth may be a feasible and cost-effective way to promote PA (Direito et al., 2017), the findings of existing reviews have been inconclusive. Some reviews have reported nonsignificant effects of mobile technology on PA levels (Flores Mateo et al., 2015), and where beneficial effects are reported, effect sizes have generally been small (Fanning et al., 2012, Bort-Roig et al., 2014, Muntaner et al., 2016, Direito et al., 2017). Additional limitations of previous reviews are the inclusion of studies where mHealth devices were used as a data collection tool rather than as an intervention in their own right (O'Reilly and Spruijt-Metz, 2013, Bort-Roig et al., 2014), and a lack of a comprehensive description of interventions and study procedures (Fanning et al., 2012). Furthermore, apart from two recent exceptions (Schoeppe et al., 2016, Direito et al., 2017), few reviews of mHealth interventions have assessed both PA and SB outcomes.

Identification of behaviour change techniques (BCTs) using standardised taxonomies is important for recognition of effective and acceptable components, to allow replication and comparison of interventions, and to facilitate further development and testing of theories (Abraham and Michie, 2008). There is also evidence that including established BCTs is associated with greater intervention effectiveness (Greaves et al., 2011). Despite this, previous reviews have concluded that many mHealth interventions lack an explicit theoretical basis (Fanning et al., 2012, Bort-Roig et al., 2014) and it remains unclear which components are most effective and accepted (Sullivan and Lachman, 2016). Identification or coding of included BCTs, and identifying the theoretical basis of existing studies are therefore important gaps to address.

As mHealth is such a rapidly progressing field due to advances in technology, studies have increased exponentially in a short space of time. Early reviews predominantly comprised studies of text messaging (SMS) interventions but the emergence of new technologies (e.g. tablets, commercial wearable activity monitors, and 'exergaming') means the evidence should be frequently reviewed in order to accurately reflect the current status. Furthermore, the use and effectiveness of mHealth interventions in specific population groups remains unclear (Schoeppe et al., 2016). It is important to consider setting or context in the evaluation of mHealth interventions as due to their complex nature, various components may produce different outcomes for different individuals in different settings (McCallum et al., 2018). Workplace mHealth interventions may differ from general interventions in terms of both intervention content and timing of effectiveness (Stephenson et al., 2017). To the authors' knowledge, there has been no previous systematic review of mHealth technology for promoting PA and reducing SB in workplace settings. A recent review of general digital health interventions in the workplace concluded that the evaluation of smartphone apps in this context is an important 'next step' for future research (Howarth et al., 2018).

Employee populations potentially have much to gain from mHealth interventions for PA and SB, yet little is known about the impact of this technology in a workplace context. Feasibility and acceptability are important aspects to consider but remain understudied and underreported (McCallum et al., 2018). This review therefore aims to provide a comprehensive synthesis of current evidence in relation to the effectiveness, feasibility and acceptability of mHealth interventions in the promotion of PA and reduction of SB in the workplace. This includes a description of intervention content in terms of common BCTs using an established behaviour change taxonomy, and a consideration of subgroup differences and the wider impact of interventions on health and related outcomes.

METHODS

Protocol and registration

The review was conducted in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Moher et al., 2009) (Additional file 1). The protocol was registered with the University of York Centre for Reviews and Dissemination PROSPERO database (CRD42017058856).

Data sources and search strategy

Searches were conducted in the following databases: MEDLINE, SPORTDiscus, Scopus, EMBASE, PsycINFO, Web of Science and the Cochrane library (including the Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), Database of Abstracts of Reviews and Effect (DARE) and Health Technology Assessment (HTA)). Studies with a publication date between January 2007 (around the time smartphones were introduced) and February 2017 were included, with no language restriction. Full updated searches were later conducted to include studies to the end of February 2018, then to the end of December 2018. A master search strategy was developed (Additional file 2) and revised for each database (see Additional file 3 for example search strategy for MEDLINE). Both free text searching and controlled vocabulary were used, including key terms such as “mHealth”, “smartphone”, “application or app”, “activity monitor or tracker”, “physical activity”, “sedentary”, “workplace” and “occupation”.

In addition, relevant studies were identified via forward and backward citation searching, including reference lists of included articles and published systematic reviews. A search of grey literature, using the same key terms and for the same time period, included dissertations and theses (ProQuest Dissertations and Theses Global), ‘mHealth Evidence’, and the ‘Fitabase’ research library (studies using the Fitbit® activity tracker).

Inclusion criteria and study selection

Both experimental (e.g. randomised controlled trials, RCTs) and quasi-experimental (e.g. pre-post uncontrolled trials) study designs were included. Studies were included if they: 1) used mHealth (including mobile phone, smartphone apps, personal digital assistants (PDAs), tablets, wearable activity monitors/trackers) as a major component of the intervention, as stated by the authors or apparent from the context of the paper, 2) included a control or comparison group (experimental studies) or pre- and post- exposure data (quasi-experimental and observational studies), 3) recruited participants in the workplace and/or the intervention was delivered in the workplace, and 4) included any measure of PA and/or SB (self-reported or objective) as a primary quantitative outcome.

Pilot and feasibility trials were included if they met the inclusion criteria. Interventions could be either standalone mHealth or multi-component (e.g. facilitated with telephone counselling). The rationale for including multi-component interventions was that many digital workplace interventions for PA and SB, as delivered in the real world, are part of multi-component health promotion programmes (Nuffield Health, 2018) and we wanted to maximise the number of studies for inclusion and scope of the review. Interventions could be designed as an exclusive workplace or a wider lifestyle intervention (i.e. where the intervention was initiated or delivered in the workplace but also included activity outside of working hours). Studies using smartphone apps for PA or SB alone or with other behaviours (e.g. diet, weight) were included.

Exclusion criteria were studies reporting web-only interventions or traditional pedometers (i.e. not able to transmit data to a consumer interface), as these fall outside the realm of mHealth technology. Interventions involving basic text messaging (SMS) alone were excluded as these have been more extensively reviewed in the past (Fanning et al., 2012), and are felt to be a different type of intervention than more advanced mHealth tools such as multimedia smartphone apps and activity monitors. Studies using mobile devices for data collection only, and studies with clinical or student populations (i.e. school, college, university) were excluded. Studies reporting only qualitative data, non-human studies, review articles and editorials, and reports published only as conference abstracts or proceedings were also excluded.

All search results were imported into EndNote X7 bibliographic software (Thompson Reuters, San Francisco, CA, USA) and duplicates removed. Two independent reviewers (SAB and AJW) screened papers for eligibility by title and abstract followed by full text screening. Disagreements were resolved through discussion and consulting a third reviewer (KM).

Data extraction

Standardised data extraction forms were completed by one reviewer (SAB) and verified by a second reviewer (AJW). Any disagreements were resolved through discussion and consulting a third reviewer (KM). The following data were captured: author; year; country; setting or workplace; study design; participants (number and characteristics); intervention description (type of mHealth technology or tool, intervention components including whether standalone mHealth or multi-component, theoretical basis, key motivational strategies or behaviour change techniques (BCTs), duration and frequency); control or comparator; study aim (i.e. improve PA and/or reduce SB); primary PA/SB outcome (including method of assessment); secondary outcomes; duration of follow-up; main study results including any relevant subgroup findings; details of acceptability, engagement and attrition. Key within- and between-group quantitative findings were summarised for each study; significant effects were $p<0.05$.

Study quality assessment

Included studies were appraised using the Effective Public Health Practice Project (EPHPP) quality assessment tool for quantitative studies (Thomas et al., 2004). This tool was developed for health promotion interventions and was selected for its application to a wide range of study designs (e.g. RCTs, cohort trials and case-control studies). The tool has demonstrated content and construct validity and both intra- and inter-rater reliability (Thomas et al., 2004, Armijo-Olivo et al., 2012).

The EPHPP quality assessment tool assesses six domains: 1) selection bias; 2) study design; 3) confounders; 4) blinding; 5) data collection methods; and 6) withdrawals and dropouts. Each study was given a rating of ‘strong’, ‘moderate’ or ‘weak’ for each domain; based on this, a global rating was then assigned for each study - ‘strong’ (no weak ratings), ‘moderate’ (one weak rating) or ‘weak’ (two or more weak ratings). Intervention integrity (proportion of participants receiving the intended intervention), fidelity of delivery (whether studies measured consistency of intervention) and appropriateness of analysis methods were also separately considered.

Two independent raters (SAB and AJW) used the tool to assess risk of bias and study quality. KM was consulted to resolve any uncertainties.

Coding of BCTs

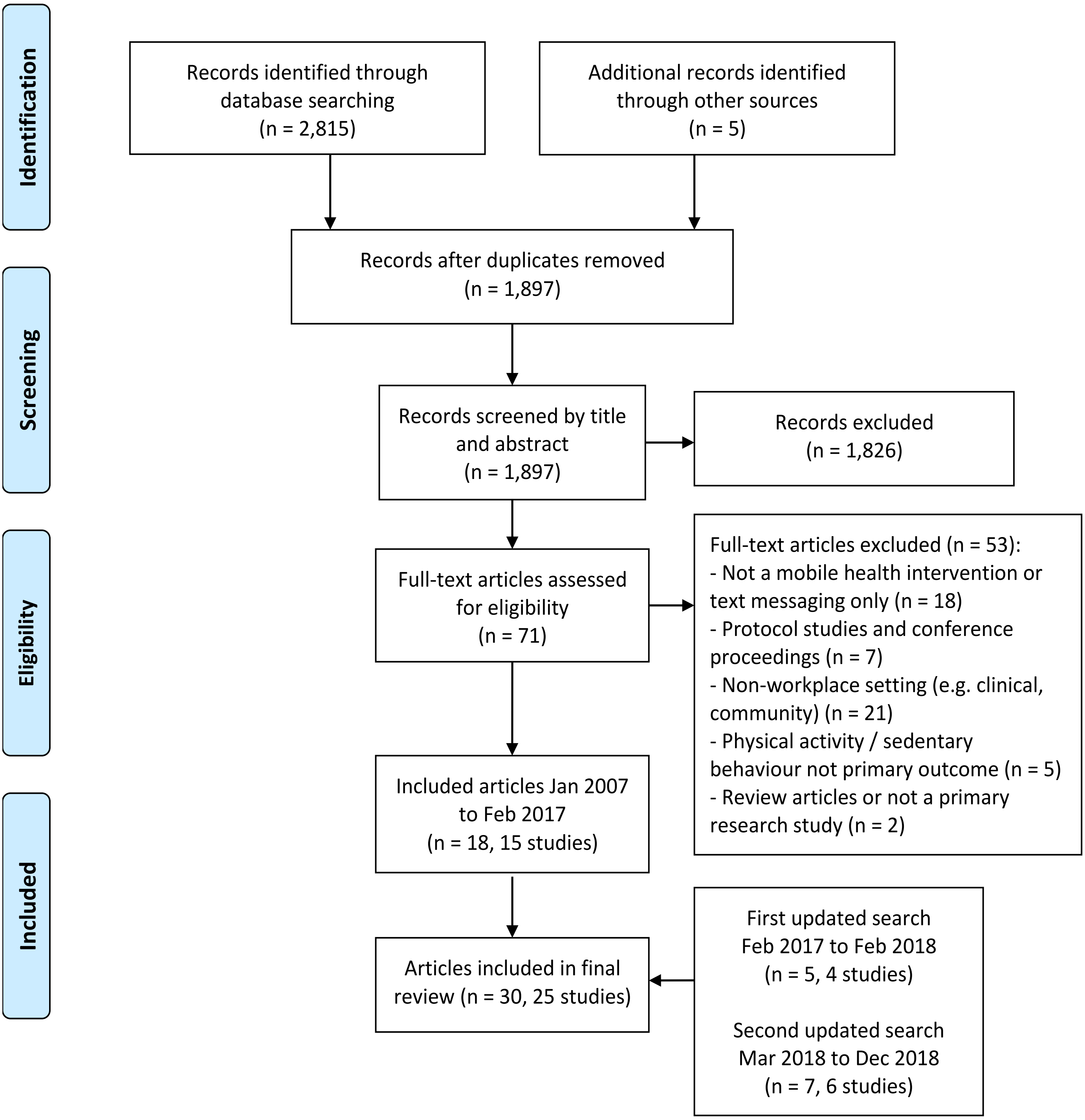
Interventions in included studies were coded for BCTs using definitions provided in the ‘Coventry, Aberdeen and London – Refined’ (CALO-RE) taxonomy for PA and healthy eating behaviours (Michie et al., 2011b). This 40-item evidence-based taxonomy was selected as it was specifically designed for PA and healthy eating behaviours, and is widely used including to characterise smartphone apps for PA (Conroy et al., 2014) and wearable activity monitors (Mercer et al., 2016). Content was coded for each intervention as a whole (i.e. mHealth and any additional components) using information from relevant results and protocol papers. Coding was completed by two independent reviewers (SAB and AJW) who were trained in Michie et al.’s Behaviour Change Taxonomy v1 (Michie et al., 2013) and consensus was reached through discussion.

RESULTS

Study selection

A flow diagram of the study selection process is shown in **Figure 1**. A total of 2,820 publications were identified in the initial searches (2,815 from databases and five from other sources). After removal of duplicates, 1,897 publication titles and abstracts were screened. The full text was obtained for 71 publications; of these, 18 publications describing 15 studies (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Ganesan et al., 2016, Jones, 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Rowe-Roberts et al., 2014, Schrager et al., 2017, Skogstad et al., 2016, Slootmaker et al., 2009, Thorndike et al., 2014, van Dantzig et al., 2013) met the criteria for inclusion. An updated search to February 2018 found an additional five publications describing four studies (Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017), and a second updated search to December 2018 found an additional seven publications for six studies (Gremaud et al., 2018, Olsen et al., 2018, Patel et al., 2018, Reed et al., 2018, Simons et al., 2018a, Simons et al., 2018b, Torquati et al., 2018), resulting in a total of 30 publications (25 studies) for inclusion in the review.

Figure 1 Flow diagram of study selection process



Study and intervention characteristics

The characteristics of the 25 included studies are summarised in **Table 2**. Eleven studies were conducted in the USA (Losina et al., 2017, Yeung et al., 2017, Jones, 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Schrager et al., 2017, Thorndike et al., 2014, Patel et al., 2018, Gremaud et al., 2018), five in Australia (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Rowe-Roberts et al., 2014, Gilson et al., 2016, Gilson et al., 2017, Olsen et al., 2018, Torquati et al., 2018), two in Canada (Neil-Sztramko et al., 2017, Reed et al., 2018), two in the Netherlands (Slootmaker et al., 2009, van Dantzig et al., 2013), one in Belgium (Simons et al., 2018a, Simons et al., 2018b), Singapore (Finkelstein et al., 2015, Finkelstein et al., 2016), Finland (Reijonsaari et al., 2009, Reijonsaari et al., 2012), Norway (Skogstad et al., 2016), and one in multiple countries (Ganesan et al., 2016). Workplace settings included academic and academic medical institutions (Jones, 2016, Koyle, 2013, Patel et al., 2016a, Schrager et al., 2017, Thorndike et al., 2014, Losina et al., 2017, Yeung et al., 2017, Gremaud et al., 2018, Patel et al., 2018), healthcare (Reed et al., 2018, Torquati et al., 2018), health insurance (Patel et al., 2016b, Rowe-Roberts et al., 2014), wellbeing improvement (Poirier et al., 2016), property and infrastructure (Brakenridge et al., 2016a, Brakenridge et al., 2016b), pension insurance (Reijonsaari et al., 2009, Reijonsaari et al., 2012), financial services (Olsen et al., 2018), road maintenance (Skogstad et al., 2016) and haulage (Gilson et al., 2016, Gilson et al., 2017). Six studies targeted multiple organisations (Finkelstein et al., 2015, Finkelstein et al., 2016, Ganesan et al., 2016, Slootmaker et al., 2009, van Dantzig et al., 2013, Neil-Sztramko et al., 2017, Simons et al., 2018a, Simons et al., 2018b). Both public and private sector organisations were represented.

The number of participants ranged from 20 in a feasibility cohort study (Neil-Sztramko et al., 2017) to over 69,000 in a large international cohort study (Ganesan et al., 2016). The majority of studies targeted sedentary, office-based employees. Sixteen of the 25 studies had a markedly higher proportion of female (≥ 60%) than male participants (Jones, 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Reijonsaari et al., 2012, Rowe-Roberts et al., 2014, Slootmaker et al., 2009, Losina et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017, Gremaud et al., 2018, Olsen et al., 2018, Patel et al., 2018, Reed et al., 2018, Torquati et al., 2018).

The most common study designs were individual randomised controlled trials (RCTs, n = 10) (Finkelstein et al., 2015, Finkelstein et al., 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Slootmaker et al., 2009, van Dantzig et al., 2013) (Gremaud et al., 2018, Patel et al., 2018) and pre-post prospective cohort studies (n = 10) (Ganesan et al., 2016, Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017, Neil-Sztramko et al., 2017, Rowe-Roberts et al., 2014, Schrager et al., 2017, Skogstad et al., 2016, Yeung et al., 2017, Olsen et al., 2018, Torquati et al., 2018). One study used a combination of these designs in two phases (Thorndike et al., 2014). Other designs included cluster RCTs (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Simons et al., 2018b), a parallel group uncontrolled randomised trial (Reed et al., 2018), and a prospective cluster trial with an asynchronous control group (Jones, 2016). Study duration varied greatly, with length of follow-up ranging from six weeks to 12 months.

Assessing the effectiveness, feasibility and/or acceptability of mHealth technology for PA promotion was the primary aim of 16 studies (Finkelstein et al., 2015, Finkelstein et al., 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Rowe-Roberts et al., 2014, Schrager et al., 2017, Skogstad et al., 2016, Thorndike et al., 2014, Losina et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017, Patel et al., 2018, Reed et al., 2018, Simons et al., 2018b). Six studies targeted both PA and SB in a single intervention (Ganesan et al., 2016, Jones, 2016, Slootmaker et al., 2009, Gilson et al., 2016, Gilson et al., 2017, Gremaud et al., 2018, Torquati et al., 2018) and three studies aimed to reduce SB (but also included PA as an outcome measure) (Brakenridge et al., 2016a, Brakenridge et al., 2016b, van Dantzig et al., 2013, Olsen et al., 2018). Although recruitment and/or delivery of the intervention took place in the workplace in all cases, 24 of the 25 studies used mHealth as a wider lifestyle intervention, including both workplace and non-workplace activity. Only one study, designed to reduce SB, was exclusively based in the workplace (van Dantzig et al., 2013).

The main mHealth tools used were wearable activity monitors or trackers (n = 11) (Finkelstein et al., 2015, Finkelstein et al., 2016, Jones, 2016, Poirier et al., 2016, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Rowe-Roberts et al., 2014, Skogstad et al., 2016, Slootmaker et al., 2009, Thorndike et al., 2014, van Dantzig et al., 2013, Losina et al., 2017, Reed et al., 2018), smartphone apps (n = 6) (Ganesan et al., 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Patel et al., 2018, Torquati et al., 2018) or a combination of the two (n = 8) (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Schrager et al., 2017, Gilson et al., 2016, Gilson et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017, Gremaud et al., 2018, Olsen et al., 2018, Simons et al., 2018b). Some studies had additional mHealth and technology intervention components, including motivational or persuasive text messaging (Koyle, 2013, Poirier et al., 2016, van Dantzig et al., 2013) or e-mails (Olsen et al., 2018), computer software or websites linked to the activity monitor (Finkelstein et al., 2015, Finkelstein et al., 2016, Jones, 2016, Poirier et al., 2016, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Skogstad et al., 2016, Slootmaker et al., 2009, Thorndike et al., 2014, van Dantzig et al., 2013, Losina et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017, Reed et al., 2018), and dedicated social media groups (Torquati et al., 2018). Eleven studies assessed mHealth as a standalone intervention (Ganesan et al., 2016, Poirier et al., 2016, Rowe-Roberts et al., 2014, Schrager et al., 2017, Slootmaker et al., 2009, van Dantzig et al., 2013, Yeung et al., 2017, Gremaud et al., 2018, Reed et al., 2018, Simons et al., 2018b, Torquati et al., 2018), whereas 14 studies used mHealth in the context of a multi-component workplace health or PA programme (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Jones, 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Skogstad et al., 2016, Thorndike et al., 2014, Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017, Neil-Sztramko et al., 2017, Olsen et al., 2018, Patel et al., 2018). Among the multi-component programmes in particular, interventions were diverse and additional components included educational materials on health and PA (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Koyle, 2013, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Gilson et al., 2016, Gilson et al., 2017, Neil-Sztramko et al., 2017, Olsen et al., 2018), managerial support (Brakenridge et al., 2016a, Brakenridge et al., 2016b), financial incentives or rewards (Finkelstein et al., 2015, Finkelstein et al., 2016, Patel et al., 2016a, Patel et al., 2016b, Skogstad et al., 2016, Thorndike et al., 2014, Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017, Patel et al., 2018), shared active workstations (Jones, 2016), online or telephone counselling (Reijonsaari et al., 2009, Reijonsaari et al., 2012, Neil-Sztramko et al., 2017), personalised feedback on activity (Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017), wellness education delivered in the workplace (Jones, 2016, Thorndike et al., 2014, Olsen et al., 2018), group-based action planning (Olsen et al., 2018), and access to personal training and nutritionists (Thorndike et al., 2014). Further detail on intervention content is given in **Table 3**.

Intervention duration ranged from one to 12 months. Frequency of delivery of the intervention components was variable but daily wear of activity monitors was encouraged in most studies. Fifteen studies reported that the intervention was based on a named behaviour change theory and/or principles of behavioural economics (Finkelstein et al., 2015, Finkelstein et al., 2016, Jones, 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, van Dantzig et al., 2013, Losina et al., 2017, Neil-Sztramko et al., 2017, Gremaud et al., 2018, Olsen et al., 2018, Patel et al., 2018, Reed et al., 2018, Simons et al., 2018a, Simons et al., 2018b, Torquati et al., 2018). A further two studies alluded to behaviour change techniques or theory in their discussion (Rowe-Roberts et al., 2014, Schrager et al., 2017), and eight studies had no clear theoretical basis (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Ganesan et al., 2016, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Skogstad et al., 2016, Slootmaker et al., 2009, Thorndike et al., 2014, Gilson et al., 2016, Gilson et al., 2017, Yeung et al., 2017). The most frequently cited behaviour change theories were the Theory of Reasoned Action (Azjen and Fishbein, 1980), the Socio-Ecological Model (Robinson, 2008), Social Cognitive Theory and Self-Efficacy (Bandura, 1997), Self-Determination Theory (Deci and Ryan, 2000), other social influence theories such as self-presentation theory (Leary, 1992) and Cialdini’s social influence strategies (Cialdini, 2001), and the Health Action Process Approach (Schwarzer et al., 2011).

A control or comparator group was present in 14 of the 25 studies (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Jones, 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Slootmaker et al., 2009, Thorndike et al., 2014, van Dantzig et al., 2013, Gremaud et al., 2018, Patel et al., 2018, Simons et al., 2018b). Of these, six could not be classed as a 'true' control group as the participants received at least a partial mHealth intervention (Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, van Dantzig et al., 2013, Gremaud et al., 2018, Patel et al., 2018), and another three studies supplied controls with wearable activity monitors for data collection (Jones, 2016, Poirier et al., 2016, Thorndike et al., 2014).

Outcome measures of PA and SB were heterogeneous. The most frequently used outcome measures for PA included daily step count, daily or weekly minutes or metabolic equivalent (MET) minutes of total activity or moderate-to-vigorous PA (MVPA). Other outcomes included exercise frequency and proportion of participants meeting step or PA goals. Studies that assessed SB commonly reported daily or weekly sedentary time, although the one study using an exclusive workplace intervention used computer activity as a proxy for SB (van Dantzig et al., 2013). Objective PA/SB outcomes were used in 20 studies (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Jones, 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Rowe-Roberts et al., 2014, Thorndike et al., 2014, van Dantzig et al., 2013, Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017, Gremaud et al., 2018, Olsen et al., 2018, Patel et al., 2018, Reed et al., 2018, Simons et al., 2018b, Torquati et al., 2018) whilst four studies relied on self-report for the primary measure of PA or SB (Reijonsaari et al., 2009, Reijonsaari et al., 2012, Schrager et al., 2017, Skogstad et al., 2016, Slootmaker et al., 2009). Ganesan and colleagues used pedometer data that was self-entered by participants (Ganesan et al., 2016).

Table 2 Characteristics of included studies

Author Year Country	Setting / workplace	Study design	Participants	Type of mHealth technology / tool[a]	Intervention	Control / comparison group(s)	Aim	Primary PA/SB outcome(s) (objective, **OB** or self-reported, **SR**)	Secondary outcome(s)	Length of follow-up
Brakenridge et al. 2016a (protocol) Brakenridge et al. 2016b (results) Australia	International property and infrastructure group (Lendlease)	Cluster RCT	n = 153 54% M, 46% F Age IG: 37.6±7.8 CG: 40.0±8.0 Office workers (at least 0.5 FTE)	Wearable activity monitor and smartphone app	Waist-worn 'LUMOback' activity monitor (LUMO Bodytech, USA) and associated smartphone app with organisational support.	Organisational support only – e.g. manager support, e-mails and educational materials.	SB	Average time per day spent sitting (work hours and overall) (**OB**, activPAL3™ accelerometer)	Average time per day spent in prolonged sitting bouts (30 minutes or more), standing and stepping Daily steps Average time period between sitting bouts Job performance, job control and work satisfaction Stress, physical and mental health-related QoL Activity monitor usage	12 months
Finkelstein et al. 2015 (protocol and baseline data) Finkelstein et al. 2016 (results) Singapore	13 organisations (various industries and government sectors)	RCT (4 arm)	n = 800 46% M, 54% F Age IG1: 35.4±8.3 IG2: 35.5±8.6 IG3: 35.5±8.4 CG: 35.6±8.6 Mostly desk-based employees (full-time)	Wearable activity monitor (and website)	Waist-worn Fitbit Zip activity monitor (Fitbit, USA) and associated website. Monetary incentives: IG1 = Fitbit only IG2 = Fitbit and charity donation IG3 = Fitbit and cash Educational booklets on PA.	No activity monitor or incentives. Educational booklets on PA only.	PA	MVPA bout minutes per week (**OB**, ActiGraph™ GT-3x+ accelerometer)	Mean daily step count % of participants meeting 70,000 weekly step goal Weight Systolic BP Cardiorespiratory fitness Quality of life Weekly step count Sedentary, light, moderate and vigorous PA (minutes/week) Participants meeting 150 minutes per week moderate PA Participants meeting 10,000 daily step target	12 months
Ganesan et al. 2016 64 countries (majority of participants from India	481 employers (private and public sector organisations) in 1481 cities	Prospective cohort (pre- and post-uncontrolled)	n = 69,219 76% M, 24% F Age 36.0±8.4	Smartphone app	Non-interactive pedometer and 'Stepathlon' mobile app (also available as website).	No control or comparison group	PA + SB	Daily steps (**SR**, pedometer data entered by participants)	Number of exercise days/week Exercise duration (<30 or ≥30 minutes/day) Sitting duration (hours/day) Weight in kilograms	100 days (approx.)

(90.2%), Australia (5%), New Zealand (1.1%) and Singapore (0.6%))			Adult employees							
Gilson et al. 2016 (baseline data and smartphone use) Gilson et al. 2017 (results) Australia	Two large Australian haulage companies	Prospective cohort (pre- and post- uncontrolled feasibility study)	n = 44 100% M, 0% F Age 47.0±10.1 Truck drivers	Wearable activity monitor and smartphone app	Wrist-worn Jawbone UP activity monitor (Jawbone, USA) used with associated smartphone app. Monetary incentives (vouchers for attaining step goals and logging diet)	No control or comparison group	PA + SB	Proportions of work time and non-work time spent physically active, sedentary and stationary+ (i.e. sitting with upper limb movement or standing) (**OB** – GENEActiv™ wrist accelerometer)	Workday diet (fruit, vegetable, saturated fat and sugar intake) Engagement with the intervention Qualitative outcomes (interviews) –driver and depot manager experiences; perceived impact of the intervention; barriers to PA	28 weeks (approx.)
Gremaud et al. 2018 USA	Academic organisation (university)	RCT	n = 146 25% M, 75% F Age IG: 40.6±11.7 CG: 40.3±11.1 Sedentary office workers (full-time)	Wearable activity monitor and smartphone app (web-based)	Waist-worn Fitbit Zip activity monitor (Fitbit, USA) used with 'MapTrek' app for gamified walking.	Activity monitor only	PA + SB	Daily steps Daily active minutes (minutes with ≥100 steps/min) (**OB** – data from Fitbit)	Bouts of sedentary behaviour (consecutive minutes with 0 steps)	10 weeks
Jones 2016 USA	Academic medical centre (Wake Forest Baptist Health)	Prospective cluster trial (with asynchronous control group)	n = 47 18% M, 82% F Age Overall mean = 50.8, range 25 to 74 years (SD not reported) Sedentary employees	Wearable activity monitor (and computer software)	Clip-on Fitbit One activity monitor (Fitbit, USA) and associated software, with wellness education. IG1 = Fitbit only IG2 = Fitbit and shared active workstations	Usual treatment (blinded activity monitors for data collection)	PA + SB	Daily steps Daily sedentary time BMI (**OB** – steps and sedentary time from Fitbit. BMI assessed objectively)	Life satisfaction Anxiety (state and trait) Health-related quality of life Self-reported sleep patterns	6 months
Koyle 2013 USA	Academic medical centre	RCT	n = 73 0% M, 100% F	Smartphone app with integrated	'Adidas miCoach' smartphone app to	Smartphone app and educational materials (same	PA	Walking distance and duration	Walking for exercise self-efficacy beliefs	6 weeks

	(University of Utah Health Care)		Age 46.5±7.6 Physically inactive employees (<150 minutes exercise per week)	accelerometer (and motivational text messages)	track walking exercise. Educational materials on PA. Tailored motivational text messages.	as intervention group). No motivational text messages.		(**OB** – smartphone app-integrated accelerometer for collection of PA data)	Likeliness of participating in other forms of PA beyond walking Height and weight (BMI) Resting pulse rate Systolic BP Qualitative experiences of the intervention (survey)	
Losina et al. 2017 USA	Academic medical centre (Brigham and Women's Hospital, Boston, Massachusetts)	Prospective cohort (pre- and post- uncontrolled feasibility study)	n = 292 17% M, 83% F Age 38±11 Sedentary, non-clinician hospital employees	Wearable activity monitor (linked with websites)	Wrist-worn Fitbit Flex activity monitor (Fitbit, USA) used with Fitbit website and study website for monitoring PA and progress. Monetary incentives (individual and team) for meeting PA goals.	No control or comparison group	PA	Average weekly minutes of MVPA Proportion of participants meeting weekly MVPA goals and CDC PA guidelines (**OB** – step data from Fitbit converted to weekly minutes of MVPA)	Fitbit wear adherence (number of weeks wearing Fitbit for ≥10 hours/day and ≥4 days/week) Participant satisfaction with programme	26 weeks (including two pre-intervention weeks)
Neil-Sztramko et al. 2017 Canada	Multiple workplaces in Greater Vancouver (nursing, emergency services, casinos and airport)	Prospective cohort (pre- and post- uncontrolled feasibility study)	n = 20 0% M, 100% F Age 42.2±8.6 Female shift workers	Wearable activity monitor and smartphone app (or website)	Wrist-worn Fitbit Flex activity monitor (Fitbit, USA) used with Fitbit app and/or website. Distance-based behavioural counselling (telephone/online)	No control or comparison group	PA	MVPA (total minutes/week and minutes/week bouts ≥10 mins) (**OB** - ActiGraph^{TM} GT-3x+ accelerometer)	Daily steps Sedentary time (minutes/week bouts ≥10 mins) Self-reported PA and sedentary time Body weight and BMI Physical and mental health-related QoL Sleep quantity and quality Feasibility outcomes: demand (reach and recruitment), implementation (delivery and resources) and acceptability (attrition and adherence to intervention, participant satisfaction).	12 weeks

Olsen et al. 2018 Australia	Financial services organisation (Brisbane)	Prospective cohort (pre- and post-uncontrolled pilot)	n = 49 31% M, 69% F[b] Age 39.5±8.7 Flexible workers (e.g. work from home one day/week)	Wearable activity monitor and smartphone app	Wrist-worn Jawbone activity monitor (Jawbone, USA) used with associated app. Group-based action planning session. Weekly e-mail reminders and resources. Healthy living seminar.	No control or comparison group	SB	Sitting time (including overall and occupational, minutes/day) (**OB** - ActiGraph™ GT-3x+ accelerometer, also self-reported sitting time assessed using adapted version of Workforce Sitting Questionnaire)	Light PA and MVPA (minutes/day, accelerometer-assessed) Self-reported PA (minutes/week, assessed using adapted version of Active Australia survey) Acceptability of the intervention (survey-assessed)	6 weeks
Patel et al. 2016a (study 1) USA	Academic organisation (University of Pennsylvania)	RCT (4 arm)	n = 281 (279 completed baseline assessment) 22% M, 78% F Age IG1: 37.1±10.9 IG2: 40.3±11.2 IG3: 41.9±11.6 CG: 39.4±12.2 Overweight and obese employees (BMI ≥27 kg/m^2)	Smartphone app with integrated accelerometer	'Moves' smartphone app (Proto Geo Oy, Finland) for step tracking. Daily feedback on steps. Monetary incentives: IG1 = gain incentive IG2 = lottery incentive IG3 = loss incentive	Smartphone app and daily feedback (as intervention group). No financial incentives.	PA	Proportion of participant-days 7000 step goal achieved during intervention (**OB** – smartphone app-integrated accelerometer)	Proportion of participant-days 7000 step goal achieved during follow-up Daily steps – intervention and follow-up	26 weeks
Patel et al. 2016b (study 2) USA	Health insurance organisation (Independence Blue Cross)	RCT (4 arm)	n = 304 23% M, 77% F Age IG1: 39.3±10.2 IG2: 38.7±10.2 IG3:41.2±10.8 CG: 43.2±10.0 Mostly sedentary employees	Smartphone app with integrated accelerometer	'Moves' smartphone app (Proto Geo Oy, Finland) for step tracking. Daily feedback on steps. Monetary incentives: IG1 = individual IG2 = team IG3 = combined	Smartphone app and daily feedback (as intervention group). No financial incentives.	PA	Proportion of participant-days 7000 step goal achieved during intervention (**OB** – smartphone app-integrated accelerometer)	Proportion of participant-days 7000 step goal achieved during follow-up Daily steps – intervention and follow-up	26 weeks
Patel et al. 2018	Academic organisation	RCT (4 arm)	n = 209 23% M, 77% F	Smartphone app with	'Moves' smartphone app (Proto Geo Oy,	Smartphone app and daily feedback (as	PA	Proportion of participant-days 7000	Proportion of participant-days 7000 step goal	26 weeks

USA	(University of Pennsylvania)		Age IG1: 41.2±11.1 IG2: 40.6±10.5 IG3: 42.9±10.3 CG: 40.0±11.0 Overweight and obese employees (BMI ≥27 kg/m^2)	integrated accelerometer	Finland) for step tracking. Daily feedback on steps. Monetary incentives: IG1 = higher frequency, smaller reward lottery IG2 = jackpot lottery IG3 = combined lottery	intervention group). No financial incentives.		step goal achieved during intervention (**OB** – smartphone app-integrated accelerometer)	achieved during follow-up Daily steps – intervention and follow-up	
Poirier et al. 2016 USA	Wellbeing improvement company (Healthways Inc)	RCT	n = 265 34% M, 66% F Age IG: 40.3±11.4 CG: 39.6±12.0 Headquarter-based employees	Wearable activity monitor (linked with website, and optional text messages)	Hip- or shoe-worn Pebble+ activity monitor (Fitlinxx Inc, USA) used with 'Walkadoo' internet-based program. Electronic messaging.	One week of blinded activity monitor wear, then instructed to maintain their usual activity routine.	PA	Daily steps (**OB** – activity monitor and website)	Proportion of participants increasing steps by 1000/day Engagement with intervention – wear time, e-mail opening and website visits	6 weeks
Reed et al. 2018 Canada	Tertiary care cardiovascula r institute (University of Ottawa Heart Institute)	Parallel-group randomised trial (3 arm)	n = 76 3% M, 97% F Age 46.3±10.9 Nurses	Wearable activity monitor (linked with website)	Ankle-worn Tractivity® activity monitor (Tractivity®, Vancouver, BC) linked with website for monitoring PA and taking part in challenges: IG1 = individual challenge IG2 = friend challenge IG3 = team challenge	No control or comparison group	PA	MVPA (weekly minutes in bouts ≥10 mins) Daily steps (**OB** – data from Tractivity® activity monitor)	Body mass (kg) BMI Waist circumference Body fat % Resting systolic BP	6 weeks
Reijonsaari et al. 2009 (protocol) Reijonsaari et al. 2012 (results)	Insurance company	RCT	n = 544 (521 included in analysis) 36% M, 64% F Age IG: 43±10.0	Wearable activity monitor (linked with website)	Belt-worn 'AM 200' activity monitor/ accelerometer (PAM BV, Netherlands) used with associated website.	Educational materials on PA. Written results of fitness tests.	PA	Weekly MET-minutes of total activity Work productivity Sickness absence (**SR** – MET-minutes from IPAQ,	Body weight (kg) Waist circumference (cm) Body fat percentage Systolic and diastolic BP (mmHg)	12 months

Finland			CG: 44±10.0 Mainly clerical employees (working ≥8 hours per week)		Educational materials on PA. Written results of fitness tests. Distance counselling (telephone/online)			productivity from QQ instrument but objective sickness absence data)	Aerobic fitness (maximal oxygen uptake, VO_2 max, ml/kg/min)	
Rowe-Roberts et al. 2014 Australia	Private healthcare and insurance company (Australian Unity group)	Prospective cohort (uncontrolled pilot)	n = 212 38% M, 62% F Age 42% under 35 35% 35-44 15% 45-54 8% 55+ Adult employees	Wearable activity monitor	Waist-worn Fitbit Ultra activity monitor (Fitbit, USA)	No control or comparison group	PA	Daily steps (**OB** – step data from Fitbit)	AUSDRISK (Australian Type 2 Diabetes Risk Assessment Tool) score Engagement with intervention (activity monitor wear) Qualitative outcomes (survey and focus groups) – experiences, engagement and activity, preferred motivational strategies	7 months
Schrager et al. 2017 USA	Academic emergency medicine residency	Prospective cohort (pre- and post-uncontrolled pilot)	n = 30 53% M, 47% F Age Median 28 years (IQR = 4.0) Physicians on a single site	Wearable activity monitor and smartphone app (or website)	Wrist-worn Fitbit Flex activity monitor (Fitbit, USA) used with Fitbit app and/or website	No control or comparison group	PA	Days per week with ≥30 minutes PA (**SR**)	Days per week with ≥10,000 steps or ≥30 minutes of active time (as measured by Fitbit at one month) Qualitative outcomes (survey) – adoption and use of device, measures of wellness, changes in PA behaviour	6 months
Simons et al. 2018a (app development and feasibility) Simons et al. 2018b (results of RCT) Belgium	Multiple workplaces in Flanders, Belgium (including retail, catering, social employment and factories)	Study 2, 2018b = Cluster RCT (study 1, 2018a was a qualitative evaluation and impact on PA/SB not reported)	n = 130 (29 clusters) 49% M, 51% F Age IG: 24.8±3.1 CG: 25.1±3.0 Lower educated (i.e. no university or college degree) working young adults, not meeting PA	Wearable activity monitor and smartphone app	Wrist-worn Fitbit Charge activity monitor (Fitbit, USA) used with 'Active Coach' app for monitoring PA	Educational brochure on PA only (generic information)	PA	Daily minutes of light, moderate and vigorous intensity PA (**OB** - ActiGraph™ GT-3x+ accelerometer)	Daily steps (from Fitbit) Self-reported context-specific PA (IPAQ) Psychosocial variables: social support, attitude (perceived benefits and barriers), self-efficacy, knowledge and intentions Engagement: usage statistics Process evaluation interviews: Opinions on Fitbit and app (e.g. usability, preferred features)	21 weeks

			guidelines at baseline (<150 minutes MVPA/week)							
Skogstad et al. 2016 Norway	Road maintenance enterprise	Prospective cohort (pre- and post-uncontrolled)	n = 121 64% M, 36% F Age M = 41.8±12.0 F = 42.6±12.5 24% road workers, 76% office workers	Wearable activity monitor (linked with website)	Wrist-worn Tappa® activity monitor/ accelerometer used with associated website (Dytt®) for step tracking. Rewards given for best performances.	No control or comparison group	PA	Weekly exercise frequency and duration (**SR**)	Aerobic fitness (maximal oxygen uptake, VO_2 max, ml/kg/min) Systolic and diastolic BP (mm Hg) Resting heart rate Lipids (total, HDL and LDL cholesterol) C-reactive protein (CRP) Glycosylated haemoglobin (HbA1c)	8 weeks (approx.)
Slootmaker et al. 2009 Netherlands	8 worksites surrounding Amsterdam (mainly office settings)	RCT	n = 102 40% M, 60% F Age IG: 32.5±3.4 CG: 31.2±3.5 Mainly office workers	Wearable activity monitor (linked with website)	Belt-worn 'AM 101' activity monitor/ accelerometer (PAM BV, Netherlands) used with associated website (PAM COACH).	Educational booklet on PA only	PA + SB	Weekly PA and sedentary time – weekly minutes of light, moderate and vigorous intensity activity and sedentary minutes (**SR** – assessed by the AQuAA questionnaire)	Self-reported determinants of PA – including behavioural intention, attitude, social influence, self-efficacy, knowledge of PA recommendations Aerobic fitness (maximal oxygen uptake, VO_2 max, ml/kg/min) Body composition – body weight and height (BMI), waist and hip circumference, skin fold thickness (% body fat)	8 months
Thorndike et al. 2014 USA	Healthcare organisation (Massachusetts General Hospital)	Phase 1 = RCT Phase 2 = team-based prospective cohort (pre- and post-uncontrolled)	n = 104 46% M, 54% F Age Mean and range (SD not reported) IG: 29 (23-36) CG: 29 (25-37) Physicians-in-training	Wearable activity monitor (linked with website)	Fitbit activity monitor (Fitbit, USA) used with Fitbit website. Gift card lottery for wearing device (phase 1) and highest steps (phase 2). Workplace initiatives: access to fitness centres, personal training and nutritionists,	Phase 1 – blinded activity monitor (no access to website). Gift card lottery and workplace initiatives (as intervention group). Phase 2 – no control or comparison group	PA	Daily step count (phase 1 median and mean steps/day, phase 2 mean steps/day) (**OB** – step data from Fitbit)	Proportion of days activity monitor was worn (i.e. compliance) Weight BMI Waist circumference Systolic and diastolic BP Lipids (total, HDL and LDL cholesterol) Use/engagement with the wider wellness programme (e.g. fitness centre, nutrition)	12 weeks

					weekly healthy lunch					
Torquati et al. 2018 Australia	Two metropolitan hospitals in Brisbane (private and public)	Prospective cohort (pre- and post- uncontrolled pilot)	n = 47 13% M, 87% F Age 41.4±12.1 Nurses and nursing managers	Smartphone app	Smartphone app for PA and diet with non-interactive pedometer and dedicated Facebook group	No control or comparison group	PA + SB	Time spent sedentary and in light activity and MVPA Daily steps (**OB** - ActiGraph™ GT-3x+ accelerometer)	Diet behaviour: Food Frequency Questionnaire (FFQ) Weight BMI Waist circumference Blood pressure Self-rated health PA and diet self-efficacy Social support Feasibility outcomes (including qualitative interviews): reach, adoption (use) and implementation	6 months
van Dantzig et al. 2013 Netherlands	Offices at various companies in Netherlands (no further detail given)	Experiment 2 = RCT (experiment 1 was a small qualitative evaluation and impact on PA/SB not reported)	n = 86 60% M, 40% F Age IG: 44.5±7.9 CG: 44.3±8.0 Sedentary office workers	Wearable activity monitor (linked with website, and persuasive text messages)	Clip-on commercial activity monitor (tri-axial accelerometer, model not stated) linked with personal web page for viewing PA data. Timely persuasive text messages on smartphones.	Activity monitor only. No text messages.	SB	Computer activity (minutes, proxy for SB) Physical activity (minutes) (**OB** – computer activity from computer software; PA from activity monitor)	Engagement with the intervention (proportion of text messages read)	6 weeks
Yeung et al. 2017 USA	Academic hospital residency (Cincinnati, Ohio)	Prospective cohort (pre- and post- uncontrolled crossover study)	n = 86 38% M, 62% F[b] Age[b] 62% 21-30 31% 31-40 5% 41-50 Internal medicine residents	Wearable activity monitor and smartphone app (or website)	Wrist-worn Fitbit Flex (Fitbit, USA) used with Fitbit app and/or website for monitoring steps (weeks 1-4 blinded, weeks 5-8 unblinded). Optional in-app activity tracking group for weeks 5-8.	No control or comparison group	PA	Daily steps (comparison of blinded vs. unblinded periods) (**OB** – step data from Fitbit)	Proportion of participants achieving ≥10,000 steps/day	8 weeks

Notes:

M = Male; F = Female; IG = Intervention Group; CG = Control/Comparison Group; FTE = Full Time Equivalent; PA = Physical Activity; SB = Sedentary Behaviour; OB = Objective; SR = Self-Reported; QoL = Quality of Life; RCT = Randomised Controlled Trial; ± or SD = Standard Deviation; MVPA = Moderate-to-Vigorous Physical Activity; BP = Blood Pressure; BMI = Body Mass Index; MET = Metabolic Equivalent; IPAQ = International Physical Activity Questionnaire; QQ = Quantity and Quality questionnaire; IQR = Interquartile Range; HDL = High Density Lipoprotein; LDL = Low Density Lipoprotein; AQuAA = Activity Questionnaire for Adolescents and Adults; CDC = Centers for Disease Control and Prevention

[a] Tools may be referred to as activity monitors or trackers in the literature; the term 'monitor' is used here for consistency

[b] Yeung et al. and Olsen et al. report gender and age of study completers only

Table 3 Summary of intervention components

Study	Intervention description Standalone mHealth (**SA**) or multi-component programme (**MC**)?	Duration of intervention	Frequency of intervention (if applicable)	Theoretical basis of intervention	Behaviour change techniques (BCTs) included (whole intervention)								
					Goal-setting (behaviour or outcome)	**Self-monitoring** (behaviour or outcome)	**Prompts / cues**	**Feedback** (behaviour or outcome)	**Rewards / incentives** (virtual or real, progress towards or achieving behaviour)	**Social support** (online or offline)	**Social comparison**	**Information on consequences of behaviour** (general or individual)	**Other BCTs**
Brakenridge et al. 2016a; Brakenridge et al. 2016b	Wearable activity monitor and smartphone app for feedback on sitting, standing, stepping, sitting breaks, posture and sleep. **MC**: organisational support (e-mails, educational materials)	12 months (although main focus, e.g. e-mails was in the first 12 weeks)	*Ad lib* wear of activity monitor and use of smartphone app. Fortnightly e-mails for first 12 weeks.	None stated	**?**	**Y**	**Y**	**Y**	**N**	**N**	**N**	**Y**	Information on where and when to perform behaviour
Finkelstein et al. 2015; Finkelstein et al. 2016	Wearable activity monitor used with website including social components. **MC:** 2 of the 3 intervention groups earned weekly incentives for step count (cash vs. charity) Educational booklets on PA.	6 months	Daily wear of activity monitor. Incentives for weekly step counts distributed every 4-6 weeks.	Economic theory and Theory of Reasoned Action	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	
Ganesan et al. 2016	'Stepathlon' mobile app (also available as website). Participants	100 days	Daily e-mails to encourage	None stated	**?**	**Y**	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	

	self-entered daily activity data and steps from pedometers. Included personalised tools for self-monitoring PA and dietary intake, quizzes, motivational e-mails, online community and chat for interactive advice / expert guidance, social networking, competitions and health information. Also gamification – race around a virtual world map. **SA**		daily activity data entry.										
Gilson et al. 2016; Gilson et al. 2017	Wearable activity monitor used with associated smartphone app for monitoring steps and diet. **MC**: Part of the 'Shifting Gears' programme. Earning points and financial rewards for attaining step goals and logging diet. Educational materials on PA and diet. Personalised feedback and guidance from researchers.	20 weeks	*Ad lib* wear of activity monitor and use of smartphone app. Personalised feedback and guidance from researcher every 4 weeks. Rewards at end of programme.	None stated	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	**N**	**Y**	Information on where and when to perform behaviour Action planning Graded tasks
Gremaud et al. 2018	Wearable activity monitor used with 'MapTrek' app (smartphone or web-based) for gamification of walking, including virtual 'avatar' and	10 weeks	Daily wear of activity monitor. *Ad lib* use of smartphone app.	Cognitive evaluation/self-determination theories and Social Cognitive Theory	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**N**	

	races, including a social competitive element and automated text messages based on PA. **SA**												
Jones 2016	Wearable activity monitor used with software installed on work computers. Included monitoring progress towards step and activity goals, competing against colleagues and earning incentives and awards for meeting step targets. **MC:** One intervention group used shared active workstations. Group wellness education delivered in the workplace.	6 months	Daily wear of activity monitor. Use of active desks for at least 30 minutes a day, 5 days a week. Monthly wellness education.	Socio-Ecological Model	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	**Y**	**Y**	Environmental restructuring
Koyle 2013	‘Adidas miCoach’ smartphone app with integrated accelerometer. Motivational text messages tailored based on initial app-delivered fitness test and walking logs. **MC:** Part of a health promotion programme including educational materials on PA.	6 weeks	*Ad lib* use of smartphone app. Weekly motivational text messages.	Self-Efficacy Theory (part of Social Cognitive Theory)	**Y**	**Y**	**Y**	**Y**	**N**	**N**	**N**	**Y**	Action planning
Losina et al. 2017	Wearable activity monitor used with websites to monitor steps and PA,	24 weeks	Daily wear of activity monitor.	Behavioural economic theory	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	**N**	**N**	Graded tasks Shaping

	including individual and team progress. **MC:** Monetary incentives (individual and team) for meeting PA goals. Personalised weekly e-mails for feedback on MVPA.		Weekly financial rewards. Weekly feedback e-mails.										
Neil-Sztramko et al. 2017	Wearable activity monitor used with associated app and/or website to monitor steps and PA. **MC**: Individualised behavioural counselling (telephone/online – all participants chose telephone).	12 weeks	*Ad lib* wear of activity monitor and use of smartphone app. 8 counselling sessions over 12 weeks.	Health Action Process Approach (HAPA)	**Y**	**Y**	**N**	**Y**	**N**	**Y**	**N**	**Y**	Barrier identification/proble m solving
Olsen et al. 2018	Wearable activity monitor used with associated app for self-monitoring, social support and prompts to reduce SB. **MC:** Group-based goal-setting, action planning and problem solving session delivered in workplace. Weekly e-mail reminders and information resources. Healthy living seminar delivered in workplace (week 4).	6 weeks	*Ad lib* wear of activity monitor and use of smartphone app. Weekly e-mails. One group-based session and one seminar.	Social Cognitive Theory	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**Y**	**Y**	Action planning Barrier identification/proble m solving Information on where and when to perform behaviour Instruction on how to perform behaviour
Patel et al. 2016a	‘Moves’ smartphone app with integrated accelerometer. Feedback on steps.	13 weeks	*Ad lib* use of smartphone app (instructed to carry phone when active).	Behavioural economics – immediate vs. delayed gratification,	**Y**	**?**	**N**	**Y**	**Y**	**N**	**N**	**N**	Action planning Prompt anticipated regret

	MC: The three intervention groups received differently framed financial incentives for achieving step goals (gain, lottery and loss).		Daily feedback and incentives.	prospect theory and regret aversion.									
Patel et al. 2016b	'Moves' smartphone app with integrated accelerometer. Feedback on steps. **MC:** The three intervention groups received differently framed financial incentives for achieving step goals (individual, team and combined).	13 weeks	*Ad lib* use of smartphone app (instructed to carry phone when active). Daily feedback and incentives.	Behavioural economics; variable reinforcement; social behaviour change theories	**Y**	**?**	**N**	**Y**	**Y**	**N**	**Y**	**N**	Action planning
Patel et al. 2018	'Moves' smartphone app with integrated accelerometer. Feedback on steps. **MC**: The three intervention groups received differently framed financial incentives for achieving step goals (high frequency small reward lottery, jackpot lottery and combined lottery).	13 weeks	*Ad lib* use of smartphone app (instructed to carry phone when active). Daily feedback and incentives.	Behavioural economics – immediate vs. delayed gratification, prospect theory and regret aversion.	**Y**	**?**	**N**	**Y**	**Y**	**N**	**N**	**N**	Action planning Prompt anticipated regret
Poirier et al. 2016	Wearable activity monitor used with website (data transmitted wirelessly) for monitoring steps. Tailored goals based on activity. Motivational electronic messages (optional). Virtual rewards, social	6 weeks	Daily wear of activity monitor. Messages up to 4 times/day (optional).	Tailored, adaptive goals based on behavioural economics and operant shaping.	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**Y**	**N**	Shaping Information about others' approval

	messaging and competitions. **SA**												
Reed et al. 2018	Wearable activity monitor linked with website for monitoring PA and steps and taking part in one of three challenges (individual, friend or team) **SA**	6 weeks	Daily wear of activity monitor. *Ad lib* use of website.	Social behaviour change theories including self-presentation theory	**?**	**Y**	**N**	**Y**	**N**	**N**	**Y**	**N**	
Reijonsaari et al. 2009; Reijonsaari et al. 2012	Wearable activity monitor used with website for goal-setting and PA monitoring. **MC:** Personalised distance counselling and support (online and telephone). Written information on fitness test results, PA and health.	12 months	Daily wear of activity monitor. Phone call or message from coach if did not log on to website every two weeks to upload PA data.	None stated	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**N**	**Y**	
Rowe-Roberts et al. 2014	Wearable activity monitor for monitoring steps. Participants were offered an optional additional device for friends/family midway through the intervention for social support and competition. **SA**	7 months	Daily wear of activity monitor.	None stated (although mentioned Transtheoretical Model in discussion)	**N**	**Y**	**N**	**N**	**N**	**Y**	**Y**	**N**	
Simons et al. 2018a; Simons et al. 2018b	Wearable activity monitor used with 'Active Coach' smartphone app for goal-setting, self-monitoring, and	9 weeks	Daily use of activity monitor and app encouraged	Self-determination theory (and BCTs identified during app development)	**Y**	**Y**	**Y**	**Y**	**N**	**Y**	**N**	**Y**	Barrier identification/problem solving Graded tasks

	tailored information to promote PA **SA**												Information on where and when to perform behaviour Instruction on how to perform behaviour
Skogstad et al. 2016	Wearable activity monitor used with associated website to monitor PA and compete against colleagues individually and in teams. Gamification – virtual internet mountain track. **MC:** Part of an organised 8-week workplace PA programme ('Dytt®'). Rewards given for best performances.	8 weeks	Daily wear of activity monitor and *ad lib* use of website. Weekly PA results posted on intranet. Rewards at end of programme.	None stated	**N**	**Y**	**N**	**N**	**Y**	**N**	**Y**	**N**	
Slootmaker et al. 2009	Wearable activity monitor used with website to monitor PA (data uploaded via docking station and software on work computers). Website provided tailored goal-setting and PA advice. Comparison of PA scores with peers. **SA**	3 months	Daily wear of activity monitor. *Ad lib* use of website.	None stated	**Y**	**Y**	**N**	**Y**	**N**	**N**	**Y**	**Y**	Action planning Graded tasks
Thorndike et al. 2014	Wearable activity monitor used with website to monitor steps, PA, weight and diet. Included gamification – virtual 'avatar' on activity monitor screen that changed size with	12 weeks (phase 1 = 6 weeks, phase 2 = 6 weeks)	Daily wear of activity monitor and *ad lib* use of website. Weekly e-mail reminders to charge device and with	None stated	**N**	**Y**	**N**	**Y**	**Y**	**N**	**Y**	**N**	

	varying level of PA / SB. Phase 1 = individual monitoring Phase 2 = team-based steps competition **MC:** Part of a 10-week, team-based worksite wellness programme ('BeFit') – included access to personal training and nutritionists. Lottery to reward device wear and attainments (highest steps).		details of gift card lottery.										
Torquati et al. 2018	Smartphone app for PA and diet (goal-setting) with non-interactive pedometer (self-monitoring) and dedicated Facebook group (social support) **SA**	3 months	*Ad lib* use of smartphone app and Facebook group	Social Cognitive Theory, goal-setting theory and control theory	**Y**	**Y**	**N**	**N**	**N**	**Y**	**Y**	**N**	
van Dantzig et al. 2013	Wearable activity monitor linked with personal web page to monitor activity patterns. Timely, persuasive text messages on smartphone during prolonged periods of sitting (detected by computer software). **SA**	6 weeks	Daily wear of activity monitor. Text messages sent every 30 minutes of uninterrupted computer activity (up to a maximum of three messages/ day).	Intervention based on four of six social influence strategies (Cialdini, 2001) – authority, commitment, consensus and scarcity	**N**	**Y**	**Y**	**Y**	**N**	**N**	**N**	**N**	

Yeung et al. 2017	Wearable activity monitor used with associated app and/or website for monitoring steps (weeks 1-4 blinded, weeks 5-8 unblinded). Optional resident-only activity tracking group to 'connect and compete' with peers for weeks 5-8. **SA**	8 weeks	Daily wear of activity monitor. *Ad lib* use of app/website.	None stated	**Y**	**Y**	**N**	**Y**	**N**	**N**	**Y optional**	**N**	Action planning

Notes:

BCT = Behaviour Change Technique; SA = Standalone mHealth; MC = Multi-Component programme; Y = Yes, included in intervention; N = No, not included in intervention; ? = unclear or difficult to determine whether included from available intervention description; PA = Physical Activity; SB = Sedentary Behaviour; MVPA = Moderate-to-Vigorous Physical Activity

Study quality

A summary of the risk of bias and quality assessment for the included studies is presented in **Table 4**. Using the EPHPP tool, only one study was judged as ‘strong’ quality (Thorndike et al., 2014). Nine studies were assigned a ‘moderate’ quality rating (Finkelstein et al., 2015, Finkelstein et al., 2016, Koyle, 2013, Patel et al., 2016a, Poirier et al., 2016, Slootmaker et al., 2009, Gremaud et al., 2018, Patel et al., 2018, Reed et al., 2018, Simons et al., 2018b) and 15 studies were given a ‘weak’ rating (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Ganesan et al., 2016, Jones, 2016, Patel et al., 2016b, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Rowe-Roberts et al., 2014, Schrager et al., 2017, Skogstad et al., 2016, van Dantzig et al., 2013, Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017, Olsen et al., 2018, Torquati et al., 2018). All except two studies (Schrager et al., 2017, Thorndike et al., 2014) were judged as ‘weak’ in terms of selection bias; participants were typically self-selected employees who volunteered to take part in a wellness programme. Representativeness and level of participation were unclear in many of the included studies.

All 25 studies used robust experimental or quasi-experimental designs. Of the 25 studies, 15 reported controlling for important confounders in their design and/or analysis. Ten studies were assigned a ‘weak’ rating in this domain due to lack of reporting or poor control of confounders in analysis (Jones, 2016, Patel et al., 2016b, Rowe-Roberts et al., 2014, Schrager et al., 2017, Gilson et al., 2017, Losina et al., 2017, Neil-Sztramko et al., 2017, Yeung et al., 2017, Olsen et al., 2018, Torquati et al., 2018). No studies received a ‘strong’ rating for blinding due to the difficulty and impracticality in blinding participants to this type of mHealth intervention. Blinding of outcome assessors was often not described, and two studies were rated as ‘weak’ in this domain as outcome assessors were reported to be unblinded (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Reijonsaari et al., 2009, Reijonsaari et al., 2012).

A ‘strong’ data collection method for the main PA/SB outcome was used by 14 studies; this included research-grade accelerometers (e.g. activPAL™, Actigraph™, GENEActiv™) (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Gilson et al., 2016, Gilson et al., 2017, Neil-

Sztramko et al., 2017, Olsen et al., 2018, Simons et al., 2018b, Torquati et al., 2018) and commercial activity monitors with high validity and reliability for the particular measure (e.g. Fitbit® used to monitor steps) (Evenson et al., 2015) (Jones, 2016, Poirier et al., 2016, Rowe-Roberts et al., 2014, Thorndike et al., 2014, Yeung et al., 2017, Gremaud et al., 2018), and the International Physical Activity Questionnaire (IPAQ) with reasonable validity and reliability (Craig et al., 2003) (Reijonsaari et al., 2009, Reijonsaari et al., 2012). Eight studies used ‘moderate’ data collection tools with either acceptable validity or reliability, including smartphone-integrated accelerometers (Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Patel et al., 2018), the Activity Questionnaire for Adolescents and Adults (AQuAA) (Chinapaw et al., 2009) (Slootmaker et al., 2009), the Tractivity® activity monitor (Reed et al., 2018), self-entered pedometer data (Ganesan et al., 2016) and step data from the Fitbit® converted to MVPA (Losina et al., 2017). Two studies used self-reported data in non-validated questionnaires (Schrager et al., 2017, Skogstad et al., 2016), and one study used computer software and an activity monitor with unreported validity and reliability (van Dantzig et al., 2013); these were therefore given a ‘weak’ data collection rating.

Withdrawals and dropouts were reported by the majority of studies (n = 24). Definitions of attrition varied between studies but it was possible to calculate attrition rates based on the number of participants failing to provide data at the final follow-up, which ranged from 0% to 74%. Only four studies (Brakenridge et al., 2016b, Ganesan et al., 2016, Gilson et al., 2017, Torquati et al., 2018) were rated as ‘weak’ in this domain due to having particularly high attrition rates of greater than 40%.

Regarding intervention integrity, most studies reported the proportion of participants receiving the allocated intervention, which was most frequently in the range of 80 to 100%. Approximately two thirds of studies reported measuring consistency of delivery or use of the intervention, with outcomes such as device wear time and interaction with technology. In the majority of studies, it was judged to be possible that participants had received an unintended intervention or this could not be determined from the reports.

Data analysis methods were generally deemed appropriate. Eight of the 13 RCTs used intention-to-treat analysis (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Patel et al., 2016a, Patel et al., 2016b, Slootmaker et al., 2009, Thorndike et al., 2014, Gremaud et al., 2018, Patel et al., 2018).

Table 4 Summary of risk of bias assessment

	Selection bias	Study design	Confounders	Blinding	Data collection method[a]	Withdrawals and dropouts	Attrition rate[b]	Global rating
Brakenridge et al. 2016a; Brakenridge et al. 2016b	W	S	S	W	S	W	IG: 62% CG: 47% Overall: 54%	**Weak**
Finkelstein et al. 2015; Finkelstein et al. 2016	W	S	S	M	S	S	IG1: 21%, IG2: 24%, IG3: 14% CG: 17% Overall: 19%	**Moderate**
Ganesan et al. 2016	W	M	S	M	M	W	IG: 47%	**Weak**
Gilson et al. 2016; Gilson et al. 2017	W	M	W	M	S	W	IG: 57%	**Weak**
Gremaud et al. 2018	W	S	S	M	S	S	IG: 0% CG: 1% Overall: 1%	**Moderate**
Jones 2016	W	M	W	M	S	M	IG1: 32%, IG2: 16% CG: 25% Overall: 24%	**Weak**
Koyle 2013	W	S	S	M	M	S	IG: 17% CG: 11% Overall: 14%	**Moderate**
Losina et al. 2017	W	M	W	M	M	S	IG: 3%	**Weak**
Neil-Sztramko et al. 2017	W	M	W	M	S	S	IG: 0%	**Weak**
Olsen et al. 2018	W	M	W	M	S	M	IG: 39%	**Weak**
Patel et al. 2016a	W	S	S	M	M	S	IG1: 3%, IG2: 8%, IG3: 4% CG: 4% Overall: 5%	**Moderate**
Patel et al. 2016b	W	S	W	M	M	S	IG1: 3%, IG2: 4%, IG3: 3% CG: 6% Overall: 4%	**Weak**
Patel et al. 2018	W	S	S	M	M	S	IG1: 5%, IG2: 5%, IG3: 7% CG: 8% Overall: 6%	**Moderate**
Poirier et al. 2016	W	S	S	M	S	S	IG: 20% CG: 17% Overall: 18%	**Moderate**
Reed et al. 2018	W	S	S	M	M	S	IG1: 8%, IG2: 0%, IG3: 4% Overall: 4%	**Moderate**

	Selection bias	Study design	Confounders	Blinding	Data collection method[a]	Withdrawals and dropouts	Attrition rate[b]	Global rating
Reijonsaari et al. 2009; Reijonsaari et al. 2012	W	S	S	W	S	M	IG: 35% CG: 32% Overall: 33%	**Weak**
Rowe-Roberts et al. 2014	W	M	W	M	S	M	IG: 34%	**Weak**
Schrager et al. 2017	M	M	W	M	W	M	IG: 35%	**Weak**
Simons et al. 2018a; Simons et al. 2018b	W	S	S	M	S	S	IG: 12% CG: 19% Overall: 15%	**Moderate**
Skogstad et al. 2016	W	M	S	M	W	S	IG: 15%	**Weak**
Slootmaker et al. 2009	W	S	S	M	M	M	IG: 26% CG: 18% Overall: 22%	**Moderate**
Thorndike et al. 2014	S	M	S	M	S	S	Phase 1 IG: 4% CG: 6% Phase 2 IG: 8% Whole trial: 13%	**Strong**
Torquati et al. 2018	W	M	W	M	S	W	IG: 74%	**Weak**
van Dantzig et al. 2013	W	S	S	M	W	S	Attrition not reported but 85/86 participants appear to have completed the trial based on final data (i.e. 1.2% attrition).	**Weak**
Yeung et al. 2017	W	M	W	M	S	S	IG: 14%	**Weak**

Notes:

S = strong; M = moderate; W = weak; IG = Intervention Group; CG = Control/Comparison Group

[a] Validity and reliability of primary outcome. For data collection method, studies were rated as 'strong' if the measure had known validity *and* reliability, 'moderate' if the measure had reasonable validity *or* reliability, and 'weak' if validity and reliability were unknown.

[b] % of participants failing to provide data at final follow-up (to nearest whole percentage)

Behaviour change techniques (BCTs)

Due to the relatively small number of studies and BCTs identified, it was not possible to determine which techniques were associated with intervention efficacy. In many cases it was difficult to determine intervention content and specific BCTs used from the available descriptions. The most frequently identified BCTs (or categories of BCTs) are shown in **Table 3**. These included self-monitoring of the behaviour or outcome (n = 22, 88% of studies), provision of feedback on the behaviour or outcome (n = 21, 84%), goal-setting for the behaviour or outcome (n = 17, 68%), social comparison (n = 14, 56%), social support (n = 12, 48%), rewards and incentives contingent on progress towards or achieving the behaviour (n = 11, 44%), and provision of information on consequences of PA and SB to the individual or in general (n = 11, 44%). Prompts and cues (n = 9, 36%) were also a common intervention component. Action planning was identified in eight studies (32%), graded tasks were described by four studies (16%), information on where and when to perform the behaviour was given in four studies (16%), and barrier identification/problem solving was used in three studies (12%). Instruction on how to perform the behaviour, shaping, and prompting anticipated regret were each used in two studies (8%). Information about others' approval and environmental restructuring were each found in only one study (4%). Individual or team competitions, and various types of gamification (such as virtual avatars and racing around a virtual landscape) were not part of the CALO-RE taxonomy but were used in several studies with smartphone apps and websites. Sixteen of the 40 BCTs listed in the CALO-RE taxonomy were not identified in any of the coded interventions.

Prompts and cues were used more frequently in interventions for SB; these were found in five of nine studies (56%) that aimed to reduce SB compared with six of 22 (27%) aiming to promote PA. Rewards and incentives were more frequently part of interventions targeting PA (11/22 studies, 50%) compared with three of nine (33%) studies that aimed to reduce SB.

Effects of interventions

Statistical methods of combining the results were not considered feasible for several reasons. There was high methodological heterogeneity with a range of different study designs, outcome measures (particularly for PA) and outcome time points. Incomplete reporting of outcome data and standard deviations precluded the calculation of reliable effect sizes. Some studies reported change in PA while others reported absolute values. In addition, several studies were either uncontrolled or did not have a 'true' control group (i.e. the comparison group received an mHealth intervention) which would have resulted in an underestimation of effect sizes. The data were therefore summarised narratively and visually. A summary of the main results for each included study is shown in **Table 5**. Impact on PA, SB and health and other related outcomes is reported separately.

Impact on physical activity

A significant increase in one or more measures of PA, over time or relative to the control or comparison group, was reported by 14 of the 25 studies (56%) (Brakenridge et al., 2016b, Finkelstein et al., 2016, Ganesan et al., 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Poirier et al., 2016, Schrager et al., 2017, Skogstad et al., 2016, Thorndike et al., 2014, Neil-Sztramko et al., 2017, Yeung et al., 2017, Gremaud et al., 2018, Patel et al., 2018). These outcomes included mean (or median) daily steps, frequency and/or duration of activity, and odds of meeting step goals. Schrager and colleagues reported a significant impact of the intervention only in participants with a low baseline activity level (Schrager et al., 2017). Six studies (24%) reported no significant impact on any PA outcome (Jones, 2016, Reijonsaari et al., 2012, van Dantzig et al., 2013, Gilson et al., 2017, Olsen et al., 2018, Simons et al., 2018b). Three studies (12%) reported reductions in PA; two uncontrolled studies reported reductions in daily steps (Reed et al., 2018), and MVPA (Torquati et al., 2018) over time, and one RCT found a reduction in light intensity PA relative to the control group, but only in a highly educated subgroup (Slootmaker et al., 2009). It was not possible to determine the impact of the intervention in two studies; in one the pre- to post- change in PA was unclear (Rowe-Roberts et al., 2014), and another (a feasibility study) did not report the statistical significance of changes as there was no reliable baseline measure of PA (Losina et al., 2017). It should be noted that five of the 14 studies that found a relative increase in PA did not have a true (i.e. non-mHealth) control group (see **Table 5**); the results suggested that one or more mHealth or complementary components had contributed to this increase, including a smartphone app (Gremaud et al., 2018), motivational text messages (Koyle, 2013) and financial incentives (Patel et al., 2016a, Patel et al., 2016b, Patel et al., 2018).

Of the 10 studies rated as 'high' or 'moderate' quality, seven (70%) reported a significant impact of the intervention on PA (Finkelstein et al., 2016, Koyle, 2013, Patel et al., 2016a, Poirier et al., 2016, Thorndike et al., 2014, Gremaud et al., 2018, Patel et al., 2018). Only four of the 11 studies (36%) using a wearable activity monitor as a single mHealth tool reported a significant absolute or relative increase in PA, compared with 10 of the 14 studies (71%) using smartphone apps or activity monitors combined with apps. Nine of the 14 studies (64%) using multi-component interventions reported a significant impact on PA (Brakenridge et al., 2016b, Finkelstein et al., 2016, Koyle, 2013, Patel et al., 2016a, Patel et al., 2016b, Skogstad et al., 2016, Thorndike et al., 2014, Neil-Sztramko et al., 2017, Patel et al., 2018) compared with five of the 11 studies (45%) that assessed standalone mHealth interventions (Ganesan et al., 2016, Poirier et al., 2016, Schrager et al., 2017, Yeung et al., 2017, Gremaud et al., 2018). There were no other discernible associations between type or length of intervention, type of workplace and impact on PA.

Significant effects on PA were reported from one month to 12 months after beginning the intervention, although only three studies reported a significant increase in PA at a time point of six months or later (Brakenridge et al., 2016b, Finkelstein et al., 2016, Schrager et al., 2017). In some cases an initial increase in PA was not sustained at later follow-up (Patel et al., 2016a, Patel et al., 2016b, Patel et al., 2018, Reed et al., 2018). In contrast, Brakenridge and colleagues reported a significant impact of the intervention at 12 months but not three months (Brakenridge et al., 2016b).

There was wide variation in effect sizes. For example, for studies reporting a significant positive impact of the intervention on mean daily step count, this ranged from a between-group difference of around 847 (95% CI 68 to 1625) (Brakenridge et al., 2016b) to 2183 (95% CI 992 to 3344) (Gremaud et al., 2018) steps per day. The large international cohort study reported the largest effect, with a mean pre-post increase of 3519 (95% CI 3484 to 3553) steps per day (Ganesan et al., 2016).

Impact on sedentary behaviour

Of the 10 studies reporting the impact of their intervention on sedentary time, only four (40%) found a significant reduction; these were a short-term wearable activity monitor and text messaging intervention in the workplace (van Dantzig et al., 2013), an activity monitor and smartphone app intervention (Gremaud et al., 2018), an activity monitor, app and behavioural counselling intervention (Neil-Sztramko et al., 2017), and a standalone smartphone app intervention (Ganesan et al., 2016). Van Dantzig and colleagues found a mean between-group difference in reduction in computer activity (a proxy for sedentary time) of 4.1 minutes, 30 minutes before and after receiving a persuasive text message (van Dantzig et al., 2013). Gremaud et al. reported a reduction of 26.6 minutes (95% CI -70.9 to -17.3) in the mean longest sedentary time in those with an activity monitor and app compared with the activity monitor only group (Gremaud et al., 2018). Neil-Sztramko et al. found a mean reduction in both objective and self-reported weekly sedentary time of 405.5 and 425.3 minutes respectively, from baseline to 12 weeks post-intervention (Neil-Sztramko et al., 2017). Ganesan and colleagues reported a mean reduction in self-reported daily sitting duration of 0.74 (95% CI 0.78 to 0.71) hours after 100 days of the smartphone app intervention (Ganesan et al., 2016).

Two studies using objective measures of sedentary time showed no significant impact of a smartphone app, pedometer and social media intervention (Torquati et al., 2018) and a multi-component programme including an activity monitor and smartphone app combined with group-based action planning and a healthy living seminar (Olsen et al., 2018). Another study found no impact of an activity monitor on self-reported sedentary time at either three or eight months follow-up (Slootmaker et al., 2009). A further two studies using objective measures showed significantly higher daily standing time (Brakenridge et al., 2016b) and lower daily sedentary time (Jones, 2016) in controls relative to the intervention group. Another study using accelerometer data reported a significant increase in the mean proportion of time spent sedentary from baseline to follow-up, but only in workday non-work time (there was a slight reduction in proportion of work time spent sedentary) (Gilson et al., 2017).

Impact on other outcomes

Of the 25 studies, 16 (64%) assessed the impact of the mHealth intervention on secondary outcomes including health and fitness, wellbeing and determinants of PA (Brakenridge et al., 2016a, Brakenridge et al., 2016b, Finkelstein et al., 2015, Finkelstein et al., 2016, Ganesan et al., 2016, Jones, 2016, Koyle, 2013, Reijonsaari et al., 2009, Reijonsaari et al., 2012, Rowe-Roberts et al., 2014, Schrager et al., 2017, Skogstad et al., 2016, Slootmaker et al., 2009, Thorndike et al., 2014, Gilson et al., 2016, Gilson et al., 2017, Neil-Sztramko et al., 2017, Reed et al., 2018, Simons et al., 2018b, Torquati et al., 2018). Eleven of these studies (69%) found an improvement in at least one outcome over time or relative to the control or comparison group (Ganesan et al., 2016, Koyle, 2013, Rowe-Roberts et al., 2014, Schrager et al., 2017, Skogstad et al., 2016, Slootmaker et al., 2009, Thorndike et al., 2014, Gilson et al., 2017, Neil-Sztramko et al., 2017, Reed et al., 2018, Torquati et al., 2018). Significant beneficial effects included weight or BMI reduction (Ganesan et al., 2016, Neil-Sztramko et al., 2017), reduced body fat percentage (Reed et al., 2018), reduced systolic blood pressure (Thorndike et al., 2014) (Reed et al., 2018), reduced resting pulse rate (Koyle, 2013), reduced total and low density lipoprotein (LDL) cholesterol (Skogstad et al., 2016) and increased high density lipoprotein (HDL) cholesterol (Thorndike et al., 2014), improved ‘AUSDRISK’ (Australian Type 2 Diabetes Risk) score (Rowe-Roberts et al., 2014), improved diet (Gilson et al., 2017, Torquati et al., 2018), improved aerobic fitness (Skogstad et al., 2016), improved self-reported health (Torquati et al., 2018) or wellness (Schrager et al., 2017), greater self-reported energy and emotional wellbeing (Neil-Sztramko et al., 2017), reduced sleep disturbance (Neil-Sztramko et al., 2017) and improved self-efficacy for walking (Koyle, 2013). However, the study by Skogstad and colleagues could not attribute the changes in aerobic fitness and cholesterol levels to changes in individual PA levels (Skogstad et al., 2016).

Slootmaker and colleagues reported a significant impact on secondary outcomes in subgroups only. This included increased awareness of PA level in overweight participants only (after three months) and reduced body weight in lower educated participants only (after eight months) (Slootmaker et al., 2009). Four studies found no impact on any secondary outcomes (Brakenridge et al., 2016b, Finkelstein et al., 2016, Jones, 2016, Simons et al., 2018b), and one RCT found a significant between-group difference in weight loss and percentage body fat, but in favour of the control group (Reijonsaari et al., 2012). Only two studies assessed work-related outcomes including work productivity and sickness absence (Reijonsaari et al., 2012) and job performance, job control and work satisfaction (Brakenridge et al., 2016b); there was no significant effect on these outcomes in the short or long term.

Subgroup findings

The most important subgroup and sensitivity findings (where applicable) for each study are reported in **Table 5**. Potential effect modifiers associated with intervention effectiveness were low baseline activity level (Finkelstein et al., 2016, Poirier et al., 2016, Schrager et al., 2017), lower education level (Slootmaker et al., 2009), African American ethnicity (Losina et al., 2017), non-obesity (Losina et al., 2017) and high risk of diabetes (Rowe-Roberts et al., 2014).

Feasibility and acceptability / additional findings

Four studies were primarily designed to assess feasibility of the intervention and/or trial methods, including measures of engagement, acceptability, attrition, demand (i.e. reach and recruitment) and implementation (i.e. delivery of the intervention) (Gilson et al., 2016, Gilson et al., 2017, Losina et al., 2017, Neil-Sztramko et al., 2017, Torquati et al., 2018). Many effectiveness studies also reported some of these outcomes, with engagement and attrition measured most frequently. Definitions of engagement and acceptability were variable. Engagement with interventions tended to be measured quantitatively using outcomes such as activity monitor wear time, usage time for apps, features used or proportion of text messages read. Acceptability was a broader concept incorporating both quantitative and qualitative measures such as participant satisfaction, perceived usability, perceived effectiveness and usefulness of the intervention (for PA/SB/other outcomes), preferred components, intentions to continue technology use, barriers to use/engagement, and adverse events. Only a small number of studies assessed qualitative experiences as reported by the participants as a measure of acceptability. For example, Rowe-Roberts and colleagues used focus groups to gain further insight into employee experiences of using the activity monitor (Rowe-Roberts et al., 2014) while Gilson and colleagues interviewed drivers and depot managers to capture experiences, insights into perceived impact of the intervention and barriers to PA (Gilson et al., 2016, Gilson et al., 2017) .

The main findings in relation to engagement, acceptability and attrition are summarised in **Table 5**. A clear decline in technology usage and engagement over time was reported by all longer duration studies (i.e. more than 12 weeks) that assessed these outcomes. Schrager and colleagues reported that only 33% of participants used their activity monitor after six months (Schrager et al., 2017), Brakenridge and colleagues reported that activity monitor use had ceased in all participants by 12 months (Brakenridge et al., 2016b), and Finkelstein and colleagues found that only around 10% of participants still wore their activity monitor after 12 months (Finkelstein et al., 2016). Common reasons for lack of engagement were broken or lost devices (Rowe-Roberts et al., 2014, Yeung et al., 2017), forgotten devices (Rowe-Roberts et al., 2014, Schrager et al., 2017, Simons et al., 2018b), lack of interest or boredom (Schrager et al., 2017, Slootmaker et al., 2009), beliefs the device was not accurate (Schrager et al., 2017), technical issues (Schrager et al., 2017, Gremaud et al., 2018, Simons et al., 2018b), fashion (Schrager et al., 2017), privacy concerns (Gilson et al., 2017), data usage costs (Gilson et al., 2017) and usability issues such as difficulty navigating the website (Slootmaker et al., 2009).

Overall, participant satisfaction was high, and employees perceived wearable activity monitor and smartphone app interventions to be an acceptable and useful method to improve PA. Of the studies aiming to reduce SB, only two included qualitative measures of acceptability (Gilson et al., 2017, Torquati et al., 2018). In one study, the activity monitor and smartphone app were perceived by drivers and depot managers as feasible, acceptable and as having a positive impact on PA and SB (Gilson et al., 2017). In contrast, a study of a smartphone app for improving diet and PA (and reducing SB) in nurses found low perceived usefulness, and interviews revealed difficulties in changing more than one behaviour at a time, and the desire for a workplace champion to implement the intervention (Torquati et al., 2018). Additional findings in relation to acceptability included individual differences in preferred motivational strategies according to levels of activity and engagement (Rowe-Roberts et al., 2014) and higher compliance with activity monitor wear with team-based competition as opposed to individual monitoring (Thorndike et al., 2014). It is also important to consider adverse events associated with technology use; in one study around 27% of participants reported at least one adverse event, typically related to reactions to wearing the activity monitor or accelerometer (Brakenridge et al., 2016b). Due to the relatively small number of studies reporting measures of acceptability, and the heterogeneity of interventions and outcomes, no associations between acceptability and intervention type or length or type of workplace were evident.

Attrition rates ranged from 0% to 74%. Predictors of loss to follow-up included female gender (Finkelstein et al., 2016), younger age (Skogstad et al., 2016) and ethnicity with lower attrition in Chinese participants (Finkelstein et al., 2016).

Table 5 Summary of main results

Study	Key findings				Subgroup findings and sensitivity analyses [a] (if applicable)	Feasibility and additional findings (including engagement, attrition and acceptability)
	Impact on PA / SB[a]	Impact on related health / other outcomes	Impact on PA / SB	Impact on other outcomes		
Brakenridge et al. 2016a; Brakenridge et al. 2016b	No significant between-group difference in PA or SB at 3 months (although the organisational support group only showed an increase in overall standing time, +14.6 minutes/day, 95% CI 2.5 to 26.8, **p = 0.018**). Significant between-group adjusted mean difference (MD) in overall daily stepping time (+20.6 minutes, 95% CI 3.1 to 38.1, **p = 0.021**) and number of daily steps (+846.5 steps, 95% CI 67.8 to 1625.2, **p = 0.033**) at 12 months – favouring the intervention group.	No significant within- or between-group findings for any health or work-related outcomes.	**↑ PA** **↑ SB (favoured control)**	**0**	Sensitivity analysis: study completers were more likely to show larger and statistically significant changes in activity at 3 months (bias from dropout of healthy participants?).	Engagement: 70.5% of participants provided with the activity monitor used it in the first 12 weeks, with mean usage of 12.1±11.6 days. Use had ceased by 12 months in all intervention participants. Acceptability: 41/153 (26.8%) participants reported at least one adverse event (e.g. reactions to activity monitor or accelerometer wear).
Finkelstein et al. 2015; Finkelstein et al. 2016	No between-group difference in MVPA (p = 0.0854) or steps (p = 0.1362) between the Fitbit only and control groups at 6 months. Cash and charity incentive groups showed higher MVPA compared with control at 6 months (cash group MD = +29 MVPA bout mins/week, 95% CI 10 to 47, **p = 0.0024**; charity group MD = +21 MVPA bout mins/week, 95% CI 2 to 39, **p = 0.0310**). At 12 months, the Fitbit and charity groups showed higher MVPA than the control (Fitbit group MD = +37 MVPA bout mins/week, 95% CI 19 to 56, **p = 0.0001**; charity group MD = +32 MVPA bout mins/week, 95% CI 12 to 51, **p = 0.0013**). The cash incentive group did not differ from the control (p = 0.1363). The only significant within-group increase in MVPA at 12 months was for the Fitbit only group (+16 mins/week, 95% CI 2 to 30, **p = 0.0301**).	No evidence for improvement in health outcomes - all intervention groups showed improvement in cardiorespiratory fitness (NET-FVO_2 max) at 6 and 12 months, but control group showed improvement at 12 months.	**↑ PA**	**0 (improved aerobic fitness in control and intervention groups)**	Subgroup analysis by baseline activity level found some differences – e.g. those in the cash and charity groups who were insufficiently active at baseline showed a significant increase in MVPA from baseline to 6 months (+22 mins/week, 95% CI 5 to 38, **p = 0.0096** and +17 mins/week, 95% CI 2 to 32, **p = 0.0231** respectively). For those who were sufficiently active at baseline, changes in MPVA were non-significant in both groups.	Engagement: 40% abandoned the Fitbit within 6 months, and by month 12 only around 10% of all participants in the intervention groups were still wearing the device. Attrition: predictors of loss to follow-up at 12 months included gender (higher attrition in females) and ethnicity (lower attrition in Chinese). Higher adherence was seen in the cash incentive group compared with the other groups (particularly at 6 months).

	The only significant between-group differences in mean daily steps were for the cash vs. control group at 6 months (+1050 steps, 95% CI 600 to 1490, **p<0.0001**) and 12 months (+500 steps, 95% CI 50 to 960, **p = 0.0289**).					
Ganesan et al. 2016	Significant increase pre- to post-intervention in mean daily steps (+3519 steps, 95% CI 3484 to 3553, **p<0.0001**), exercise days/week (+0.89 days, 95% CI 0.87 to 0.92, **p<0.0001**) and odds of exercising ≥30 minutes/day (1.65, 95% CI 1.61 to 1.68, **p<0.0001**). Significant decrease pre- to post-intervention in mean sitting duration (-0.74 hours/day, 95% CI -0.78 to -0.71, **p<0.0001**).	Significant pre- to post-intervention reduction in weight (-1.45kg, 95% CI -1.53 to -1.38, **p<0.0001**).	**↑ PA** **↓ SB**	**↓ weight**	Subgroup analysis by gender, year cohort, geographic region and income group – no significant differences for any of the main outcomes. Men showed greater weight loss than women (-1.63kg, 95% CI -1.72 to -1.54 compared with -0.74kg, 95% CI -0.91 to -0.57). Predictors of weight loss included increase in step count, increase in exercise days and decrease in sitting duration.	
Gilson et al. 2016; Gilson et al. 2017	Non-significant increase in mean proportions of work time spent physically active from baseline to post-intervention and follow-up (+1%, 7 minutes/day). Non-significant decrease in mean proportions of work time spent sedentary at post-intervention (-6%) and follow-up (-9%). Significant *increase* in mean proportion of workday non-work time spent sedentary baseline to follow-up (**p = 0.007**) and decrease in mean proportion of workday non-work time stationary+ baseline to post-intervention (**p = 0.037**) and follow-up (**p<0.033**). 65% of participants showed positive changes in PA (and at least one dietary choice) at follow-up.	Significant increase in workday fruit (**p = 0.023**) and vegetable (**p= 0.024**) consumption by one serving/day at end of programme.	**0 PA** **↑ SB (workday non-work time only)**	**↑ diet (fruit and vegetable intake)**	N/A	Engagement: 26/44 (59%) participants used the activity monitor. Use for step count monitoring remained constant but dietary logging significantly declined from baseline to study completion. Attrition: moderately high - only 19/44 (43%) participants completed the study. Acceptability: Barriers to technology use included technical issues, data usage costs and privacy concerns. From qualitative interviews, the overall intervention was perceived as feasible, acceptable and as having positive impact on PA by drivers and depot managers. The mHealth component was perceived to have a greater impact on behaviour than financial incentives.
Gremaud et al. 2018	Relative to control (Fitbit-only group) after the start of the intervention, the smartphone app users showed an immediate increase in	Not studied	**↑ PA** **↓ SB**	**Not studied**	Sensitivity analyses accounting for wear compliance, and excluding	Engagement: compliance with Fitbit wear declined over time, but app users

	mean daily steps of 2183 (95% CI 992 to 3344). Daily active minutes similarly increased by 12.8 (95% CI 6.3 to 19.3). However, participants' steps declined during the study period overall. The mean longest bouts of sedentary time decreased by 26.6 minutes (95% CI -70.9 to -17.3) in the intervention group relative to control.		**(but control group received mHealth intervention)**		data collected following a bug in the app platform did not alter the significance of the findings.	were more likely to wear the Fitbit daily than the Fitbit-only group. Low attrition: 144/146 (99%) participants provided follow-up data (although short-term study). Acceptability: participants rated the app as easy and enjoyable to use and useful for increasing PA. The main reported barrier to technology use was Fitbit battery issues (8/48, 17% of respondents).
Jones 2016	Between-group difference in % change in mean daily steps neared significance: IG1 (activity monitor only) = +9% IG2 (activity monitor and active desk) = -17% CG = -15% (p = 0.06, favouring the activity monitor only group) Significant *increase* in sedentary time in intervention groups compared with control: IG1 = +255.5 minutes/day, 95% CI 127.5 to 383.5 IG2 = +353.2 minutes/day, 95% CI 219.1 to 487.3 CG = 0 minutes/day, 95% CI -58.6 to 58.5 (**p<0.0001**)	No significant within- or between-group findings for BMI, sleep or any psychosocial outcomes.	**0 PA** **↑ SB (favoured control)**	**0**	N/A	Engagement: participants reporting follow-up data wore the Fitbit activity monitor for a mean of 177 of 210 days (84% adherence). Attrition: dropouts were similar to study completers in terms of baseline activity, weight, gender and ethnicity.
Koyle 2013	The 'control' group showed a significant within-group reduction in mean weekly walking duration from week 1 to week 6 (-50.3 minutes, **p<0.001**). The reduction in the intervention group was non-significant (-20.8 minutes/week, p = 0.99). After controlling for baseline activity level, the intervention group walked more minutes than the control but a significant difference was found only at week 6 (β = 38.21, **p = 0.03**).	Significant pre- to post-intervention increase in self-efficacy beliefs (walking self-efficacy scale) for both groups: Intervention 85.6±12.1 to 90.0±10.8 (**p = 0.0003**) and control 78.3±14.7 to 87.3±11.6 (**p = 0.0288**). Significant decrease in resting mean pulse rate for intervention group only: 70.0±11.0 to 63.2±10.2 (**p = 0.038**).	**↑ PA (but control group received mHealth intervention)**	**↑ self-efficacy** **↓ resting pulse rate**	N/A	Acceptability: Participants in both groups commented positively on their experiences of taking part and found the study motivating. Text messages were reported as helpful. Many participants felt stronger and/or reported having lost weight as a result of taking part. At the end of the trial, all but one person chose to continue using the app.

		No increase in likeliness to participate in other PA. No significant within- or between-group findings for weight, BMI or systolic BP.				
Losina et al. 2017	Average weekly duration of MVPA increased from 54±64 minutes in two weeks pre-intervention to 62±89 minutes post-intervention (statistical significance of changes not reported). 86% of participants met either their weekly PA goal or CDC PA guidelines for ≥6/24 weeks. 52% met either their PA goal or CDC guidelines for ≥12/24 weeks.	Not studied	**? (statistical significance of PA changes not reported and pre-intervention measure was contaminated by use of Fitbit)**	**Not studied**	Subgroup analysis by baseline physical ability, ethnicity and obesity found some differences. Those able to walk a mile at baseline (vs. those unable, **p = 0.010**), African Americans (vs. all other ethnicities, **p = 0.016**) and non-obese participants (vs. obese, **p = 0.018)** met PA guidelines more consistently throughout the programme.	Engagement: 63% of participants were classed as adherent Fitbit wearers (i.e. wearing Fitbit ≥4 days/week for ≥20 weeks). Wear declined over time (e.g. 94% adherent wearers after one month vs. 62% after 6 months). Acceptability: two thirds of participants were satisfied with the programme; 79% indicated they would participate again.
Neil-Sztramko et al. 2017	Significant increase in mean total MVPA (+110.3 minutes/week, **p<0.01**) and significant increase in mean daily steps (+1488.7, **p<0.01**) from baseline to post-intervention. Significant reductions in objective mean sedentary time (-405.5 minutes/week, bouts ≥10 mins, **p = 0.02**) and self-reported mean sedentary time (-425.3 minutes/week, **p<0.01**) from baseline to post-intervention.	Significant pre- to post-intervention reductions in weight (mean change -0.9kg, **p = 0.03**) and BMI (mean change -0.3kg/m^2, **p = 0.04**). Significant improvements in some domains of health-related quality of life: energy/fatigue (**p = 0.01**) and emotional wellbeing (**p = 0.04**). Significant improvements in sleep disturbances (**p = 0.04**) and day dysfunction due to sleepiness (**p = 0.04**).	**↑ PA** **↓ SB**	**↓ weight and BMI** **↑energy / fatigue and emotional wellbeing score** **↓ sleep disturbance and day dysfunction due to sleepiness**	N/A	High engagement with Fitbit: all participants reported using the Fitbit. Of the 18 who provided Fitbit data, the device was worn 94.5% of the total study period. Attrition: all participants completed the study. High acceptability: 16/19 (84.2%) participants were very or somewhat satisfied with the intervention. Demand: high demand for participation. Recruitment to time and target was exceeded. Implementation: technical issues were common; 5 (25%) Fitbits were returned. A further two devices were lost.
Olsen et al. 2018	No significant changes in sedentary time from pre- to post-intervention: Mean change in accelerometer-assessed sedentary minutes/day +0.08 (95% CI -30 to +30). Mean change in total self-reported sitting time in office -56 mins/day (95% CI -128.5 to	Not studied	**0 PA** **0 SB**	**Not studied**	N/A	Moderate attrition: only 30 of the 49 (61%) recruited participants provided some post-intervention data. Acceptability: overall acceptability of the intervention was high. 90% of participants were satisfied or very

	+17.0) and when working at home +20.5 (95% CI -64.5 to 105.5). No significant changes in accelerometer-assessed or self-reported PA.					satisfied with the programme and 83% agreed that the tracker was a useful tool for behaviour change.
Patel et al. 2016a	No significant between-group differences in mean daily steps (within-group changes not reported). Proportion of participant-days 7000 step goal was achieved was significantly higher for loss incentive group compared with control for the 13 week intervention period (MD = +0.16, 95% CI 0.06 to 0.26, **p = 0.001**). This effect was not sustained at follow-up after incentives were removed.	Not studied	**↑ PA (but control group received mHealth intervention)**	**Not studied**	Sensitivity analysis: adjusting for device and different methods of accounting for missing data did not affect the significance of the findings.	Low attrition: only 5% of participants did not complete the study.
Patel et al. 2016b	Compared with control during the intervention period, the mean proportion achieving the 7000 step goal was significantly higher for the combined incentive group only (MD = +0.17, 95% CI 0.07 to 0.28, **p<0.001**). The combined incentive group also had higher mean daily steps than the control at the end of the intervention period (MD = +1446 steps, 95% CI 448 to 2444, **p = 0.005**). No significant differences were sustained at follow-up.	Not studied	**↑ PA (but control group received mHealth intervention)**	**Not studied**	Sensitivity analysis: adjusting for device and missing data did not affect the significance of the findings.	Low attrition: only 4% of participants failed to complete the study.
Poirier et al. 2016	Mean daily steps pre- to post-intervention increased for the intervention group (+309 steps/day, ±1874) and decreased for the control group (-661 steps/day, ±1824). MD = 970 steps/day, **p<0.001**. The proportion of participants achieving an increase of 1000 steps/day was significantly greater in the intervention group (29.9%) than the control (16.4%), **p = 0.018**.	Not studied	**↑ PA**	**Not studied**	Subgroup analysis by baseline activity level: Sedentary group (<5000 steps/day): Mean change +594±1558 steps/day in intervention group vs. +47±1299 steps/day in control group, **p = 0.04**. Low-to-somewhat active group (5000 to 9999 steps/day): Mean change of -110±2106 steps/day in intervention group vs. -1286±1783 steps/day in control group, **p = 0.004**.	High engagement with intervention: Participants wore the activity monitor on 78.6% of days (33/42) on average; e-mails were opened on 21.9% of days (9.2/42); website visits occurred every 3.6 days (11.8/42). 130/133 (97.7%) intervention participants still wore the activity monitor, opened e-mails and/or visited the website after 6 weeks. Attrition: moderately low (around 82% provided complete follow-up data) although short term study. Participants with complete outcome data were similar

					Sensitivity analysis: included some study non-completers – statistically significant between-group difference in mean daily steps remained (**p<0.001**).	to those without in terms of baseline PA level, ethnicity, income and education.
Patel et al. 2018	Compared with control during the intervention period, the (unadjusted) mean proportion achieving the 7000 step goal was significantly higher for the combined lottery incentive group only (0.38 vs. 0.26). The adjusted odds ratio for achieving the goal (combined vs. control) was 3.00 (95% CI 1.28-7.02, **p = 0.012**). No significant differences were sustained at follow-up, after incentives were removed. No significant between-group differences in mean daily steps (within-group changes not reported).	Not studied	**↑ PA (but control group received mHealth intervention)**	**Not studied**	Sensitivity analysis: adjusting for device and missing data did not change the significance of the findings.	Low attrition: only 6% of participants did not complete the study.
Reed et al. 2018	Initial increase in MVPA but significant decline from week 2 to week 6 (i.e. post-intervention), **p<0.05**. Significant decline in daily steps from baseline to week 6 (**p<0.05**). There were no significant between-group differences in either MVPA (p = 0.292) or steps (p = 0.333).	Within-group significant reductions in % body fat (-0.8±4.8, **p = 0.015**) and resting systolic BP (-2.6±8.8 mm Hg, **p = 0.019**). No significant within- or between-group changes in body mass, BMI or waist circumference.	**↓ PA (steps only, and no control group)**	**↓ % body fat** **↓ systolic BP**	N/A	Engagement: Participants wore the activity monitor for at least 10 hours/day for 31/42 intervention days on average (overall compliance rate 74%). Wear declined over time (e.g. average of 6.0±1.9 days per week at baseline compared with 3.5±3.0 days in week 6). Low attrition: 72/75 (96%) participants completed the study.
Reijonsaari et al. 2009; Reijonsaari et al. 2012	No significant within- or between-group differences in PA: 6-month between-group adjusted MD = -365 weekly MET-minutes, 95% CI -733 to 3; 12-month between-group adjusted MD = -207 weekly MET-minutes, 95% CI -531 to 116 (negative values favour control).	No significant between-group difference in productivity (adjusted MD in QQ score at 6 months = 1.3, 95% CI -2.0 to 4.7 and adjusted MD at 12 months = -1.1, 95% CI -4.9 to 2.8). Between-group difference for change in weight and % body fat favoured control (adjusted MD for weight change at 12 months = -0.5kg, 95% CI -1.0 to 0.0; adjusted MD for	**0 PA**	**↑ weight (favoured control)** **↑ % body fat (favoured control)**	Subgroup analyses by gender, job characteristics, age, self-rated baseline PA level and sickness absence days in the past year did not modify the results. Adherence to the intervention did not mediate sickness absence (MD between adhering and non-adhering subgroups was 0.0 days, 95% CI -1.2 to 0.9).	Engagement: decline in engagement (use of website, communication with coaches) over time, particularly in the last 6 months.

		change in % body fat = -0.6%, 95% CI -1.0 to -0.2).			Sensitivity analysis: adjusting for missing data did not affect the results.	
Rowe-Roberts et al. 2014	Findings only reported descriptively. Average daily steps reported by AUSDRISK score at beginning and end of trial; 'high' baseline score participants that moved to 'low' risk at the end of the study had the highest average daily steps at the end of the study (12,294). Average overall daily steps were: High risk group: 8588 Medium risk group: 7836 Low risk group: 7878	23% of participants reduced their AUSDRISK score over 7 months.	**?** **(pre- to post-change in PA unclear and no control group)**	↓ **AUSDRISK score**	Results stratified by AUSDRISK score – device seemed more effective for those at high risk of diabetes (see left).	Engagement: overall low engagement with the activity monitor – average monthly dropout rate was 15% and only 36% of participants were still using the device at the end of the study. High baseline diabetes risk participants showed the highest level of engagement: Mean number of months engaged with the activity monitor = 5.7 for high risk; 4.4 for medium risk; 4.2 for low risk Acceptability: low engagement was predominantly driven by device issues, e.g. broken, lost, forgotten devices. Individual differences in preferred motivational strategies, e.g. inactive/unengaged participants preferred games whereas active/engaged participants preferred 'goal-oriented functionalities', e.g. smart reminders and normative information about appropriate PA levels.
Schrager et al. 2017	No significant overall change in PA level. Self-reported median (IQR) days/week of ≥30 minutes PA: Baseline 2.5 (1.9) 1 month 2.8 (1.5) 6 months 3.0 (2.0) ($p = 0.67$ for change baseline to 1 month; $p = 0.36$ for change baseline to 6 months) PA monitor-measured median (IQR) days/week ≥10,000 steps or ≥30 minutes PA: Baseline 2.5 (1.9) 1 month 2.5 (2.7) ($p = 0.69$ for change baseline to 1 month)	18/30 (60%) participants described a positive impact on their wellness after one month of activity monitor use.	**↑ PA** **(low baseline activity level only)**	**↑ wellness (qualitative report only)**	Subgroup analysis by baseline activity level and device use. Significant increase in self-reported median (IQR) days/week of ≥30 minutes PA for the most inactive (n = 10): Baseline 1.5 (0.9) 1 month 2.4 (1.2) 6 months 2.0 (2.0) (**$p = 0.04$** for change baseline to 1 month; **$p = 0.04$** change baseline to 6 months) No significant between- or within-group differences in PA level for those who used the activity monitor for 6 months and those who discontinued use prior to the study end.	Engagement: decline in engagement over time - 67% continued to use the device after one month, but only 33% still used their device after 6 months. Acceptability: barriers to use included forgetfulness, not wanting to wear the device, boredom, beliefs it was not accurate, technical issues and fashion.

Simons et al. 2018a; Simons et al. 2018b	No significant between-group differences in any of the objective or self-reported PA outcomes at post-intervention or follow-up (PA decreased over time in the intervention and control groups).	No significant impact on any self-reported psychosocial variables.	**0 PA**	**0 (improved knowledge of PA guidelines in control and intervention groups)**	N/A	Engagement: Decline in engagement over time, with significant reductions in Fitbit wear, viewing graphs in the app and reading notifications. Attrition: 110/130 (85%) participants provided primary outcome data at follow-up. Acceptability: The majority of participants rated the Fitbit and app as self-explanatory (36/51, 71%), user friendly (40/51, 78%), and interesting (34/51, 67%), but few found the tips and facts motivating (10/41, 24%), used them to be physically active (8/41, 20%) and believed they were tailored to their lifestyle (7/41, 17%). Barriers to technology use included technical problems and forgetting to wear or charge the Fitbit.
Skogstad et al. 2016	Significant increase in self-reported frequency of PA from baseline to follow-up (**p = 0.001**). % exercising at baseline: 37% ≤ 1 day/week 47% 2-3 times/week 15% ≥ 4 times/week % exercising at follow-up: 13% ≤ 1 day/week 58% 2-3 times/week 28% ≥ 4 times/week Half of participants reported increased PA frequency at follow-up. Mean increase in daily low intensity PA (e.g. walking): 13.7±29.4 minutes for men 13.7±17.2 minutes for women. Mean increase in daily high intensity PA (e.g. jogging: 8.3±18.2 minutes for men 8.6±14.6 minutes for women	Significant improvement in maximal oxygen uptake (+2.8 ml/kg/min, 95% CI 1.4 to 4.3, **p = 0.00022**). Significant reduction in total cholesterol (-0.12 mmol/L, 95% CI -0.22 to -0.01, **p = 0.032**) and LDL cholesterol (-0.13 mmol/L, 95% CI -0.22 to -0.04, **p = 0.0034**). Significant *increase* in diastolic BP (+1.67 mm Hg, 95% CI 0.23 to 3.12, **p = 0.024**). Despite the significant improvements in health outcomes (aerobic fitness, total and LDL cholesterol), further analysis could not attribute these to the change in individual PA levels.	**↑ PA**	**↑ aerobic fitness** **↓ total cholesterol** **↓ LDL cholesterol** **↑ diastolic BP**	Subgroup analysis by education. Education was an effect modifier for diastolic BP, total and LDL cholesterol: Only workers with low education showed a significant increase in diastolic BP (+4.4 mm Hg, 95% CI 2.03 to 6.86, **p = 0.0004**). Only workers with high education showed a significant decrease in total cholesterol (-0.21 mmol/L, 95% CI -0.08 to -0.34, **p = 0.0015**) and LDL cholesterol (-0.22 mmol/L, 95% CI -0.12 to -0.32, **p = 0.0001**).	Attrition: Participants lost to follow-up differed from study completers – younger, lower HDL and higher CRP (and mostly men and blue collar workers). Acceptability: perceived impact on other outcomes - 12 participants reported improved nutritional habits at follow-up, 3 increased quality of sleeping, 4 reduced or quit smoking and 2 reduced alcohol intake.

Slootmaker et al. 2009	No overall significant between-group difference in sedentary time or PA (mins/week) at 3 months or 8 months follow-up. 3 month between-group difference (adjusted for gender, age, education and BMI), β and 95% CI: Sedentary time: 10 (-435 to 455) Light intensity PA: -129 (-337 to 79) Moderate intensity PA: -13 (-89 to 63) Vigorous intensity PA: -6 (-75 to 62)	No overall significant intervention effect on any secondary outcomes (aerobic fitness, determinants of PA and body composition) at 3 or 8 months.	**↓ PA (light-intensity, highly educated only)** **0 SB**	**↑ awareness of PA level (overweight only)** **↓ body weight (lower educated only)**	Subgroup analysis by education, adherence to programme and BMI. Education was an effect modifier for PA; higher educated intervention participants showed a significant reduction in light intensity PA at 3 months compared with control: adjusted difference in mins/week, β and 95% CI = -349 (-632 to -66), **p = 0.02**. The proportion of participants aware of their adherence to PA recommendations increased among overweight participants in the intervention group compared with control at 3 months (adjusted OR = 16.4, 95% CI 1.3 to 214, **p = 0.02**). There was a reduction in body weight among the lower educated intervention participants compared with control at 3 months (adjusted difference, β = -1.6kg, 95% CI -2.8 to -0.4, **p = 0.01**). Higher engagement/adherence to the programme did not result in increased PA.	Engagement: majority of participants engaged with the intervention; 73% wore the activity monitor regularly and the website was used almost once a week. Acceptability: barriers to technology use included lack of interest and difficulty finding items on the website. 74% of activity monitor users read the tailored advice, of whom 39% found it unappealing.
Thorndike et al. 2014	At the end of phase 1, there was no significant difference between the intervention and control groups in median daily steps overall (intervention 6369 vs. control 6063, p = 0.16) or mean daily steps on days the monitor was worn (intervention 7886±3622 vs. control 7600±3492, p = 0.63).	Significant overall reduction in systolic BP from baseline to end of study: 121±15.4 mm Hg to 117±12.6 mm Hg (**p = 0.004**). Significant overall increase in HDL cholesterol from baseline to end of study: 57±14.7	**↑ PA (team-based competition only)**	**↓ systolic BP** **↑ HDL cholesterol**	The authors compared mean daily steps during inpatient rotations with outpatient rotations during the whole study; physicians were significantly more active during outpatient rotations (difference of 648 steps, **p<0.001**).	Engagement: compliance with wearing the activity monitor was significantly higher in phase 2 than phase 1: 77% vs. 60%, **p<0.001**.

	Mean daily steps were significantly higher in phase 2 (team competition) than phase 1 (individual monitoring) for those assigned to the control group in phase 1 (7971 vs. 7567, **p = 0.002**) but not for those in the intervention group (7832 vs. 7739, p = 0.13).	mg/dL to 61±15.7 mg/dL (**p<0.001**). No significant change in diastolic BP, weight, BMI, waist circumference, total or LDL cholesterol.				
Torquati et al. 2018	Significant reduction in daily MVPA from baseline to 3 months, median (IQR): 19.1 (24.6) to 13.3 (13.9) minutes/day (**p = 0.01**). Near significant reduction in daily MVPA from 3 months to 6 months: median (IQR): 13.3 (13.9) to 12.5 (13.4) minutes/day (p = 0.07). Significant reduction in mean daily steps from baseline to 3 months (8,496±2,528 to 8,136±2,395, **p = 0.04**). No significant changes in sedentary time from baseline to 3 months (p = 0.17) or from 3 months to 6 months (p = 0.64).	Significant increase in daily fruit and vegetable intake from baseline to 3 months (**p = 0.04**). Significant improvement in self-rated health from month 3 to month 6 (**p<0.05**). No significant changes in weight, BMI, waist circumference, blood pressure, PA or diet self-efficacy or social support.	**↓ PA** **0 SB**	**↑ diet (fruit and vegetable intake)** **↑ self-rated health**	Subgroup analysis with participants with complete data only did not change the significance of the findings (although MVPA increased at month 6 following an initial reduction).	Engagement: Low engagement with the smartphone app, with 68.4% using it less than once per month or never. PA goals were set infrequently, and social components were not used. Attrition was high, with only 12/47 (26%) attending the 6 month follow-up. Acceptability: Participants reported that changing both PA and diet at the same time was challenging. Interviews revealed low perceived usefulness of the smartphone app. Overall reach was poor (13% of potential participants were reached and 9.4% were willing to take part). Implementation: Participants required more frequent contacts with researchers or a workplace champion.
van Dantzig et al. 2013	A significantly higher reduction in computer activity (mean minutes of activity, 30 minutes before and after receiving (virtual) text message) was observed in the intervention group compared with control: Intervention group reduction of 10 minutes vs. control group reduction of 5.9 minutes, **p<0.001**. Non-significant within- or between-group change in PA (average value during 5 minute interval before and after a message) (p-value not reported).	Not studied	**0 PA** **↓ SB** **(but control group received mHealth intervention)**	**Not studied**	Engagement with the intervention was explored with subgroup analysis. There was no significant impact of proportion of text messages read on computer activity (p>0.10) and no significant interaction between proportion read and time (p>0.10), i.e. receiving messages led to breaks but content was not important.	Engagement: an average of 46% (SD = 34.6) of the total number of text messages sent were read.
Yeung et al. 2017	Significant increase in median daily steps from blinded to unblinded intervention period (**p = 0.001**):	Not studied	**↑ PA**	**Not studied**	Subgroup analysis by occupation. Surgical residents had significantly higher steps	Moderate engagement with Fitbit: device wear ranged from 91% in those volunteering for the activity tracking

	Median (IQR) Weeks 1-4 (blinded): 7260 (2410) steps/day Weeks 5-8 (unblinded): 8266 (3306) steps/day Significant increase in number of participants achieving an average of ≥10,000 steps/day (**p = 0.04**): Weeks 1-4 (blinded): n = 9 (12%) Weeks 5-8 (unblinded): n = 17 (23%)				than non-surgical residents (**p = 0.018),** however differential impact of the intervention was unclear. Participants who elected to join the activity tracking group had higher median daily steps than those who did not (7938 vs. 7442, **p = 0.042**), however the direction of the effect is unclear (may be due to prior increased motivation).	group to 51% in those not volunteering (unblinded period). A decline in engagement was seen over the 12 weeks in the latter group only. Attrition: high adherence with 74/86 (86%) participants completing the study. Acceptability: The most common barrier to participation was loss of the Fitbit (7/86, 8%).

Notes:

IG = Intervention Group; CG = Control/Comparison Group; PA = Physical Activity; SB = Sedentary Behaviour; MVPA = Moderate-to-Vigorous Physical Activity; MD = Mean Difference; BMI = Body Mass Index; BP = Blood Pressure; ± = Standard Deviation; N/A = Non-Applicable; AUSDRISK = Australian Type 2 Diabetes Risk Assessment Tool; IQR = Interquartile Range; LDL = Low Density Lipoprotein; HDL = High Density Lipoprotein; CRP = C-Reactive Protein; OR = Odds Ratio; CDC = Centers for Disease Control

↑ = significant increase in outcome ↓ = significant decrease in outcome **0** = no significant change in outcome (between-group or within-group)

[a] Significant p-values (where reported in included papers) are in **bold**

DISCUSSION

While the methodological quality of many of the included studies was weak, based on this review there is reasonable evidence that mHealth interventions in workplace settings are a potentially effective and feasible method for increasing PA. There is some evidence that they may also be effective in reducing SB. However, findings are mixed and effect sizes are small, particularly for the impact on SB and in the longer term.

A significant increase in PA, either over time or relative to the control or comparison group, was observed in 56% (14/25) of studies, and in a higher proportion of studies rated as 'high' or 'moderate' quality (7/10, 70%). The findings in relation to SB were less clear, with only 40% (4/10) of studies reporting a significant reduction in sedentary time, and a further three studies reporting relative increases in certain measures of sedentary time. It may be that reducing sedentary time at work leads to corresponding increases in time spent sedentary outside of work; this demonstrates the importance of holistic interventions that take both work and non-work contexts into account (Gilson et al., 2017).

It is important to assess feasibility in addition to effectiveness of complex interventions such as mHealth (Craig et al., 2008, Bowen et al., 2009). Many studies included measures of engagement with the intervention, and the vast majority showed a decline in engagement over time. It is not yet clear whether this disengagement from the technology is detrimental to behaviour change or if sustained behaviour change can be achieved without continued used of the mHealth tool. Future studies could draw comparisons with, and learn from, eHealth interventions to reduce SB and increase PA in the workplace, such as the studies by Mainsbridge et al. (2014) (Mainsbridge et al., 2014), Pedersen et al. (2014) (Pedersen et al., 2014) and Irvine et al. (2011) (Irvine et al., 2011). Only a small number of studies included qualitative measures of acceptability such as interviews to explore participants' experiences, mechanisms of behaviour change and reasons for the decline in engagement. Future studies should focus on these areas. There also appears to be a need for more standardised definitions, assessment and reporting of engagement and acceptability in the field of mHealth (McCallum et al., 2018).

The findings generally concur with the existing evidence for potential effectiveness and acceptability, most prominent in the short term, reported in reviews of mHealth interventions for PA and SB in non-workplace contexts (Fanning et al., 2012, Bort-Roig et al., 2014, Muntaner et al., 2016, Schoeppe et al., 2016, Direito et al., 2017). Due to considerable heterogeneity and the small number of high quality studies, it was not possible to draw any definitive conclusions on the relative effectiveness or acceptability of different types of interventions, although there was some evidence that wearable activity monitors alone, and standalone mHealth interventions with no additional 'offline' components, were less likely to result in increased PA. Similarly, previous reviews have suggested that multi-component interventions may be more effective than standalone mHealth interventions (Schoeppe et al., 2016, Nuffield Health, 2018).

This review is the first to focus on mHealth technology for the promotion of PA and reduction of SB in workplace interventions. A recent systematic review by Stephenson and colleagues which assessed the impact of computer-based, mobile and wearable technologies on SB suggested that the effects of workplace interventions may be more prominent than non-workplace interventions at medium term follow-up (Stephenson et al., 2017). While there was insufficient data to test this hypothesis in the present review, this highlights the potential importance of setting and the possibility of differential results.

There was a small amount of evidence to suggest that mHealth for PA and SB may be more effective for more sedentary employees (Finkelstein et al., 2016, Poirier et al., 2016, Schrager et al., 2017) and those with lower levels of education (Slootmaker et al., 2009). There may also be differential effectiveness according to health status at baseline (Rowe-Roberts et al., 2014), BMI and ethnicity (Losina et al., 2017). Future studies should aim to clarify which subgroups are likely to benefit most from workplace mHealth interventions. The acceptability and impact of mHealth for underrepresented groups such as shift workers, who experience unique barriers to PA and may have an elevated risk of cardiovascular disease (Puttonen et al., 2010), diabetes (Gan et al., 2015) and obesity (Sun et al., 2018), should also be explored further.

The review found some evidence for a positive impact on health and wellbeing outcomes (physiological and psychological) of mHealth interventions for PA and SB. It is recommended that future studies investigate the wider impact on health and wellbeing in addition to measures of 'organisational wellness' such as productivity, sickness absence and economic analyses, which were included as outcomes in only a small minority of studies. Most studies included in this review focused on workplaces in developed countries, with many based in academic and healthcare organisations. There will be a need for more diverse samples in a greater range of workplace settings as mHealth becomes more prevalent.

A 'weak' quality rating was assigned to a high proportion (15/25, 60%) of studies. Selection bias and lack of blinding were the weakest areas overall, although these are common issues in workplace and mHealth interventions (Cancelliere et al., 2011, Muntaner et al., 2016). Many studies lacked a true control group or did not include a reliable measure of baseline activity. Studies were highly heterogeneous in terms of methodology and outcomes, and some studies used data collection methods for the primary PA or SB outcome with below satisfactory validity and reliability. The mHealth tool itself may be an efficient method for data collection, for example most commercial activity monitors provide a real-time, objective, valid and reliable measure of step count (Evenson et al., 2015). This will be an important advance for studies that currently rely on self-reported data. There is also a need for improved reporting and consistent use of outcome measures to facilitate future synthesis of findings and meta-analyses. Combined with the relatively small number of included studies and mostly small sample sizes, these factors make it difficult to draw definitive conclusions regarding the impact of mHealth on PA and SB.

The most frequently used mHealth interventions were wearable activity monitors or trackers and/or smartphone apps. However, interventions were highly heterogeneous in terms of both mHealth and additional content, frequency, duration and mode of delivery. Similar to previous reviews of mHealth for PA and SB, the most commonly identified BCTs included self-monitoring, feedback, goal-setting and social comparison (Stephenson et al., 2017, Schoeppe et al., 2016, Sullivan and Lachman, 2016). Several studies incorporated rewards (virtual or real) and social support in their interventions.

Prompts and cues were more frequently used to target SB; this BCT was also frequent in the workplace interventions reviewed by Stephenson et al., compared with non-workplace interventions (Stephenson et al., 2017). However, descriptions of interventions and BCTs were unclear or incomplete in many cases and it was not possible to determine with confidence which specific techniques were incorporated.

Future studies should aim for more transparent reporting of intervention content and specification of embedded BCTs, to facilitate identification of the most impactful and acceptable intervention components. There may also be a need for new behaviour change taxonomies specifically for mHealth interventions, for example to include in-app competitions, various types of gamification, virtual avatars, and to distinguish between virtual and real rewards. It was also apparent that many interventions did not have a strong theoretical or evidence basis. It has been suggested that new behaviour change models and theories may be needed to account for the interactive, dynamic and adaptive nature of mHealth interventions (Riley et al., 2011).

Long-term impact and acceptability of mHealth technology is still unclear. There is a need for studies with longer duration of follow-up, further qualitative investigation of reasons for the substantial decline in engagement over time and subsequently how engagement may be maximised. Mixed methods studies will be particularly valuable to elucidate the feasibility and acceptability of mHealth to promote PA and reduce SB in a workplace setting, as well as determining the longer term impact on outcomes.

Strengths and limitations

Strengths of this review are that it was conducted in accordance with PRISMA guidelines (Moher et al., 2009), the robust nature of the search strategy, study selection and data extraction process, and the systematic assessment of study quality using the EPHPP tool (Thomas et al., 2004). The review comprehensively included a range of study types, with a combination of quantitative and mixed methods studies. This enabled synthesis of findings related to acceptability and engagement in addition to intervention effectiveness. The identification of BCTs using an established taxonomy will facilitate comparison of interventions and possible future replication. The review is the first to consider studies of mHealth for PA and SB that were conducted in a workplace setting.

The main limitations are that meta-analysis could not be performed due to the relatively small number of included studies, heterogeneity of methods and outcomes and incomplete reporting. The high proportion of studies rated 'weak' for methodological quality limits confidence in the findings. Furthermore, the possibility of publication bias should be recognised.

Summary of recommendations for future research

According to the findings of this review, it is recommended that future studies:

- Use larger samples in more diverse workplace settings (outside of academia and healthcare), include underrepresented groups such as shift workers, and consider behaviour both within and outside of the workplace
- Report more fully intervention components (including the identification of BCTs using established taxonomies such as CALO-RE) and outcomes
- Focus on SB in addition to PA, and use objective and efficient data collection methods (including the mHealth tool itself) to capture this data
- Where practicable, include a no-intervention control (experimental studies) or at least a reliable baseline measure of PA/SB (for quasi-experimental studies)
- Consider the wider impact on health and wellbeing, and work-related outcomes such as productivity and sickness absence
- Use mixed and qualitative methods to explore short- and long-term impact, feasibility and acceptability, including participants' experiences, reasons for the decline in engagement with mHealth technology, mechanisms of behaviour change, and the relationship between engagement and intervention effectiveness
- Capture data on adverse events associated with mHealth technology use

- Explore further the relative impact and feasibility of standalone mHealth and multi-component interventions, including those combined with other online and offline components
- Explore subgroup differences, including which interventions and components/BCTs are most acceptable and impactful, and for whom

CONCLUSION

There is reasonable evidence to support the use of mHealth in the promotion of PA in workplace interventions. Despite low methodological quality, early studies have demonstrated feasibility, acceptability and potential effectiveness of mHealth based interventions in a workplace context. The longer term impact, and the impact on SB, are less clear. There is a clear need for new high quality, mixed methods studies with better reporting of interventions and outcomes, in order to explore the reasons for decline in engagement over time and the longer term potential of mHealth in workplace interventions for promoting PA and reducing SB.

Abbreviations

BCT = Behaviour Change Technique
EPHPP = Effective Public Health Practice Project
mHealth = mobile health
MET = Metabolic Equivalent
MVPA = Moderate-to-Vigorous Physical Activity
PA = Physical Activity
RCT = Randomised Controlled Trial
SB = Sedentary Behaviour

Contributorship

SAB conceptualised and designed the study with support from AJW, KM, LP and JH. SAB conducted the searches. SAB and AJW completed screening, data extraction, quality assessment and BCT coding. SAB drafted and wrote the manuscript; AJW, KM, LP and JH provided critical feedback. All authors read and approved the final version.

2.4 Summary

This systematic review of 25 experimental and quasi-experimental studies found evidence to support the use of mHealth technology (wearable activity monitors and smartphone apps) to promote PA in workplace settings. This was the first published review of mHealth for PA and SB in this context. In accordance with the evidence from existing reviews of mHealth in non-workplace contexts (Fanning et al., 2012, Bort-Roig et al., 2014, Muntaner et al., 2016, Schoeppe et al., 2016, Direito et al., 2017), there was reasonable evidence that mHealth interventions are feasible, acceptable and potentially effective in increasing PA, at least in the short term. This review contributed a more comprehensive synthesis than earlier reviews, incorporating evidence from a range of study designs and considering not only effectiveness of interventions, but also findings in relation to feasibility and acceptability, including factors such as engagement, attrition and participant satisfaction. Behaviour change techniques (BCTs) were coded using the CALO-RE taxonomy (Michie et al., 2011b) to systematically characterise interventions with minimal bias.

The review also revealed remaining gaps in evidence in this field. There was insufficient evidence to draw any firm conclusions on long-term impact and acceptability, although most longer-term studies found a decline in engagement with mHealth technology over time. There was also limited and inconsistent evidence for the impact of mHealth on SB in workplace settings. Few studies had qualitatively assessed feasibility and acceptability, and the wider impact on health, wellbeing and productivity was unclear. The included studies were predominantly set in universities and healthcare organisations and there were no studies of mHealth interventions for PA or SB in the police force. The gaps identified were used to help to define the specific research questions for the PAW-Force study, which are given in **Chapter 1**.

Many of the included studies were of weak methodological quality according to the EPHPP quality assessment tool. It was recognised that some of the aspects of study quality were universal and unavoidable issues, such as potential selection bias resulting from recruitment of volunteers in the workplace, and the infeasibility of blinding participants to the mHealth interventions. Nevertheless, the EPHPP assessment criteria were considered in the design of the PAW-Force study, and the limitations of the studies identified in the review were considered and used to guide methodology for the study. The intervention was clearly defined and reported with coding of BCTs, and outcomes were comprehensively reported to facilitate meta-analysis for future systematic reviewers. A valid and reliable primary outcome was selected, with a reliable baseline measure. In addition, potential confounders were considered in analysis of the results (i.e. capture of data on factors influencing PA levels). The next chapter will include a description of the intervention and methods.

CHAPTER 3 THE PAW-FORCE STUDY: INTERVENTION AND METHODS

In this chapter, the methods used in the PAW-Force study are outlined. This includes a description of the study design, setting and recruitment, the intervention, data collection and outcomes, and data analysis methods.

3.1 Study design

3.1.1 Overview and rationale for study design

While the intervention was decided by the Devon and Cornwall Police, the study design was developed in consultation with academics and key stakeholders in the police force. Decisions were made based on the existing evidence base in addition to strategic and practical considerations for the police force. After extensive discussion, a quasi-experimental single group pre- and post- pilot study design was considered more appropriate than an experimental approach (e.g. a randomised controlled trial, RCT), for the following reasons:

- Based on gaps in existing evidence, it was deemed necessary to consider a range of aspects including feasibility and acceptability of both the intervention and study methods. It was agreed that a feasibility or pilot study design was most appropriate to meet this need. A key aim was to explore the feasibility of outcomes and potential impact, rather than definitive effectiveness.
- Organisational leaders had already taken the decision to promote the Bupa Boost app widely across the police forces involved. This was an issue resulting from the ‘real-world’ nature of the study, which made controlling which individuals did and did not receive the intervention infeasible. In addition, randomising between sites would not have been feasible as the sites (Plymouth Basic Command Unit (BCU) and North Dorset territorial region) were diverse socio-demographically and geographically. Randomising within sites raised the issue of possible contamination; officers and staff could be partially exposed to the intervention and encouraged by their colleagues.
- There was a need for a consistent, objective measure of physical activity (PA) for all participants. If a control group had worn a Fitbit® activity monitor they would be receiving some degree of intervention. There was

questionable value in including a comparison or control group that was not a ‘true’ control group, as several previous workplace mobile health (mHealth) intervention studies had done (such as Gremaud et al., 2018, Patel et al., 2016a, Patel et al., 2016b, Patel et al., 2018, and van Dantzig et al., 2013). The use of accelerometers was considered for capturing PA and sedentary time, but this posed logistical issues and was considered burdensome for police officers, who already had to wear the Fitbit® in addition to the policing uniform and equipment.

- Non-RCT designs are increasingly being advocated for mHealth studies; these are considered more appropriate and efficient given the rapidly changing and complex nature of mHealth technology (Dallery et al., 2013, Pham et al., 2016, McCallum et al., 2018, Arigo et al., 2019).

3.1.2 Why and how mixed methods were used

A mixed methods approach is recommended by the UK Medical Research Council (MRC) for the evaluation of complex interventions (Moore et al., 2014) and is considered important when assessing feasibility (Bowen et al., 2009). This approach was taken to capture quantitative outcomes in relation to impact of the technology, and more qualitative aspects such as feasibility and acceptability of the intervention. A fuller picture and more comprehensive understanding was produced through the incorporation of both quantitative and qualitative data.

The key dimensions of mixed methods designs include timing of methods (simultaneous or sequential), importance or priority, when quantitative and qualitative data are integrated, and function of integration (Bryman, 2006). In the present study, a simultaneous design was used with quantitative and qualitative data collected concurrently, and equal importance was given to the two types of data. Quantitative and qualitative data were integrated at the individual participant level during framework analysis, and to a greater extent at the dataset level during interpretation of overall findings (described in **section 3.5.3**). There were multiple functions of integration. Quantitative data were used to inform sampling for subsequent qualitative interviews, for example to ensure that participants with a range of baseline activity levels (and those who had shown varying degrees of engagement with the intervention and behaviour change) were represented. Qualitative data were used to expand on, explain, and triangulate with quantitative findings. The two main types of triangulation, confirmation and complementarity (Small, 2011), were used (see **section 3.5.3**).

3.1.3 Ethical approval and study registration

Ethical approval for the PAW-Force study was granted by the University of Exeter Medical School Research Ethics Committee in March 2017 (Ref. Mar17/B/116).

The study protocol was registered with ClincalTrials.gov in May 2017 (Ref. NCT03169179).

3.2 Setting, recruitment and sampling

3.2.1 Setting

The study was conducted in two sites - Plymouth BCU, Devon and Cornwall Police (an urban site) and North Dorset territorial region, Dorset Police (a more rural site). Most police officers and staff who worked at Plymouth BCU were based at one of four sub-sites (Crownhill, Charles Cross, Devonport and Plympton/Plymstock). The main work streams within Plymouth BCU included local investigation, local/neighbourhood policing and response policing. There were several additional work teams within the BCU including intelligence, communications and administration. North Dorset was composed of four main sub-sites (Blandford, Gillingham, Shaftesbury and Sturminster Newton). The policing functions covered by these sites included local/neighbourhood policing and response policing.

3.2.2 Recruitment

Recruitment took place from April to May 2017. The study was advertised using a range of methods, including:

- posters in the workplace;
- the Devon & Cornwall and Dorset Police intranets;

- e-mails throughout the organisations; and
- staff bulletins.

Special constables, who did not have access to the company intranet, received information on the study via the ‘DutySheet’ software program. Targeted e-mails were also sent via the organisations’ ‘admin hubs’ to ensure police staff were included. The relevant Communications teams for the Devon and Cornwall and Dorset Police forces assisted in promoting the study during the recruitment phase. Officers and staff who were interested in participating expressed their interest to the researcher (via telephone or e-mail), and were provided with an information sheet (see **Appendix 4a**) and a link to complete the online screening and consent forms (see **Appendix 4b**).

All participants were volunteers. Sampling methods used were a combination of initial opportunistic sampling (participants were selected on a ‘first come first served’ basis) and later purposive maximal variation sampling to ensure representation of the various occupational groups (police officers, police staff, Police Community Support Officers or PCSOs, and special constables). Snowballing via word of mouth was also used to complement the maximal variation sampling method.

Inclusion criteria were those who expected to be employed within the police force for the duration of the study, and who owned (or had access to) a smartphone or tablet that was compatible with the Bupa Boost app (Apple or Android 4.0.3 or higher), with Bluetooth and internet access. The only exclusion criterion was severe limited mobility, i.e. those who would be physically unable to increase their step count over the duration of the study.

Following completion of screening and consent forms, eligible participants were contacted with further details and information on when and where to collect the Fitbit® device. Those who were ineligible (or who failed to provide consent) were informed of this and thanked for their interest.

3.2.3 Sample size calculation

It is important to justify sample size in feasibility and pilot trials (Billingham et al., 2013). The number of participants to be recruited was based on feasibility and availability of resources, the sample size required for statistical significance, and estimated attrition. Although the aim of the study was not to establish effectiveness, power calculations were performed as a starting point to determine an appropriate sample size. Power calculations were based on effect sizes reported in similar existing studies (see **Appendix 3** for details of calculations). The expected effect sizes for the primary quantitative outcome variable (mean daily step count) and key secondary outcomes (weekly minutes of moderate-to-vigorous physical activity (MVPA) and SF-12 Health Survey score) were considered and the largest calculated sample size was taken. Sample size calculations were performed in Stata version 14.0 (StataCorp, 2015) using the paired t-test comparing two correlated means. Calculations were based on 80% power and the 5% significance level (p = 0.05).

A sample size of 128 was shown to be sufficient to detect a mean increase in daily step count of 300 (standard deviation, SD = 1000, correlation 0.3) or 500 (SD = 2000, correlation 0.5). This was also sufficient to detect changes in the secondary outcomes being considered. It was difficult to predict the expected attrition rate, as dropout rates in existing mHealth studies in the workplace have been highly variable, ranging from 0% to 74% (Buckingham et al., 2019). A minimum target sample size of 150 was initially specified to allow for 15% attrition. Due to higher than expected interest in the study and an ample supply of Fitbit® devices, the sample size was further increased to approximately 180 (allowing for 25 to 30% attrition).

Of the 190 participants screened for eligibility, 182 began the study. Further detail on the numbers of participants recruited is given in **Chapter 4**.

3.2.4 Sampling of interviewees

Potential participants indicated their willingness to be interviewed on the consent form for the main study (see **Appendix 4b**). Of the 182 eligible participants who began the study, 175 (96%) provided consent to be interviewed. A purposive sample from this group was selected for baseline interviews, with maximal variation according to age, gender and occupation. Both sites (Plymouth BCU and North Dorset) were represented. Participants interviewed at baseline (n = 10) were invited to take part in subsequent interviews (post-intervention and follow-up). However, due to reasons of availability, a need to ensure representation of those with a range of activity levels, and a need to ensure representation of those who had shown various levels of engagement with the intervention and behaviour change, some additional interviewees were purposively selected for later interviews. In total, 32 interviews were conducted with 16 participants. Further detail on the number and characteristics of interviewees is given in **Chapter 4**. As described in **section 3.1.2**, quantitative data were used to help to ensure a varied sample for qualitative data collection. The quantitative data included baseline step count, step changes through the study and usage data (including Fitbit® wear time and use of the Bupa Boost app).

3.3 The intervention and study procedures

3.3.1 Overview of the intervention and study process

As illustrated in **Figure 2**, participants received a 12-week intervention with 8 months to follow-up. Following collection of their Fitbit Charge 2® activity monitor from the relevant administrative office at each site, participants were instructed to wear the device on their wrist for seven consecutive days at baseline (week 0) while continuing to maintain their usual activity levels. The screen was covered by a sticker during this baseline week, and participants were asked not to log in to the Fitbit® app during this time; this helped to ensure a valid pre-intervention measure of the primary quantitative outcome (daily step count).

Following completion of the first (baseline) questionnaire, participants entered the intervention phase, where they began to use the Fitbit® (with the screen uncovered) together with the Bupa Boost app. The 12-week intervention was divided into an 'individual' and a 'social' phase. During weeks 1 to 6, participants were instructed to only use the 'individual' features of the Fitbit® and Bupa Boost app. These included:

- goal-setting (to increase their daily step count and any other PA goals that were important to them);
- self-monitoring;
- receiving feedback on progress via the app;

- earning virtual rewards for their achievements (wellness points and badges); and
- access to the ‘Bupa library’ within Bupa Boost, for self-help information on maintaining a healthy lifestyle.

During weeks 7-12, participants entered the ‘social’ phase, where they were encouraged to link up with their colleagues within the Bupa Boost app. In addition to the ‘individual’ features, they were able to:

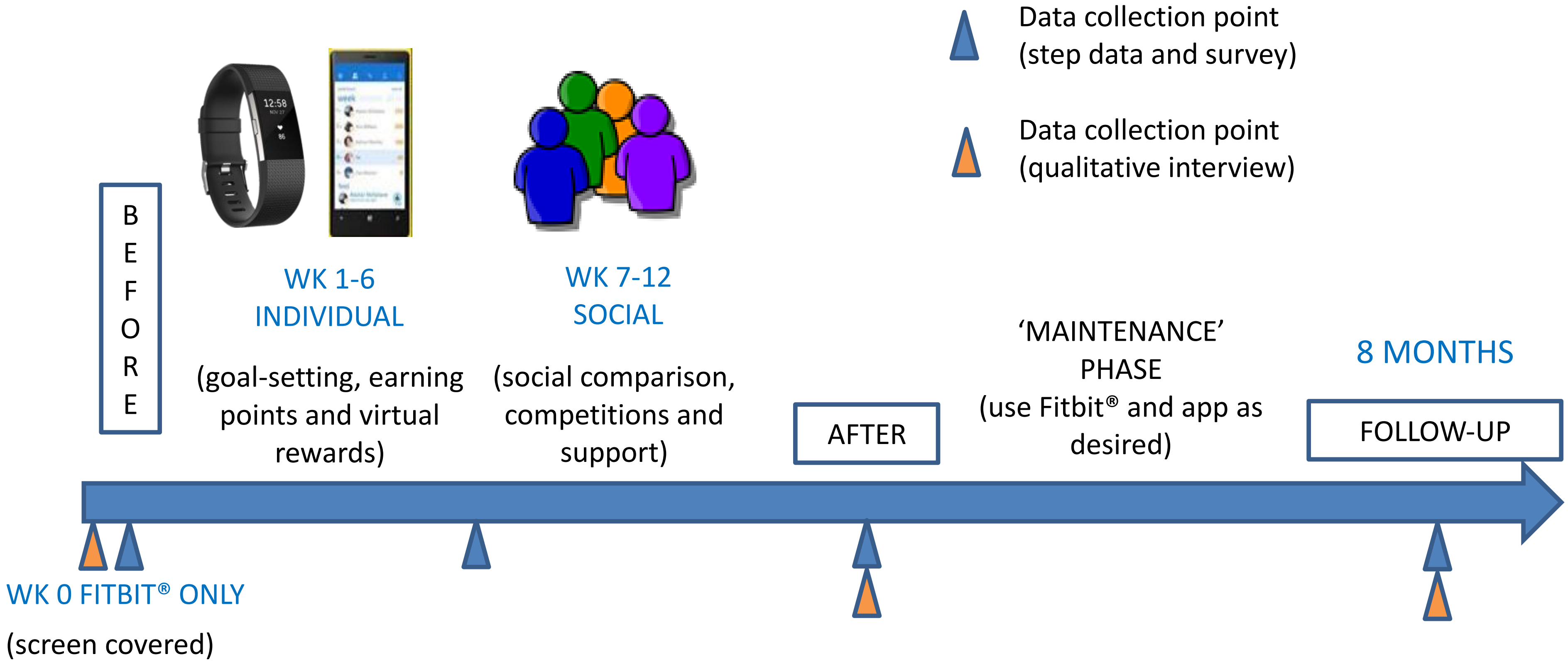
- compare themselves with their colleagues via a social feed;
- compete with their colleagues (in individual challenges and/or working as part of a team in company challenges); and
- give and receive social support through virtual ‘likes’ and messages.

While participants were also able to use the Fitbit® app in the intervention phase, they were encouraged to only use the Bupa Boost app for the social features. At the end of the 12 weeks, there was a five month ‘maintenance phase’ during which participants continued to use the Fitbit® and Bupa Boost app as and when they desired. A detailed description of the intervention (including component features and inherent behaviour change techniques or BCTs) is provided in **section 3.3.3**.

As shown in **Figure 2**, step data were collected at four time points – baseline (week 0), mid-intervention (week 6), post-intervention (week 12) and follow-up (month 8). Participants were asked to complete an online survey at the same four time points; this captured data on self-reported PA and sedentary time, physical and mental health-related quality of life, perceived stress and perceived productivity. The surveys also included quantitative and qualitative outcomes in relation to engagement and acceptability of the intervention (such as perceived usability and usefulness). Step data were downloaded by participants and uploaded to the survey or e-mailed to the researcher. Outcomes are described fully in **section 3.4.3**.

Interviews were conducted at three time points – baseline (week 0), post-intervention (week 12) and follow-up (month 8). Interview methods are described in detail in **section 3.4.4**.

Figure 2 Overview of the PAW-Force study process



3.3.2 Study management

The study was managed remotely, with regular written instructions sent via e-mail by the researcher. These included details of how to register, wear and charge the Fitbit®, how to obtain and upload step count data, how to use the Bupa Boost app and how to connect it to the Fitbit®, and where to find further information or support with technical issues. Participants also received written guidance on setting 'SMART' goals (Specific, Measurable, Achievable, Relevant and Time-bound). Details of instructions sent at various time points are included in **Appendix 4e**.

While there were no formal 'facilitators' in the study, at times when the researcher was unavailable, participants were signposted to a designated health and wellbeing champion at each site (a police inspector in Plymouth BCU and a PCSO in North Dorset), who provided assistance and support with queries. Queries included those related to study participation and technical issues with the Fitbit® device or Bupa Boost app. A supply of spare Fitbits®, chargers and straps was also available from the administrative offices at each site; these were distributed to individuals as and when needed. It is important to note that the researcher and health and wellbeing champions had a supportive role only, and did not play an active role in delivering the intervention. The PAW-Force intervention was intended as a self-directed and standalone mHealth intervention, delivered with minimal in-person input, as this was the least resource-intensive and most practical approach.

3.3.3 The PAW-Force intervention: Description and theoretical basis

3.3.3.1 Description and coding of the intervention

Systematic reviews of mHealth studies in workplace and non-workplace settings have reported that interventions are often poorly described (Direito et al., 2017, Howarth et al., 2018, Buckingham et al., 2019) and frequently lack a clear theoretical basis (Fanning et al., 2012, Bort-Roig et al., 2014). Explicit reporting of complex interventions and their theoretical components is a key recommendation of the MRC and the National Institute for Health and Care Excellence (NICE) (National Institute for Health and Care Excellence, 2007, Craig et al., 2008). This is necessary to understand which components of the intervention may be most impactful, to enable comparison and replication of interventions, and to facilitate further development and evaluation of interventions and theories (Craig et al., 2008, Abraham and Michie, 2008).

The PAW-Force intervention (i.e. the Fitbit® and Bupa Boost app) was 'ready-made', that is provided by the Devon and Cornwall Police Force in conjunction with Bupa rather than designed or developed by the research team, and therefore was not based on any particular overarching theory. Despite this, care was taken to ensure that its behavioural components were clearly identified and reported.

The intervention was coded according to Michie and colleagues' 'Coventry, Aberdeen and London – Refined' (CALO-RE) taxonomy, a standardised classification system of BCTs for PA and healthy eating behaviours (Michie et al., 2011b). A BCT is defined as a theory-based component that is designed to change behaviour. The CALO-RE taxonomy was selected as it was specifically designed for PA and healthy eating behaviours, and has been used widely including in interventions of smartphone apps and wearable activity monitors (Conroy et al., 2014, Mercer et al., 2016). The CALO-RE taxonomy also contains a relatively small number of BCTs (n = 40), which were felt to be more relevant compared with other available taxonomies, such as the Behaviour Change Taxonomy version 1, which contained 93 more generic BCTs (Michie et al., 2013). Furthermore, the CALO-RE taxonomy had been used to code interventions in the systematic review (see **Chapter 2**).

The subjective nature of the coding process was an important consideration. Although clear definitions of BCTs are given by Michie and colleagues, some subjectivity in judgement cannot be avoided. Other researchers have noted that the identification and coding of BCTs in wearable activity monitors may depend on how the user experiences the device and which features are discovered (Mercer et al., 2016). To minimise subjective biases, the PAW-Force intervention was independently coded by two researchers (S. Buckingham and A.J. Williams) who were trained in the Behaviour Change Taxonomy version 1 (no separate training programme exists for the CALO-RE taxonomy) and who had used both the Fitbit® and Bupa Boost app. For the small number of BCTs where disagreement existed, consensus was reached through discussion. The specific BCTs identified within the Fitbit® and Bupa Boost app and the associated components or features are detailed in **Table 6**.

Table 6 Coded BCTs and intervention components within the Fitbit® activity monitor and Bupa Boost app

Specific behaviour change technique (BCT, CALO-RE taxonomy)	Included in Fitbit®? (S = Social phase only)	Component or feature	Included in Bupa Boost app? (S = Social phase only)	Component or feature
1. Provide information on consequences of behaviour in **general**	N		Y	Self-help information on PA and health (Bupa library)
2. Provide information on consequences of behaviour to the **individual**	N		N	
3. Provide information about others' approval	Y***** (S)	Messaging with friends (approval or disapproval).	Y (S)	Messaging with colleagues (approval or disapproval). 'Likes' for others' achievements.
4. Provide normative information about others' behaviour	Y***** (S)	Compare weekly steps and achievements with friends within app.	Y (S)	Social feed – compare daily PA and goals achieved with colleagues. Leader board – compare wellness points and achievements with colleagues (weekly, monthly, all time).
5. Goal setting (behaviour)	Y	Set activity goals in app, e.g. 10,000 daily steps, 30 daily active minutes, 250 steps per hour.	Y	Set general or specific PA/fitness goals (choose from suggested list or custom, e.g. walk to work). Also nutrition, mindfulness and relaxation (if desired).
6. Goal setting (outcome)	Y	Set weight goals in app (if desired)	N	
7. Action planning	Y	Activity goals specified in terms of context, frequency, duration or intensity – e.g. 10,000 daily steps.	Y	Set specific activity goals in app – e.g. walk to work on three days per week.
8. Barrier identification / problem solving	N		N	
9. Set graded tasks	Y	Small goals of 250 steps per hour – helps to meet overall daily step goal.	N	
10. Prompt review of behavioural goals	Y	Detail of goals achieved given in app. Goals can be edited at any time.	Y	Tick off goals in app when achieved. Goals achieved are listed in 'activity feed'. Goals can be edited at any time.
11. Prompt review of outcome goals	Y	Weight goals may be reviewed and edited	N	

Specific behaviour change technique (BCT, CALO-RE taxonomy)	Included in Fitbit®? (S = Social phase only)	Component or feature	Included in Bupa Boost app? (S = Social phase only)	Component or feature
		at any time (if desired)		
12. Provide rewards contingent on effort or progress towards behaviour	Y	Virtual badges for PA progress and 'lifetime achievements' (e.g. the 'London Underground' badge for walking 402 kilometres in total)	Y	Earn wellness points for progress towards behaviour (e.g. earn more points for higher daily steps). Virtual badges for PA progress and 'lifetime achievements'.
13. Provide rewards contingent on successful behaviour	Y	Virtual badges and trophies for achieving goals.	Y	Wellness points and virtual badges for goals achieved (and goal streaks).
14. Shaping	Y	Graded (virtual) rewards for greater achievements over time. Goal streaks within challenges*	Y	Graded (virtual) rewards for greater achievements over time. Goal streaks.
15. Prompting generalisation of a target behaviour	N		N	
16. Prompt self-monitoring of behaviour	Y	Monitoring of steps, distance, calories burned, floors climbed, active minutes (also biofeedback - sleep and heart rate)	Y	Monitoring of steps and activity duration (via data from Fitbit®)
17. Prompt self-monitoring of behavioural outcome	Y	Monitoring of resting heart rate and weight (if desired)	N	
18. Prompting focus on past success[1]	N		N	
19. Provide feedback on performance	Y	Feedback on progress towards goals and goals achieved (in-app and notifications). Includes visual feedback – colour changes to green when activity goals are met. Feedback is real-time and personalised. Can also review past reports of PA and sedentary time in app (also heart rate and sleep if desired)	Y	Feedback on progress towards goals and goals achieved (in-app and notifications). Real-time and personalised. View past reports of PA, goals achieved, and previous wellness points and badges.
20. Provide information on **where and when** to perform the behaviour	N		N	

Specific behaviour change technique (BCT, CALO-RE taxonomy)	Included in Fitbit®? (S = Social phase only)	Component or feature	Included in Bupa Boost app? (S = Social phase only)	Component or feature
21. Provide instruction on how to perform the behaviour	N		N	
22. Model / demonstrate the behaviour	N		N	
23. Teach to use prompts / cues	Y	Smart notifications for PA and SB, e.g. reminders to move when 250 steps per hour not reached	Y	Prompts/notifications to remind the individual when a goal has not been achieved.
24. Environmental restructuring	N		N	
25. Agree behavioural contract	N		N	
26. Prompt practice	N		N	
27. Use of follow-up prompts	N	(Prompts are consistent in frequency throughout the intervention and maintenance phase)	N	(Prompts are consistent in frequency throughout the intervention and maintenance phase)
28. Facilitate social comparison	Y***** (S)	Compare weekly steps and achievements with friends within app. Social challenges.	Y (S)	Social feed. Leader board. Individual and group competitions / challenges.
29. Plan social support / social change	Y***** (S)	Supportive messaging with friends.	Y (S)	Supportive messaging with colleagues. 'Likes' for others' achievements.
30. Prompt identification as role model / position advocate	N		N	
31. Prompt anticipated regret	N		N	
32. Fear arousal	N		N	
33. Prompt self-talk	N		N	
34. Prompt use of imagery	N		N	
35. Relapse prevention / coping planning	N		N	
36. Stress management / emotional control training	Y	Relaxation exercises within the Fitbit® (guided breathing)	N	
37. Motivational interviewing	N		N	
38. Time management	N		N	
39. General communication skills training	N		N	
40. Stimulate anticipation of future rewards	Y	User is aware from the outset that (virtual) rewards will be given for progress and achievements.	Y	View available badges and achievements / activities needed to earn them.

***** = minimal use in this intervention (participants were encouraged to focus on the Bupa Boost app for these behaviour change techniques) [1] Review of past reports of PA and sedentary time is not classed as focus on past success as this BCT relates to behaviour preceding the intervention according to the CALO-RE taxonomy

In summary, the intervention included 20 of a possible 40 unique BCTs according to the CALO-RE taxonomy. Five BCTs were included in the Fitbit® (or Fitbit® app) only; these were: **goal-setting for outcome; set graded tasks; prompt review of outcome goals; prompt self-monitoring of outcome; stress management training**. One BCT (**provision of information on consequences of behaviour in general**) was included in the Bupa Boost app only. A further 14 BCTs were included in both the Fitbit® and Bupa Boost app:

- **Provide information about others' approval (social phase only)**
- **Provide normative information about others' behaviour (social phase only)**
- **Goal-setting (behaviour)**
- **Action planning**
- **Prompt review of behavioural goals**
- **Provide rewards contingent on effort or progress towards behaviour**
- **Provide rewards contingent on successful behaviour**
- **Shaping**
- **Prompt self-monitoring of behaviour**
- **Provide feedback on performance**
- **Teach to use prompts / cues**
- **Facilitate social comparison (social phase only)**
- **Plan social support / social change (social phase only)**
- **Stimulate anticipation of future rewards**

Participants were encouraged to use the Fitbit® with the Bupa Boost app. The Fitbit® app and Bupa Boost app were linked via a tick box within Bupa Boost, which enabled data on steps and activity duration to be automatically extracted and included in an 'activity feed' in the Bupa Boost app. It was necessary to download the Fitbit® app to set up the Fitbit®, link it to Bupa Boost, access help features etc., but participants were encouraged to focus mainly on using the Bupa Boost app, particularly for the social phase in weeks 7 to 12. This was to maximise consistency of the intervention received between participants. Similarly, participants logged on to the Fitbit® website to download their step data files but received no intervention content through the website.

3.3.3.2 Basis for 'individual and 'social' phases

The 'individual' and 'social' phases (described in **section 3.3.1**) had a practical and theoretical basis. While it was initially planned to explore the various intervention components in a larger number of groups or phases, this was not feasible as access to specific features could not be restricted within the Bupa Boost app. Social influence theories emphasise the effects of social factors on behaviour, for example Bandura's social cognitive theory states that learning and behaviour change occur as a direct result of observing others (Bandura, 1997). It was therefore hypothesised that the 'social' app features, which included social support, competition and comparison in addition to all of the 'individual' features, would have a greater perceived and observed impact on PA than the individual features alone. The Socio-Ecological Model (described in **Chapter 4**) also proposes that the more layers that are targeted by an intervention, the more likely behaviour change is to occur, i.e. targeting both the individual and interpersonal layers will have a greater impact than the individual layer alone (Bronfenbrenner, 1989, Stokols, 1996, Robinson, 2008). A number of previous studies had provided support for the power of social processes in mHealth for behaviour change. For example, King and colleagues found that a socially framed app (including social support, modelling, normative feedback and competition) was associated with greater increases in PA and greater reductions in sedentary time than an analytically framed app and an affectively framed app with no social features (King et al., 2016). However, other studies had found no differences in PA between app-based individual and social feedback (Harries et al., 2016). In the PAW-Force study, it was feasible for participants to experience all of the individual features within the Bupa Boost app (weeks 1 to 6) before linking up with their colleagues within the app, which automatically gave them access to all of the social features (weeks 7 to 12).

3.3.3.3 Proposed logic model for the intervention

The MRC recommends producing a logic model to improve conceptualisation and understanding of complex interventions and their causal pathways (Moore et al., 2014). It is recommended that the logic model should capture the core components of process evaluation and how they interact with each other, including implementation (what is delivered and how), mechanisms of impact (how the intervention works, including participant responses, mediators and consequences), outcomes (short, intermediate and longer term), and the wider context in which the intervention is delivered (Moore et al., 2014).

Prior to the start of the PAW-Force study, a logic model was developed and is shown in **Figure 3**.

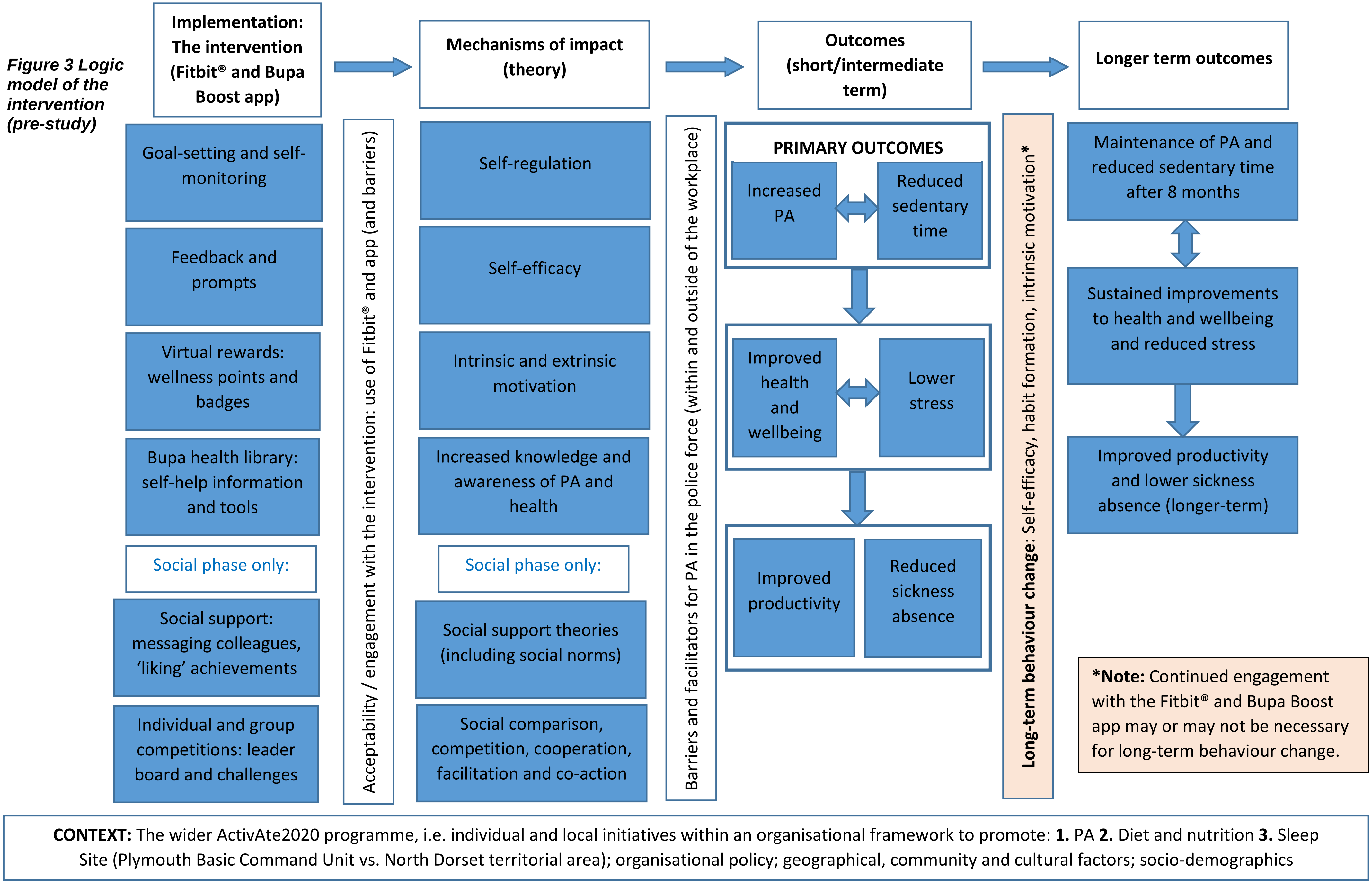


Figure 3 Logic model of the intervention (pre-study)

In this logic model, mechanisms may apply to more than one intervention component and vice versa. For example, self-efficacy may be enhanced through both goal-setting and feedback. Similarly, goal-setting and self-monitoring may lead to improved knowledge and awareness and greater self-efficacy. We aimed to clarify the preferred and most impactful intervention components and associated mechanisms of impact in this study.

It is important to consider barriers and facilitators related to engagement with the intervention (which will inhibit or facilitate the mechanisms of impact, for example lack of technological skills may prevent use) and barriers and facilitators for PA within the police force, both within and outside of the workplace (which may impact on mechanisms and outcomes, and potentially also implementation). These may be at individual, social and organisational levels. We aimed to elucidate these barriers and facilitators and develop a deeper understanding of behaviour change mechanisms in this context. Contextual factors, i.e. anything external to the intervention such as individual socio-demographics, organisational policy, or geographical site, may impact on (or be impacted by) implementation, mechanisms and outcomes. We aimed to improve understanding of the wider context of PA and sedentary behaviour (SB) in the police force, and the impact of context on the success (or failure) of the intervention.

Some outcomes may be directly influenced through the intervention (for example reduced stress through setting mindfulness and relaxation goals in the Bupa Boost app) or indirectly influenced through increased PA (Emerson et al., 2017). There is a proposed negative association between PA and sedentary time – as PA increases, sedentary time is expected to reduce and vice versa. It was proposed that there will be a similar negative association between health and wellbeing and stress, and that these outcomes would in turn lead to improved productivity (Guo et al., 2015) and reduced sickness absence (Baicker et al., 2010). It was hypothesised that reduced sickness absence may include both fewer sickness episodes and shorter durations of absence. It was proposed that a positive reinforcing loop between the outcomes of PA and sedentary time and health and wellbeing and reduced stress may become apparent in the longer term. It was postulated that longer-term behaviour change mechanisms may include improved self-efficacy, habit formation and intrinsic motivation. Outcomes in the short and longer term were explored quantitatively and qualitatively.

Key assumptions of this model were that there was a need to improve PA levels and reduce SB within the police force, that the infrastructure and resources available would be sufficient to run a pilot study of the intervention, and that it would be feasible to recruit participants to the study to explore these processes.

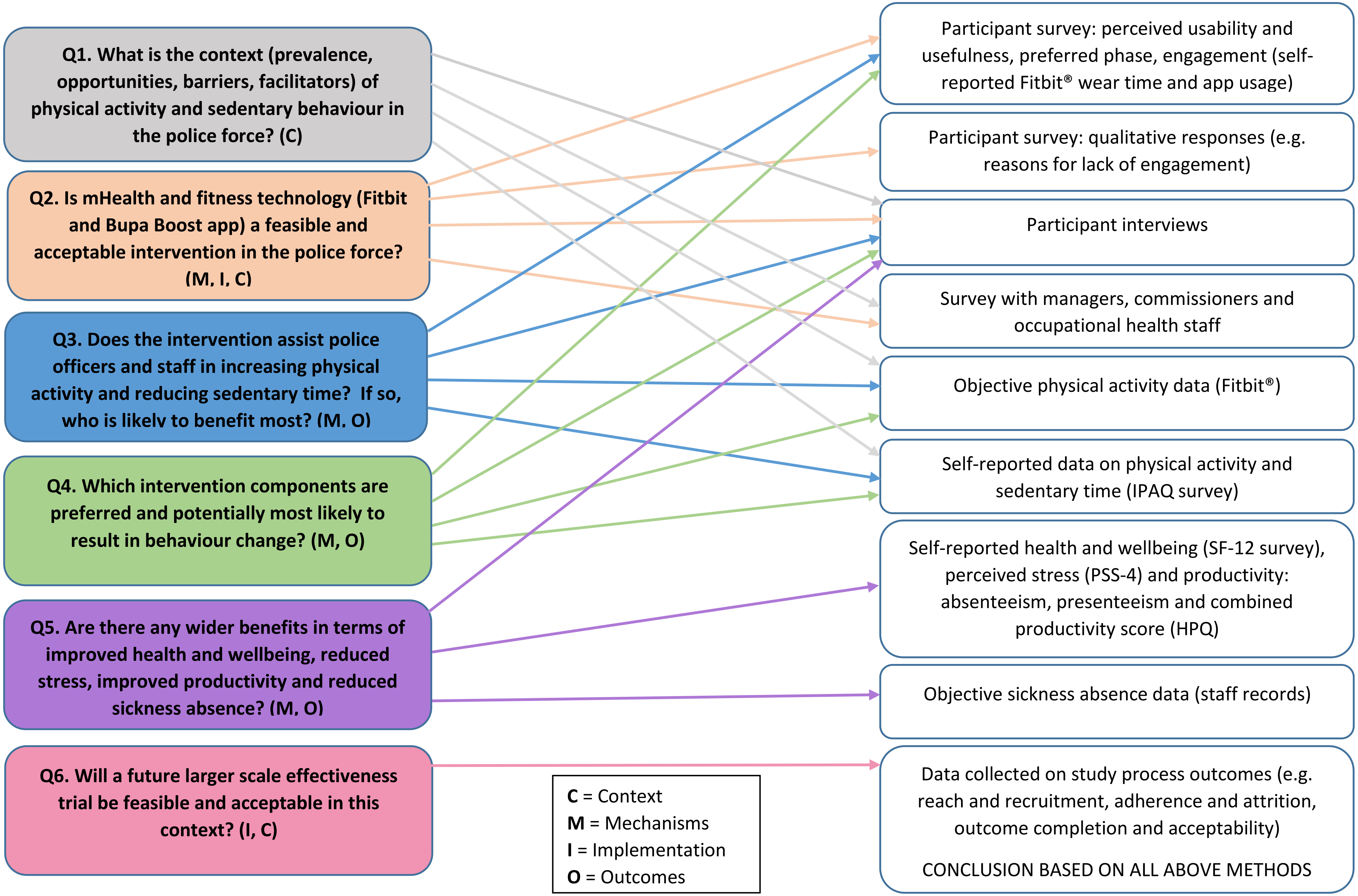
A revised logic model based on the study findings is presented in **Chapter 8**.

3.4 Data collection and outcomes

3.4.1 Overview of research questions and data sources

The research questions and methods used in the process and outcome evaluation of the PAW-Force study are shown in **Figure 4**. The six main research questions were each addressed using a combination of quantitative and qualitative data sources, including objectively measured step count, data from online surveys, participant interviews, a separate survey with police managers, commissioners and occupational health staff, and sickness absence data. Each of these data sources is presented below.

Figure 4 Principal research questions and methods in the PAW-Force study: Process and outcome evaluation



3.4.2 Objective physical activity data (Fitbit®)

The primary quantitative outcome was (change in) mean daily step count as recorded by the Fitbit®. This objective measure was selected for its known validity and reliability. Evidence from systematic reviews indicates that Fitbit® activity monitors have high validity for the outcome of step count when compared with research-grade accelerometers in both laboratory and 'free-living' situations (Evenson et al., 2015, Feehan et al., 2018). High inter-device reliability for recording of steps has also been indicated (Evenson et al., 2015, Dontje et al., 2015).

Step data were collected at baseline (week 0), mid-intervention (week 6), post-intervention (week 12) and follow-up (month 8). Participants were instructed to use the Fitbit® website to download their step data from the previous seven days (automatically collected by the activity monitor) at each time point. They then either uploaded their step data (as a CSV or Excel file) to the online survey or sent it directly via e-mail to the researcher. Mean daily step count from the previous seven days was calculated by the researcher for each participant (see **section 3.5.1.1**).

3.4.3 Online surveys

At the same time as uploading their step count data (in week 0, week 6, week 12 and month 8), participants were asked to complete an online questionnaire relating to their PA levels, sedentary time, health-related quality of life, perceived stress and self-perceived productivity. Surveys were administered and managed using LimeSurvey (Limesurvey GmbH, 2003), an open source online survey application. Data were transferred securely to the researcher using a unique study ID.

Before being administered to participants, the survey was piloted with 10 non-participants (friends and colleagues of the researcher). The purpose of this was to obtain feedback on questionnaire content and structure (which was used to make some small improvements to readability and format), and to estimate the time it would take each participant to complete the questionnaire (approximately 10-15 minutes).

The baseline questionnaire was available for only five days to ensure a timely and simultaneous start of the study for all participants, whereas each subsequent questionnaire was available for approximately 10 days. A maximum of two reminders at each data collection point was sent to participants who had not completed the survey.

The following questionnaires and data/outcomes were included:

- *International Physical Activity Questionnaire (IPAQ) short version (week 0, week 6, week 12, month 8)*
  The need to capture activities in addition to step count, and also sedentary time, was recognised. Despite providing an accurate measure of steps, studies had shown that Fitbit® activity monitors had unknown or poor validity for other types of PA and may overestimate moderate and vigorous PA while underestimating sedentary time (Reid et al., 2017, Dominick et al., 2016). The IPAQ was selected to capture self-reported PA and sedentary time. The questionnaire has known validity and reliability in a range of populations internationally (Craig et al., 2003). The IPAQ captured data on PA during the previous seven days which coincided with the data collection period for steps. Clear definitions and visual examples of moderate and vigorous activities were given, based on Ainsworth's compendium of physical activities (Ainsworth et al., 2011), to assist participants with completion. The IPAQ outcomes included:

    - Total PA (in minutes/week and MET-minutes/week*****)
    - Moderate-to-vigorous PA or MVPA (MET-minutes/week*****)
    - Sedentary time (estimated hours spent sedentary on a typical weekday)
    - The IPAQ also provided a categorical outcome for 'low', 'moderate' or 'high' activity level based on the other outcomes (definitions are given in **Chapter 4**).

  *******Note: MET-minutes/week is the usual outcome from the IPAQ (Craig et al., 2003). The scoring spread sheet used (Cheng, 2016) did not allow calculation of minutes/week of MVPA.*

- *SF-12 Health Survey (week 0, week 6, week 12, month 8)*

  The SF-12 Health Survey (Ware et al., 1996) was used to assess generic health and wellbeing including physical and mental health-related quality of life. The SF-12 was selected for its concise nature, suitability for use in a healthy population, and known validity (Okonkwo et al., 2010). In this 12-item questionnaire, participants were asked to recall and report how they felt physically and emotionally during the previous four weeks.

  The outcomes included:

  - Physical Component Score (scale of 0-100*****)
  - Mental Component Score (scale of 0-100*****)

  *******Note: Higher scores indicate higher health-related quality of life*

- *Perceived Stress Scale (PSS-4) (week 0, week 6, week 12, month 8)*

  The 4-item version of the Perceived Stress Scale (PSS-4) was selected as it has been widely used with known validity and reliability (Cohen et al., 1983, Cohen and Williamson, 1988, Warttig et al., 2013). The PSS was designed as a measure of perceived stress in community samples and has been previously used in the police force, for example to study the impact of a resilience training intervention in an urban police department (Ramey et al., 2016). Study participants provided Likert-scale responses based on how they had felt during the previous four weeks. The outcome was a single stress score on a scale of 0 to 16 (higher scores indicate higher perceived stress).

- *Self-perceived productivity from the absenteeism and presenteeism questions of the World Health Organisation (WHO) Health and Work Performance Questionnaire (HPQ) (week 0, week 6, week 12, month 8)*

  The absenteeism and presenteeism questions of the WHO's Health and Work Performance Questionnaire (HPQ) provided summary measures of perceived productivity in the workplace. The HPQ has been validated in a range of occupations (Kessler et al., 2003) and has been reported to have excellent validity, reliability and sensitivity to change (Kessler et al., 2004). Participants were asked to report their absenteeism and presenteeism (i.e. working whilst unwell) during the previous four weeks. The outcomes included:

    - Absolute absenteeism (hours lost per month)
    - Relative absenteeism (absenteeism relative to expected hours of work)
    - Absolute presenteeism (self-rated work performance)
    - Relative presenteeism (self-rated work performance relative to colleagues)
    - Combined productivity score (relative absenteeism and relative presenteeism)

- *Socio-demographic, occupational and health data (week 0 only)*
  The following data were collected at baseline:
    - Age (years)
    - Gender
    - Ethnicity
    - Marital status
    - Living circumstances (e.g. living alone) and children
    - Dog ownership (known to be an important independent determinant of PA and SB and therefore an important variable to include and control for in PA trials (Dall et al., 2017))
    - Postcode and urban/rural residence
    - Education
    - Occupational details including:
        - Organisation, site and work team
        - Occupation/rank
        - Whether role is mainly active/sedentary/equally active and sedentary
        - Years of service in the police force
        - Working hours and shift work
        - Income category (pro-rata if part-time)
    - Smoking status
    - Alcohol use
    - Presence/absence of a chronic health condition (and name of condition)
    - Previous use of a wearable activity monitor and/or health or fitness app

- *Engagement with the intervention (week 6, week 12 and month 8, with wear and usage questions additionally included in week 0)*
    - Wear time for Fitbit® device (self-reported frequency and duration)
    - Usage time for Bupa Boost app (self-reported frequency and duration)
    - Reasons for disengagement with the Fitbit® and/or Bupa Boost app
    - Goals set (including whether goals were set and details of goals set)
    - Which features of the Bupa Boost app were used
    - Perceived usability and usefulness of the Fitbit® and Bupa Boost app (Likert scale responses, *week 6 and week 12 only*)
- *Reporting of whether any specific events impacted on activity levels during the previous week, and specification of these (e.g. illness, annual leave) (week 0, week 6, week 12, month 8)*
- *Details of participation in any new PA schemes or activities run by the police force since the beginning of the study (i.e. possible co-interventions) (month 8 only)*

A copy of all surveys used in the study is included in **Appendix 5**.

3.4.4 Participant interviews

Semi-structured interviews were conducted at three time points – prior to the intervention (week 0) to build an understanding of participants' expectations, and at post-intervention (week 12) and follow-up (month 8) to gain more in-depth information on experiences throughout the study. Thirty-two interviews were conducted with 16 participants (see **section 3.2.4** for details of selection of interviewees and **Chapter 4** for full numbers and characteristics). A theory-based topic guide was developed for each interview (see **Appendix 6**). As qualitative research is an iterative process (Dicicco-Bloom and Crabtree, 2006), earlier interviews were used to inform the topic guide for subsequent interviews. The interviews at each time point had a different purpose, although some overlapping themes were explored through all three interviews. The key themes explored in each interview are shown in **Table 7**.

Table 7 Interviews and key themes explored

Interview	Themes
Pre-intervention (week 0)	Prior experiences and expectations of mHealth / fitness technology
Post-intervention (week 12)	Short-term engagement with the intervention Experiences and short-term behaviour change
Follow-up (month 8)	Longer-term engagement and experiences Maintenance of PA levels Experience of study participation
All interviews	Wider context of PA and SB in the police force (including workplace PA initiatives, use, barriers and suggestions) Barriers and facilitators for PA Barriers and facilitators for technology use

With the exception of one face-to-face interview, all interviews were conducted via telephone. This was the most practical method for reasons of convenience for interviewees; most interviews took place during work time and could easily be rearranged at short notice. Telephone interviews were also more feasible for the interviewer, as participants were based in geographically dispersed and remote locations. There is evidence that well designed and planned telephone interviews can yield useful and high quality data, of the same quantity, nature and depth as face-to-face interviews (Taylor et al., 1998, Sturges and Hanrahan, 2004). All interviews were audio-recorded and transcribed verbatim, with consent given before the interview began. Field notes were taken during or immediately after each interview and used to guide analysis. The number of interviews was based on data saturation; when no new themes were apparent at each time point, no further interviews were conducted. Interviews ranged from 10 to 27 minutes' duration.

For reasons of availability of resources and to ensure consistency, all interviews were conducted by a single researcher. Reflexivity and transparency are essential in qualitative research interviews (Dicicco-Bloom and Crabtree, 2006), as the characteristics of the researcher and his or her relationship with participants may influence study conduct and interpretation of findings (Tong et al., 2007). In the PAW-Force study, the interviewer was a white female PhD student from a health research background, with some knowledge but little experience of the policing occupation. Prior to interviews being conducted, interviewees were personally unknown to the researcher, and all contact had been via e-mail or telephone. Interviewees were fully aware of the purpose of the research, and participated voluntarily. While subjectivity is unavoidable in qualitative research, bias was minimised by avoiding leading questions, maintaining a neutral stance, and only answering questions about the findings of the research at the end of the follow-up interview. To ensure a reliable understanding of interview content, verbal summaries were given by the interviewer during and at the end of interviews, and confirmed with interviewees.

3.4.5 Survey with managers, commissioners and occupational health staff

Workplace interventions are more likely to be successful with support from staff at all organisational levels including higher-level managers (Hendriksen et al., 2016b, Brand et al., 2017). A separate short online survey (also administered using LimeSurvey) was conducted with a sample of managers, commissioners and occupational health and wellbeing staff within the Devon and Cornwall and Dorset Police forces. The aim of this survey was to explore the perceived feasibility, acceptability and cost-effectiveness of the intervention for the wider workforce.

Potential participants were identified from the Devon and Cornwall and Dorset Police web pages and through key contacts made during the study. A purposive sample was selected to ensure representation of a range of occupations and sites. Participants were approached via e-mail, given an information sheet (see **Appendix 4c**) and asked to complete the online consent form and survey (see **Appendix 4d** and **Appendix 5e**) if they agreed to take part. Twenty-three individuals were invited to take part in the survey; of these, there were 10 respondents.

The survey contained mainly open-ended questions relating to perceived usefulness of mHealth and fitness technology in the police force, perceived benefits for individual staff and the organisation, perceived cost-effectiveness, and (in keeping with the theme of context) other recommendations for policies or strategies to encourage PA and reduce SB in the police force. Survey respondents were asked to disclose their organisation and occupational role but demographic data such as age and gender were not collected as this was felt to be less relevant for this part of the study.

3.4.6 Sickness absence data

As part of the consent form for the PAW-Force study (see **Appendix 4b**), participants indicated whether they agreed to allow their individual sickness absence records to be accessed for the purpose of the research. Consent was provided by 180 of the 182 eligible participants who began the study (99%). The data for those who had provided consent were requested from the Performance and Analysis teams of the Devon and Cornwall and Dorset Police forces. Data were requested for recorded sickness absence for each individual during two time periods – pre-intervention (June 2016 to May 2017) and post-intervention (June 2017 to May 2018). Details of total days lost (the total duration of illness) and duty days lost (work days lost due to illness) were provided for each sickness episode. These were used to calculate three main quantitative outcomes, first for each individual then at an aggregate level for all participants. The outcomes were total days lost, duty days lost and total number of sickness episodes (pre-intervention and post-intervention). Reasons for sickness absence were also collected and aggregated. Due to issues of data quality and completeness, sickness absence data were only available for the Plymouth BCU site.

3.4.7 Feasibility of study methods

To answer research question 6 (Will a future larger scale effectiveness trial be feasible and acceptable in this context?) (see **Figure 4**, **section 3.4.1**), the following outcomes were assessed quantitatively and qualitatively:

- Reach and recruitment, including:
    - Reach and recruitment as perceived by interviewees
    - Actual numbers recruited to time and target
    - Representativeness of the sample (occupation, age, gender and ethnicity) compared with the wider police forces using organisational data

- Adherence and attrition
    - Percentage of participants who did not complete the study. Attrition was defined as the percentage of participants who consented to participate and began the intervention but failed to provide any data at 8 month follow-up.
    - Participant-reported reasons for dropout
- Comparison of study completers and non-completers
    - Study completers (i.e. those who provided data for one or more outcomes at the 8-month follow-up) were compared with non-completers (i.e. those providing no data at 8-month follow-up). Independent samples t-tests and Chi-squared tests were used to compare the two groups by socio-demographic and occupational characteristics, health, and baseline measures of quantitative outcomes (objective and self-reported PA, sedentary time, physical and mental health-related quality of life, perceived stress and productivity).
- Feasibility and acceptability of data collection procedures and outcomes
    - Proportions/percentages of participants providing data on objective and self-reported PA
    - Acceptability of data collection methods (including burden of questionnaire completion) as perceived by interviewees
    - Feasibility of collection of secondary data (i.e. sickness absence data)
- Acceptability of study participation
    - Any other general or specific views on study participation (as perceived by interviewees)

3.4.8 Additional feasibility data

A consideration of unexpected mechanisms or consequences is recommended as part of process evaluation and feasibility (Moore et al., 2014). Such outcomes have been underreported in the mHealth literature. Although not considered major outcomes (and therefore not included in **Figure 4**), participants were asked to report any adverse events or issues experienced with the Fitbit® and Bupa Boost app during the study. This included negative physical and psychological consequences, and issues such as malfunctioning Fitbit® devices. These were recorded on a spread sheet and contributed to the assessment of feasibility of the intervention.

3.5 Data analysis

3.5.1 Quantitative analysis

A statistical analysis plan was produced prior to the start of the study. A combination of descriptive and inferential analysis was used to explore the context of PA and SB in the police force at baseline (Q1), to assess the impact on PA, sedentary time and secondary outcomes (Q3 and Q5), and to examine which intervention components (individual versus social) were preferred and potentially most impactful (Q4). Descriptive analysis was also used to quantitatively explore aspects such as use of or engagement with the Fitbit® and Bupa Boost app (part of Q2).

Where inferential analysis was used, the estimated effects (e.g. means or mean differences) were reported with standard deviations, 95% confidence intervals and p-values. A p-value <0.05 was considered statistically significant and all tests were two-tailed. Quantitative analysis was performed using Microsoft Excel and Stata version 15.0 (StataCorp, 2017).

3.5.1.1 Preparation of data for analysis

Mean daily step count was calculated for each participant at each outcome assessment point (week 0, week 6, week 12 and month 8). As previous researchers had done (Wang et al., 2016, Poirier et al., 2016), a minimum wear criterion was applied so that mean daily steps were only calculated for participants who had worn the Fitbit® for five or more of the previous seven days including at least one weekend day, and where more than 500 steps per day were recorded. Adjusted averages were calculated for some participants; where they stated there had been a specific event (such as illness) impacting on their activity level for a particular day, this day was excluded from analysis. If baseline step data were unavailable, subsequent step data were excluded from analysis.

Scoring protocols were followed for all validated questionnaires (IPAQ, SF-12, PSS-4 and HPQ). IPAQ outcomes were calculated using a pre-designed Excel spread sheet (Cheng, 2016). Spread sheets were created by the researcher for scoring of the SF-12, PSS-4 and HPQ.

3.5.1.2 Baseline data analysis

Descriptive analysis (including mean, standard deviation, range and proportions/percentages) was used for key demographics and baseline values of all primary and secondary outcomes. Tabular and graphical summaries (histograms, bar charts, box plots) were produced for key variables to assess their distribution.

Subgroup differences were explored in baseline mean daily step count, self-reported PA and sedentary time, in addition to secondary outcomes including health-related quality of life, perceived stress and perceived productivity. Independent samples t-tests were used for binary variables (e.g. gender) and one-way analysis of variance (ANOVA) was used for variables with more than two categories (e.g. age group, occupation). Where the results of the ANOVA were significant, Tukey post-hoc t-tests were performed for each group (only significant values were reported for Tukey comparisons). The following assumptions were met for t-tests and the one-way ANOVA: continuous dependent variable; independent observations; two or more categories for the independent variable; approximate normal distribution of the dependent variable for each category with no significant outliers; and approximately equal variance between groups (checked using Bartlett's test for equal variances).

Correlations between baseline values of all continuous outcomes were explored. Associations between mean daily steps and self-reported PA and sedentary time were explored, in addition to correlations between mean daily steps and secondary outcomes (health-related quality of life, perceived stress and perceived productivity). The assumptions for the parametric Pearson's product-moment correlation were met, including related pairs and no significant outliers. Linearity and homoscedasticity were checked using scatterplots. All variables followed an approximately normal distribution. The Pearson's r correlation coefficients for the sample were reported, and were interpreted according to Cohen's guidance; a correlation coefficient of 0.1 was considered small, 0.3 moderate, and 0.5 large (Cohen, 1988).

3.5.1.3 Assessment of impact of the intervention

Changes in primary and secondary outcomes from baseline (week 0) to mid-intervention (week 6), post-intervention (week 12) and follow-up (month 8) were explored using paired t-tests, correlations and regression.

Paired t-tests were used to compare baseline values of primary outcomes (mean daily steps, self-reported PA and sedentary time) and secondary outcomes (health-related quality of life, perceived stress and perceived productivity) with mid-intervention, post-intervention, and follow-up values. The assumptions for paired t-tests were met, including continuous dependent variables and independent observations. The normality of the distribution for changes in steps, self-reported PA and sedentary time (and also for changes in secondary outcomes) was explored visually using histograms and box plots. Each variable of interest was approximately normally distributed, thus allowing parametric paired t-tests to be used. The steps change from baseline to mid-intervention appeared to follow a less normal distribution, so the non-parametric Wilcoxon matched-paired signed rank test was also used; however as the p-values were very similar (and applying the central limit theorem to assume that the variable is approximately normally distributed), only the results for the parametric tests were reported.

Subgroup analysis was performed to explore differences in changes in mean daily steps, self-reported PA and sedentary time, and some secondary outcomes, by baseline activity level (<10,000 vs. ≥10,000 steps/day), age, gender and occupation. A sensitivity analysis was carried out to control for external factors that might have affected participants' activity levels throughout the study. This analysis was performed with the mean daily steps outcome only, and included only participants who reported that no specific events had affected their PA level in the week prior to completing the survey.

Correlations were used to explore associations of changes in mean daily steps with changes in self-reported PA and sedentary time, and with changes in secondary outcomes (health-related quality of life, perceived stress and perceived productivity). The procedures and assumptions for correlation analysis as described in **section 3.5.1.2** were followed. Correlations were explored separately for participants with a baseline daily step count of <10,000 and ≥10,000.

Exploratory regression was carried out to investigate which participant characteristics may be associated with potential intervention effectiveness in both the short and longer term. Both linear and logistic regressions were used with change in mean daily steps as the dependent variable. Linear regression was first carried out with change in steps as a continuous outcome. All assumptions were met; normal distribution and homoscedasticity of residuals (or error terms) were checked using histograms of residuals and plots of residuals against fitted values. For logistic regression, the dependent variable was dichotomised to increase in steps (1) and decrease or no change (0). The assumptions of independent observations and low multicollinearity of predictor variables were met. All predictor variables were also dichotomised for ease of interpretation of odds ratios.

Regression models were developed using a forward selection process. The relationships between predictor and dependent variables were checked using univariate regression (for linear regression) and Chi-squared tests (for logistic regression). All variables with a p-value <0.2 were included in the model. All potential interactions between predictor variables were considered. Gender and police force were included even when not significant predictors of the change in steps as they were important control variables (and gender was significantly associated with steps at baseline). As the results of the linear and logistic regressions were similar in terms of significance, only the logistic regression models were reported.

To assess differences in the relative impact of the individual and social phase, the change in PA outcomes (mean daily step count, self-reported PA and sedentary time) between week 0 and week 6 was compared with the change in these outcomes between week 0 and week 12 using paired t-tests. Subgroup differences (baseline activity level, age, gender and occupation) were again explored. Differences in preferred phase by baseline activity level, age, gender, occupation and education were explored using Chi-squared tests.

Multiple hypothesis testing was considered as a potential issue in analysis; however given the relatively small sample size and exploratory nature of the study, Bonferroni correction was not felt to be appropriate. There are also problems with such corrections in that the likelihood of type 2 errors is also increased (Perneger, 1998).

3.5.1.4 Missing data

It was assumed that missing data in the PAW-Force study were missing at random, i.e. systematic differences in missing and observed values may be explained by differences in observed data (Sterne et al., 2009). Available case analysis or pairwise deletion was used and was a preferred method to listwise deletion to maximise the amount of data available for analysis. Imputation methods were considered unnecessary for a study of this type (exploratory/feasibility). Multiple imputation is a time- and resource-intensive procedure that requires the support of a statistician and would be considered more important in an effectiveness trial, while single imputation methods (such as last observation carried forward) may be unreliable, subject to various biases and may result in an underestimation of variability (Sterne et al., 2009).

Participants who had withdrawn from the study before the 8-month follow-up assessment were asked for consent to use the data collected to date, and were included in analysis (if consent was given).

3.5.1.5 Sickness absence data

Analysis of sickness absence data (Plymouth BCU only) was performed in Excel. Descriptive statistics on number of sickness episodes, total days lost and duty days lost were produced for the study sample. A comparison was made between one year pre-intervention (June 2016 to May 2017) and one year post-intervention (June 2017 to May 2018), with the calculation of differences in frequency and percentage changes between the two time periods. A Chi-squared test was used to statistically explore the numbers of participants with one or more sickness episodes versus no sickness episodes by time period. Frequencies were calculated for categories of reasons for sickness absence (e.g. respiratory disorders, psychological disorders) and differences and percentage differences were compared pre- to post-intervention. Participants who had not been employed by Devon and Cornwall Police for the whole two-year period June 2016 to May 2018 were excluded from analysis. Participants who had withdrawn from the study were included (with consent) provided they were still employed by Devon and Cornwall Police at the end of May 2018.

3.5.2 Qualitative analysis

3.5.2.1 Participant interviews

Interviews were analysed thematically using the Framework Method. The Framework Method was developed by Ritchie and Spencer (Ritchie and Spencer, 1994) and is an increasingly popular, systematic method for the management and analysis of qualitative interview data (Gale et al., 2013). The defining feature of the Framework Method is the matrix output consisting of rows (participants or cases), columns (codes) and cells containing summaries of data. The Framework Method has many advantages including its clearly defined, logical procedure, rigour and transparency, and flexibility to incorporate additional data such as quantitative data and field notes (Gale et al., 2013). Field notes included information on participants' background and occupation, the interviewer's initial summary of the interview and early perceptions of themes, in addition to questions or topics for exploration.

The Framework Method was selected for the PAW-Force study as it provided an at-a-glance summary of findings and key quotes for each participant, facilitated comparisons between participants by theme (e.g. barriers to PA), and also allowed assessment of changes over time (by participant and by theme). As the Framework Method is not associated with a specific epistemological or theoretical approach, it allowed interview data to be analysed both inductively (data-driven) and deductively (theory-driven). For example, for the theme of perceived impact of the intervention on PA and sedentary time, analysis was deductively guided by the COM-B model of behaviour change (Michie et al., 2011a). In contrast, other themes (or sub-themes) were inductively identified from the data, for example many interviewees talked about the impact of stress and sleep on their wellbeing.

Analysis was carried out in NVivo version 11 (QSR International, 2015), with framework matrices imported into Excel for editing. The main stages of analysis included:

1. Transcription of each interview. All interviews were transcribed verbatim. Some interviews were transcribed directly by the researcher (S. Buckingham) while others were transcribed by an external company (GoTranscript®). All transcripts were double-checked for accuracy by the researcher and corrections made as necessary.
2. Familiarisation with the interview. Transcripts were read and re-read in conjunction with field notes.
3. Coding – working through each interview in turn, interesting content (words, short phrases, sentences or paragraphs) was assigned a descriptive or conceptual label. The manual coding method in NVivo was used for this process.
4. Developing a working analytical framework. Based on coding of the first few transcripts, the resulting codes were used to form a working analytical framework in the structure of a 'tree diagram' in NVivo. This was an iterative process, with refinements made based on coding of subsequent interviews until no new codes were generated. Codes were reordered and grouped into categories during this process. An example of an analytical framework (for baseline interviews) is given in **Figure 5**.
5. Applying the analytical framework. The remaining transcripts were indexed, i.e. content was coded according to the existing categories and codes.
6. Charting data into the analytical framework. A framework matrix based on the content of interview transcripts was produced in NVivo before being imported into Excel. Each code was a separate column, and each participant formed a separate row. The cells of the matrix were filled with summaries of the data (including interview content and field notes) and key quotations from interviews. As previous researchers had done (Heath et al., 2012), quotes were underlined and assigned a rating based on usefulness (Q/QQ/QQQ). Some quantitative participant data were included in the framework matrix (see **section 3.5.3**). Socio-demographic data (age, gender, occupation) were also included to facilitate exploration of potential subgroup differences. As an example, a section of the framework matrix for the 8-month follow-up interviews is shown in **Figure 6**.
7. Interpreting the data. The matrix was reviewed with comparisons being made between participants (or cases) by codes, and within participants across

codes. The codes were used to produce sub-themes, followed by broader themes. This process was influenced by theory (deductive) and concepts generated from the data (inductive).

Figure 5 Analytical framework of codes/categories for baseline interviews

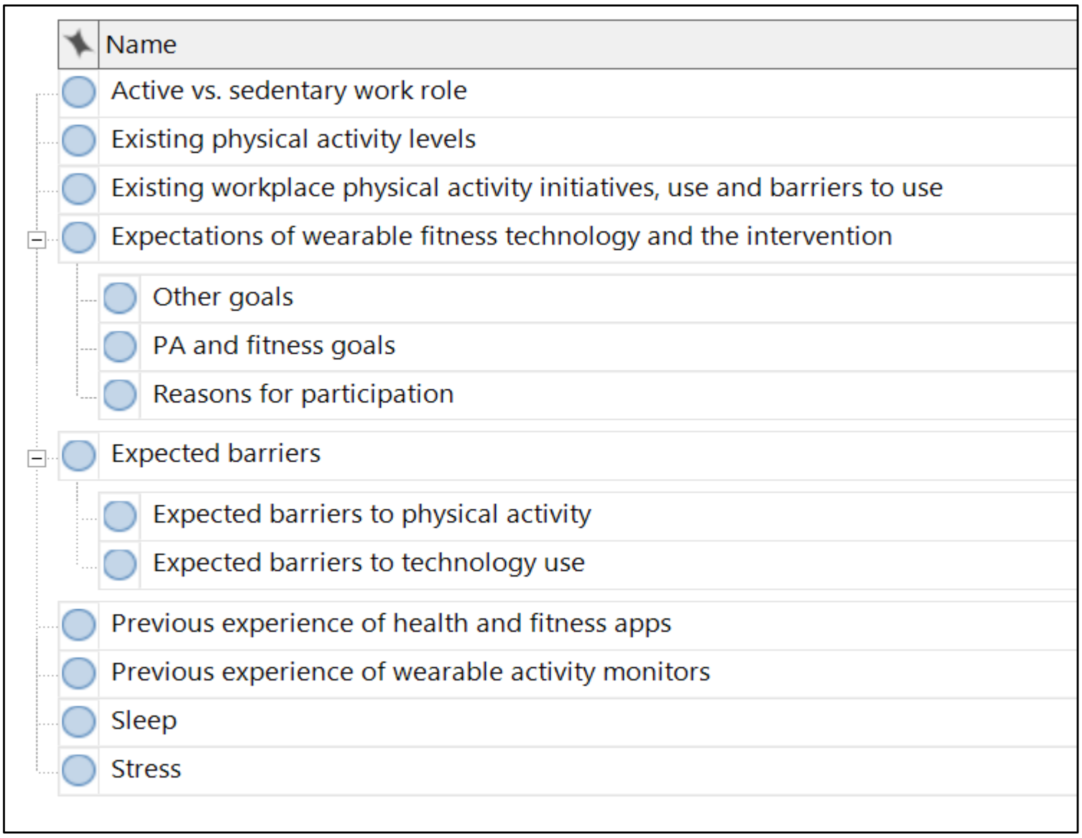


Figure 6 Section of framework matrix for 8-month follow-up interviews

	A : Continued engagement and use of Fitbit	B : Continued engagement and use of Bupa Boost
5 : DO10. Follow-up (first interview) Phase = 3 Gender = Male Age = 40+ Occupation = Police officer 7938 steps, moderately active, 931 weekly MET-mins, 10 hours sedentary **(increase in steps, activity unchanged, slight increase in sedentary time since baseline)**	Still highly engaged with Fitbit, wearing 24/7: "I've pretty much had the Fitbit attached to my wrist from the beginning. It's only left my wrist to charge for about an hour or so. It's on there pretty much permanently. It's a great bit of kit." **QQQ** Plans to continue to wear after study. Mainly uses to monitor steps and sleep.	Still using Bupa Boost daily, particularly motivated by the social features: "I'm updating the Bupa Boost pretty much every day. We've got a good little group going in [site] which just keeps everybody motivated. We've got some folks in there that are of all different sort of ranges and abilities."**QQ** Engagement with Bupa Boost improved over time: "It took a while to get used to. I'd probably say a couple of weeks just to understand the bits that I like."**QQ**
8 : PL73. Follow-up Phase = 3 Gender = Female Age = 18-39 Occupation = Police staff 11535 steps, moderately active, 2026 weekly MET-mins, 9 hours sedentary **(large increase in steps, increase in activity, and reduction in sedentary time since baseline**)	Still wearing Fitbit 24/7 - "every day, all the time". Plans to continue using Fitbit and associated app. Now uses the run function (i.e. exercise tool in addition to daily wear): "It's quite good because I could record how far I'm going, how fast I'm going and compare it to the last time I went so that's quite useful."**Q** Competes with friends (rather than colleagues) within Fitbit app.	No longer uses Bupa Boost app, prefers social features of Fitbit app and competes against friends rather than colleagues.

A different framework matrix was produced for each time point (baseline, post-intervention and 8-month follow-up), but matrices were cross-referenced during analysis. As only six individuals took part in all three interviews (for reasons of participant availability), detailed longitudinal analysis at an individual participant level was not carried out. However, factors such as changes in levels of engagement with the intervention and PA behaviours over time were considered for all interviewees using available qualitative and/or quantitative data.

Each interview had a specific purpose (as shown in **Table 7**, **section 3.4.4**); baseline interviews assessed expectations, post-intervention interviews explored short-term behaviour change, and follow-up interviews focused on longer-term engagement and maintenance of PA. Some themes (or sub-themes) were therefore unique at a particular time point (for example, long-term engagement in the 8-month follow-up interview only) while others were common themes across different interviews (for example, workplace PA initiatives including use, barriers and suggestions).

While sufficient resources were not available to enable coding by two or more researchers, the framework matrices and themes generated were checked by an independent researcher (Dr. C. Guell). This is one of the recommended ways to improve rigour in qualitative research (Tong et al., 2007).

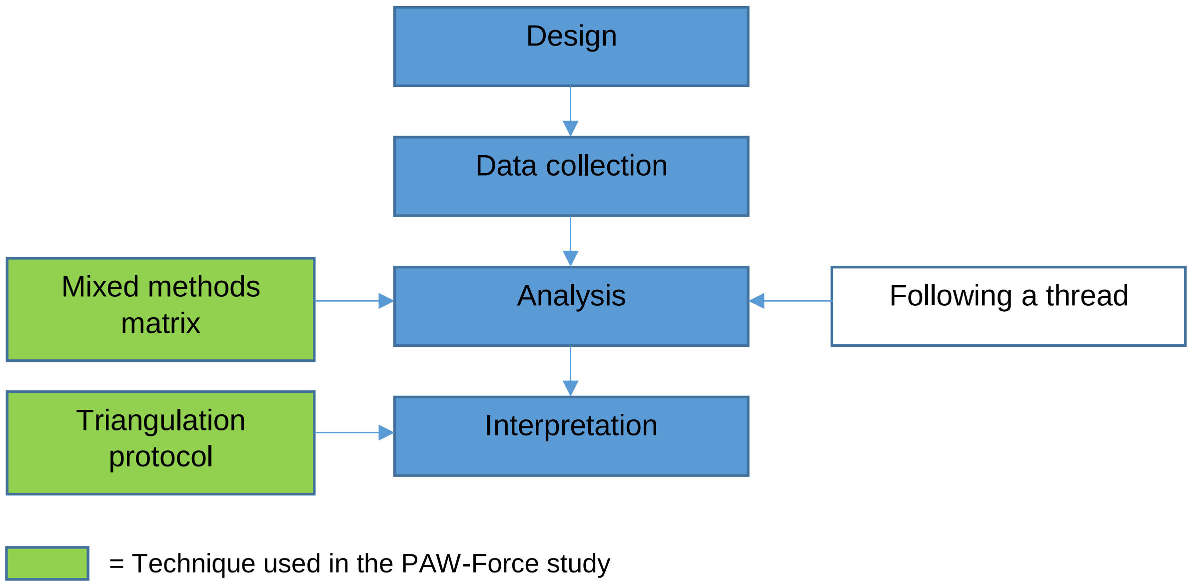
3.5.2.2 Survey responses

In addition to interview data, some qualitative questions were included in the participant surveys and the survey with managers, commissioners and occupational health staff. Responses to these questions were analysed thematically, following the guidance of Braun and Clarke (Braun and Clarke, 2006). As with framework analysis, themes were identified both inductively and deductively. Some content analysis was undertaken, for example reasons for disengagement with the Fitbit® and Bupa Boost app were categorised and frequencies reported. For the managers' survey, results were reported according to the survey questions with themes identified within each question.

3.5.3 Mixed methods analysis

Quantitative and qualitative data were collected and analysed concurrently throughout the study. The stages at which quantitative and qualitative data may be integrated in a mixed methods study (design, data collection, analysis and interpretation) and three key techniques for integrating data (mixed methods matrix, following a thread and the triangulation protocol) are shown in **Figure 7** (O'Cathain et al., 2010). In the PAW-Force study, data were integrated at the analysis phase at an individual participant level within the framework matrix, and at the interpretation phase at the level of the dataset with triangulation of findings.

Figure 7 Stages and techniques for integrating data in mixed methods research (adapted from O'Cathain et al., 2010)



At an individual participant level, quantitative data for each individual interviewee were integrated into the framework matrix. Mean daily step count, self-reported PA and sedentary time, and quantitative changes in activity levels over time, were included. This integration of quantitative outcomes with the qualitative themes enabled a more detailed and comprehensive analysis and produced a more complete picture. For example, the mixed methods matrix showed that those who were less active at baseline generally showed higher engagement with the intervention and higher PA levels in the short and longer term.

Quantitative and qualitative data were also integrated at the level of the dataset, with key themes interpreted from both types of data together. Qualitative findings were used to expand on, explain, and triangulate with (confirm and/or complement) quantitative data. An example of expansion was for the theme of perceived usability and usefulness of the Fitbit® and Bupa Boost app, where in addition to the quantitative survey ratings, interviews provided in-depth views on specific positive and negative aspects of the technology. As an example of explanation, interviews helped to explain the quantitative observation that self-reported PA increased significantly for study participants overall while daily steps did not significantly change (participants focused on a range of activities which would not necessarily have been recorded by the Fitbit® as steps). Confirmation involves the use of multiple data sources to explore a phenomenon, to verify the findings from one type of data with those from another (Small, 2011). In the PAW-Force study, qualitative data on perceived impact on PA confirmed the quantitative finding that the intervention seemed to be of greatest benefit to those who were less active at baseline. Complementarity is the use of mixed methods to explore related but different facets of a phenomenon, to yield an enriched understanding (Greene et al., 1989). An example of complementarity was for the theme of feasibility of study methods, where quantitative and qualitative data provided complementary information on different aspects of feasibility such as recruitment rates and perceived reach.

As part of the triangulation process, discrepancies between the quantitative and qualitative findings were considered. There was one main instance where the findings did not appear to agree (the apparent discrepancy between survey-assessed and interview-reported health and wellbeing outcomes). The findings were reconciled or reinterpreted and alternative explanations were sought. It was also recognised that further research may be required to explore such complex phenomena.

The interpretation of quantitative and qualitative findings together was also checked by a second researcher to improve rigour. A summary of integrated quantitative and qualitative findings according to the six overarching research questions is given in **Chapter 8**.

3.6 Summary

The PAW-Force intervention and study methods were fully described in this chapter. Although the intervention (Fitbit® and Bupa Boost app) was not designed by the research team, 20 evidence-based behaviour change techniques were identified. A mixed methods approach was adopted for the study, combining a range of quantitative (both objective and self-reported) and qualitative data sources to answer six key research questions. The study findings will be presented in the following chapters.

CHAPTER 4 PARTICIPANTS, KEY BASELINE FINDINGS AND CONTEXT

4.1 Introduction

The aims of this chapter are to fully describe the study sample in terms of socio-demographic and occupational characteristics, physical activity (PA) and sedentary behaviour (SB) (both within and outside of the workplace), and health and wellbeing at baseline, and to build a comprehensive picture of the context of the research. The chapter includes quantitative and qualitative survey data in addition to qualitative data from interviews.

An understanding of context is critical in complex interventions and is one of the three key components of process evaluation together with implementation and mechanisms of impact (Craig et al., 2008, Moore et al., 2014, Moore et al., 2015). Context includes "anything external to the intervention that may act as a barrier or facilitator to its implementation, or its effects" (Moore et al., 2015, p2). Context may also influence mechanisms of impact and can explain why the intervention may work differently (or have differing levels of acceptability) for different individuals or in different settings. Ease of interpretation and generalisability of findings will also result from a good understanding of context (Moore et al., 2015).

The importance of context and external factors are also recognised in several influential behaviour change theories. For example, the 'COM-B model' (described in **Chapter 6**) includes physical and social opportunity as determinants of behaviour (Michie et al., 2011a), and the organisational and community environment are important influences on behaviour in the Socio-Ecological Model (Bronfenbrenner, 1989, Stokols, 1996, Lee et al., 2017). As specified in the logic model for the PAW-Force intervention (**Figure 3**), contextual influences may include the wider ActivAte2020 programme and organisational policy surrounding PA, health and wellbeing in the police force, barriers and facilitators to PA, differences by socio-demographics or worksite (Plymouth BCU or North Dorset territorial region), and geographical, community or cultural factors.

The next three sections of this chapter (**sections 4.2 to 4.4**) include a description of the flow of participants through the study, and description of the main socio-demographic and occupational characteristics of study participants, followed by a similar description of the subsample of interview participants. This is followed by five sections (**sections 4.5 to 4.9**) of overarching themes relating to baseline findings and context which are based on interview themes combined with quantitative findings. The themes are: active versus sedentary work roles; physical activity and sedentary time (including subgroup differences); health, wellbeing, stress and productivity (and associations with physical activity); opportunities, barriers and facilitators for physical activity in the police force; suggestions for promoting physical activity and reducing sedentary time in the workplace. The chapter concludes with a discussion of the main findings which are related to the existing literature (**section 4.10**).

4.2 Participant flow diagram

The number of participants enrolled in the study and participant flow from initial consent to completion of the final follow-up questionnaire are summarised in **Figure 8**. In brief, 190 police officers and staff completed the online screening and consent form; eight of these were ineligible and excluded for reasons shown in **Figure 8**. These reasons included: being on holiday or away from the workplace at the start of the study; having no access to a mobile device that is compatible with the Fitbit® and Bupa Boost app; expecting to leave the police force during the study period; having a medical operation during the study; working in a Bluetooth-controlled environment therefore unable to use the Fitbit® or a smartphone in the workplace. The 182 individuals who were eligible to participate in the study were provided with a Fitbit® activity monitor and were asked to complete the baseline questionnaire after wearing the (blinded) device for a week. Two participants did not complete this questionnaire and were subsequently excluded. Seven participants officially withdrew through the 8 month study period (reasons for withdrawal are discussed further in **Chapter 7**). An additional proportion did not complete the questionnaires at each data collection point, i.e. 19/178 (11%) at week 6, 25/176 (14%) at week 12, and 30/173 (17%) at month 8 (note some participants completed earlier but not later questionnaires and vice versa). Levels of completion for the various outcomes are discussed further in **Chapter 7**. The overall participant retention rate from beginning the study to 8-month follow-up was 143/182 (79%).

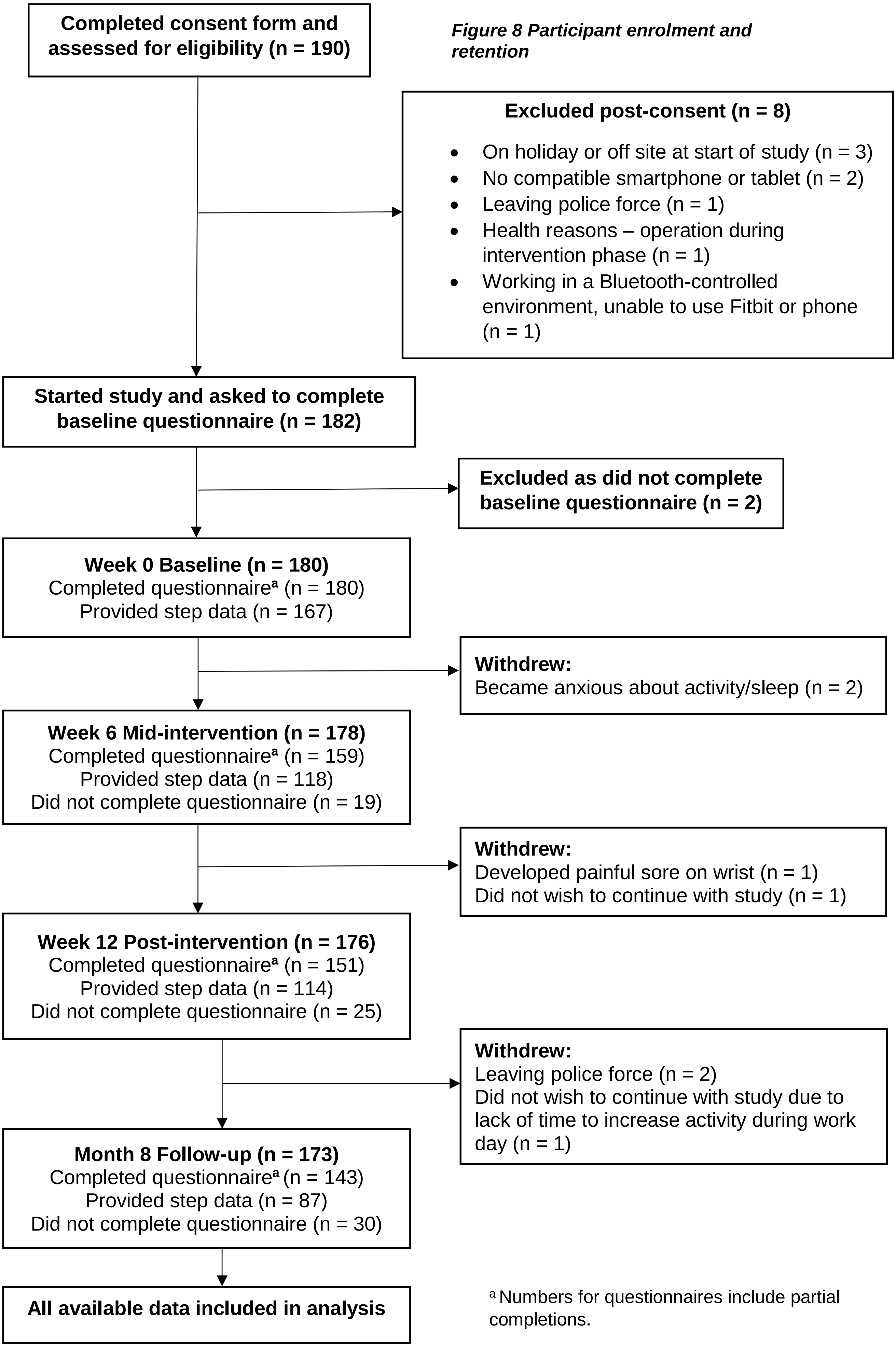


Figure 8 Participant enrolment and retention

[a] Numbers for questionnaires include partial completions.

4.3 Characteristics of study participants

The socio-demographic and occupational characteristics of the 180 participants providing baseline questionnaire data are provided in **Table 8**. The age of participants ranged from 19 to 64 years, with a mean age of 39.3±9.6 years. Of the 180 participants, 107 (59%) were male. The majority of participants (n = 177, 98%) reported themselves as any White ethnicity. The sample was diverse in terms of marital status (112/180 or 62% were married or in a civil partnership) and education (57/180 or 32% were educated to degree level or above). Only 16% (n = 29) of participants lived alone. Of the 180 participants, 68 (38%) were dog owners. Approximately half (54%, n = 96) of the participants reported their main residence as urban, 24% (n = 43) as suburban and 23% (n = 41) as rural.

The majority of study participants (n = 128, 71%) served the Devon and Cornwall Police Force; all of these were based at the various sites within the city of Plymouth Basic Command Unit (BCU) (as shown in **Table 9**). The remainder (n = 52, 29%) were employed by Dorset Police (mostly rural North Dorset sites) at the time of enrolment (see **Table 10**). There was representation from all of the work teams across the two forces, including local investigation, local policing/neighbourhood, response policing, and administration or other teams (see **Table 11**). The representativeness of the study sample (number, occupation, gender, age and ethnicity) compared to the Plymouth BCU and North Dorset police populations as a whole is discussed in **Chapter 7**.

Approximately 63% (n = 114) of study participants were police officers, and their rank ranged from constable to higher level officers (inspector, chief inspector or superintendent). The majority of participating officers were of lower rank (87/114 or 76% constable, 23/114 or 20% sergeant), reflecting the higher proportions of lower rank officers within this population. Around 20% (n = 36) of those recruited were police staff, and 17% (n = 30) were either employed in a Police Community Support Officer (PCSO) role or were in a voluntary special constable role (due to the small number of special constables this category was grouped with PCSOs to protect individual anonymity).

On average, participants had served the police force (including the current force and any previous forces) for 12.1±8.0 years. Half (50%, n = 90) of those recruited had an annual salary of less than £30,000. The vast majority of participants (n = 167, 93%) were working for 30 or more hours per week at the time of enrolment. A high proportion of participants (n = 144, 80%) were shift workers. Of these, 30 (21%) worked night shifts and 59 (41%) worked rotating shifts.

Table 8 Participant characteristics: socio-demographic and occupational

Study variables	Participated in study (n = 180)
Age in years, mean (SD)	39.3 (9.6)
Male, n (%)	107 (59 %)
Ethnicity, n (%)	
White	177 (98 %)
Marital status, n (%)	
Single (never married or civil partnered)	40 (22 %)
Married or civil partnership	112 (62 %)
Divorced, separated or widowed	26 (14 %)
Prefer not to say	2 (1 %)
Living circumstances, n (%)	
Alone	29 (16 %)
With children under 18	96 (53 %)
With other adults	145 (81 %)
Dog owner, n (%)	68 (38 %)
Main residence, n (%)	
Urban (city or town)	96 (54 %)
Suburban	43 (24 %)
Rural (including rural village, hamlet or isolated dwelling)	41 (23 %)
Highest level of education, n (%)	
Lower secondary school (GCSE, CSE, O-level, Standard Grade, Intermediates)	38 (21 %)
Upper secondary school (AS or A-level, Scottish Highers)	44 (24 %)
Professional or technical qualification (below degree)	41 (23 %)
University / college degree	48 (27 %)
Postgraduate (masters / PhD)	9 (5 %)
Police force, n (%)	
Devon & Cornwall Police (Plymouth Basic Command Unit)	128 (71 %)
Dorset Police (North Dorset)	52 (29 %)
Occupation, n (%)	
Police officer	114 (63 %)
Police community support officer (PCSO) or special constable	30 (17 %)
Police staff	36 (20 %)
Rank, n (%) (officers only, n = 114)	
Constable	87 (76 %)
Sergeant	23 (20 %)
Inspector, chief inspector or superintendent	4 (4 %)
Years of police force service, mean (SD)	12.1 (8.0)
Working 30 or more hours per week, n (%)	167 (93 %)
Annual salary	
< £30,000	90 (50 %)
≥£30,000	86 (48 %)
Prefer not to say	4 (2 %)
Shift work, n (%)	144 (80 %)
Type of shift[a] (shift workers only, n = 144)	
Morning (early)	95 (66 %)
Afternoon (late)	96 (67 %)
Night	30 (21 %)
Rotating	59 (41 %)

[a] Note: Some participants worked more than one type of shift. SD = Standard Deviation

Table 9 Work site: Plymouth Basic Command Unit (n = 128)

Site	No. of participants	% of participants
Charles Cross	54	42
Crownhill	54	42
Devonport	10	8
Plympton or Plymstock	5	4
Other	5	4

Table 10 Work site: North Dorset (n = 52)

Site	No. of participants	% of participants
Blandford	36	69
Gillingham	3	6
Shaftesbury	4	8
Sturminster Newton	3	6
Other	6	12

Note: Percentages may not total 100 due to rounding

Table 11 Police force and work team: Number and % of participants (percentages by police force)

	Work team				
	Local investigation	**Local policing/ neighbourhood**	**Response policing**	**Administration and other**	**Total**
Devon & Cornwall Police (Plymouth BCU)	23 (18 %)	26 (20 %)	33 (26 %)	46 (36 %)	128 (100 %)
Dorset Police (North Dorset)	0 (0 %)	20 (38 %)	25 (48 %)	7 (13 %)	52 (100 %)
Total	23 (13 %)	46 (26 %)	58 (32 %)	53 (29 %)	180 (100 %)

4.4 Characteristics of interview participants

A sub-sample of participants was selected from the study sample (see **Chapter 3** for sampling procedure) to participate in interviews at three stages - pre-intervention (week 0), post-intervention (week 12) and follow-up (month 8). In total, 32 interviews were conducted with 16 participants. Only six individuals took part in interviews at all three time points, for reasons of participant availability. The main characteristics of the interview participants are presented in **Table 12**. As specified in the selection criteria, interview participants were diverse in terms of age, gender, occupation and site. The characteristics of interview participants also closely matched the overall sample of study participants. Approximately 63% (n = 10) of interviewees were male and 37% (n = 6) were female. The age of interviewees ranged from 20 to 58 years, with a mean age of 40.8±12.6 years. The sample comprised nine police officers (56%), five members of police staff (31%) and two PCSOs (13%). One member of police staff also had a special constable role. Of the 16 interviewees, 11 (69%) were based at Plymouth BCU and five (31%) were at the North Dorset sites.

Table 12 Characteristics of interview participants

Interviewee ID	Gender	Age category	Occupation (at baseline)	Interviews completed (1 = pre-intervention (week 0), 2 = post-intervention (week 12), 3 = follow-up (month 8))
1	M	18-39	Police staff	1
2	M	18-39	Police officer (constable)	1,2,3
3	M	18-39	Police officer (sergeant)	1,2,3
4	M	40+	Police officer (inspector)	1,2,3
5	M	40+	Police staff	1
6	M	40+	PCSO	1,2,3
7	F	18-39	PCSO	1,2
8	F	18-39	Police officer (constable)	1,2,3
9	F	40+	Police staff	1,3
10	F	40+	Police officer (constable)	1,2,3
11	M	18-39	Police staff and special constable	2
12	M	40+	Police officer (constable)	2
13	F	18-39	Police officer (sergeant)	2,3
14	F	18-39	Police staff	2,3
15	M	18-39	Police officer (constable)	3
16	M	40+	Police officer (inspector)	3

Interviewees also varied in their baseline activity levels and experienced varying levels of engagement with the intervention and changes in their activity over time. For example, baseline activity ranged from around 5,000 steps per day to over 15,000 steps per day, and from ‘low activity’ to ‘high activity’ level according to the

International Physical Activity Questionnaire (IPAQ). Perceived and actual changes in activity levels of interviewees and the overall study sample will be discussed further in the chapters that follow.

4.5 Active versus sedentary work roles

When asked, “How active is your role?” in the baseline questionnaire, 58% of participants (n = 105) reported their work role as mainly sedentary. Only 17% (n = 30) stated they were mainly active in their role, and 25% (n = 45) reported they spent approximately equal amounts of time active and sedentary. The interview findings generally reflected the sedentary nature of the role of many officers and staff, although there was considerable variation in levels of PA involved in different occupations and work streams. All of the police staff interviewed were sedentary, and mostly office-based. For example:

> *"It is completely office-based. You sit at a desk with your phone. You haven’t got much movement at all."*
>
> (Police staff, male, age 18-39)

> *"It’s a sedentary role. I work on the radio dispatching officers, and also take and make phone calls. So I'm usually sat down."*
>
> (Police staff, female, age 18-39)

One member of police staff had a role that was mainly vehicle-based and sedentary:

> *“It involves a little bit of walking about but it’s not a very active role.”*
>
> (Police staff, male, age 40+)

The active nature of the role of police officers varied according to work stream. For example, interviewees reported that response officers (approximately one third of the study sample) tended to be more active than local investigation and local policing teams:

> *“…maybe I will stay where I am or go off into a different work stream. That will have a bearing to how active I am at work. So I could be going back to*

response where, just by nature of the work I will be doing more walking, or I could be sat behind a desk here."

(Police constable, male, age 18-39)

One interviewee worked in the field of cybercrime and reported that most of the work was computer-based and sedentary with occasional travelling involved. Activity levels also appeared to vary by rank; higher level officers (such as those with a supervisory role and inspectors) were usually more office-based than lower level officers:

"I think in my day-to-day job I am sort of sat behind a desk all day. I can just be sat behind the computer listening and talking on the phone and doing emails all day… I have to deal with all the admin."

(Police sergeant, male, age 18-39)

"I have to spend a fair bit more time at a computer, managing administrative tasks."

(Police inspector, male, age 40+)

PCSOs were more active during the working day as they were typically on foot patrol:

"We certainly do more walking than say, a police officer, because we walk round to give a visible presence to members of the public."

(PCSO, male, age 40+)

Both urban and rural PCSOs were active, although both roles involved travelling by car. One interviewee commented on the changing nature of the role and suggested that it was less active than previously:

"The role has changed over the years… we mainly do walk on our patch but recently because of the reduction in numbers we do have vehicles as well. So I now drive to my patch and then do foot patrol around the area."

(PCSO, male, age 40+)

Interviewees were generally more active outside of working hours, and activity varied widely in type and intensity. For example, activities ranged from dog walking to cycling, running and high-intensity interval training. A police inspector reported ‘bursts’ of exercise in between sedentary periods:

> *“One of the problems with me is I do tend to have physical bursts like at the gym or activity and then a lot of time sat down.”*
>
> (Police inspector, male, age 40+)

One police constable reported ‘overcompensating’ for the sedentary role with higher activity outside of work:

> *“I am just conscious of like how much, eight hours, I am sat at my desk. It is not good for your posture… I am aware that we are so sedentary that I try and overcompensate with so much training.”*
>
> (Police constable, female, age 18-39)

4.6 Physical activity and sedentary time

Activity levels at baseline varied greatly, with daily step counts ranging from 3,797 to 20,819. The distribution of baseline steps was approximately normal, with a slight positive skew towards higher step counts. The overall mean daily step count for participants at baseline was 10,555±3,259 (95% confidence interval or CI: 10,057 to 11,053) and the median was 10,319 (interquartile range: 8,171 to 15,279). As **Table 13** shows, male officers and staff were significantly more active than females (p = 0.005). The mean daily step count for males was 11,132±3,105 (95% CI: 10,516 to 11,749) compared with 9,692±3,315 for females (95% CI: 8,884 to 10,501).

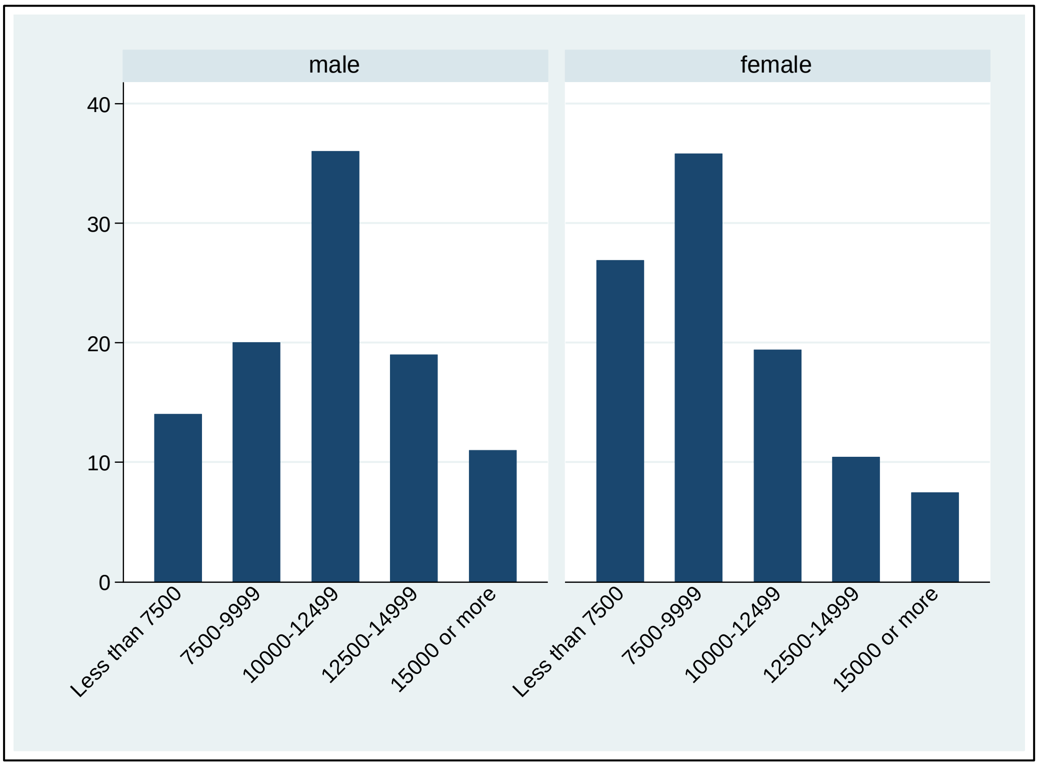
Table 13 Baseline mean daily step count (overall and by gender)

Gender	Mean (SD)	95% CI	Range	Median	Interquartile range
Male (n = 100)	11,132 (3,105)	10,516 to 11,749	4,873 to 20,110	11,151	9,118 to 12,736
Female (n = 67)	9,692 (3,315)	8,884 to 10,501	3,797 to 20,819	8,709	7,381 to 11,544
Overall (n = 167)	10,555 (3,259)	10,057 to 11,053	3,797 to 20,819	10,319	8,171 to 15,279

Note: Step data were missing for 13 participants. Mean daily steps were calculated for participants who had worn the Fitbit on ≥5 of the previous 7 days including at least one weekend day.
SD = Standard Deviation; 95% CI = 95% Confidence Interval
t-test for difference in means male vs. female: $t_{(165)} = 2.86$, **p = 0.005**

As **Figure 9** shows, a lower percentage of females were in the higher step count categories (i.e. 10,000 daily steps or greater). As **Table 14** shows, 46% of the participants (34% of males and 63% of females) failed to meet the 10,000 average daily steps target that is typically used in wearable activity monitors and apps and has been recommended by the American Heart Association (Strath et al., 2013).

Figure 9 Baseline mean daily step categories by gender



Note: Step data were missing for 13 participants. Mean daily steps were calculated for participants who had worn the Fitbit on ≥5 of the previous 7 days including at least one weekend day.

Table 14 Participants meeting 10,000 daily step target at baseline

Baseline mean daily step count	**Number and % of participants**		
	Male	**Female**	**All**
Less than 10,000	34 (34 %)	42 (63 %)	76 (46 %)
10,000 or higher	66 (66 %)	25 (37 %)	91 (54 %)
Total	100 (100 %)	67 (100 %)	167 (100 %)

Note: Step data were missing for 13 participants. Mean daily steps were calculated for participants who had worn the Fitbit on ≥5 of the previous 7 days including at least one weekend day.

There were some significant differences in mean daily steps at baseline according to age category (p = 0.011). With the exception of the 50+ age group, the older age categories recorded more steps per day on average than the younger participants. The 40 to 49 year age group had a significantly higher daily step count at baseline than those aged 18 to 29 years (11,721±3,419 versus 9,439±;3,126; p = 0.021) (see **Table 15**). Despite the more active role reported by PCSOs and special constables, the baseline step count of this group did not differ significantly from police officers or police staff (p = 0.091) (see **Table 16**). This suggests that police officers are more active outside of working hours, as reported in the interviews. The higher daily step count in the 40 to 49 year age category was apparent across all occupations (see **Figure 10**).

Table 17 shows mean daily step count according to rank; although sergeants had higher mean daily steps than constables (11,924±3,507 versus 10,702±2,998), there were no statistically significant differences between the groups (p = 0.241).

Table 15 Baseline mean daily steps by age category (n = 167)

Age (years)	Mean (SD)	95% CI	Range
18 - 29 (n = 25)	9,439 (3,126)	8,149 to 10,729	5,087 to 18,323
30 - 39 (n = 67)	10,474 (3,081)	9,722 to 11,225	3,797 to 19,626
40 - 49 (n = 48)	11,721 (3,419)	10,728 to 12,714	6,563 to 20,819
50 + (n = 27)	9,715 (3,016)	8,522 to 10,908	5,279 to 15,729

Note: Step data were missing for 13 participants. Mean daily steps were calculated for participants who had worn the Fitbit on ≥5 of the previous 7 days including at least one weekend day.
SD = Standard Deviation; 95% CI = 95% Confidence Interval
One-way ANOVA: $F_{(3, 163)}$ = 3.82, **p = 0.011**
Tukey post-hoc tests: difference between 18-29 and 40-49 age groups: **p = 0.021,** difference between 40-49 and 50+ age groups: **p = 0.047**

Table 16 Baseline mean daily steps by occupation (n = 167)

Occupation	Mean (SD)	95% CI	Range
Police Community Support Officers (PCSOs) and special constables (n = 27)	9,583 (2,763)	8,490 to 10,676	5,087 to 15,815
Police officers (n = 107)	10,955 (3,117)	10,357 to 11,552	4,873 to 20,819
Police staff (n = 33)	10,054 (3,888)	8,675 to 11,432	3,797 to 18,323

Note: Step data were missing for 13 participants. Mean daily steps were calculated for participants who had worn the Fitbit on ≥5 of the previous 7 days including at least one weekend day.
SD = Standard Deviation; 95% CI = 95% Confidence Interval
One-way ANOVA: $F_{(2, 164)}$ = 2.44, p = 0.091

Figure 10 Baseline mean daily steps by occupation and age (n = 167)

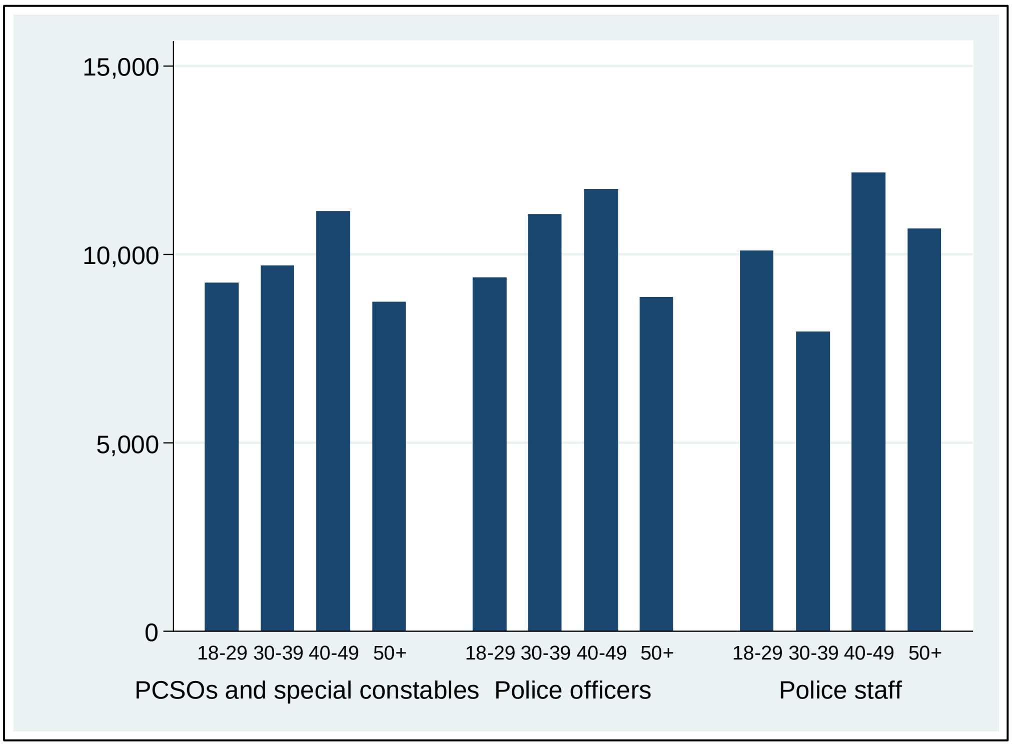


Table 17 Baseline mean daily steps by rank (officers only)

Rank	Mean (SD)	95% CI	Range
Constable (n = 80)	10,702 (2,998)	10,035 to 11,369	4,873 to 20,819
Sergeant (n = 23)	11,924 (3,507)	10,407 to 13,440	6,234 to 20,110
Inspector and higher (n = 4)	10,436 (2,661)	6,202 to 14,669	6,563 to 12,599

Note: Mean daily steps were calculated for participants who had worn the Fitbit on ≥5 of the previous 7 days including at least one weekend day.
SD = Standard Deviation; 95% CI = 95% Confidence Interval
One-way ANOVA: $F_{(2, 104)} = 1.44$, p = 0.241

Participants with 10 or more years of police service (including the current force and any previous forces) had a significantly higher daily step count at baseline than those with less than 10 years' service (10,953±2,979 versus 9,805±3,638; $t_{(165)} = -2.19$, p = 0.030). There were also differences in baseline step count according to marital status; participants who were married or in a civil partnership had a significantly higher daily step count than those who were single, divorced, separated or widowed (11,140±3,149 versus 9,470±3,253; $t_{(163)} = -3.23$, p = 0.002). There were no significant differences in mean daily step count at baseline according to police force (p = 0.426), presence of a health condition (p = 0.193), smoking status (p = 0.197) or dog ownership (p = 0.810).

Tables 18 to 21 summarise the self-reported PA and sedentary time of participants at baseline, overall and by gender (according to the IPAQ). The mean overall level of PA was 170.4 total minutes/week (95% CI: 154.8 to 186.0) or 3,182.1 MET-minutes/week (95% CI: 2,810.3 to 3,553.9). The mean level of moderate-to-vigorous physical activity (MVPA) was 1,718.6 MET-minutes/week (95% CI: 1,449.5 to 1,987.6). Activity levels varied greatly, with some participants doing as little as 10 minutes of total PA per week, and no MVPA. The mean self-reported sedentary time on a typical weekday was 6.41 hours (95% CI: 5.98 to 6.85).

As shown in **Tables 18 to 21**, male participants had a significantly higher level of total activity than females at baseline, including weekly minutes (p = 0.009) and weekly MET-minutes (p = 0.022). Male participants also had a significantly higher level of MVPA than females at baseline (p = 0.003). There were no gender differences in weekday sedentary time (p = 0.493).

Table 18 Baseline total physical activity (weekly minutes)

Gender	Mean (SD)	95% CI	Range
Male (n = 107)	187.5 (116.4)	165.2 to 209.8	10.0 to 540.0
Female (n = 73)	145.4 (83.8)	125.8 to 164.9	10.0 to 410.0
Overall (n = 180)	170.4 (106.2)	154.8 to 186.0	10.0 to 540.0

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval
t-test for difference male vs. female: $t_{(178)}$ = 2.66, **p = 0.009**

Table 19 Baseline total physical activity (weekly MET-minutes)

Gender	Mean (SD)	95% CI	Range
Male (n = 107)	3,537.5 (2,857.5)	2,989.9 to 4,085.2	148.5 to 16,398.0
Female (n = 73)	2,661.1 (1,844.4)	2,230.8 to 3,091.5	66.0 to 8,253.0
Overall (n = 180)	3,182.1 (2,527.8)	2,810.3 to 3,553.9	66.0 to 16,398.0

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval
t-test for difference male vs. female: $t_{(178)}$ = 2.31, **p = 0.022**

Table 20 Baseline moderate-to-vigorous physical activity (weekly MET-minutes)

Gender	Mean (SD)	95% CI	Range
Male (n = 107)	2,056.1 (2,032.7)	1,666.5 to 2,445.7	0.0 to 12,240.0
Female (n = 73)	1,223.8 (1,349.4)	909.0 to 1,538.7	0.0 to 7,560.0
Overall (n = 180)	1,718.6 (1,829.5)	1,449.5 to 1,987.6	0.0 to 12,240.0

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval
t-test for difference male vs. female: $t_{(178)} = 3.07$, **$p = 0.003$**

Table 21 Baseline self-reported weekday sedentary time (hours per day)

Gender	Mean (SD)	95% CI	Range
Male (n = 107)	6.29 (3.07)	5.70 to 6.88	1.00 to 15.00
Female (n = 73)	6.60 (2.76)	5.95 to 7.24	1.50 to 13.00
Overall (n = 180)	6.41 (2.94)	5.98 to 6.85	1.00 to 15.00

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval
t-test for difference male vs. female: $t_{(178)} = -0.69$, $p = 0.493$

According to the IPAQ classification, although the majority of participants were moderately or highly active at baseline, a small proportion (24/180, 13%) were classed as ‘low activity level’ (see **Table 22**). This meant that they were failing to meet public health recommendations of approximately 150 minutes (or approximately 600 MET-minutes) of moderate intensity activity per week (World Health Organisation, 2019).

Table 22 Baseline activity category (according to IPAQ classification) (n = 180)

Category	Number and % of participants
Low	24 (13 %)
Moderate	62 (34 %)
High	94 (52 %)

Note:

High = Vigorous-intensity activity on at least 3 days achieving a minimum total physical activity of at least 1500 MET-minutes/week, or 5 or more days of any combination of walking, moderate-intensity or vigorous-intensity activities achieving at least 3000 MET-minutes/week in total

Moderate = At least 20 minutes vigorous-intensity activity per day for 3 or more days/week, or at least 30 minutes moderate-intensity activity per day for 5 or more days/week, or 5 or more days of any combination of walking, moderate-intensity or vigorous-intensity activities achieving at least 600 MET-minutes/week in total

Low = Failing to meet criteria for ‘high’ or ‘moderate’ categories

As **Table 23** shows, mean daily steps at baseline had a moderately strong (Cohen, 1988) and highly significant association with the self-reported measures of PA, including total weekly PA in both minutes (r = 0.31; p<0.001) and MET-minutes (r = 0.41; p<0.001) and weekly MET-minutes of MVPA (r = 0.41; p<0.001). There were no statistically significant associations between self-reported weekday sedentary time and any of the measures of PA (objective or self-reported) (all p-values >0.05).

Table 23 Correlations between baseline mean daily steps, self-reported physical activity outcomes and sedentary time

	Mean daily steps	**Total physical activity (weekly minutes)**	**Total physical activity (weekly MET-minutes)**	**Moderate-to-vigorous physical activity (weekly MET-minutes)**	**Weekday sedentary time (hours per day)**
Mean daily steps	1.00 n = 167				
Total physical activity (weekly minutes)	0.31*** n = 167	1.00 n = 180			
Total physical activity (weekly MET-minutes)	0.41*** n = 167	0.82*** n = 180	1.00 n = 180		
Moderate-to-vigorous physical activity (weekly MET-minutes)	0.41*** n = 167	0.66*** n = 180	0.88*** n = 180	1.00 n = 180	
Weekday sedentary time (hours per day)	-0.10 n = 167	-0.09 n = 180	-0.13 n = 180	-0.04 n = 180	1.00 n = 180

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*****p<0.05 ******p<0.01 *******p<0.001

4.7 Health, wellbeing, stress and productivity (and associations with physical activity)

Of the 180 participants, 118 (66%) had no self-reported health conditions at baseline. The presence of one or more health conditions was reported by 33% (n = 60) of participants. The health conditions reported, with proportions as a percentage of the total sample, are shown in **Table 24**. The most frequently reported conditions were asthma (16/180, 9%), depression or anxiety (15/180, 8%) and type 1 or type 2 diabetes (7/180, 4%). Musculoskeletal disorders were reported by several participants, including osteoarthritis or rheumatoid arthritis (5/180, 3%), chronic back pain (5/180, 3%) and other connective tissue or musculoskeletal disorders (5/180, 3%). The ‘other specified’ health conditions reported by 6% (10/180) of participants included visual impairment, hypo- and hyper-thyroidism and other psychological conditions.

Table 24 Health conditions reported at baseline

Health condition	Number and % reporting (n = 180)
Asthma	16 (9 %)
Depression or anxiety	15 (8 %)
Diabetes - type 1 or 2	7 (4 %)
Arthritis (osteo or rheumatoid)	5 (3 %)
Chronic back pain	5 (3 %)
Other connective tissue or musculoskeletal disorder	5 (3 %)
Neurological disorder	4 (2 %)
Digestive disorder	4 (2 %)
Hypertension	4 (2 %)
Heart disease or circulatory disorder	3 (2 %)
Other specified health condition	10 (6 %)
Prefer not to say	2 (1 %)

Note: Some participants reported more than one health condition

The majority of the 180 participants (n = 123, 68%) had never smoked, while 28% (n = 51) were ex-smokers. Only six participants (3%) were current smokers. Frequency of alcohol consumption ranged from ‘never’ to ‘more than once a week’. Approximately equal numbers of participants consumed alcohol less than once a week (n = 83, 47%) and more than once a week (n = 95, 53%).

Physical health-related quality of life at baseline, as measured by the SF-12 Physical Component Score (PCS), was higher than the norm (average) score of approximately 50 (Utah Department of Health, 2001), with participants scoring 54.25±6.78 overall. Mental health-related quality of life was slightly lower than the norm, with a mean score of 47.56±9.26 (see **Table 25**). Quality of life scores were wide ranging, particularly for the Mental Component Score (MCS), where the lowest score was 17.71 and the highest score was 67.58. There was some evidence that female participants had a significantly lower mental health-related quality of life than male participants (mean 45.82±10.58 versus 48.73±8.10; p = 0.043). There were no significant differences by occupation (i.e. police officer vs. police staff vs. PCSO/special constable) in either PCS (p = 0.082) or MCS (p = 0.419) scores. Participants employed by the Dorset Police force had a near significantly higher mean MCS score than those employed by Devon and Cornwall Police (49.64±6.36 vs. 46.71±10.11; p = 0.059).

Table 25 Baseline SF-12 scores (for health-related quality of life) overall and by gender

	Physical Component Score (PCS)			Mental Component Score (MCS)		
	Mean (SD)	**95% CI**	**Range**	**Mean (SD)**	**95% CI**	**Range**
Male (n = 103)	54.11 (6.49)	52.84 to 55.38	35.58 to 68.27	48.73 (8.10)	47.15 to 50.32	22.22 to 62.48
Female (n = 69)	54.47 (7.24)	52.73 to 56.21	25.36 to 70.27	45.82 (10.58)	43.27 to 48.36	17.71 to 67.58
Overall (n = 172)	54.25 (6.78)	53.23 to 55.27	25.36 to 70.27	47.56 (9.26)	46.17 to 48.96	17.71 to 67.58

Note: Higher scores indicate higher quality of life; minimum possible score 0 and maximum possible score 100.
SD = Standard Deviation; 95% CI = 95% Confidence Interval
t-test for difference, male vs. female, Physical Component Score: $t_{(170)} = -0.34$, p = 0.736
t-test for difference, male vs. female, Mental Component Score: $t_{(170)} = 2.04$, **p = 0.043**

Table 26 summarises perceived stress levels at baseline, as assessed by the PSS-4. The mean PSS-4 score was 4.85±3.19, comparable to the norm of 4.8±3.0 reported in service occupations (Cohen and Williamson, 1988). There were no significant differences in perceived stress levels by gender (p = 0.263), occupation (p = 0.746) or police force (p = 0.332).

Table 26 Baseline 4-item Perceived Stress Scale (PSS-4) scores overall and by gender

	Mean (SD)	**95% CI**	**Range**
Male (n = 103)	4.63 (2.94)	4.06 to 5.21	0 to 12
Female (n =69)	5.19 (3.52)	4.34 to 6.04	0 - 13
Overall (n = 172)	4.85 (3.19)	4.37 to 5.33	0 - 13

Note: Higher scores indicate higher perceived stress; minimum possible score 0 and maximum possible score 16.
SD = Standard Deviation; 95% CI = 95% Confidence Interval
t-test for difference, male vs. female: $t_{(170)} = -1.12$, p = 0.263

The theme of stress in the workplace arose in the qualitative interviews, with most interviewees experiencing moderate levels of stress during their working day. Experienced stress was reported to result from acute exposure to stressful situations (such as anti-social behaviour, conflict and aggression) and more chronic issues including increasing pressure of work and demands on time:

“Even as a PCSO we deal with conflict on a regular basis, it might not be so high level as with the police but we are the ‘first responder’ if you like. You do get aggression and obviously your levels of adrenaline then increase. Normally it can be dealt with without any violence or anything like that and the adrenaline is not then used in your body.”

(PCSO, male, age 40+)

“I do have those kind of adrenaline rush dynamic responses to jobs as well so I tend to find my job is a little bit 0 – 60 in zero seconds.”

(Police inspector, male, age 40+)

“There is so much pressure for us in our department… I think that we do get affected by stress levels in here.”

(Police constable, female, age 18-39)

Interviewees recognised the associations between physical and mental health and the importance of PA to counteract the harmful impact of stress. For example:

"I suffer a little bit with stress and anxiety and I think I’ve recognised that it is primarily linked to diet and fitness."

(Police inspector, male, age 40+)

“It [physical activity] gets it out of your system, doesn’t it? It just makes you feel better. I can switch off quite easily when I am running on a treadmill.”

(Police constable, female, age 40+)

One police officer perceived an association between sedentary time and stress levels, and expected the intervention to help to reduce stress:

“I think that it [the intervention] will go towards reducing some stress levels as well, because I can just be sat behind the computer listening and talking on the phone and doing e-mails all day.”

(Police sergeant, male, age 18-39)

Sleep was perceived by interviewees as an important component of health and wellbeing. Several police officers reported problems with quality and quantity of sleep, which were recognised as a consequence of stress and shift work:

"I used to sleep eight hours straight through and didn't wake up in the night but I can't remember the last time I did that."

(Police sergeant, male, age 18-39)

"I think that we do get affected by stress levels in here. All of us in here have problems sleeping."

(Police constable, female, age 18-39)

"I think as you are older you take longer to recover from shifts… I know my sleep pattern compared with younger ones on the section is completely different. I average about five and a half hours' sleep and I am sort of fidgeting and fussing and waking up for about two of those, so it is not quality sleep, not at all."

(Police constable, female, age 40+)

The quantitative findings (shown in **Table 27** and **Table 28**) generally supported the perceived associations between PA, health-related quality of life and perceived stress. Mean daily steps at baseline had a weak but significant positive correlation with physical health-related quality of life as assessed by the SF-12 ($r = 0.16$; $p<0.05$), and a weak non-significant positive correlation with mental health-related quality of life ($r = 0.13$; $p = 0.108$). There was a weak correlation (nearing significance) between mean daily steps and perceived stress ($r = -0.15$; $p = 0.052$), with higher daily steps associated with a lower PSS-4 score (i.e. lower stress). The Physical and Mental Component scores of the SF-12 were negatively correlated, i.e. higher self-reported physical health was associated with lower self-reported mental health ($r = -0.32$; $p<0.001$).

For self-reported PA outcomes measured via the IPAQ, there was a weak but significant positive correlation between total weekly physical activity (MET-minutes) and mental health-related quality of life as assessed by the SF-12 ($r = 0.15$; $p<0.05$). A similar correlation was found between total weekly physical activity (MET-minutes) and perceived stress ($r = -0.18$; $p<0.05$), i.e. higher activity levels were associated with lower perceived stress. There was also a significant positive correlation between MVPA (weekly MET-minutes) and physical health-related quality of life as assessed by the SF-12 ($r = 0.25$; $p<0.01$). Although there were no significant associations between weekday sedentary time and PCS or perceived stress, the correlation between sedentary time and MCS neared significance (r = -0.15; p = 0.055), with greater sedentary time associated with lower mental health-related quality of life.

Table 27 Correlations between baseline mean daily steps and measures of wellbeing (physical and mental health-related quality of life and perceived stress)

	Mean daily steps	**SF-12 Physical Component Score (PCS)[1]**	**SF-12 Mental Component Score (MCS)[1]**	**Perceived Stress Scale (PSS-4) score[2]**
Mean daily steps	1.00 n = 167			
SF-12 Physical Component Score (PCS)[1]	0.16***** n = 164	1.00 n = 172		
SF-12 Mental Component Score (MCS)[1]	0.13 n = 164	-0.32******* n = 172	1.00 n = 172	
Perceived Stress Scale (PSS-4) score[2]	-0.15 n = 164	0.16***** n = 171	-0.73******* n = 171	1.00 n = 172

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*****p<0.05 ******p<0.01 *******p<0.001
[1] Higher scores indicate higher quality of life
[2] Higher scores indicate higher perceived stress

Table 28 Correlations between baseline self-reported physical activity, sedentary time, and measures of wellbeing (physical and mental health-related quality of life and perceived stress)

	Total physical activity (weekly minutes)	**Total physical activity (weekly MET-minutes)**	**Moderate-to-vigorous physical activity (weekly MET-minutes)**	**Weekday sedentary time (hours per day)**	**SF-12 Physical Component Score (PCS)[1]**	**SF-12 Mental Component Score (MCS)[1]**	**Perceived Stress Scale (PSS-4) score[2]**
Total physical activity (weekly minutes)	1.00 n = 180						
Total physical activity (weekly MET-minutes)	0.82*** n = 180	1.00 n = 180					
Moderate-to-vigorous physical activity (weekly MET-minutes)	0.66*** n = 180	0.88*** n = 180	1.00 n = 180				
Weekday sedentary time (hours per day)	-0.09 n = 180	-0.13 n = 180	-0.04 n = 180	1.00 n = 180			
SF-12 Physical Component Score (PCS)[1]	0.10 n = 172	0.15 n = 172	0.25** n = 172	-0.03 n = 172	1.00 n = 172		
SF-12 Mental Component Score (MCS)[1]	0.08 n = 172	0.15* n = 172	0.05 n = 172	-0.15 n = 172	-0.32*** n = 172	1.00 n = 172	
Perceived Stress Scale (PSS-4) score[2]	-0.08 n = 172	-0.18* n = 172	-0.07 n = 172	0.12 n = 172	0.16* n = 171	-0.73*** n = 171	1.00 n = 172

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*****p<0.05 ******p<0.01 *******p<0.001
Orange shading is used to show correlations of most relevance.
1 Higher scores indicate higher quality of life
2 Higher scores indicate higher perceived stress

Table 29 shows the mean and range of values for self-reported absenteeism, presenteeism and productivity scores, as assessed by the Health and Work Performance Questionnaire (HPQ) at baseline. Males had slightly higher self-perceived productivity overall than females, but the difference was non-significant (mean combined productivity score 1.12±0.36 vs. 1.02±0.36; p = 0.080). There were no significant occupational differences in any of the HPQ outcomes.

Correlations between mean daily steps at baseline and HPQ scores are shown in **Table 30**. There was an unexpected weak but significant negative correlation between mean daily steps and relative presenteeism (r = -0.20; p<0.01), and also between steps and the combined relative absenteeism and relative presenteeism score (r = -0.20; p<0.05), i.e. participants with higher daily steps rated their work performance as lower relative to colleagues, and had lower productivity. Subgroup analysis showed that this negative correlation was only significant for male participants (steps and relative presenteeism r = -0.32, p<0.01; steps and combined productivity score r = -0.32, p<0.01) and police officers (steps and relative presenteeism r = -0.21, p<0.05). There were no significant correlations between any of the measures of self-reported PA (or sedentary time) and any of the HPQ outcomes (all p-values >0.05).

Table 29 Baseline Health and Work Performance Questionnaire (HPQ) scores: self-reported absenteeism, presenteeism and productivity (n = 174)

	Mean (SD)	**95% CI**	**Range**
Absolute absenteeism[1] (hours lost per month)	6.56 (39.61)	0.63 to 12.48	-80 to 228
Relative absenteeism[2]	0.03 (0.21)	-0.001 to 0.06	-0.50 to 0.76
Absolute presenteeism[3]	78.62 (13.18)	76.65 to 80.59	30 to 100
Relative presenteeism[4]	1.10 (0.26)	1.06 to 1.14	0.38 to 2.00
Combined productivity score (relative absenteeism and relative presenteeism)[5]	1.08 (0.36)	1.02 to 1.13	0.24 to 2.33

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval
1 Higher score = higher absenteeism (negative score = worked more than expected)
2 Higher score = higher absenteeism, relative to expected hours (negative score = worked more than expected)
3 Higher score = lower lost performance, i.e. higher productivity
4 Higher score = higher performance or productivity relative to colleagues
5 Higher score = higher productivity

Table 30 Correlations between baseline mean daily steps and HPQ scores (self-reported absenteeism, presenteeism and productivity)

	Mean daily steps	Absolute absenteeism[1] (hours lost/month)	Relative absenteeism[2]	Absolute presenteeism[3]	Relative presenteeism[4]	Combined productivity score (relative absenteeism and relative presenteeism)[5]
Mean daily steps	1.00 n = 167					
Absolute absenteeism[1] (hours lost/month)	0.08 n = 166	1.00 n = 174				
Relative absenteeism[2]	0.10 n = 166	0.93*** n = 174	1.00 n = 174			
Absolute presenteeism[3]	-0.08 n = 166	-0.12 n = 174	-0.13 n = 174	1.00 n = 174		
Relative presenteeism[4]	-0.20** n = 166	-0.13 n = 174	-0.16* n = 174	0.45*** n = 174	1.00 n = 174	
Combined productivity score[5] (relative absenteeism and relative presenteeism)	-0.20* n = 166	-0.66*** n = 174	-0.72*** n = 174	0.36*** n = 174	0.79*** n = 174	1.00 n = 174

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*p<0.05 **p<0.01 ***p<0.001
Orange shading is used to show correlations of most relevance.
[1] Higher score = higher absenteeism (negative score = worked more than expected)
[2] Higher score = higher absenteeism, relative to expected hours (negative score = worked more than expected)
[3] Higher score = lower lost performance, i.e. higher productivity
[4] Higher score = higher performance or productivity relative to colleagues
[5] Higher score = higher productivity

4.8 Opportunities, barriers and facilitators for physical activity in the police force

Barriers and facilitators to PA in the police force, whether within or outside of the workplace, are likely to influence the success of the intervention. An understanding of these is an important component of context and a key element of process evaluation (Moore et al., 2014, Moore et al., 2015). This section focuses on the interview findings in relation to existing initiatives and opportunities for PA in the police force, and barriers and facilitators to PA.

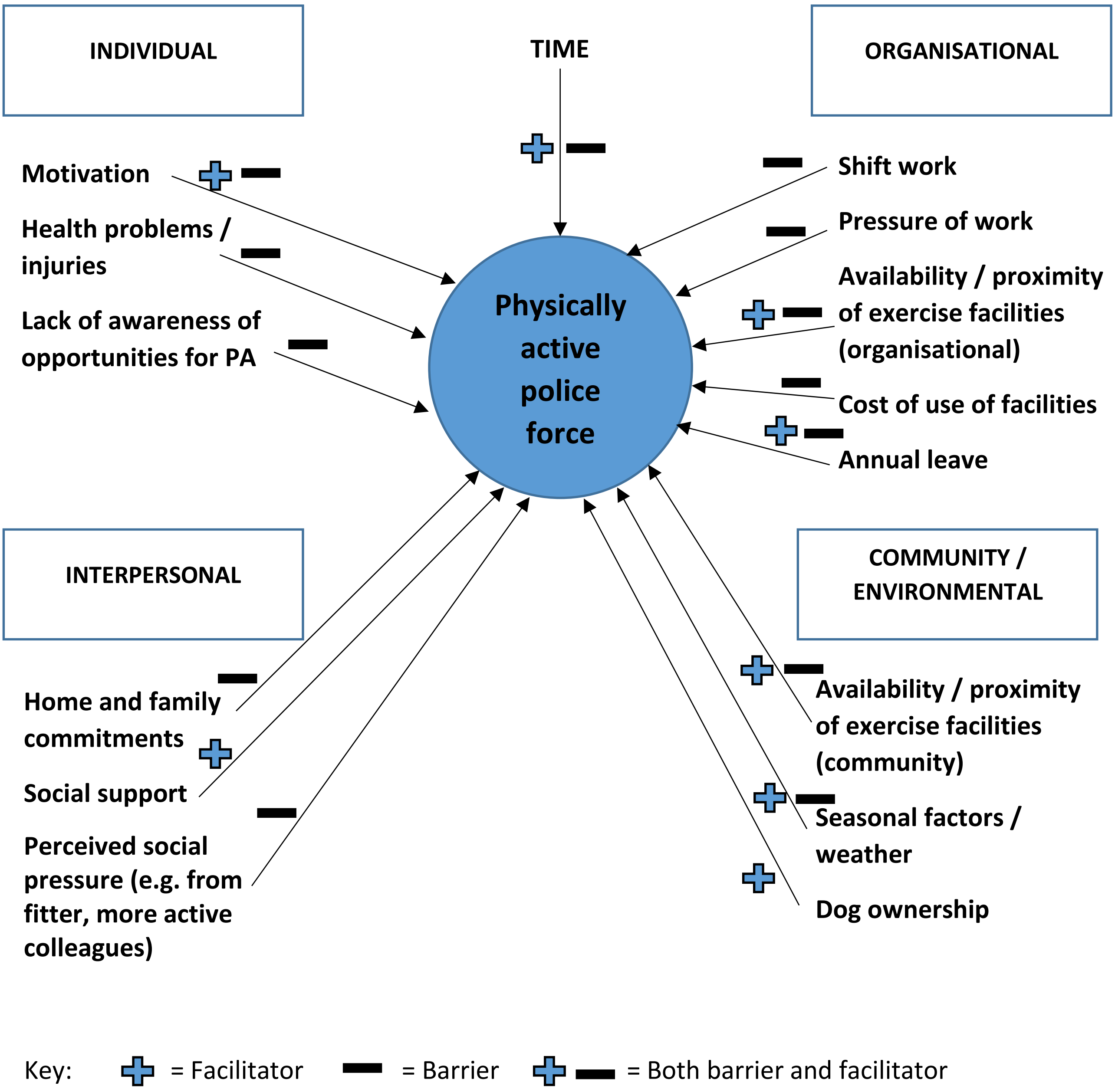
Interviewees reported several existing opportunities for being physically active within the police force. These included police force gyms, exercise classes (e.g. a lunchtime fitness club), and sports teams such as football, rugby and 'indoor rowing'. The extent of use and uptake of these opportunities varied between individuals. For example, some interviewees used the police force gyms on a daily basis, while others reported all of their activity took place in their own time, outside of working hours:

> *"I come into work at half past five in the morning, train in the gym until seven, get ready for work and come upstairs."*
>
> (PCSO, male, age 40+)

> *"I don't take part in any physical activity programmes at work. I actually go to my local gym which is in a town only a couple of miles down the road."*
>
> (PCSO, female, age 18-39)

Originating from Bronfenbrenner's Ecological Theory of Human Development (Bronfenbrenner, 1989) the Socio-Ecological Model has since been widely applied to public health (Lee et al., 2017). The model provides a conceptual framework for recognising the multi-level influences on a person's behaviour, including individual, interpersonal or social, organisational, community, and policy factors (Lee et al., 2017). **Figure 11** presents the main barriers and facilitators to PA for the police force using a Socio-Ecological framework. This is based solely on interview findings, including pre-intervention, post-intervention and follow-up interviews. The main barriers and facilitators are categorised in four levels – individual, interpersonal, organisational and community/environmental. These will now be discussed in turn, beginning with the organisational level barriers, which were most frequently mentioned and covered in depth by interviewees.

Figure 11 Barriers and facilitators for physical activity in the police force: A Socio-Ecological representation



The majority of officers and staff interviewed reported they were less active on office-based work days than off-duty days. Pressure of work and staff shortages was a frequently stated barrier, which resulted in a lack of time for PA, and for taking breaks in sedentary time during the working day:

"The only way that you could promote exercise and health and wellbeing during the working day, is by having more staff to allow people to have a break and we don't have that luxury at the moment."

(Police sergeant, female, age 18-39)

"Well, I think where we are at [small police site] there would be a good place to walk… but it's the pressure of the job which doesn't allow you to take maybe a proper lunch break. Quite often, I will literally sit at my desk and eat my lunch."

(Police constable, female, age 18-39)

Shift work was a key organisational level barrier to PA, and important to consider given the high proportion of shift workers (80%) in the study sample. Night shifts in particular were associated with tiredness and lack of energy for PA:

"I'd probably say the biggest hurdle, I find, is working night shifts. When you're working nights, it's so exhausting. It just absolutely annihilates your energy levels – particularly being at the wrong side of 40! I find I get home, go to bed at about eight o'clock in the morning… then by about midday, one o'clock, that's it, I can't sleep anymore and you're just exhausted the whole time. I might be able to muster up a dog walk, but in terms of going for a run or anything like that, it'd probably be counterproductive. You give up trying to do stuff on your night shift."

(Police inspector, male, age 40+)

Most officers and staff mentioned shift work as a barrier to using the police force gyms, and reported that when working hours were outside of the typical 9.00am to 5.00pm pattern it was difficult to fit exercise in to the working day. Early, late and night shifts were all associated with a lack of time for PA, and shift work was the main reason stated by some for not using the facilities provided by the police force:

"It's difficult with timings. After an early shift I'm so shattered, the last thing I want to do is go to the gym. Then before a late shift I'm walking the dog, so I can't come in any earlier to do it. I'm not going to do it at 10 o'clock at night. Then the only other shift I do is night, so... I'm not sure that I would use it."

(Police staff, female, age 18-39)

"A couple of people I work with go to the gym when they finish on a late [shift]… but the last thing I want to do at midnight is think about the gym!"

(Police staff, female, age 40+)

A lack of time for PA when working shifts was a particular problem for those with family commitments. For example:

"I used to go to the gym three times a week religiously… then moving up to different work - shift work - my hours have completely changed and the nature of the work is different and I just don't have chance to go to the gym as I finish at six o'clock if we finish on time. So if I went to the gym after work I could easily be getting home at half past seven, eight, and when you have got a missus at home who wants to eat, and family life, it is just not conducive."

(Police constable, male, age 18-39)

Rotating or changing shifts, experienced by 41% of shift workers in the study sample, were perceived as a particular barrier to maintaining a regular exercise routine, as they prevented attendance at weekly sports or fitness classes:

"I used to do martial arts before I started the staff role. But because of the six on, four off shift pattern I couldn't really justify it. Everything runs in weekly patterns and I wasn't always working the same days each week."

(Police staff, male, age 18-39)

"It is just getting involved and finding the time really. It is difficult on a shift pattern to fit it in. If I had the time I would consider doing something within the police, but you go one week and you can't go for four, and you go again and can't go for six… continuing going, that's the problem."

(Police staff, male, age 18-39)

Although shift work was clearly a key barrier to PA, a minority of those interviewed managed to overcome this:

"Shift work is always a pain, if you do lates, but for me if I'm on late shift I'll come in before work and use the gym and go in then. If I'm on day shift I will get up early and go before I start."

(PCSO, male, age 40+)

Lunchtime fitness clubs were offered by one of the police forces in the study, but there were perceived inequalities in access to this facility. Interviewees reported that this opportunity depended greatly on work stream and shift pattern, with some participants feeling excluded from this as a result:

> *"I have got a feeling there was something being set up at one of the stations but it is not something that I would be able to go and do because it was on a certain day on certain lunch breaks. It does exclude quite a lot of people, particularly response people because obviously they don't get meal breaks very often, and again it is away from the station... so it is not something that I could do."*
>
> (Police constable, female, age 40+)

> *"The one thing that I have found in the workplace is that it is really role dependent on the opportunity for physical exercise... when you have officers, perhaps in the response function, who are 24/7 shifts, it literally is from the word go, it's absolutely crazy. They don't have those same opportunities."*
>
> (Police sergeant, female, age 18-39)

In addition to perceived work pressures and shift work, another major organisational barrier to PA was lack of availability of or proximity to exercise facilities. A lack of exercise facilities in the workplace was reported by several interviewees, and was a particular issue with small, rural police stations (i.e. the North Dorset sites):

> *"We are a bit limited where we work. There's not many options to exercise."*
>
> (Police sergeant, male, age 18-39)

> *"The nearest force gym is half an hour's drive away… there are a couple of other schemes but it is right the other side of the county so we are a bit limited here. I do it out of my own time really as we just don't have the facilities up this way."*
>
> (PCSO, female, age 18-39)

This issue was not exclusive to rural police stations, however, and was also reported by a police officer working in a more remote urban site:

> *"We're in a remote station so we've got literally nothing around us. We haven't got the privilege of having a Gucci gym like Headquarters."*
>
> (Police constable, female, age 18-39)

Interviewees also perceived some of the existing gym facilities as poor quality. For example, a lack of personal coaching and support and overcrowding were mentioned:

> *"The police gyms are shocking. There is no personal coaching… it has just not got what I need."*
>
> (Police constable, female, age 18-39)

> *"I think the facilities at [large city police station] aren't brilliant. I have never used the gym there. There is one in [small city police station] which I have used but it's tiny; you find if there is more than two of you in there it becomes very crowded."*
>
> (Police constable, male, age 40+)

Cost was initially stated as a barrier to use of police force gyms; some sites had a supplementary fee in addition to the monthly subscription, but this fee was removed by the end of the study:

> *"There's a gym at work. They've just actually got rid of the membership fees that people happen to pay additionally, on top of the monthly fee. For me, now, I'm like, brilliant, because you can jump in and out of different gyms as you travel around the force for work, whereas before you'd have to pay an individual membership for each gym. So you wouldn't be able to use it but they've got rid of it which is amazing, now I can just go and use the gym without worrying about that."*
>
> (Police sergeant, female, age 18-39)

Interviewees reported that annual leave influenced their activity levels. For some, PA was lower (e.g. beach holidays), whereas others reported being more active when on annual leave (e.g. active holidays):

> *"When I was on leave, I tended not to do so much, I was just relaxing. I think that's probably a healthy part of-, you can still be active, but you do need to relax sometimes."*
>
> (PCSO, male, age 40+)

> *"I went on a holiday to New York and we were doing about ten miles a day walking around."*
>
> (Police constable, male, age 18-39)

Individual level factors influencing PA included motivation, health problems and injuries, and lack of awareness of opportunities for PA in the workplace. Motivation appeared to vary greatly between participants. There was evidence that some officers and staff were highly intrinsically motivated at the beginning of the study, whilst others felt that lack of motivation was preventing them from being active:

> *"There won't be anything that stops my training. I've got to be doing something… it is just really important for me. I love it!"*
>
> (Police constable, female, age 18-39)

> *"The main problem is myself really, trying to get myself motivated."*
>
> (Police staff, female, age 40+)

> *"Going down the gym is a major problem for me. I know I should go and I don't. I have got a gym membership but it is not used that often."*
>
> (Police constable, male, age 40+)

One police officer believed the lack of motivation and energy for PA she experienced was exacerbated by the sedentary work role:

> *"… tonight I will be going to the gym even if it is just for half an hour, but I generally find that when I have been sitting round all day I don't feel like doing it."*
>
> (Police constable, female, age 40+)

The same individual recognised that motivation also varied from month to month:

> *"I feel better because I did dry January. In that particular month, I can honestly say I felt much better because I was more health conscious, more everything conscious. February, I slipped off again… but then March, I said, 'I'm going to get back into it again'. I had a down month and then I picked up again. It's pits and stops with me."*
>
> (Police constable, female, age 40+)

Health problems and injuries were stated as short- and long-term barriers to PA. Musculoskeletal disorders were common, experienced by around 9% of the study sample overall (see **section 4.7, Table 24**). These were often experienced by, although not limited to, older staff and officers:

> *"As you get older, things don't always work as well as they did. I have got what I've got. My knees are knackered, my back plays up from time to time, but that's been an ongoing thing for the last 15 years. Fair wear and tear, I think my doctor called it."*
>
> (Police staff, male, age 40+)

One interviewee had had a hip replacement at a young age, which restricted the type of activity he could do:

> *"… it has been fine since but I have to be a bit careful with it. They have advised me not to run and things like that, so I am a bit more limited."*
>
> (Police sergeant, male, age 18-39)

In some cases, there seemed to be a lack of awareness of opportunities for PA in the police force.

> *"If there are any [workplace PA opportunities] I am not aware of them."*
>
> (Police constable, female, age 40+)

> *"I'm not aware of any [opportunities] to be fair. There probably are, it's just me not necessarily looking into it."*
>
> (Police constable, male, age 18-39)

One interviewee perceived a reduction in opportunities for PA in the police force over time:

> *"I wouldn't say there are many opportunities for me to keep fit at work… all the days of sport within the police are few and far between now really - we don't get any time off to do that sort of stuff."*
>
> (Police sergeant, male, age 18-39)

It is important to consider individual differences in preferences for PA in the workplace. While some interviewees were keen to use facilities provided by the police force, others preferred to keep their work and home lives separate:

"I come to a private gym. I don't really like to stay at work to use the gyms at work. I like to get away from work when I'm done."

(Police inspector, male, age 40+)

"I think because I work with police officers I don't want to then train with them."

(Police constable, female, age 18-39)

"The main reason I don't [use the force gym]… is because I have a thing about keeping work life and home life separate. When it comes to the end of the day or even in the beginning, I don't want to be spending extra time training at work when I can go home and perhaps go to the gym with my other half or go horse-riding or do exercise away from the work environment, because we're here for such a long time anyway. I just find it a bit healthier to exercise away from the work environment."

(Police sergeant, female, age 18-39)

The main interpersonal level barriers to PA were home and family commitments and perceived social pressure. Lifestyle factors were also influential. One interviewee stated that she had to fit in activity with her partner's work as well as her own. Others reported that they needed to prioritise family life over exercise, whether looking after children or caring for an elderly relative:

"I got married four years ago and then we had a child as well and we renovated a house and we had another child recently… so a lot of my priorities have changed. I used to get to the gym about three or four times a week and now I am lucky if I go down there once a week."

(Police sergeant, male, age 18-39)

"I've got older family to look after, I've got Dad to look after, a husband to look after and time is sparse, it really is."

(Police constable, female, age 40+)

For the less active and older interviewees, perceived social pressure from fitter, younger, more active colleagues discouraged use of the facilities provided by the police force. For example:

"My colleagues are all very fit, and they're a lot younger than me. I've tried different gyms and it's the people that are, 'Well is that all you've done?'… that puts me off."

(Police staff, female, age 40+)

Social support, including encouragement from or exercising with a partner or friend, was seen as a facilitator for some interviewees:

"My partner's been trying to encourage me [to be more active]."

(Police staff, female, age 18-39)

"My partner and I, we both think we need to do something to get more active, because now he starts an office job... he's saying now that he's getting frustrated because while he used to be out and about and a lot of physical activity, he's not getting that in work now. We're trying to think of things that we can do together because we're not particularly fit."

(Police staff, female, age 40+)

There were also barriers and facilitators at the community and environmental level. These included availability of or proximity to exercise facilities outside of the workplace, seasonal factors/weather, and dog ownership. Some interviewees in more rural and remote areas found it easier to access their local community gym than police force gyms:

"I'm at a very small station. I have to go to my own gym near where I live. We haven't got the facilities at the workplace."

(PCSO, female, age 18-39)

Seasonal factors and the weather were frequently mentioned through the study as having an impact on activity levels. Participants reported being less active in very hot weather or wet weather, and some reported being less active during winter:

"I think through the winter months, you just don't feel - well, I don't feel like going out anyway, but once the sun is shining it's easier."

(Police staff, female, age 40+)

Of the entire study sample, 38% were dog owners and dog walking was frequently reported as a preferred way to be physically active. One interviewee felt strongly that dog ownership was an important facilitator for sustaining her activity level in the longer term, and also helped her to overcome barriers to PA such as poor weather:

> *"The dog is pretty good motivation. You can tell everybody to get a dog! There's been a massive increase in my activity since we've had him because obviously if it's raining and horrible, the temptation is going to be just to stay inside and not be active at all, but when you've got a dog looking at you wanting to go out you have to, don't you… I still walk the dog in the snow. I would say that has made a massive difference."*
>
> (Police staff, female, age 18-39)

Time (or usually lack of time) was perceived as a major influencer of PA in the police force. This was an overarching theme that did not fit into any particular 'level' of the Socio-Ecological Model. Lack of time was usually as a result of home and family commitments and/or work commitments:

> *"I am working and you just try to fit everything in, don't you… and fail!"*
>
> (Police constable, female, age 40+)

> *"I need more hours in the day! That's the only thing that's stopping me, just not having enough time to let my body recover."*
>
> (Police constable, female, age 18-39)

4.9 Suggestions for promoting physical activity and reducing sedentary time in the workplace

This section presents the main suggestions for increasing PA and reducing sedentary time in the police force. These are based on the barriers and facilitators discussed above, and include interviewees' own recommendations for a more active police force.

1. **Designated time for physical activity/fitness/wellness during work hours**
   Designated time for PA during the working week was the most common suggestion made by those interviewed. Many felt strongly about this, and

believed that dedicated time should be a core element of the role, given the requirement and expectation for the fitness of officers. For example:

> *"I think like the services, the Navy, the Army etcetera, there should be fitness time built into the routine. Even if it's just half an hour… I don't agree with making us do a fitness test and then not giving us any time to get fit."*
>
> (Police constable, female, age 40+)

> *"We don't get allocated time at work to go to the gym, to go out for a run, which I strongly feel the force should, especially as we have a fitness test to pass. Especially in our line of work with the stress and the high intensity situations that we go into, you need to burn that excess adrenaline off or else that has adverse effects to your health as well. For those reasons, I think the force should implement dedicated time for us to have a period of time where we can go out and do some phys, even if it's just a couple of hours, just to go out and do it. The main thing, for me, is for the job to create time for us to actually go out and be physically active."*
>
> (Police constable, male, age 18-39)

The consensus was that workplace PA opportunities (such as the lunchtime fitness sessions) should be equally accessible to all work streams. Interviewees suggested that this time should be flexible and tailored to individual preferences and needs (e.g. according to health and physical capabilities), and could include an option for general 'wellness' time:

> *"What you've got to do to improve people's health is to find an interest, a sport, an activity, which will help <u>them</u>***** to improve their health. I think that what you've got to do is find everybody something to help them improve."*
>
> (PCSO, male, age 40+) Note: ***** = emphasised by interviewee

> *"They [the police force] spoke about bringing in like an hour a week or a month that you could just go and do some kind of physical exercise. They say these things, but nothing gets put into place. But wellness is a big umbrella, and will mean different things to different people. Some people would like meditation time. I'd probably like training time."*
>
> (Police constable, female, age 18-39)

2. **Improved access to (and quality of) police force gyms**

   As stated previously, those working in more rural and remote sites desired more facilities for PA in the workplace. It was clear that such facilities would be used:

> *"If there was literally something in the next office, I'd be on it right before work, after work, lunchtime…"*
>
> (Police constable, female, age 18-39)

> *"If they ever have activities up further this way, I would be really up for it."*
>
> (PCSO, female, age 18-39)

One interviewee suggested that if funds did not allow more police force gyms to be established, links with local colleges could be used to improve access to sports facilities. Improved gym facilities to include larger gyms, personal coaching and support were desired by some officers and staff.

3. **Personal feedback and support following the annual fitness test**

   Recommendations included further support regarding the annual fitness test, including provision of an annual, optional personal fitness record, and greater individual feedback and advice:

> *"If there was a personal record of your current fitness, and that was measured every year, I think that would focus the mind to either remain fit or to get fitter… but there's nothing like that at the moment. A personal fitness record, which would be optional… I think people are focused if you know you're going to be measured on it as well, especially if you see that date come up again, you'd potentially try and better it."*
>
> (Police sergeant, male, age 18-39)

> *"The remedial stuff is very, very non-existent. It is a case of giving you a bit of advice… I would suggest that the help and advice isn't really there at any level. I am not saying that everybody gets a personal fitness instructor but what I am saying is it should be a bit more proactive. Because I think unfortunately that the police don't like to go there too much with that sort of thing. They don't like to impinge on - they don't like to say to people - you know what, you are overweight,*

you need to lose weight, you need to be a bit more active. If that was put in place a lot quicker then people would respond to it I think. It is a very, very sort of controversial and delicate area to go into really."

(Police constable, male, age 40+)

4. Increase the difficulty of the fitness test

There was a general perception that the current annual fitness test was too basic, and should be increased in difficulty to improve fitness levels:

> *"Since they got rid of the old sort of test… there has been a noted drop away in, I believe anyway, fitness levels and because the test is now very basic… you have to do a 15 metre shuttle run, you are literally running, if you can call it that, for three minutes 25 seconds. I am not entirely sure what that represents really."*
>
> (Police constable, male, age 40+)

5. Focus on the needs of the least active officers and staff

Regardless of their own activity levels, interviewees believed that future wellbeing and PA promotion programmes should focus on those who were most in need of intervention, and that a wide range of sports should be included to appeal to more officers and staff. For example:

> *"For me who enjoys sport… it's easy because I like doing it. For somebody that doesn't, and there are a lot of people in the police force I believe who don't like it, I think the force needs to encourage more sport, more activities, not just going to the gym because some people hate the gym. They've got to look at finding other interests that'll make people more active."*
>
> (PCSO, male, age 40+)

One interviewee suggested implementing a 'closed time' for police force gyms, where the facilities could be used exclusively by the least active staff. This might help to overcome the perceived social barriers:

> *"The gyms I've been to, I don't like because it's too posy. It'd be nice having something that caters for people of all ages and all abilities… In some places you've got to be a health freak or a fitness freak to even start and that puts me off. I'd like to go somewhere that I know somebody else is going to be huffing and puffing and not be able to hold a conversation because they can't breathe when they're walking. I know we've got a gym at work but the ones I've seen that go in there*

are the ones that do half marathons, they do triathlons and they do this and they do that. If they had say, a closed time at the gym where the ones that are not in the peak of fitness… because everybody works shift work… everybody just goes in as and when, but if there was like a designated time, say on Wednesday afternoon or something like that, for the ones that need motivation and support from others without feeling patronised..."

(Police staff, female, age 40+)

6. **More health and wellbeing champions to organise group activities and workplace competitions**

Several interviewees suggested there was a need for more health and wellbeing champions and organised activities in the workplace, as currently this role was performed by only a small number of officers:

> *"I know that there is something in place but there just isn't the staff to do it. I have mentioned [wellness team sergeant] because he has got a very small team around him and he is trying to encourage people to be more well, which is great, but he just doesn't have the capability to get round to everybody."*
>
> (Police constable, male, age 40+)

Interviewees recommended workplace PA competitions and challenges with real incentives and rewards (the use of virtual rewards is discussed later in **Chapter 6**). Another suggestion was to continue to promote awareness and participation in any national police-led initiatives for PA, such as the annual 'Snowdonia Seven' challenge.

7. **Improved awareness of existing opportunities for PA**

It was clear from the interviews that many officers and staff were unaware of all the existing PA opportunities available to them, such as exercise classes, force sports teams and national initiatives such as the 'Love2Ride' cycling to work scheme. Involvement of the communications teams within the respective forces could be used to facilitate this.

8. **Greater support from managers and supervisors to increase PA and encourage breaks in sedentary time**

Officers and staff wished for greater support from higher level staff to encourage PA, breaks in sedentary time, and health and wellbeing in general. For example:

> *"I think supervisors could make a difference in ensuring that their staff get regular meal breaks and for those that are at a desk, reminding them to get up and walk around a little bit. I know it's people's responsibility but it's too easy once you get ensconced in a job to not do it."*
>
> (Police constable, female, age 40+)

Interviewees with supervisory responsibilities were supportive of this:

> *"If you were… in a role where you're primarily in the office, your line manager, or I certainly would for my team, give them the time to say, look, if you want to use your lunch break to go to the gym, or you want an extra half an hour, because you want to use a gym, then I'd look at that and make sure it's suitable."*
>
> (Police sergeant, female, age 18-39)

Officers and staff desired more input from management than just written information:

> *"My issue with management is they're very good at providing advice in writing… how to eat well and all the rest of it. When you're in a front line role, it's very difficult to eat properly. When you're backing colleagues it's very difficult to find the time and because I'm a bit older, your body takes longer to recover as well from everything."*
>
> (Police constable, female, age 40+)

9. **(Optional) use of mHealth / wearable fitness technology**

Many interviewees believed mHealth technology may be an effective intervention to promote PA in the police force, but that its use should be optional. Views on acceptability of the intervention are discussed in detail in **Chapter 6**.

4.10 Summary / Discussion

In summary, the study sample (n = 180) was diverse in terms of both socio-demographic and occupational characteristics. The participants comprised a large proportion of shift workers (80% of the total sample). Interview participants were also diverse and representative of the overall study sample; this is a strength of the research.

Overall, both the quantitative and qualitative findings demonstrated the need for an intervention to promote PA and reduce sedentary time in the police force. The majority of study participants (58%) reported their role as mainly sedentary and mean self-reported sedentary time on a typical weekday was 6.41 hours. This compares with approximately five hours per day for the average European employee (Nuffield Health, 2018). Spending longer than six hours a day sedentary is of concern as this level has been associated with a significantly higher risk of mortality in adults (Patel et al., 2010). The officers and staff interviewed were generally more active outside of work hours. This confirms the findings of existing international studies, such as a study in the USA that used objective PA data to show that policing was primarily a sedentary occupation, with police officers more active on off-duty days (Ramey et al., 2014). While participants in the present study were moderately active overall, almost half were doing less than 10,000 steps per day at baseline, and a small proportion (around 13%) were classed as 'low activity level' according to the IPAQ (i.e. failing to meet public health recommendations of approximately 150 minutes or 600 MET-minutes of moderate PA per week – World Health Organisation, 2019). This is comparable to a UK survey of over 5,000 police officers and staff using the same outcome, with 14% of females and 10% males in this category (Gibson et al., 2018). There may therefore be a need for PA interventions in the police force population at a national and international level.

A novel finding was individual differences in activity level based on factors such as occupation, work stream, gender, marital status and age. In general, PCSOs and special constables tended to be more active during the working day, while police officers tried to compensate for their sedentary role with more exercise outside of work. Officers working within the response policing function were generally more active than those in local investigation and local policing (neighbourhood) teams.

This was also reported in a study of Polish police officers (Soroka and Sawicki, 2014). Male officers and staff had higher objective and self-reported levels of PA than females; this is a typical finding in the general adult population (Althoff et al., 2017). Those who were married or in a civil partnership had higher daily steps than those who were single, divorced or widowed; this is in concordance with epidemiological studies that show higher PA amongst those who are married or cohabiting (Koyanagi et al., 2018). Contrary to the findings of previous studies that have shown that police officers become less active as they progress through their career (Lagestad and Van Den Tillaar, 2014), the most active age group was 40 to 49 years, who were significantly more active than those aged 18 to 29. In addition, those with 10 or more years of police service had higher daily steps. Younger police officers and staff, and those earlier in their career, may therefore need to be targeted in future PA interventions.

The significant correlation between the objective outcome of steps from the Fitbit® and self-reported measures of PA adds support to the validity of these measures. The use of multiple outcomes is a strength of the study, and it is recommended that future studies follow this example.

One third of the study participants reported a health condition at baseline, the most prevalent being asthma, depression and anxiety, diabetes and musculoskeletal conditions. There was no evidence that the prevalence of these conditions was higher than in the general population (Asthma UK, 2019, Public Health England, 2018), however it is difficult to draw definitive conclusions based on a small sample. It is also possible that some individuals in this sample had undiagnosed physical or mental health conditions. Cardiovascular disease and metabolic syndrome have been previously reported as highly prevalent within the police force (Franke et al., 2002, Hartley et al., 2011, Anderson et al., 2016). Clinical data on risk factors for these conditions, such as cholesterol and body mass index, were not collected in the present study due to resource limitations, but could be used as baseline measures and outcomes in future trials.

High levels of depression have previously been associated with the policing occupation, with a prevalence almost double that of the general population (Hartley et al., 2011). The SF-12 survey findings in the present study suggested lower than expected mental health-related quality of life, with a mean MCS of 47.56, compared with the US norm score of 50 (Utah Department of Health, 2001). There was some evidence that male participants had better mental health than female participants, and that participants based at the more rural North Dorset sites had better mental health than the urban Plymouth BCU officers and staff; this is worthy of further investigation.

The mean PSS-4 score was almost identical to the previously reported norm for service occupations (Cohen and Williamson, 1988). Some participants reported very high levels of stress and it was clear from the interviews that both acute and chronic stress were commonly associated with the policing role. High stress levels are a common finding in studies of the police force (Gershon et al., 2002, Ramey et al., 2016). As perceived stress has been independently associated with cardiovascular disease in the law enforcement occupation (Franke et al., 2002), interventions to reduce stress are clearly required.

The links between PA and health were clear from both the qualitative and quantitative findings. Interviewees perceived that higher levels of PA (and lower levels of SB) were linked with better physical and mental health and lower stress. Correlations between objective and self-reported measures of PA, physical and mental health-related quality of life and stress supported this perception. The correlation between self-reported sedentary time on a typical weekday and SF-12 MCS score neared significance, with higher sedentary time associated with worse mental health. The positive associations between PA and both physical and mental health are well documented (Penedo and Dahn, 2005). Although these relationships are likely complex and bi-directional (Hiles et al., 2017, Liao et al., 2017), the findings indicate the potential for interventions to improve PA, reduce sedentary time and consequently improve health and wellbeing. The negative correlation between the SF-12 Physical and Mental Component scores at baseline was an anomalous finding. The reasons why lower physical health may be associated with better mental health (and vice versa) should be explored in future studies, to find out if this is an unexplained characteristic of this sample or more widely apparent in the police force.

Although not a core component of this research, several interviewees reported sleep problems, which were commonly perceived as resulting from high stress and shift work. Sleep is an important aspect of health and wellbeing, and should be addressed in future mHealth studies, particularly given the high prevalence of shift workers in the police force. Sleep is discussed later in **Chapter 6** in relation to the wider impact of the intervention.

Perceived productivity was not positively correlated with PA at baseline; in fact there was an unexpected negative correlation between mean daily step count and relative presenteeism for male participants and police officers. This finding is not in line with previous studies reporting higher levels of work performance in those who are more active (such as Pronk et al., 2004). The impact of the intervention on productivity, and the usefulness of the HPQ absenteeism and presenteeism questions as an assessment tool, are discussed in later chapters.

The interviews provided a detailed insight into the wider context of the intervention including existing opportunities for, and barriers and facilitators to PA in the police force. There were several existing opportunities for PA in the Devon and Cornwall and Dorset Police forces, including workplace gyms, exercise classes and sports teams. The extent to which interviewees used these facilities varied greatly. Barriers and facilitators to PA in this context were classed into four main categories according to the Socio-Ecological Model. These included individual (e.g. motivation), interpersonal (e.g. social support), organisational (e.g. shift work) and community/environmental (e.g. season) level factors. Several of these factors have been cited in studies of adults in non-occupational settings, including health problems (Kurti et al., 2015), childcare duties (Kurti et al., 2015), lack of time (Kurti et al., 2015), the weather (Kurti et al., 2015, Welch et al., 2018), season (O'Connell et al., 2014), dog ownership (Peel et al., 2010, Dall et al., 2017) and availability of exercise facilities (Cleland et al., 2015). Some barriers identified in the present study may be particularly pertinent to the policing occupation, such as shift work, work pressures, lack of availability or access to exercise facilities in the workplace, and perceived social pressure from more active colleagues. While previous survey studies have recognised lack of time, work pressures and lack of access to exercise facilities as barriers to PA in the police force (Soroka and Sawicki, 2014, Lagestad and Van Den Tillaar, 2014), there have been few, if any, in-depth qualitative studies of factors influencing PA levels in this population. This is therefore an important novel contribution to the literature. Shift work was a major barrier to initiating and sustaining PA, and was associated with tiredness, a lack of time, and was a key reason given by some for not making use of the exercise facilities provided by the police force. Changing shifts prevented regular participation in sports or exercise classes. The interviews revealed important individual differences (e.g. in preferences for PA in or outside of the workplace) and differences by police force (e.g. lack of availability of exercise facilities was a larger issue with remote and rural police stations). The Socio-Ecological Model was selected for its recognition of multiple levels of influence (including external influences) on behaviour, its applicability to PA and SB (Essiet et al., 2017, O'Donoghue et al., 2016), and its appropriateness in an occupational setting with the organisational layer component. While the Socio-Ecological Model has previously been used to characterise factors impacting on PA and SB in an occupational setting (Morris et al., 2018), it has not been previously applied to PA in the police force; this is an important methodological contribution.

Recognition of key influences on PA in the police force led to suggestions for promoting PA and reducing sedentary time in this context. The main recommendations made by interviewees were designated time for PA (or general wellness) during work hours, improved access to (and quality of) police force gyms, changes to the annual fitness test (including increased difficulty and provision of individual feedback on performance), further support from managers and supervisors (including encouraging breaks in sedentary time), and more workplace champions and organised activities. A need to focus on the needs of the least active officers and staff was recognised, in addition to a need to improve awareness of existing opportunities for PA in the workplace. Overall, there appeared to be a need for multi-level interventions that are suitable for, and accessible by, the diverse population of officers and staff.

In summary, this chapter has provided a comprehensive picture of the context (a key component of process evaluation) of PA, sedentary time, and health and wellbeing in the police force. The baseline findings demonstrated the need for the intervention and highlighted individual differences to give an indication of who may benefit most. The findings surrounding health and wellbeing in the police force add to and extend existing literature, for example emphasising the issues of stress and relatively low mental health-related quality of life, which may be more prevalent amongst female and urban police officers and staff. The importance of multiple external influences on PA and SB was recognised, and should be a principal consideration in the design and evaluation of interventions.

Using findings from qualitative interviews, and taking a Socio-Ecological approach, our understanding of key barriers and facilitators for PA in this context was furthered, making a novel contribution to the literature. Suggestions for the promotion of PA and reduction of sedentary time may inform future organisational policies. Combined with mixed methods results in later chapters, these findings will help to explain success (or otherwise) of the PAW-Force intervention and may be used to enhance the design and targeting of future interventions within the police force.

CHAPTER 5 IMPACT OF THE PAW-FORCE INTERVENTION

5.1 Quantitative exploratory analysis of primary outcomes

5.1.1 Introduction and overview

In this section, the quantitative outcome data are used to explore the impact of the PAW-Force intervention on physical activity (PA, including objectively measured daily step count and self-reported activity) and sedentary time. Assumptions and methods used for data analysis are described fully in **Chapter 3**.

The section begins with a summary of mean daily steps, self-reported PA and sedentary time at the various outcome time points. Changes in these primary outcomes for the various phases of the study are then presented, followed by a consideration of the relative impact of the individual and social phases. Subgroup and sensitivity analyses are included throughout this section. For changes in steps only, the results of an exploratory regression analysis on participant characteristics are presented. Correlations between changes in the primary outcome variables are then explored. The section ends with a summary and discussion of the key findings in relation to the existing literature.

5.1.2 Summary of primary outcomes (steps, self-reported physical activity and sedentary time)

Table 31 provides an overview of the absolute mean values for steps, self-reported PA and sedentary time at each time point – baseline (week 0), mid-intervention (week 6), post-intervention (week 12) and follow-up (month 8).

Table 31 Physical activity (PA) and sedentary time - absolute values at four time points (all participants, all available data)

Outcome	**Mean (SD) n = number of observations** **Baseline (week 0)**	**Mid-intervention (week 6)**	**Post-intervention (week 12)**	**Follow-up (month 8)**
Step count (mean steps/day)[1]	10,555 (3,259) n = 167	10,717 (3,260) n = 118	10,706 (3,180) n = 114	10,191 (2,921) n = 87
Total PA (minutes/week)	170.4 (106.2) n = 180	191.5 (115.7) n = 157	192.3 (116.2) n = 151	192.5 (112.4) n = 143
Total PA (MET-minutes/week)	3,182.1 (2,527.8) n = 180	3,555.3 (2,770.2) n = 157	3,648.7 (2,643.3) n = 151	3,555.1 (2,710.6) n = 143
Moderate-to-vigorous PA (MVPA) (MET-minutes/week)	1,718.6 (1,829.5) n = 180	1,932.4 (2,000.3) n = 157	2,097.4 (1,960.3) n = 151	2,164.6 (2,291.0) n = 143
Sedentary time (hours on a typical weekday)	6.41 (2.94) n = 180	6.18 (3.26) n = 157	6.29 (3.43) n = 151	6.03 (3.04) n = 143

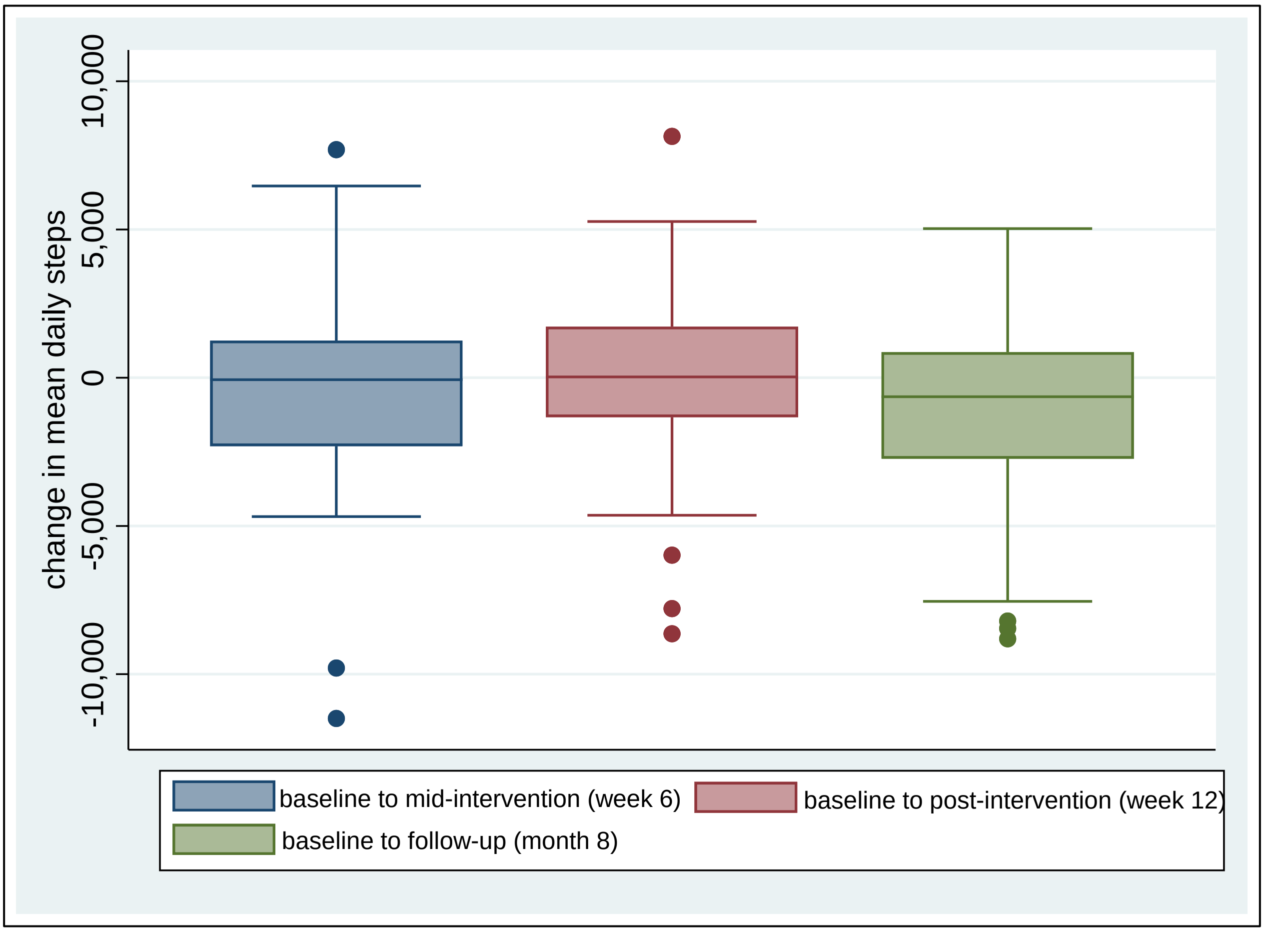
Note: SD = Standard Deviation

[1] Mean daily steps were calculated for participants who had worn the Fitbit on ≥5 of the previous 7 days including at least one weekend day.

5.1.3 Step count

Figure 12 shows the distribution of changes in mean daily steps at each time point for participants overall.

Figure 12 Box plot of change in mean daily steps from baseline to mid-intervention (week 6, n = 118), baseline to post-intervention (week 12, n = 114) and baseline to follow-up (month 8, n = 87)



Overall, as summarised in **Table 32** there was no significant change in mean daily step count from baseline to mid-intervention (week 6) or post-intervention (week 12). However, when split into subgroups, there was a significant increase in steps for female participants at week 12 (mean increase of 636 steps/day, 95% confidence interval or CI: +66 to +1,205; p = 0.030) compared with a non-significant change in males (mean decrease of 327 steps/day, 95% CI: -1,047 to +393; p = 0.368). There was a significant increase for participants with a mean daily step count less than 10,000 at baseline, at week 6 (mean increase of 1,041 steps/day, 95% CI: +370 to +1,712; p = 0.003) and week 12 (mean increase of 1,028 steps/day, 95% CI: +417 to +1,639; p = 0.001). Police community support officers (PCSOs) and special constables showed an increase in steps at week 6 (mean increase of 1,294 steps/day, 95% CI: +123 to +2,464; p = 0.032) and week 12 (mean increase of 1,121 steps/day, 95% CI: +124 to +2,118; p = 0.030). Step changes were non-significant for police officers and police staff (p>0.05).

Table 32 Change in mean daily step count baseline to mid-intervention (week 6) and baseline to post-intervention (week 12)

Group	Mean change from baseline (week 0)					
	Mid-intervention (end of individual phase, week 6)			**Post-intervention (end of social phase, week 12)**		
	Mean change (SD) n = number of observations	**95% CI**	**p-value**	**Mean change (SD) n = number of observations**	**95% CI**	**p-value**
All participants	-214 (2,830) n = 118	-730 to +302	0.413	+78 (2,599) n = 114	-404 to +561	0.748
Male	-621 (3,128) n = 70	-1,367 to +125	0.101	-327 (2,927) n = 66	-1,047 to +393	0.368
Female	+380 (2,228) n = 48	-267 to +1,027	0.243	+636 (1,961) n = 48	+66 to +1,205	**0.030**
Age						
Under 40	-412 (2,586) n = 64	-1,058 to +234	0.207	-209 (2,434) n = 63	-822 to +404	0.498
40+ years	+21 (3,103) n = 54	-826 to +868	0.961	+434 (2,773) n = 51	-346 to +1,213	0.270
Baseline activity level						
<10000 steps/day	+1,041 (2,286) n = 47	+370 to +1,712	**0.003**	+1,028 (2,128) n = 49	+417 to +1,639	**0.001**
≥10000 steps/day	-1,045 (2,863) n = 71	-1,723 to -367	**0.003**	-637 (2,706) n = 65	-1,308 to +33	0.062
Occupation						
Police officers	-603 (2,960) n = 78	-1,271 to +64	0.076	-144 (2,750) n = 76	-773 to +484	0.649
PCSOs and special constables	+1,294 (2571) n = 21	+123 to +2,464	**0.032**	+1,121 (2,130) n = 20	+124 to +2,118	**0.030**
Police staff	-282 (1,980) n = 19	-1,237 to +672	0.542	-141 (2,217) n = 18	-1,243 to +962	0.791

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold.

As **Table 33** shows, there was an overall significant reduction in mean daily step count from baseline to 8-month follow-up (mean decrease 888 steps/day, 95% CI:

-1,518 to -258; p = 0.006). However, when split into subgroups, there was no significant reduction in steps for female participants (p = 0.803), for those aged 40 years or older (p = 0.281) or for PCSOs and special constables (p = 0.518) or police staff (p = 0.257). Similar to the mid- and post-intervention findings, participants with a mean baseline daily step count of 10,000 or less showed a significant increase in steps (mean increase 810 steps/day, 95% CI: +115 to +1,506; p = 0.024) from baseline to 8-month follow-up.

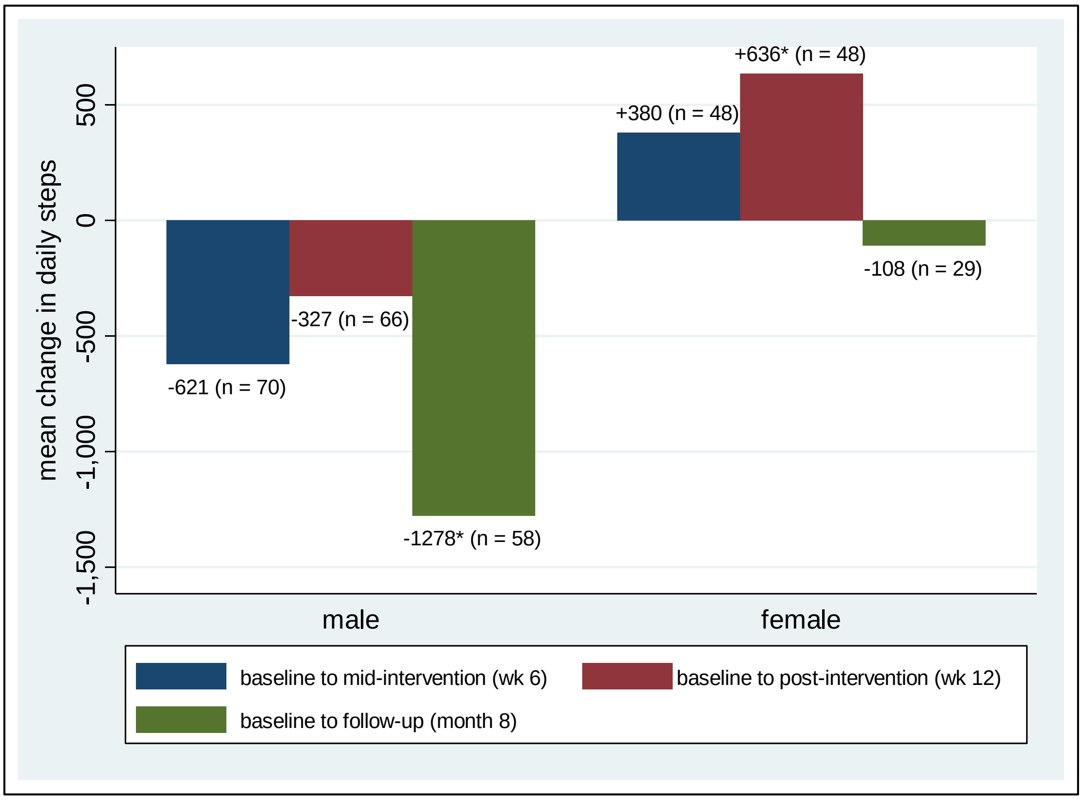
Table 33 Change in mean daily step count baseline to 8-month follow-up

Group	Number of observations	Change baseline to follow-up (8 months)		
		Mean change (SD)	**95% CI**	**p-value**
All participants	87	-888 (2,956)	-1,518 to -258	**0.006**
Male	58	-1,278 (3,174)	-2,113 to -444	**0.003**
Female	29	-108 (2,319)	-990 to +774	0.803
Age				
Under 40	49	-1,152 (2,851)	-1,971 to -333	**0.007**
40+ years	38	-548 (3,090)	-1,564 to +468	0.281
Baseline activity level				
<10,000 steps/day	33	+810 (1,961)	+115 to +1,506	**0.024**
≥10,000 steps/day	54	-1,926 (2,993)	-2,743 to -1,109	**<0.001**
Occupation				
Police officers	61	-1,125 (3,121)	-1,925 to -326	**0.007**
PCSOs and special constables	10	+421 (1,977)	-994 to +1,835	0.518
Police staff	16	-801 (2,721)	-2,251 to +648	0.257

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold

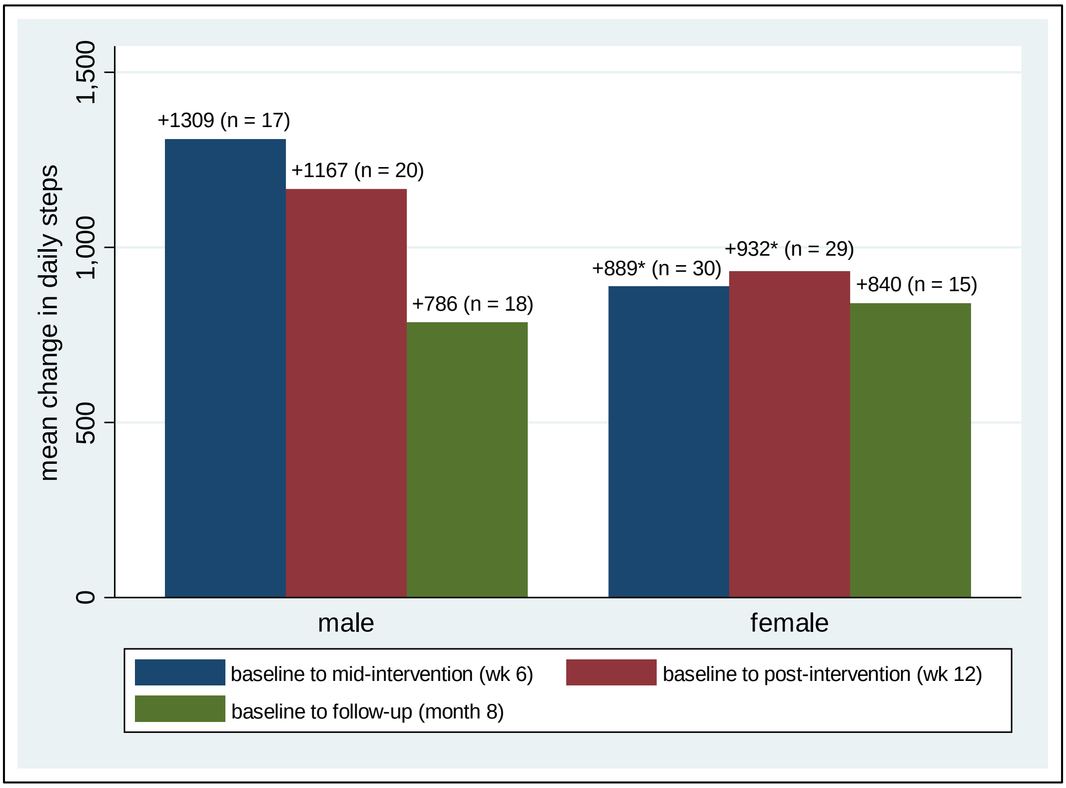
Figure 13 shows the differential changes in mean daily step count that were apparent for male and female participants, including the significant increase in steps for females at week 12 and the significant reduction for males at 8-month follow-up. For participants who were less active at baseline (i.e. mean daily steps <10,000/day), the gender differences were less pronounced, with both males and females showing an increase in steps at all time points (see **Figure 14**).

Figure 13 Changes in mean daily step count by gender (baseline to week 6, baseline to week 12, and baseline to 8-month follow-up) - all participants



* = significant change (i.e. p<0.05)

Figure 14 Changes in mean daily step count by gender (baseline to week 6, baseline to week 12, and baseline to 8-month follow-up) - participants with baseline step count <10,000/day



* = significant change (i.e. p<0.05)

A sensitivity analysis controlling for self-reported events affecting PA level (e.g. illness and annual leave) removed the overall significant reduction in mean daily steps that was observed between baseline and 8-month follow-up (p = 0.126) (see **Table 34**). The changes in mean daily steps from baseline to week 6 and baseline to week 12 remained non-significant (p = 0.467; p = 0.444 respectively).

Table 34 Sensitivity analysis for change in mean daily steps including participants reporting no events affecting physical activity level

Mean change from baseline (week 0)								
Mid-intervention (end of individual phase, week 6)			**Post-intervention (end of social phase, week 12)**			**Follow-up (8 months)**		
Mean change (SD) n = number of observations	**95% CI**	**p-value**	**Mean change (SD) n = number of observations**	**95% CI**	**p-value**	**Mean change (SD) n = number of observations**	**95% CI**	**p-value**
-272 (2,520) n = 46	-1,021 to +476	0.467	+262 (2,561) n = 57	-418 to +941	0.444	-765 (3,178) n = 42	-1,755 to +225	0.126

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold

As described in **Chapter 3**, exploratory regression analysis was carried out to investigate which participant characteristics may be associated with potential intervention effectiveness in both the short and longer term, and whether the previous subgroup findings remained significant when controlling for various predictors. Both linear and logistic regressions were carried out with change in mean daily step count as a continuous and binary (i.e. increase in steps versus decrease or no change) dependent variable respectively. As the results for the two methods were similar only the logistic models are presented here for ease of interpretation. A logistic regression model for the change in steps from baseline to week 12 was developed using forward selection and is shown in **Table 35**. The estimated model had a good fit (log likelihood = -69.82; likelihood ratio (LR) Chi-square = 18.25; p = 0.006). All possible interactions between predictor variables were considered but none were found to be significant.

When controlling for age, years of service, education, gender and police force, participants with a baseline step count of less than 10,000/day had a higher odds of an increase in mean daily step count (OR 2.94, 95% CI: 1.24 to 6.95) than those with a baseline step count of 10,000/day or higher, from baseline to week 12. When controlling for baseline steps and the above predictors, participants aged over 40 years had a higher odds than younger participants of an increase in steps during this period (OR 2.83, 95% CI: 1.11 to 7.24).

Table 35 Logistic regression model for increase in mean daily step count from baseline to week 12 (post-intervention) (n = 114)

Variable	Odds ratio	Standard error	p-value	95% CI
low baseline steps (i.e. <10,000/day)	2.94	1.29	**0.014**	1.24 to 6.95
age over 40 years	2.83	1.36	**0.030**	1.11 to 7.24
less than 10 years police force service	2.48	1.22	0.065	0.94 to 6.49
professional or higher education	0.48	0.21	0.087	0.21 to 1.11
male gender	0.81	0.35	0.632	0.35 to 1.91
employed by Devon & Cornwall Police Force	0.77	0.36	0.573	0.31 to 1.91
constant[1]	0.65	0.42	0.504	0.19 to 2.28

Note: 95% CI = 95% Confidence Interval. p-values where significant (i.e. <0.05) are highlighted in bold.
Variables are presented in the order they were entered into the regression model (forward selection).
All predictor variables are binary (variables in the table are coded as 1).
For overall model: log likelihood = -69.82; likelihood ratio (LR) Chi-square (6 degrees of freedom, d.f.) = 18.25; **p = 0.006**.
[1] expected value when all predictor variables equal zero

A logistic regression model for the change in steps from baseline to 8-month follow-up was developed in the same way and is presented in **Table 36**. The estimated model had a good fit (log likelihood = -42.87; likelihood ratio (LR) Chi-square = 29.74; p<0.001).

Table 36 Logistic regression model for increase in mean daily step count from baseline to 8-month follow-up (n = 87)

Variable	Odds ratio	Standard error	p-value	95% CI
low baseline steps (i.e. <10,000/day)	6.23	3.52	**0.001**	2.06 to 18.87
less than 10 years police force service	5.59	3.78	**0.011**	1.49 to 21.04
no previous activity monitor use (prior to study)	2.68	1.63	0.106	0.81 to 8.84
male gender	0.50	0.31	0.260	0.15 to 1.67
age over 40 years	2.64	1.71	0.136	0.74 to 9.42
professional or higher education	0.42	0.24	0.130	0.13 to 1.29
employed by Devon & Cornwall Police Force	0.41	0.26	0.155	0.12 to 1.40
constant[1]	0.25	0.24	0.158	0.04 to 1.72

Note: 95% CI = 95% Confidence Interval. p-values where significant (i.e. <0.05) are highlighted in bold.
Variables are presented in the order they were entered into the regression model (forward selection).
All predictor variables are binary (variables in the table are coded as 1).
For overall model: log likelihood = -42.87; likelihood ratio (LR) Chi-square (7 degrees of freedom, d.f.) = 29.74; **p < 0.001**.
[1] expected value when all predictor variables equal zero

From baseline to 8-month follow-up, when controlling for age, gender, years of service, education, police force and previous activity monitor use, participants with a baseline step count of less than 10,000/day had a higher odds of an increase in mean daily step count (OR 6.23, 95% CI: 2.06 to 18.87). When controlling for baseline steps and the above predictors, those with less than 10 years of police service had a higher odds of an increase in steps compared with those employed for 10 or more years (OR 5.59, 95% CI: 1.49 to 21.04).

When baseline steps were controlled for, gender became a non-significant predictor of an increase in steps from baseline to week 12, and from baseline to 8-month follow-up. Occupation was also a non-significant predictor of an increase in steps at each time point when controlling for baseline steps.

5.1.4 Self-reported physical activity and sedentary time (IPAQ outcomes)

Table 37 shows the mean changes in self-reported PA and sedentary time from baseline to mid-intervention (week 6) and baseline to post-intervention (week 12), as assessed by the International Physical Activity Questionnaire (IPAQ). From baseline to week 6, there was a significant overall increase in self-reported total PA (mean increase of 27.8 minutes/week (95% CI: +10.9 to +44.7; p = 0.001) or 460.3 MET-minutes/week (95% CI: +71.3 to +849.3; p = 0.021)). There was also a significant increase in moderate-to-vigorous physical activity (MVPA), with a mean increase of 271.9 MET-minutes/week (95% CI: +6.3 to +537.6; p = 0.045) from baseline to week 6. From baseline to week 12, there was a significant overall increase in self-reported total PA (mean increase of 22.7 minutes/week (95% CI: +4.8 to +40.6; p = 0.013) or 465.4 MET-minutes/week (95% CI: +106.7 to +824.1; p = 0.011)) and MVPA (mean increase of 402.9 MET-minutes/week (95% CI: +129.9 to +676.0; p = 0.004)). There were no significant overall changes in weekday sedentary time (all p-values >0.05).

Table 37 Change in self-reported physical activity (PA) and sedentary time baseline to mid-intervention (week 6) and baseline to post-intervention (week 12) - all participants

Outcome	Mean change from baseline (week 0)					
	Mid-intervention (end of individual phase, week 6)			**Post-intervention (end of social phase, week 12)**		
	Mean change (SD) n = number of observations	**95% CI**	**p-value**	**Mean change (SD) n = number of observations**	**95% CI**	**p-value**
Total PA (minutes/week)	+27.8 (107.1) n = 157	+10.9 to +44.7	**0.001**	+22.7 (111.2) n = 151	+4.8 to +40.6	**0.013**
Total PA (MET-minutes/week)	+460.3 (2,467.5) n = 157	+71.3 to +849.3	**0.021**	+465.4 (2,230.9) n = 151	+106.7 to +824.1	**0.011**
MVPA (MET-minutes/week)	+271.9 (1,685.1) n = 157	+6.3 to +537.6	**0.045**	+402.9 (1,698.0) n = 151	+129.9 to +676.0	**0.004**
Sedentary time (hours on a typical weekday)	-0.12 (3.36) n = 157	-0.65 to +0.41	0.651	-0.03 (3.45) n = 151	-0.59 to +0.52	0.906

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold

From baseline to 8-month follow-up, there was a near significant increase in total PA (mean increase 18.6 minutes/week, 95% CI: -0.1 to +37.2; p = 0.052) and a significant increase in MVPA (mean increase 420.5 MET-minutes/week, 95% CI: +56.4 to +784.6; p = 0.024) (see **Table 38**). Weekday sedentary time decreased slightly at each time point but the changes were non-significant (all p-values >0.05).

Table 38 Change in self-reported physical activity (PA) and sedentary time baseline to 8-month follow-up - all participants

Outcome	Number of observations	Change baseline to follow-up (8 months) Mean change (SD)	95% CI	p-value
Total PA (minutes/week)	143	+18.6 (113.0)	-0.1 to +37.2	p = 0.052
Total PA (MET-minutes/week)	143	+317.7 (2,522.0)	-99.2 to +734.6	p = 0.134
MVPA (MET-minutes/week)	143	+420.5 (2,202.6)	+56.4 to +784.6	**p = 0.024**
Sedentary time (hours on a typical weekday)	143	-0.24 (2.99)	-0.73 to +0.26	p = 0.344

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold

For the less active participants (defined as those with a mean daily step count of less than 10,000 at baseline), there was a significant increase in all of the self-reported measures of PA from baseline to week 6, including total PA (mean increase 39.2 minutes/week (95% CI: +16.0 to +62.3; p = 0.001) or 799.3 MET-minutes/week (95% CI: +3,21.1 to +1,277.5; p = 0.001)) and MVPA (mean increase 497.2 MET-minutes/week (95% CI: +158.8 to +835.7; p = 0.005)). Larger effects were seen from baseline to week 12, with a mean increase in total PA of 49.6 minutes/week (95% CI: +24.8 to +74.4; p<0.001) or 1,093.8 MET-minutes/week (95% CI: +598.8 to +1,588.8; p<0.001) and a mean increase in MVPA of 795.5 MET-minutes/week (95% CI: +393.3 to +1,197.7; p<0.001) (see **Table 39**).

The increases in activity remained significant for this group at 8-month follow-up; a mean increase from baseline in total PA of 41.6 minutes/week (95% CI: +15.9 to 67.3; p = 0.002) or 1,067.6 MET-minutes/week (95% CI: +457.7 to 1,677.5; p<0.001) was reported, with a mean increase in MVPA of 839.7 MET-minutes/week (95% CI: +290.2 to 1,389.1; p = 0.003) (see **Table 40**).

For this group of participants, there were no significant changes in sedentary time from baseline to week 6 (p = 0.064), baseline to week 12 (p = 0.961) or baseline to 8-month follow-up (p = 0.860).

Table 39 Change in self-reported physical activity (PA) and sedentary time baseline to mid-intervention (week 6) and baseline to post-intervention (week 12) - participants with baseline step count <10,000/day

Outcome	Mean change from baseline (week 0)					
	Mid-intervention (end of individual phase, week 6)			Post-intervention (end of social phase, week 12)		
	Mean change (SD) n = number of observations	**95% CI**	**p-value**	**Mean change (SD) n = number of observations**	**95% CI**	**p-value**
Total PA (minutes/week)	+39.2 (93.5) n = 65	+16.0 to +62.3	**0.001**	+49.6 (97.8) n = 62	+24.8 to +74.4	**<0.001**
Total PA (MET-minutes/week)	+799.3 (1,930.0) n = 65	+3,21.1 to +1,277.5	**0.001**	+1,093.8 (1,949.3) n = 62	+598.8 to +1,588.8	**<0.001**
MVPA (MET-minutes/week)	+497.2 (1,365.8) n = 65	+158.8 to +835.7	**0.005**	+795.5 (1,583.7) n = 62	+393.3 to +1,197.7	**<0.001**
Sedentary time (hours on a typical weekday)	+0.90 (3.83) n = 65	-0.05 to +1.85	0.064	-0.02 (3.73) n = 62	-0.97 to +0.93	0.961

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold

Table 40 Change in self-reported physical activity (PA) and sedentary time baseline to 8-month follow-up - participants with baseline step count <10,000/day

Outcome	Number of observations	Change baseline to follow-up (8 months)		
		Mean change (SD)	**95% CI**	**p-value**
Total PA (minutes/week)	60	+41.6 (99.6)	+15.9 to 67.3	**0.002**
Total PA (MET-minutes/week)	60	+1,067.6 (2361.0)	+457.7 to 1,677.5	**<0.001**
MVPA (MET-minutes/week)	60	+839.7 (2126.8)	+290.2 to 1,389.1	**0.003**
Sedentary time (hours on a typical weekday)	60	-0.07 (3.17)	-0.89 to +0.75	0.860

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold

5.1.5 Components of the intervention (individual versus social)

In this section, the 'individual' (weeks 1 to 6) and 'social' (weeks 7 to 12) phases of the study are compared to explore their relative impact on PA and sedentary time. The two phases were characterised by the use of different components of the intervention, i.e. features of the Fitbit® (and its associated app) and the Bupa Boost app. In brief, the main components of the individual phase were goal-setting, self-monitoring and feedback within the Fitbit® and Bupa Boost apps, and earning points and obtaining virtual rewards within Bupa Boost. The social phase additionally included linking with colleagues for social comparison, individual and team competitions, and social support within the Bupa Boost app. **Chapter 3** contains a full description of the intervention and its associated behaviour change techniques (BCTs).

As shown in **Table 41** and **Table 42** there were no significant differences in apparent impact of the intervention between the individual and social phases of the study, for mean daily step count ($p = 0.433$), self-reported total weekly minutes ($p = 0.786$) or MET-minutes ($p = 0.738$) of PA, weekly MET-minutes of MVPA ($p = 0.354$), or weekday sedentary time ($p = 0.604$). No significant subgroup differences were observed between the two phases of the intervention by gender, age group, baseline activity level or occupation. However, although non-significant, most groups showed a higher mean increase (or lower decrease) in steps during the social phase relative to the individual phase. The only exception to this was the police staff group, who showed a mean reduction of 554 steps in the social phase relative to the individual phase. It is difficult to draw definitive conclusions regarding occupational subgroup differences however, due to the small numbers in the police staff ($n = 14$) and PCSO/special constable ($n = 16$) groups providing step data in both phases.

Table 41 Comparison of individual and social phase: change in mean daily step count

Group	Number of observations[1]	Mean change individual phase (baseline to week 6) (SD)	Mean change social phase (baseline to week 12) (SD)	Mean difference individual vs. social phase[2] (SD)	95% CI for difference	p-value for difference
All participants	95	-284 (2,685)	-68 (2,679)	216 (2,666)	-328 to +759	0.433
Male	54	-797 (2,824)	-599 (3,042)	198 (3,003)	-621 to +1018	0.630
Female	41	+393 (2,357)	+631 (1,929)	238 (2,179)	-449 to +926	0.488
Age						
Under 40	51	-500 (2,740)	-392 (2,506)	107 (2,673)	-644 to +859	0.775
40+ years	44	-33 (2,629)	+308 (2,849)	341 (2,683)	-475 to +1,156	0.404
Baseline activity level						
<10,000 steps/day	38	+1,033 (2,267)	+1,072 (368)	39 (2,675)	-841 to +918	0.929
≥10,000 steps/day	57	-1,161 (2,598)	-828 (2,679)	333 (2,677)	-377 to +1,044	0.351
Occupation						
Police officers	65	-580 (2,713)	-263 (2,913)	318 (2,779)	-371 to +1,006	0.360
PCSOs and special constables	16	+940 (2,776)	+1,414 (1,882)	474 (2,990)	-1,120 to +2,067	0.536
Police staff	14	-305 (2,176)	-859 (1,565)	-554 (1,512)	-1,427 to +319	0.194

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval
p-values where significant (i.e. <0.05) are highlighted in bold.
1 Complete case analysis (i.e. only participants with outcome data at baseline, 6 weeks and 12 weeks included)
2 Positive values indicate higher relative increase (or lower decrease) in steps during social phase

Table 42 Comparison of individual and social phase: change in self-reported physical activity (PA) and sedentary time (all participants)

Outcome	Number of observations[1]	Mean change individual phase (baseline to week 6) (SD)	Mean change social phase (baseline to week 12) (SD)	Mean difference individual vs. social phase[2] (SD)	95% CI for difference	p-value for diff.
Total PA (minutes/ week)	143	+23.6 (108.6)	+20.9 (113.3)	-2.7 (118.7)	-22.3 to +16.9	0.786
Total PA (MET-minutes/ week)	143	+363.7 (2,493.6)	+434.0 (2,274.3)	+70.3 (2,507.7)	-344.3 to +484.8	0.738
MVPA (MET-minutes/ week)	143	+231.7 (1,720.8)	+371.0 (1,722.3)	+139.4 (1,790.8)	-156.7 to +435.4	0.354
Sedentary time (hours on a typical weekday)	143	-0.06 (3.38)	+0.08 (3.47)	+0.14 (3.23)	-0.39 to +0.67	0.604

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval
p-values where significant (i.e. <0.05) are highlighted in bold.
[1] Complete case analysis (i.e. only participants with outcome data at baseline, 6 weeks and 12 weeks included)
[2] Positive values indicate higher relative increase (or lower decrease) in outcome during social phase

5.1.6 Correlations between changes in steps, self-reported physical activity and sedentary time

As **Table 43** shows, from baseline to week 12 (post-intervention), there was a moderate significant positive correlation between change in mean daily step count and change in self-reported total PA in weekly minutes (r = 0.30; p<0.01) and weekly MET-minutes (r = 0.35; p<0.001). There was also a moderate significant positive correlation between change in mean daily step count and change in MVPA (r = 0.30; p<0.01). There was no significant correlation between change in self-reported weekday sedentary time and changes in steps or any of the self-reported PA outcomes.

Table 43 Correlations between changes in steps, self-reported physical activity (PA) and sedentary time baseline to post-intervention (week 12) - all participants

	Mean daily steps	Total PA (weekly minutes)	Total PA (weekly MET-minutes)	MVPA (weekly MET-minutes)	Weekday sedentary time (hours per day)
Mean daily steps	1.00 n = 114				
Total PA (weekly minutes)	0.30** n = 114	1.00 n = 151			
Total PA (weekly MET-minutes)	0.35*** n = 114	0.80*** n = 151	1.00 n = 151		
MVPA (weekly MET-minutes)	0.30** n = 114	0.67*** n = 151	0.85*** n = 151	1.00 n = 151	
Weekday sedentary time (hours per day)	0.005 n = 114	-0.11 n = 151	-0.04 n = 151	0.01 n = 151	1.00 n = 151

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*p<0.05 **p<0.01 ***p<0.001

From baseline to 8-month follow-up (see **Table 44**), there was a slightly stronger positive correlation between change in mean daily step count and changes in self-reported total PA in weekly minutes (r = 0.41; p<0.001) and weekly MET-minutes (r = 0.48; p<0.001), and between change in mean daily step count and change in MVPA (r = 0.41; p<0.001). During this period, there was also a significant negative correlation between change in self-reported weekday sedentary time and changes in total PA in weekly minutes (r = -0.28; p<0.001) and weekly MET-minutes (r = -0.20; p<0.05), i.e. as PA increased sedentary time was reduced. The correlation between change in sedentary time and change in MVPA neared significance (r = -0.16, p = 0.064). There was no significant correlation between changes in sedentary time and steps (r = -0.13; p = 0.241).

Table 44 Correlations between changes in steps, self-reported physical activity (PA) and sedentary time baseline to 8-month follow-up - all participants

	Mean daily steps	**Total PA (weekly minutes)**	**Total PA (weekly MET-minutes)**	**MVPA (weekly MET-minutes)**	**Weekday sedentary time (hours per day)**
Mean daily steps	1.00 n = 87				
Total PA (weekly minutes)	0.41*** n = 87	1.00 n = 143			
Total PA (weekly MET-minutes)	0.48*** n = 87	0.77*** n = 143	1.00 n = 143		
MVPA (weekly MET-minutes)	0.41*** n = 87	0.66*** n = 143	0.86*** n = 143	1.00 n = 143	
Weekday sedentary time (hours per day)	-0.13 n = 87	-0.28*** n = 143	-0.20* n = 143	-0.16 n = 143	1.00 n = 143

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*****p<0.05 ******p<0.01 *******p<0.001

5.1.7 Summary / Discussion

Figure 15 and **Figure 16** illustrate the statistical significance of changes in the primary outcomes at the various time points of the study.

Figure 15 Summary of significance of changes in primary outcomes at various time points: all participants

Outcome	Mean change from baseline (week 0)		
	Mid-intervention (week 6)	Post-intervention (week 12)	Follow-up (month 8)
Step count (mean steps/day)	0	0	- [1]
Total PA (minutes/week)	+	+	0
Total PA (MET-minutes/week)	+	+	0
MVPA (MET-minutes/week)	+	+	+
Sedentary time (hours on a typical weekday)	0	0	0

+ = positive impact

- = negative impact

0 = no significant impact

Note: [1] No impact with sensitivity analysis

Figure 16 Summary of significance of changes in primary outcomes at various time points: less active participants, i.e. baseline step count <10,000/day

Outcome	Mean change from baseline (week 0)		
	Mid-intervention (week 6)	Post-intervention (week 12)	Follow-up (month 8)
Step count (mean steps/day)	+	+	+
Total PA (minutes/week)	+	+	+
Total PA (MET-minutes/week)	+	+	+
MVPA (MET-minutes/week)	+	+	+
Sedentary time (hours on a typical weekday)	0	0	0

+ = positive impact

- = negative impact

0 = no significant impact

Overall, there was no statistically significant change in mean daily step count from baseline to week 6 (mid-intervention) or week 12 (post-intervention). There was a significant reduction in mean daily step count from baseline to 8-month follow-up; however, in a sensitivity analysis controlling for self-reported external factors affecting participants' PA levels (such as illness and annual leave), the reduction became non-significant. This shows the influence of factors outside of the intervention (such as annual leave, the weather and seasonal factors) causing short- and long-term fluctuations in activity; these are discussed in depth in **Chapter 4**. It is therefore important to capture (and control for where possible) data on such factors in future studies.

A reactivity effect is one potential explanation for the lack of a significant change in steps. Reactivity is defined as a change in behaviour as a result of being monitored (Matevey et al., 2006). It should be considered that participants in the PAW-Force study might have consciously or subconsciously increased their PA level at baseline in response to wearing the Fitbit®. However, studies have found little evidence for this effect when 'sealed' devices with no visual feedback are used (Matevey et al., 2006, Tudor-Locke et al., 2011a). Reactivity is therefore unlikely to have had a large impact on baseline activity levels. It appears that overall participants maintained, rather than increased, their daily steps through the course of the study, and the qualitative findings supported this observation (see **Chapter 6**).

In contrast to the findings for step count, self-reported PA showed an overall increase throughout the study. There was a significant mean increase in total weekly PA of 22.7 minutes or 465.4 MET-minutes, and a significant mean increase in MVPA of 402.9 MET-minutes/week, from baseline to week 12. These changes were largely maintained to 8-month follow-up, with a mean (near significant) increase from baseline in total weekly PA of 18.6 minutes or 317.7 MET-minutes, and a significant mean increase in MVPA of 420.5 MET-minutes/week. The findings suggest that many participants might have increased their activity in ways that were not captured in Fitbit-recorded steps, such as strength training, cycling or swimming. This is discussed further in the light of the qualitative findings in **Chapter 6**. The possibility that these results may in part be a consequence of social desirability bias arising from the self-reported nature of the outcome assessment tool should be considered. However, there is little evidence that this constitutes a significant problem with self-reported PA, particularly for surveys completed online. For example, a large longitudinal study (with a representative sample of the Dutch population) investigated social desirability with the web-based short-form IPAQ, and found no association between self-reported PA or sedentary time and social desirability as assessed by the Marlowe-Crowne Scale (Crutzen and Göritz, 2011).

There was an indication of a small but non-significant reduction in self-reported weekday sedentary time from baseline to week 6, week 12 and 8-month follow-up. At the 8-month follow-up, the mean reduction in sedentary time from baseline was 0.24 hours (approximately 15 minutes/day). Some similar studies that aimed to reduce sedentary time in the workplace using mHealth have reported increases in sedentary time in intervention participants relative to controls (Brakenridge et al., 2016b, Jones, 2016). While it is reassuring that there were no significant increases in sedentary time over the course of this study, the intervention appeared to have a limited impact on sedentary behaviour (SB). These findings are discussed with reference to the qualitative findings (and the Socio-Ecological Model) in **Chapter 6**.

There was evidence of subgroup differences in changes in mean daily step count throughout the study, with the strongest predictor being baseline activity level. Participants with a baseline step count of less than 10,000/day appeared to benefit most, with a significant mean increase from baseline to week 12 of 1,028 steps/day. This group also showed the greatest increases in self-reported PA, with a mean increase in total weekly PA of 49.6 minutes or 1,093.8 MET-minutes, and a mean increase of 795.5 MET-minutes/week of MVPA from baseline to week 12. On the whole, these changes were maintained for this group at the 8-month follow-up, with significant mean increases from baseline of 810 steps/day, 41.6 minutes or 1,067.6 MET-minutes/week of total PA and 839.7 MET-minutes/week of MVPA. While the aim of the study was not to precisely estimate effect sizes, the increases in steps were similar in magnitude to those reported by three previous randomised trials of workplace mHealth interventions, where the post-intervention between-group differences in mean daily steps were around 850 (Brakenridge et al., 2016b), 970 (Poirier et al., 2016) and 2,000 (Gremaud et al., 2018) respectively.

It is important to consider the clinical significance of the findings in addition to the statistical significance. Increasing daily steps has been shown to have important clinical benefits, including reduced morbidity and mortality. Benefits reported by longitudinal cohort studies include improved insulin sensitivity and reduced adiposity (Dwyer et al., 2011) and a reduced risk of cardiovascular events (Yates et al., 2014). A large prospective cohort study of Australian adults by Dwyer and colleagues (2015) found a linear relationship between changes in daily steps and all-cause mortality. According to this study, increases of 1,000 steps per day (i.e. similar in magnitude to the changes shown by the less active participants in the PAW-Force study) are associated with a significantly reduced risk of mortality of approximately 6% (adjusted hazard ratio 0.94; 95% CI: 0.90 to 0.98) (Dwyer et al., 2015).

The finding that those who are most in need of the intervention may benefit most is promising, particularly given that the greatest health gains from increasing PA are realised by the least active section of the population (Blair and Connelly, 1996, Arem et al., 2015). There is strong evidence for a curvilinear relationship between PA and health outcomes (Warburton and Bredin, 2017). Even small increases in MVPA (to levels lower than recommended by the public health guidelines) have been shown to have clinically relevant health benefits, including reduced morbidity and mortality from cardiovascular disease, cancer and other chronic conditions (Arem et al., 2015, Warburton and Bredin, 2017). For example, compared with individuals that report no leisure time PA, a 20% lower overall mortality risk has been observed in those performing less than the public health recommended minimum, with a 31% lower mortality risk at one to two times the recommended minimum (Arem et al., 2015). The relatively large increases in PA levels of the less active officers and staff in the PAW-Force study that were sustained after 8 months (e.g. mean increase in total weekly PA of approximately 42 minutes/week from baseline to month 8) are therefore likely to confer significant health benefits.

Similar to the PAW-Force study, other mHealth studies have found that less active subgroups have shown the largest increases in PA. For example, an activity monitor and app intervention was associated with significant improvements in PA levels only in those who did not meet the Centers for Disease Control and Prevention recommended weekly PA level at baseline (Schrager et al., 2017). Similarly, a study of the ‘Pokemon GO’ app reported that the largest increases in daily steps were observed in those with the lowest baseline activity levels (Xian et al., 2017).

According to these findings, less active subgroups should be targeted in future interventions.

It appeared that the intervention had a differential impact on activity levels of male and female participants in both the short and longer term; for example only females showed a significant increase in mean daily steps from baseline to week 12 while only males showed a significant reduction from baseline to 8-month follow-up. However, these differences seemed to be largely driven by gender differences in baseline activity level (see **Chapter 4**). Similarly, there initially appeared to be a differential impact based on occupation, with only the PCSO and special constable group showing significant increases in mean daily steps from baseline to week 12; again, this effect became non-significant when controlling for steps at baseline.

In addition to baseline activity level, age over 40 years was a significant predictor of an increase in steps in the short term (baseline to week 12). This suggests that older participants may initially benefit most from the intervention in this context. Having less than 10 years of police force service also predicted an increase in steps in the longer term (baseline to 8-month follow-up). There were no significant differences in step changes by police force or occupation when controlling for baseline activity level and other predictors such as age, gender, education, years of police service and previous activity monitor use. Potential subgroup differences in the impact of mHealth interventions for PA are poorly understood, one reason for this being insufficient sample sizes (e.g. King et al., 2016). The findings that less active officers and staff, older members of the police force, and those with fewer years of policing service may benefit most from this type of intervention are novel and worthy of further investigation. Future effectiveness trials of mHealth interventions should explore further the magnitude of impact on PA levels for different subgroups.

There were no significant differences between the individual and social phases of the study in step count, self-reported PA or sedentary time. The lack of a significant difference in activity between the two phases does not provide support for social theories of behaviour change and contradicts the findings of some similar studies of mHealth interventions. For example, a randomised controlled trial by King and colleagues found a socially framed app (including social support, normative feedback, modelling, group-based collaboration and competition) was associated with greater increases in MVPA and greater reductions in sedentary time than an analytically framed app and an affectively framed app with no social features (King et al., 2016). In contrast, a trial by Harries and colleagues (Harries et al., 2016) found that there was no significant difference in daily steps in an app-based intervention between an individual and a social feedback group. Some evidence suggests that social influence may be more important for sustaining longer-term engagement with mHealth technology (Hamari and Koivisto, 2015, Du et al., 2016) and subsequent behaviour change (Hamari and Koivisto, 2015, Fitzgerald and McClelland, 2017). This might explain the lack of a difference in activity between the individual and social phases during the initial 12 week intervention period in the present study. An alternative explanation may be low engagement with the Bupa Boost app. This will be discussed further in relation to findings on preferred phase and integrated with qualitative data in **Chapter 6**. There was no evidence of subgroup differences in the relative impact of the individual and social phases, by age, gender, occupation or baseline activity level.

Changes in the PA outcomes (objectively recorded daily step count and self-reported total weekly PA and weekly MVPA) were moderately and significantly correlated throughout the study. This suggests that in general, participants increased their daily steps and other types of activity concurrently. This finding also adds support to the validity and reliability of self-reported measures of PA, such as the IPAQ, when compared with a validated objective measure of activity. Self-reported measures may have higher completeness than objective measures; this is discussed later in **Chapter 7**. Changes in self-reported weekday sedentary time were not significantly correlated with changes in objective step count or self-reported PA from baseline to week 12 (post-intervention). However, there was a significant negative correlation between changes in sedentary time and self-reported total PA from baseline to 8-month follow-up, i.e. as PA increased, sedentary time was reduced, and vice versa. This suggests that the two behaviours are related (particularly in the longer term) and shows potential to address both PA and SB in a single workplace intervention. This has been achieved by a small number of existing mHealth intervention studies such as an international, multi-organisational cohort study of the 'Stepathlon' app by Ganesan and colleagues; the researchers found that after 100 days of use, the app successfully led to increased PA and reductions in sitting duration (Ganesan et al., 2016).

These results should be interpreted with caution as the study was an uncontrolled pilot, which was not designed to provide definitive effect sizes or to assess efficacy. The use of multiple hypothesis testing should also be mentioned as a potential limitation; this may be associated with the potential for false positives or type 1 errors (Sedgwick, 2014). Nevertheless, the findings are valuable in that they indicate the overall potential beneficial impact of the intervention in both the short and longer term, and suggest who may benefit most (i.e. those who are less active, and potentially also older officers and staff, and those with fewer years of policing service). While it may be argued that regression to the mean might have contributed to the observed increases in daily step count amongst the less active participants, this is not believed to be a major issue. The consistent results found across multiple time points, multiple outcomes (i.e. both objective and self-reported PA), and agreement with the qualitative findings (discussed in **Chapter 6**) increase our confidence in the validity of the results. The validity of the findings was further improved by capturing data on factors influencing PA levels. The within-subject approach taken to compare the individual and social phases of the study removed between-subjects variability; alternatives to traditional randomised controlled trials (RCTs) including single case designs (e.g. n-of-1 trials) are increasingly being advocated in mHealth research (Dallery et al., 2013, Pham et al., 2016, McCallum et al., 2018, Arigo et al., 2019).

5.2 Quantitative exploratory analysis of secondary outcomes

5.2.1 Introduction and overview

In this section, the impact of the PAW-Force intervention on secondary outcomes is examined using quantitative data, including health and wellbeing, perceived stress, perceived productivity in the workplace and sickness absence. Changes in health and wellbeing, perceived stress and productivity are first presented, and assessed for statistical significance over the various phases of the study. This is followed by an exploration of the associations between changes in steps, self-reported PA and sedentary time, and health, wellbeing and productivity. A pre- and post-intervention comparison of objective sickness absence data for the Plymouth Basic Command Unit (BCU) site is then presented. The section concludes with a summary and discussion of the key findings with reference to the relevant literature. Subgroup analyses are included throughout this section.

5.2.2 Health and wellbeing, stress and productivity

The absolute mean values for the self-reported secondary outcomes of physical and mental health-related quality of life (as assessed by the SF-12 Health Survey), perceived stress (as assessed by the Perceived Stress Scale, PSS-4) and productivity (assessed by the absenteeism and presenteeism questions of the Health and Work Performance Questionnaire, HPQ) at each time point are shown in **Table 45**.

Table 45 Self-reported secondary outcomes - absolute values at four time points (all participants, all available data)

Outcome	Mean (SD) n = number of observations Baseline (week 0)	Mid-intervention (week 6)	Post-intervention (week 12)	Follow-up (month 8)
SF-12 Physical Component Score (PCS)[1]	54.25 (6.78) n = 172	53.98 (7.05) n = 148	54.26 (6.22) n = 146	53.85 (7.07) n = 139
SF-12 Mental Component Score (MCS)[1]	47.56 (9.26) n = 172	48.10 (9.23) n = 148	48.46 (10.00) n = 146	49.72 (7.96) n = 139
Perceived Stress Scale (PSS-4) score[2]	4.85 (3.19) n = 172	4.75 (3.07) n = 147	4.63 (3.17) n = 144	4.24 (3.05) n = 139
Health and Work Performance Questionnaire (HPQ) scores:				
Absolute absenteeism (hours lost/month)	6.56 (39.61) n = 174	12.21 (39.19) n = 148	16.21 (56.81) n = 146	7.89 (36.17) n = 139
Relative absenteeism[3]	0.03 (0.21) n = 174	0.07 (0.33) n = 148	0.04 (0.78) n = 146	-0.003 (0.68) n = 139
Absolute presenteeism[4]	78.62 (13.18) n = 174	77.43 (14.34) n = 148	78.36 (14.39) n = 146	78.49 (13.02) n = 139
Relative presenteeism[5]	1.10 (0.26) n = 174	1.08 (0.31) n = 148	1.07 (0.28) n = 146	1.09 (0.40) n = 139
Combined productivity score[6] (relative absenteeism and relative presenteeism)	1.08 (0.36) n = 174	1.02 (0.42) n = 148	1.06 (0.92) n = 146	1.10 (0.76) n = 139

Note: SD = Standard Deviation
[1] Higher scores indicate higher quality of life; minimum possible score 0 and maximum possible score 100.
[2] Higher scores indicate higher perceived stress; minimum possible score 0 and maximum possible score 16.
[3] Higher score = higher absenteeism, relative to expected hours (negative score = worked more than expected)
[4] Higher score = lower lost performance, i.e. higher productivity
[5] Higher score = higher performance or productivity relative to colleagues
[6] Higher score = higher productivity

Overall, there were no statistically significant changes in physical or mental health-related quality of life, perceived stress, or any of the HPQ outcomes (absenteeism, presenteeism and combined productivity score) from baseline to week 6 (mid-intervention) or baseline to week 12 (post-intervention) (see **Table 46**).

From baseline to 8-month follow-up, there was a significant improvement in mental health-related quality of life (mean increase in SF-12 Mental Component Score or MCS 1.75 points, 95% CI: +0.28 to +3.23; p = 0.020) (see **Table 47**). Changes in physical health-related quality of life, perceived stress and all of the HPQ outcomes (absenteeism, presenteeism and combined productivity score) were non-significant (all p-values >0.05).

Table 46 Change in self-reported secondary outcomes baseline to mid-intervention (week 6) and baseline to post-intervention (week 12) - all participants

Outcome	Mean change from baseline (week 0)					
	Mid-intervention (end of individual phase, week 6)			Post-intervention (end of social phase, week 12)		
	Mean change (SD) n = number of observations	**95% CI**	**p-value**	**Mean change (SD) n = number of observations**	**95% CI**	**p-value**
SF-12 Physical Component Score (PCS)[1]	-0.02 (6.75) n = 145	-1.13 to +1.09	0.966	+0.15 (6.76) n = 144	-0.96 to +1.27	0.786
SF-12 Mental Component Score (MCS)[1]	+0.08 (7.21) n = 145	-1.11 to +1.26	0.897	+0.26 (8.91) n = 144	-1.21 to +1.72	0.730
Perceived Stress Scale (PSS-4) score[2]	+0.02 (2.59) n = 144	-0.41 to +0.45	0.923	-0.11 (2.80) n = 142	-0.57 to +0.36	0.654
Health and Work Performance Questionnaire (HPQ) scores:						
Absolute absenteeism (hours lost/month)	+3.72 (54.84) n = 147	-5.21 to +12.66	0.412	+8.51 (67.31) n = 146	-2.50 to +19.52	0.129
Relative absenteeism[3]	+0.03 (0.39) n = 147	-0.04 to +0.09	0.431	+0.00 (0.79) n = 146	-0.13 to +0.13	0.997
Absolute presenteeism[4]	-1.16 (12.36) n = 147	-3.17 to +0.86	0.259	-0.48 (15.19) n = 146	-2.96 to +2.01	0.704
Relative presenteeism[5]	-0.02 (0.33) n = 147	-0.07 to +0.03	0.423	-0.03 (0.31) n = 146	-0.09 to +0.02	0.171
Combined productivity score[6] (relative absenteeism and relative presenteeism)	-0.06 (0.52) n = 147	-0.14 to +0.03	0.185	-0.19 (0.96) n = 146	-0.18 to +0.14	0.816

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used.
p-values where significant (i.e. <0.05) are highlighted in bold
[1] Higher scores indicate higher quality of life
[2] Higher scores indicate higher perceived stress
[3] Higher score = higher absenteeism, relative to expected hours
[4] Higher score = lower lost performance, i.e. higher productivity
[5] Higher score = higher performance or productivity relative to colleagues
[6] Higher score = higher productivity

Table 47 Change in self-reported secondary outcomes baseline to 8-month follow-up - all participants

Outcome	Number of observations	Change baseline to follow-up (8 months)		
		Mean change (SD)	95% CI	p-value
SF-12 Physical Component Score (PCS)[1]	137	-0.23 (7.30)	-1.47 to +1.00	0.708
SF-12 Mental Component Score (MCS)[1]	137	+1.75 (8.73)	+0.28 to +3.23	**0.020**
Perceived Stress Scale (PSS-4) score[2]	137	-0.41 (3.12)	-0.94 to +0.12	0.128
Health and Work Performance Questionnaire (HPQ) scores:				
Absolute absenteeism (hours lost/month)	139	-0.07 (49.38)	-8.35 to +8.21	0.987
Relative absenteeism[3]	139	-0.04 (0.69)	-0.15 to +0.08	0.510
Absolute presenteeism[4]	139	-0.43 (14.74)	-2.90 to +2.04	0.730
Relative presenteeism[5]	139	-0.03 (0.39)	-0.09 to +0.04	0.401
Combined productivity score[6] (relative absenteeism and relative presenteeism)	139	+0.01 (0.77)	-0.12 to +0.14	0.896

Note: SD = Standard Deviation; 95% CI = 95% Confidence Interval. Pairwise deletion used
p-values where significant (i.e. <0.05) are highlighted in bold
[1] Higher scores indicate higher quality of life
[2] Higher scores indicate higher perceived stress
[3] Higher score = higher absenteeism, relative to expected hours
[4] Higher score = lower lost performance, i.e. higher productivity
[5] Higher score = higher performance or productivity relative to colleagues
[6] Higher score = higher productivity

5.2.3 Associations between changes in physical activity, sedentary time and secondary outcomes

5.2.3.1 Changes in steps and changes in health and wellbeing and perceived stress

As **Table 48** shows, there were no significant associations between change in mean daily steps and changes in physical or mental health-related quality of life or changes in perceived stress from baseline to week 12 (post-intervention). The SF-12 Physical Component Score (PCS) was negatively correlated with the MCS, i.e. improvements in physical health-related quality of life were associated with a decline in mental health-related quality of life (r = 0.50; p<0.001). The PCS was also weakly but significantly correlated with perceived stress, i.e. improvements in physical health-related quality of life were associated with increasing stress levels (r = 0.19; p<0.05). The MCS was moderately negatively correlated with the perceived stress score, i.e. improvements in mental health were associated with reduced stress levels (r = -0.46; p<0.001).

Table 48 Correlations between change in mean daily steps and changes in health and wellbeing and perceived stress baseline to post-intervention (week 12) - all participants

	Mean daily steps	**SF-12 Physical Component Score (PCS)[1]**	**SF-12 Mental Component Score (MCS)[1]**	**Perceived Stress Scale (PSS-4) score[2]**
Mean daily steps	1.00 n = 114			
SF-12 Physical Component Score (PCS)[1]	0.05 n = 112	1.00 n = 144		
SF-12 Mental Component Score (MCS)[1]	-0.03 n = 112	-0.50*** n = 144	1.00 n = 144	
Perceived Stress Scale (PSS-4) score[2]	0.003 n = 109	0.19* n = 141	-0.46*** n = 141	1.00 n = 142

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*****p<0.05 ******p<0.01 *******p<0.001
[1] Higher scores indicate higher quality of life
[2] Higher scores indicate higher perceived stress

From baseline to 8-month follow-up, there was a weak but significant positive association between change in mean daily steps and change in physical health-related quality of life (r = 0.25; p<0.05) (see **Table 49** and **Figure 17**). The scatterplot (**Figure 17**) shows that participants that reduced their daily steps at follow-up tended to experience poorer physical health (and to some extent, participants who increased their steps experienced improved physical health).

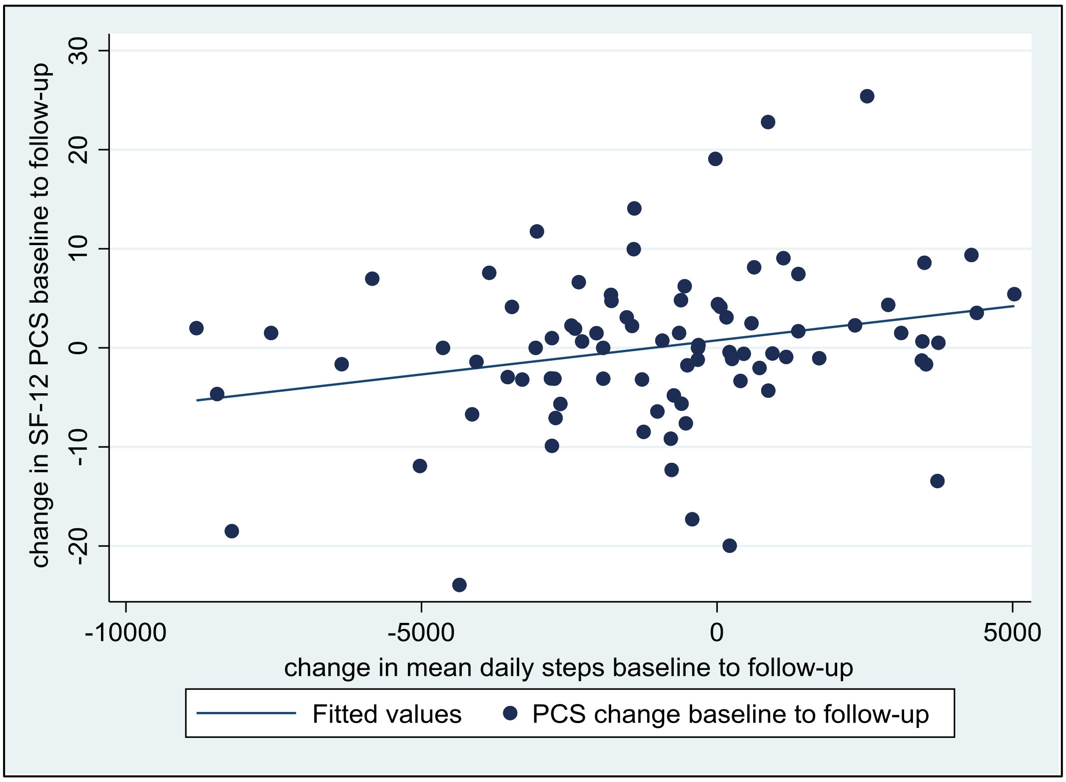
Despite the overall improvement in mental health-related quality of life at the 8-month follow-up, this was not significantly associated with the change in mean daily steps (r = -0.06; p = 0.584). There was no significant association between change in mean daily steps and change in perceived stress from baseline to 8-month follow-up (r = -0.10; p = 0.351). As before, the SF-12 PCS was negatively correlated with the MCS (i.e. improvements in physical health-related quality of life were associated with a decline in mental health-related quality of life) (r = -0.39; p<0.001). The MCS was again moderately negatively correlated with the PSS-4 score, i.e. improvements in mental health were associated with reduced levels of perceived stress (r = -0.49; p<0.001).

Table 49 Correlations between change in mean daily steps and changes in health and wellbeing baseline to 8-month follow-up - all participants

	Mean daily steps	SF-12 Physical Component Score (PCS)[1]	SF-12 Mental Component Score (MCS)[1]	Perceived Stress Scale (PSS-4) score[2]
Mean daily steps	1.00 n = 87			
SF-12 Physical Component Score (PCS)[1]	0.25* n = 86	1.00 n = 137		
SF-12 Mental Component Score (MCS)[1]	-0.06 n = 86	-0.39*** n = 137	1.00 n = 137	
Perceived Stress Scale (PSS-4) score[2]	-0.10 n = 85	0.17 n = 136	-0.49*** n = 136	1.00 n = 137

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*p<0.05 **p<0.01 ***p<0.001
[1] Higher scores indicate higher quality of life
[2] Higher scores indicate higher perceived stress

Figure 17 Scatterplot of changes in mean daily steps and SF-12 Physical Component Score (PCS) baseline to 8-month follow-up



A subgroup analysis with the less active participants (i.e. those with a mean daily step count of less than 10,000 at baseline) found that there were no significant correlations between change in mean daily steps and changes in health and wellbeing or perceived stress. This lack of association was seen from baseline to week 12 and baseline to 8 months (all p-values >0.05).

5.2.3.2 Changes in steps and changes in productivity

There were no significant associations between change in mean daily steps and changes in self-reported absenteeism, presenteeism or productivity (as assessed by the HPQ) from baseline to week 12, or from baseline to 8-month follow-up (see **Table 50** and **Table 51**). The negative correlation between change in steps and change in the combined relative absenteeism and relative presenteeism score from baseline to 8 months neared significance (r = -0.20, p = 0.060), i.e. participants with the greatest increases in steps appeared to show a greater decline in productivity. This association was non-significant for participants with a baseline daily step count of less than 10,000 (r = -0.24; p = 0.177; n = 33).

Table 50 Correlations between change in steps and changes in self-reported absenteeism, presenteeism and productivity baseline to post-intervention (week 12) - all participants

	Mean daily steps	**Absolute absenteeism (hours lost/month)**	**Relative absenteeism[1]**	**Absolute presenteeism[2]**	**Relative presenteeism[3]**	**Combined productivity score[4] (relative absenteeism and relative presenteeism)**
Mean daily steps	1.00 n = 114					
Absolute absenteeism (hours lost/month)	0.003 n = 113	1.00 n = 146				
Relative absenteeism[1]	0.02 n = 113	0.60*** n = 146	1.00 n = 146			
Absolute presenteeism[2]	0.03 n = 113	-0.23** n = 146	-0.13 n = 146	1.00 n = 146		
Relative presenteeism[3]	-0.01 n = 113	-0.23** n = 146	-0.14 n = 146	0.54*** n = 146	1.00 n = 146	
Combined productivity score[4] (relative absenteeism and relative presenteeism)	-0.02 n = 113	-0.58*** n = 146	-0.96*** n = 146	0.20* n = 146	0.393*** n = 146	1.00 n = 146

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*****p<0.05 ******p<0.01 *******p<0.001
Orange shading is used to show correlations of most relevance.
[1] Higher score = higher absenteeism, relative to expected hours
[2] Higher score = lower lost performance, i.e. higher productivity
[3] Higher score = higher performance or productivity relative to colleagues
[4] Higher score = higher productivity

Table 51 Correlations between change in steps and changes in self-reported absenteeism, presenteeism and productivity baseline to 8-month follow-up - all participants

	Mean daily steps	**Absolute absenteeism (hours lost/month)**	**Relative absenteeism[1]**	**Absolute presenteeism[2]**	**Relative presenteeism[3]**	**Combined productivity score[4] (relative absenteeism and relative presenteeism)**
Mean daily steps	1.00 n = 87					
Absolute absenteeism (hours lost/month)	0.18 n = 87	1.00 n = 139				
Relative absenteeism[1]	0.14 n = 87	0.50*** n = 139	1.00 n = 139			
Absolute presenteeism[2]	-0.02 n = 87	-0.16 n = 139	-0.07 n = 139	1.00 n = 139		
Relative presenteeism[3]	-0.18 n = 87	-0.13 n = 139	-0.08 n = 139	0.40*** n = 139	1.00 n = 139	
Combined productivity score[4] (relative absenteeism and relative presenteeism)	-0.20 n = 87	-0.50*** n = 139	-0.81*** n = 139	0.23** n = 139	0.63*** n = 139	1.00 n = 139

Note: Pearson correlation coefficient and number of observations. Pairwise deletion used.
*****p<0.05 ******p<0.01 *******p<0.001
Orange shading is used to show correlations of most relevance.
[1] Higher score = higher absenteeism, relative to expected hours
[2] Higher score = lower lost performance, i.e. higher productivity
[3] Higher score = higher performance or productivity relative to colleagues
[4] Higher score = higher productivity

5.2.3.3 Changes in self-reported physical activity, sedentary time, and health, wellbeing, perceived stress and productivity

Similar to the findings with steps, no statistically significant Pearson's r correlations were observed between change in self-reported PA (or sedentary time) and changes in health and wellbeing outcomes (physical and mental health-related quality of life) or perceived stress from baseline to post-intervention (week 12). From baseline to 8-month follow-up, there were no significant associations between changes in self-reported PA and changes in the health and wellbeing outcomes or perceived stress. There was a small significant negative correlation between change in sedentary time and change in the SF-12 PCS from baseline to follow-up (r = -0.18; p = 0.039; n = 137), i.e. reductions in sedentary time were associated with improvements in physical health-related quality of life.

There were no significant associations between changes in self-reported PA or sedentary time and changes in any of the HPQ outcomes (absenteeism, presenteeism and combined productivity score) from baseline to week 12. From baseline to 8-month follow-up, there was a significant positive correlation between change in total PA (weekly MET-minutes) and change in absolute absenteeism ($r = 0.22$; $p = 0.010$; $n = 139$), i.e. greater increases in PA were associated with greater increases in absenteeism. A similar positive association was observed between change in MVPA (weekly MET-minutes) and change in absolute absenteeism ($r = 0.22$, $p = 0.010$, $n = 139$). This association was not observed for relative absenteeism or for any of the other HPQ outcomes.

5.2.4 Sickness absence

Sickness absence data were not available for one of the study sites (North Dorset). Data were available for Plymouth BCU and the results of a basic analysis for this site only are presented. However, this site was also affected by data quality issues, so the findings must be interpreted with caution.

Table 52 shows the differences in the number of sickness episodes, duty days lost, and total days lost between the one year pre-intervention period (June 2016 to May 2017) and the one year post-intervention period (June 2017 to May 2018). There was a small increase in the total number of sickness absence episodes (+8%) and a larger increase in the number of duty days lost and total days lost from pre- to post-intervention. Duty days lost increased by approximately 39% and total days lost increased by approximately 29%.

The numbers of participants with one or more sickness episodes and no sickness episodes for each time period are shown in **Table 53**. The results of a Chi-squared test showed that there were no significant differences in observed and expected frequencies, i.e. there was no evidence to suggest that sickness absence had increased significantly from pre- to post-intervention ($\chi^2 = 0.32$, $p = 0.573$). As there was no significant change in the number of sickness episodes, no inferential analysis was performed on duty days lost and total days lost.

Table 52 Changes in number of sickness episodes, duty days lost and total days lost from pre- to post-intervention

	Number of sickness episodes	**Duty days lost**	**Total days lost**
June 2017-May 2018 (post-intervention)	124	714	1019
June 2016-May 2017 (pre-intervention)	115	515	793
Difference (post-pre)	+9	+199	+226
% difference	+8	+39	+29

Note: Only study participants employed by Devon & Cornwall Police between June 2016 and May 2018 are included (n = 106). One participant did not give consent for data use and is therefore excluded.

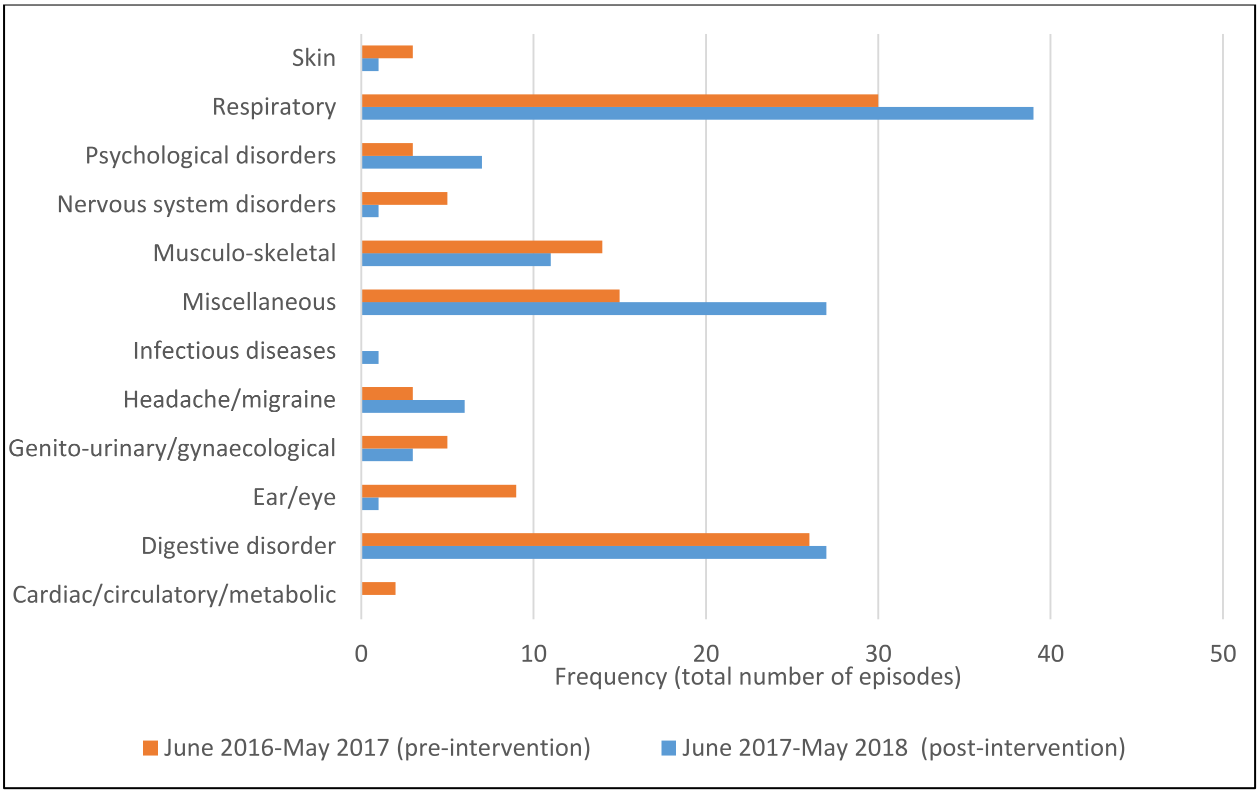
Table 53 Frequencies of participants with one or more vs. no sickness episodes by time period (percentages by row)

	Number (%) of participants with one or more sickness episodes	**Number (%) of participants with no sickness episodes**	**Total**
June 2017-May 2018 (post-intervention)	63 (59)	43 (41)	106 (100)
June 2016-May 2017 (pre-intervention)	67 (63)	39 (37)	106 (100)
Total	130 (61)	82 (39)	212 (100)

Note: Only study participants employed by Devon & Cornwall Police between June 2016 and May 2018 are included (n = 106). One participant did not give consent for data use and is therefore excluded. Percentages may not total 100 due to rounding. Chi-square (1 degree of freedom, d.f.) = 0.32; $p = 0.573$.

Reasons for sickness absence for the 106 study participants in the pre- and post-intervention periods are shown in **Figure 18**. These did not appear to change considerably; the four most common illness categories both pre- and post-intervention were respiratory conditions, digestive disorders, miscellaneous conditions, and musculoskeletal disorders. Some conditions appeared to decrease in frequency as a reported cause of sickness absence (e.g. cardiac/circulatory/metabolic), while others increased in frequency (e.g. respiratory).

Figure 18 Reasons for sickness absence: comparison of pre- and post-intervention periods (n = 106)



5.2.3 Summary / Discussion

Figure 19 summarises the statistical significance of changes in the secondary outcomes (physical and mental health-related quality of life, perceived stress and perceived productivity) at the different time points.

Figure 19 Summary of significance of changes in secondary outcomes at various time points: all participants

Outcome	Mean change from baseline (week 0)		
	Mid-intervention (week 6)	Post-intervention (week 12)	Follow-up (month 8)
SF-12 Physical Component Score (PCS)	**0**	**0**	**0**
SF-12 Mental Component Score (MCS)	**0**	**0**	**+**
Perceived Stress Scale (PSS-4) score	**0**	**0**	**0**
Health and Work Performance Questionnaire (HPQ) scores:			
Absolute absenteeism	**0**	**0**	**0**
Relative absenteeism	**0**	**0**	**0**
Absolute presenteeism	**0**	**0**	**0**
Relative presenteeism	**0**	**0**	**0**
Combined productivity (relative absenteeism and presenteeism)	**0**	**0**	**0**

+ = positive impact

- = negative impact

0 = no significant impact

Overall, there were no significant changes in physical or mental health-related quality of life, perceived stress or productivity from baseline to week 6 (mid-intervention) or from baseline to week 12 (post-intervention). In the longer term, there was a small but statistically significant improvement in mental health-related quality of life (as measured by the SF-12 Mental Component Score, MCS), with a mean increase of 1.75 points from baseline to 8-month follow-up. The mean score at the 8-month follow-up was 49.72, which was closer to the US norm of 50 (Utah Department of Health, 2001). It is unclear whether the observed improvement in MCS is a clinically important change with health benefits for officers and staff. Studies aiming to estimate the minimum clinically important difference for the SF-12 outcomes have mainly involved clinical populations such as Parkinson's disease (Shulman et al., 2010) and post-surgery patients (Clement et al., 2014). A study of individuals with low back pain suggested that changes in MCS of 3.77 and in PCS of 3.29 are clinically relevant (Diaz-Arribas et al., 2017), but this may be too high for a ‘healthy’ population such as the police force.

The increase in MCS was not correlated with increases in objective or self-reported PA. This contrasts with previous findings where increases in light intensity PA have been shown to correlate with improvements in the SF-36 MCS (Takayanagi et al., 2018). The results therefore suggest that the improvement in mental health may be due to another component of the intervention (such as setting relaxation and mindfulness goals in the Bupa Boost app) or possibly external influences, rather than as a result of increased activity levels. Another explanation might be that participants’ mental health benefited over the longer term from the increased social support and camaraderie resulting from continued use of the social features in the Fitbit® and Bupa Boost app. This is discussed further together with survey and interview findings in **Chapter 6**.

There were no significant changes (either statistical or clinically relevant) in physical health-related quality of life, perceived stress or productivity from baseline to 8-month follow-up. However, although there was no significant overall increase in either daily step count or physical health-related quality of life, there was a significant positive association between changes in steps and the SF-12 Physical Component Score from baseline to 8 months. Participants who reduced their daily steps showed greater declines in physical health, and to some extent participants who increased their steps showed greater improvements in physical health. There was a corresponding small but significant negative correlation between changes in self-reported weekday sedentary time and physical health-related quality of life during this period, adding further support that physical health may improve as a result of increased PA and reduced SB within a workplace intervention. Despite this, for the less active subgroup, there was no significant association between changes in steps and physical health-related quality of life; however, it is important to reiterate that this may in part be due to the small sample size in subgroup analysis which reduced power to detect an association.

Findings in relation to the wider health impacts of workplace wellness interventions with a PA component have been mixed. The literature suggests that potential benefits may include improved self-perceived wellness (Schrager et al., 2017), enhanced psychological wellbeing (Abdin et al., 2018), improved physical and mental health-related quality of life (Emerson et al., 2017), lower perceived stress levels (Emerson et al., 2017), reduced weight and body mass index (Ganesan et al., 2016, Yu et al., 2017, Freak-Poli et al., 2013a), lower blood pressure (Thorndike et al., 2014), reduced resting pulse rate (Koyle, 2013) and improved fasting plasma glucose (Freak-Poli et al., 2013a). In contrast, some studies have reported no or even negative effects of workplace PA interventions, such as relative increases in weight and body fat compared with a control group (Reijonsaari et al., 2012). Effect sizes for health and wellbeing benefits are generally small, and evidence may be limited and of poor quality (Freak-Poli et al., 2013b). Researchers have consequently advised that employers should have ‘modest expectations’ about the potential health benefits of such programmes implemented in real-world workplace settings (Yu et al., 2017). It may be that beneficial health impacts of PA interventions do not become apparent until several months or years later and future studies may therefore require a longer follow-up.

Furthermore, it may be difficult to determine causality between PA and health and wellbeing, including the direction of the relationship. For example, are those who are physically active more likely to experience improved physical health, or are those with better physical health more likely to engage in PA? It is likely that the association is bi-directional and dynamic (Hiles et al., 2017, Liao et al., 2017).

The negative correlation between changes in the Physical and Mental Component Scores of the SF-12 was an unexpected finding in the present study. This correlation was apparent at baseline (as reported in **Chapter 4**) and remained significant throughout the study. This shows that good physical health is not necessarily associated with good mental health in this sample, and should be explored further.

The lack of an observed association between changes in PA (or sedentary time) and productivity in workplace interventions is not uncommon. While a significant positive association between PA and productivity outcomes has been reported in observational studies (e.g. job performance and quality of work - Pronk et al., 2004), intervention studies for PA (Gram et al., 2012) and SB (Brakenridge et al., 2016b) have found no significant impact on productivity, even when the intervention was otherwise successful. Presenteeism is a component of the HPQ and was used to calculate the overall productivity score; there were no significant changes in presenteeism throughout the study. The evidence is far from established in this field, although there is some preliminary indication that workplace wellness programmes are capable of producing improvements in presenteeism (Cancelliere et al., 2011). As with the impact on health and wellbeing, longer follow-up may be needed to detect improvements in presenteeism and productivity in the longer term. Standardised measures of presenteeism for use in future intervention trials have been called for (Cancelliere et al., 2011).

There was an unexpected positive correlation between changes in self-reported PA and self-reported absolute absenteeism from baseline to 8 months, suggesting that officers and staff who are more active have higher levels of absenteeism. There was also a near significant negative correlation between change in steps and change in HPQ productivity score during this period. These findings may be of concern to managers and commissioners, for whom productivity and the economic benefits of PA interventions are likely to be both expected outcomes (see **Chapter 7**) and perceived as of central importance to supporting such initiatives (Buckley et al., 2015, Morris et al., 2018).

One possible explanation is that these results may be a consequence of using the HPQ as an assessment tool, which does not differentiate between sickness absence and factors such as annual leave; for example, if more participants were on leave at follow-up this could explain the higher ‘absenteeism’. Relative absenteeism may be a more suitable measure of absenteeism than the absolute values, as it allows proportional comparison of workers who vary in full-time equivalents (Kessler et al., 2004). There were no significant correlations between changes in PA and relative absenteeism.

It was not feasible to collect objective sickness absence data for both study sites, due to lack of availability of this data for North Dorset. In addition, both sites were affected by data quality issues. It is recommended that the organisations ensure that sickness absence data are complete, accurate and collected in a timely manner. Procedures for reporting sickness absences should be standardised between the Devon and Cornwall and Dorset Police forces if possible. Taking these issues into consideration, basic analysis of sickness data for the Plymouth BCU site found that there was no significant change in the total number of sickness absence episodes from pre- to post-intervention. Furthermore, the number of duty days and total days lost due to sickness appeared to increase in frequency. This contrasts with previous studies that have shown that more physically active employees (van den Heuvel et al., 2005, Proper et al., 2006) including police officers (Steinhardt et al., 1991) have lower levels of absenteeism. Reasons for sickness absence also remained fairly consistent in frequency over time, with respiratory, digestive and musculoskeletal disorders as the most frequent causes of absence. It must be kept in mind that these findings are tentative, however, due to the data quality issues and the wide range of factors that are known to impact on sickness absence (Korlin et al., 2009, Airaksinen et al., 2018). Few mHealth intervention studies have included objective sickness absence as an outcome; this is an important gap to address in future research.

As with the PA outcomes, the above findings need to be carefully interpreted in the context of an uncontrolled study design (see **Chapter 8** for a full discussion of strengths and limitations). However, the results are valuable in terms of exploring the impact (and mechanisms of action) of the intervention on health and wellbeing, perceived stress and productivity up to 8 months, and the potential impact on sickness absence up to 12 months, after beginning the intervention. There were also some notable findings in relation to the appropriateness of outcome measures for future trials, such as the HPQ to assess productivity, and the need for high quality sickness absence data. Larger sample sizes and longer follow-up may be needed to fully explore and clarify associations between PA, sedentary time, health and wellbeing, stress, productivity and sickness absence. The key findings in this section are discussed with reference to the qualitative findings in **Chapter 6**.

CHAPTER 6 ACCEPTABILITY OF THE INTERVENTION FOR POLICE OFFICERS AND STAFF

6.1 Introduction and overview

In this chapter, data from the surveys and interviews (described in **Chapter 3**) are combined to build a picture of the acceptability of the PAW-Force intervention for police officers and staff. Acceptability has varying definitions in mobile health (mHealth); however it includes aspects such as perceived usability, perceived usefulness, engagement or use, barriers to engagement, preferred intervention components, and negative impact or adverse events (Buckingham et al., 2019). These aspects, particularly qualitative measures of acceptability and participant experiences, in both the short and longer term, are generally understudied and underreported in the current mHealth literature (McCallum et al., 2018).

The chapter begins with an exploration of the previous experiences and expectations of mHealth for study participants (**section 6.2**). In **section 6.3**, engagement (i.e. intervention use) is outlined, including perceived usability and perceived usefulness and reasons for lack of engagement. The perceived impact of the intervention on physical activity (PA) and sedentary time is explored in **section 6.4**, with reference to preferred intervention components and mechanisms of behaviour change. The perceived impact on secondary outcomes (e.g. health and wellbeing) is then considered (**section 6.5**). In **section 6.6**, findings related to adverse events and unintended consequences of intervention use are presented. The chapter concludes with a summary and discussion of the key findings (**section 6.7**).

6.2 Previous experiences and expectations of mHealth and fitness technology

Engagement and experience of mHealth technology may be influenced by users' previous experiences and expectations (Carter et al., 2018). This section briefly focuses on participants' previous experiences of mHealth and fitness technology, views on PA and reasons for participation, and expectations of the intervention including expected barriers to technology use. Quantitative and qualitative data from baseline questionnaires and interviews are combined.

6.2.1 Previous use and experiences of mHealth and fitness technology

Table 54 shows the proportions of participants who reported having used a wearable activity monitor and/or a health or fitness app prior to taking part in the study. Just over half of all participants (98/180, 54%) had previously used a health or fitness app, compared with 37% (66/180) who had worn an activity monitor. Of the total sample, 31% (55/180) had used both mHealth tools, while 39% (71/180) had used neither. The most commonly used activity monitors were earlier versions of the Fitbit®, the Apple Watch®, Garmin® running watches and the (now discontinued) Jawbone UP®. Frequently used health and fitness apps included those associated with these activity monitors, in addition to apps such as MyFitnessPal® to track diet and PA, and Strava® for running and cycling.

Table 54 Previous use of mHealth and fitness technology: Number and % of participants (percentages of total number of participants)

Previous activity monitor use	**Previous health or fitness app use**		
	No	***Yes***	**Total**
No	71 (39 %)	43 (24 %)	114 (63 %)
Yes	11 (24 %)	55 (31 %)	66 (37 %)
Total	82 (46 %)	98 (54 %)	180 (100 %)

The baseline interviews confirmed the range of experience levels with mHealth and fitness technology. Some participants had no experience of this technology (*"It's all new to me"*), others had some familiarity from friends and family members with activity monitors or apps, while others were regular users (*"I am used to tracking… I am used to logging everything"*).

6.2.2 Views on PA and reasons for participation

It was clear from the interviews that the study participants had a range of levels of motivation at baseline and views on being physically active. Some lacked motivation for exercise and had no interest in sport:

> *"I can't get my head round people enjoying sports and gyms and things like that. I've tried it but didn't enjoy it."*
>
> (Police staff, female, age 40+)

Others were highly intrinsically motivated at the beginning of the study, i.e. they exercised for personal enjoyment and obtained satisfaction from the behaviour itself (Deci and Ryan, 2000):

> *"I've always been interested in sport. I love it and I have always loved it."*
>
> (PCSO, male, age 40+)

The Transtheoretical Model is an early psychological model of behaviour change devised by Prochaska and DiClemente (Prochaska and DiClemente, 1983). The model is well established and has been widely used to conceptualise changes in lifestyle behaviours including PA (De Leon et al., 2014, Walsh et al., 2017). A core aspect of this model is the 'stages of change', where it is proposed that individuals move through various stages in changing their behaviour – precontemplation (not even thinking about changing), contemplation (thinking about changing), preparation (preparing to change), action (enacting changes) and maintenance (sustaining the behaviour) (Prochaska and DiClemente, 1983). At the beginning of the PAW-Force study, most participants appeared to be at the later 'stages of change' according to the Transtheoretical Model, with many at the 'preparation' stage. They had volunteered to participate, were ready to change their behaviour and preparing to make the necessary changes in the near future. For example:

> *"I would like to do more. My fitness levels aren't what they were a few years ago. I'm preparing to start a fitness plan with my wife."*
>
> (Police sergeant, male, age 18-39)

Despite varying levels of motivation, all of those interviewed were enthusiastic about taking part in the study and using the Fitbit® and Bupa Boost app. The reasons for participating in the study, as given by interviewees at baseline, were as follows:

- Improving PA level or trying a new type of exercise
- Reducing sedentary time at work
- Improving fitness
- Comparing one's activity level with colleagues
- Maintaining or improving health and wellbeing
- Wanting to lose weight or improve appearance

- Reducing stress
- Improving sleep
- Improving confidence and enjoyment of life
- Desire to learn about mHealth and fitness technology

Although the reasons were varied they centred on improving or maintaining PA or reducing sedentary time. Those who perceived themselves to be sufficiently active at baseline were keen to reduce their sedentary time:

> *"I know that I do enough fitness. I am just conscious of like how much, eight hours, I am sat at my desk."*
>
> (Police constable, female, age 18-39)

Many reported 'secondary' reasons for taking part, such as wanting to improve their health and wellbeing or their physical appearance:

> *"I am taking part for my health because I know that if I look better I will feel better. I am hoping it is just going to have this knock-on effect for this forever feeling like, ugh, my God, you know. I know I have got to do it... for my health, my appearance and general wellbeing I suppose really."*
>
> (Police constable, female, age 40+)

One interviewee was motivated to participate by the results of a recent health and fitness screening in the workplace:

> *"I wanted to do this trial because… somebody came in [to work] and did a fitness check. After this health check I thought, no, I need to do something."*
>
> (Police staff, female, age 40+)

6.2.3 Expectations of mHealth and fitness technology / the intervention

Regardless of their own experience with mHealth technology, some common expectations were reported by interviewees. The expected functions of the technology were enabling self-monitoring and awareness of PA and sedentary time, goal-setting, social influence, and motivation.

All of those interviewed talked about the technology giving information on their activity levels (and other health measures such as heart rate and sleep). Some interviewees reported that this was the component that most appealed to them as they liked to see evidence of their behaviour. For example:

> *"I like numbers and I like data. If I can see data and improvements and tangible results because I don't necessarily see results in my body… I like to see evidence that I have been working hard. I am looking forward to getting that out of everyday use by using an activity tracker."*
>
> (Police inspector, male, age 40+)

Some participants believed that self-monitoring and increased awareness would be sufficient to provide motivation to increase their activity levels:

> *"It's a good way like anything, it's like when you go on a diet or when you do anything… anything that makes you think about what you're doing. I'm hoping that what it might do then is just make me make that effort to do a little bit more."*
>
> (Police staff, male, age 40+)

> *"I would have thought wearing it would make you be more active, especially with the displays of how many steps you are doing, graphs and so on, and it would motivate you to increase your steps."*
>
> (Police constable, male, age 18-39)

One police constable expected the technology to allow her to see the 'bigger picture' by monitoring progress over time:

> *"… because if I just look at it day to day I probably wouldn't remember… it would be nice to have a longer period of analysis."*
>
> (Police constable, female, age 18-39)

Some expected the goal-setting component of the intervention to work well for them, and many talked about the importance of small, achievable goals. For one police officer, goal-setting appealed to his competitive nature:

> *"I am one if I have got a target or plan I will go for it because I am quite competitive and will try and achieve something, so I thought it would probably work for me."*
>
> (Police sergeant, male, age 18-39)

Other interviewees expected to be more motivated by the social influence aspects of the intervention. This is a form of extrinsic motivation, where a behaviour is performed for external rewards (or to avoid punishment) rather than for its own enjoyment (Deci and Ryan, 2000). For one member of police staff with a 'low' baseline activity level (according to the International Physical Activity Questionnaire, IPAQ), both social accountability and social support were expected to be important extrinsic motivators:

> *"I thought this should give me motivation… if nobody's sort of, like, there watching me, I might not do it… It's when you've got the support of people around you, that encourages you."*
>
> (Police staff, female, age 40+)

Social support outside of, in addition to within, the app was expected to influence behaviour:

> *"I am hoping people will be talking about it and encouraging each other on."*
>
> (Police constable, female, age 40+)

Overall, the expectation of those interviewed was that the intervention would be effective for increasing or maintaining PA and/or reducing sedentary time, although one interviewee recognised the importance of individual differences in the impact of interventions:

> *"Not everything works for everybody. Whether it will make a difference I don't know, but I'd like to think it would."*
>
> (Police constable, female, age 40+)

Three of the 10 participants interviewed at baseline reported that they expected to encounter barriers to use of the Fitbit® or Bupa Boost app. All of these reported lack of technological skills as a potential problem (for example, *"I am a bit of a technophobe"* or, *"Phones are for making phone calls and texting; that's about the limit of it with me."*). Support for setting up and using the technology was perceived as necessary by these individuals, and all reported having a family member who was able to help (for example, *"I have got a son at home who, if I get stuck, can help me sort it out"*).

6.3 Engagement, perceived usability and perceived usefulness

Engagement with digital behaviour change interventions is of central importance and is a prerequisite for intervention effectiveness (Perski et al., 2017). However, the 'law of attrition', which refers to the discontinuation of eHealth application use over time (Eysenbach, 2005), is still considered one of the greatest challenges in relation to eHealth technology (Kohl et al., 2013). Declining engagement and discontinuation of the use of mHealth technology over time is extensively reported in the literature (Afshin et al., 2016, Buckingham et al., 2019). While many studies have been relatively short term and the levels of engagement vary greatly between studies, attrition rates of up to 100% have been reported in workplace mHealth studies of one year duration (for example, Brakenridge et al., 2016b).

Within this context, this section includes a quantitative and qualitative overview of engagement with the intervention. Quantitative measures of engagement are first presented, including frequency and duration of use of the Fitbit® and the Bupa Boost app, features used and goals set. Quantitative survey findings and interview findings are then combined to build a picture of the usability and usefulness of the intervention, as perceived by participants. Reasons for lack of engagement (which were closely associated with usability and usefulness) are then explored as an additional dimension of acceptability, again combining survey and interview findings.

6.3.1 Intervention usage, features used and goals set

Overall, engagement with the Fitbit®, as measured by self-reported wear time, was high in both the short and longer term, although there was a small decline over time, as shown in **Table 55**. At baseline, all 180 participants had worn the Fitbit® for at least four of the previous seven days, for a mean of 6.9±0.4 days and 22.6±2.8 hours/day. After six weeks (mid-intervention), 98% (156/159) of responding participants were wearing the Fitbit® (the percentage of the whole sample wearing the device was unknown). Mean wear time after six weeks was 6.7±1.0 days/week and 22.3±3.3 hours/day. After 12 weeks (post-intervention), 97% (146/151) of responding participants were wearing the Fitbit®, for a mean of 6.6±1.0 days/week and 22.0±3.7 hours/day. At the 8-month follow-up, 83% (119/143) of survey respondents were wearing the Fitbit® (one of these was still using the device but had not worn it during the previous seven days). At this time point, the mean wear time was 6.5±1.1 days/week for 21.4±4.1 hours/day. Approximately 17% of participants (24/143) reported that they had stopped wearing the Fitbit® after 8 months.

Table 55 Summary of Fitbit® wear time (self-reported) at the various outcome time points

Time point	**n[1]**	**Days worn during the previous 7 days**		**Hours per day worn during the previous 7 days**	
		Mean (SD)	**Range**	**Mean (SD)**	**Range**
Baseline (week 0)	180	6.9 (0.4)	4 to 7	22.6 (2.8)	7 to 24
Mid-intervention (week 6)	156	6.7 (1.0)	1 to 7	22.3 (3.3)	5 to 24
Post-intervention (week 12)	146	6.6 (1.0)	1 to 7	22.0 (3.7)	6 to 24
Follow-up (month 8)	118	6.5 (1.1)	1 to 7	21.4 (4.1)	7 to 24

Note: SD = Standard Deviation; n = number of participants who reported wearing the Fitbit
1 Participants who reported wearing the Fitbit for 0/7 days or less than one hour/week are not included in the table.

Engagement with the Bupa Boost app (as measured by self-reported usage time) was lower throughout the study, and a more rapid decline in engagement over time was observed (see **Table 56**). After six weeks (mid-intervention), 65% (104/159) of responding participants reported using the Bupa Boost app, compared with 60% (91/151) after 12 weeks (post-intervention). At the 8-month follow-up, only 27% (39/143) of responding participants reported that they were still using the app (three of these were still using the app occasionally but had not logged in during the previous seven days and are therefore not included in the table). Despite the high non-usage rate, the mean usage time (frequency and duration) did not change significantly during the study. There was a small increase in mean usage frequency from 4.5±2.4 days/week at week 6 to 5.1±2.1 days/week at week 12 (remaining at 5.1±2.3 days/week at month 8). Throughout the study, the majority of users of the Bupa Boost app logged in for one to five minutes per day.

Table 56 Summary of Bupa Boost use (self-reported) at the various outcome time points

Time point	n[1]	Days used during the previous 7 days		Time per day used during the previous 7 days n (%)				
		Mean (SD)	Range	Less than one minute	1-5 minutes	5-15 minutes	15-30 minutes	30-45 minutes
Baseline (week 0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Mid-intervention (week 6)	104	4.5 (2.4)	1 to 7	9 (9)	64 (62)	27 (26)	4 (4)	0 (0)
Post-intervention (week 12)	91	5.1 (2.1)	1 to 7	7 (8)	49 (54)	26 (29)	8 (9)	1 (1)
Follow-up (month 8)	36	5.1 (2.3)	1 to 7	7 (19)	21 (58)	6 (17)	0 (0)	2 (6)

Note: SD = Standard Deviation; n = number of participants who reported using the Bupa Boost app
[1] Participants who reported they had not logged in to the app within the previous 7 days are not included in the table.

The features of the Bupa Boost app used by participants at each time point are detailed in **Table 57**. Goal-setting and earning virtual rewards (i.e. wellness points and badges) were used most frequently throughout the study; these ‘individual’ features were reported to be used by 89% (93/104) of participants at week 6, 74% (77/104) at week 12, and 45% (29/65) at month 8. The second most frequently used feature was group challenges (i.e. use of the leader board to compare and compete with colleagues); 41% (43/104) of participants at week 12 and 28% (18/65) at month 8 reported to have used this feature. Company and team challenges were reported to have been used by 25% (26/104) of participants at week 12 and 15% (10/65) at month 8, and use of the social feed was reported by 28% (29/104) of participants at week 12 and 12% (8/65) at month 8. The Bupa library including self-help information, and supportive messages to colleagues, were used less frequently, with only 8% (5/65) and 6% (4/65) of participants respectively reporting having used these features at 8 months. There was an overall decline in the use of features over time. Acceptability of these features as perceived by interviewees is discussed in **section 6.4**.

Table 57 Features used in Bupa Boost (self-reported at each time point)

Time point	Total number of responses	Setting own goals, earning wellness points and badges n (%)	Bupa library and self-help information n (%)	Group challenges / leader board to compete with colleagues alone n (%)	Company or team challenges (competing with colleagues as a team) n (%)	Social feed (to view progress of colleagues) n (%)	Messages to colleagues (motivation and encouragement) n (%)
Mid-intervention (week 6)	104	93 (89)	14 (13)	N/A	N/A	N/A	N/A
Post-intervention (week 12)	104	77 (74)	9 (9)	43 (41)	26 (25)	29 (28)	5 (5)
Follow-up (month 8)	65	29 (45)	5 (8)	18 (28)	10 (15)	8 (12)	4 (6)

Note: n = number of participants using feature
N/A = participants were instructed not to use feature until the social phase (weeks 7 to 12)

Between the beginning of the study and week 6, approximately 50% (80/159) of participants reported having set PA or fitness goals in addition to increasing their daily step count. This included goals set within the Fitbit® and/or the Bupa Boost app. It was evident that many participants had used action planning, i.e. detailed goal-setting to specify when, in which situation and/or where to perform the behaviour (Michie et al., 2011b). The Bupa Boost app provided suggested goals and also allowed goals to be created or customised. Examples of goals set were:

- Go to the gym three times per week
- Walk to work four days per week
- Lift weights twice a week
- Walk the dog for longer than 30 minutes each day
- Swim at least once a week
- Cycle 100 miles per week

While the majority of participants using Bupa Boost used the app to manage their PA and fitness (e.g. 93/104 or 89% at week 6), a proportion used it to manage additional health and wellbeing behaviours. This included goal-setting in relation to nutrition (39/104, 38% at week 6), mindfulness (27/104, 26% at week 6) and relaxation (20/104, 19% at week 6). Examples of nutrition, mindfulness and relaxation goals respectively were, ‘drink more water’, ‘no screens after 10.00pm’, and ‘spend time outdoors six days per week’.

6.3.2 Perceived usability and usefulness

The Technology Acceptance Model is an early theory that was proposed to explain and predict user acceptance of information technologies (Davis, 1989). According to this model, perceived ease of use and perceived usefulness are the two main determinants of acceptance and use of a new technology (Davis, 1989). The model has since been applied to health information technology and while some modifications have been suggested, there is evidence that these two factors remain valid and important for predicting acceptance (Holden and Karsh, 2010). In the PAW-Force study, participants were asked to rate the perceived usability and usefulness of the Fitbit® and Bupa Boost app on a Likert-type scale in both the mid-intervention (week 6) and post-intervention (week 12) surveys. The questions were based on participants’ perceptions of ease of use and whether the technology helped them to be more physically active.

The results are shown in **Table 58** and **Table 59**. The Fitbit® was rated as easier to use than Bupa Boost in both the mid-intervention (week 6) and post-intervention (week 12) surveys. For example, the mean usability rating of the Fitbit® at week 12 was 4.7 compared with 3.6 for Bupa Boost (where 5 was the highest possible score and 1 was the lowest). Ratings for usability of the Bupa Boost app ranged from 1 to 5, compared with 3 to 5 for the Fitbit®, indicating that some participants perceived major usability issues with Bupa Boost. The ratings for usefulness (i.e. whether participants perceived the technology had helped them to increase their PA level) were slightly lower for both the Fitbit® and Bupa Boost app, but most agreed that the Fitbit® had been useful (mean rating 3.9) compared with a mean rating of 3.2 for Bupa Boost. Ratings did not change substantially from week 6 to week 12.

Table 58 Perceived usability of the Fitbit® and Bupa Boost app

Time point	Fitbit usability rating			Bupa Boost app usability rating		
	n	**Mean (SD)**	**Range**	**n**	**Mean (SD)**	**Range**
Mid-intervention (week 6)	158	4.8 (0.5)	3 to 5	128	3.5 (1.2)	1 to 5
Post-intervention (week 12)	147	4.7 (0.5)	3 to 5	118	3.6 (1.2)	1 to 5

Note: SD = Standard Deviation; n = number of responses
Participants were asked: "On a scale of 1 to 5, where 5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree, 1 = strongly disagree, how much do you agree that the *Fitbit/Bupa Boost* was easy to use?"

Table 59 Perceived usefulness of the Fitbit® and Bupa Boost app

Time point	Fitbit usefulness rating			Bupa Boost app usefulness rating		
	n	**Mean (SD)**	**Range**	**n**	**Mean (SD)**	**Range**
Mid-intervention (week 6)	158	3.9 (1.1)	1 to 5	128	2.9 (1.3)	1 to 5
Post-intervention (week 12)	147	3.9 (1.0)	1 to 5	117	3.2 (1.3)	1 to 5

Note: SD = Standard Deviation; n = number of responses
Participants were asked: "On a scale of 1 to 5, where 5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree, 1 = strongly disagree, how much do you agree that the *Fitbit/Bupa Boost* helped you to be more physically active?"

Interview data and qualitative survey data supported these quantitative results. All of the participants that were interviewed stated the Fitbit® was easy to use and navigate (for example, *"It's a great little watch… it's really easy to navigate around"*). These views were shared by participants who reported low technological familiarity and skills. For example:

> *"I think, overall, it is very simple to use, even for people like me."*
>
> (Police constable, female, age 40+)

All of the interviewees reported that the Fitbit® had either met or exceeded their expectations, for example:

> *"It's probably better than I expected it to be. I thought I'd get bored of it really quickly, and give up on it. I thought it wouldn't be particularly accurate. But it was better than I expected it to be for sure."*
>
> (Police staff, female, age 18-39)

As further evidence for high perceived acceptability of the Fitbit®, several interviewees talked about purchasing a newer model of the device for themselves (*"I've been looking at upgrading the Fitbit"*), recommending it to family, friends and colleagues (*"A lot of people in the station are talking about getting them as well now"*), or even buying one for a family member (*"I'm considering getting one for my dad because he's had a stroke and he's not as active as he used to be"*).

In contrast, interviewees reported relatively lower usability and usefulness of the Bupa Boost app:

> *"I found it very awkward to use. It wasn't user friendly at all."*
>
> (Police constable, male, age 18-39)

> *"I found it [Bupa Boost app] was a little bit complicated really, and it wasn't that simple to follow like the Fitbit is."*
>
> (Police constable, male, age 40+)

There was a general link between perceived usability and usefulness and engagement with (i.e. use of) the intervention. This is summarised by the following participant:

> *"The Bupa Boost app, I used very occasionally because I didn't particularly find it a very user friendly or useful app."*
>
> (Police staff, female, age 18-39)

For some, perceived usability and usefulness of Bupa Boost improved over time as use and familiarity increased:

> *"It [Bupa Boost app] took a while to get used to. I'd probably say a couple of weeks just to understand the bits that I like."*
>
> (Police inspector, male, age 40+)

> *"Regarding Bupa Boost, yes, I do like that app, not as much as say, the Fitbit, only because I didn't know how to use it at the beginning. I started playing around with it and then I started realising how to use it… I still use it now, so, that must mean I do like Bupa Boost. I've definitely grown to it."*
>
> (PCSO, female, age 18-39)

Interviewees who initially expected difficulty in using the technology tended to experience more problems with usability, particularly with the Bupa Boost app.

Some overcame this with help from family members, but for others their perceived lack of technological skills remained a barrier, for example:

> *"I'm not technically-minded. It's probably another reason for not automatically going to it [the Bupa Boost app]. I don't tend to use apps on my phone massively. I just use what I have to."*
>
> (Police constable, female, age 40+)

The qualitative survey and interview data provided further insight into the specific aspects (positive and negative) of usability and usefulness of the Fitbit® and Bupa Boost, which are summarised in **Figure 20**. The Fitbit® was praised for its user interface and easy to navigate menus, in contrast to the Bupa Boost app where several spoke negatively of its user interface and reported it was difficult to find colleagues within the app. The Fitbit® was also perceived as convenient to charge (*"The battery was good, I probably charged it once a week and that was about it"*). Participants liked the range of features of the Fitbit® (such as monitoring heart rate and sleep in addition to PA), and found that it was small and lightweight yet durable. Some interviewees mentioned that it was convenient and compatible with the policing uniform and role:

> *"It's good to wear with the police uniform because it's black and it's quite subtle."*
>
> (Police sergeant, male, age 18-39)

> *"Practicality-wise it's quite useful for specials as it's quite a durable little thing, we can have a watch on our wrist without worrying too much about it getting broken by some lovely person who wants to tackle someone or something [laughs]. What I had before was an iWatch and I didn't want to wear that out and about on the streets. It's been nice having something a bit more durable for certain times."*
>
> (Police staff / special constable, male, age 18-39)

Waterproofing was the most commonly suggested improvement to usability and usefulness of the Fitbit®:

> *"I don't like the fact that it's not waterproof. I go to the beach a few days in the summer, surfing, and playing on the beach and in the sea. I'm in the sea for maybe four or five hours. It seems ironic that you've got to take off an activity*

tracker. It seems like it's almost not fit for purpose, if it's not made to be waterproof and you can't wear it in the sea... My kids have got their Garmin one on. I've got to take off my Fitbit."

(Police inspector, male, age 40+)

"I kayak so I don't use it in those situations which is a shame really."

(Police constable, male, age 40+)

In addition to improved suitability for water-based activities, other suggested changes to the Fitbit® were greater suitability and accuracy for a wider range of activities including boxing, gym exercise, and horse riding. Some interviewees perceived that their true PA level was not reflected in their step count as recorded by the Fitbit®:

"I've been doing boxing training… the only limitation is I can't wear it for that because of the gloves. I can't get the gloves on with it - which is a shame because my activity would have shown I had achieved a lot more if I just had that on at the time."

(Police sergeant, male, age 18-39)

"One of the complaints that people say is, look, I go in the gym, I work really hard, but it doesn't record that as a step. I can see that; you're not really stepping. But it almost looks like you're not doing any exercise. Some people are doing 5,000 steps a day, but they go in the gym for two hours and that's not recorded."

(PCSO, male, age 40+)

"I horse ride once a week. I think the only thing, wearing it horse riding and even though it's physical exercise, it tells me I've gone about 10K when I haven't, I've been on the back of a horse!"

(Police sergeant, female, age 18-39)

One interviewee suggested that the Fitbit® could be modified or adapted for those who work night shifts. As the current cut-off for measuring daily steps is midnight to midnight, this is designed for those with a typical 9 to 5 work pattern. This might result in feelings of discouragement when monitoring steps on a daily basis. This may be an important consideration when designing mHealth technologies for shift workers:

> *"Sometimes… my body clock isn't midnight to midnight. My body is seven in the morning until seven in the morning. If I do a night shift, I might sleep most of the day, then do a night shift, it will read for that day only 3,000 steps. Of course, I'm going to be awake another 12 hours yet. If you're a night worker, the data gives you a midnight cut-off, even though you're going to be awake for another 10 hours."*
>
> (Police inspector, male, age 40+)

Despite the comparatively lower perceived usability of Bupa Boost, participants recognised its main advantages were flexible and personalised goal-setting and its applicability to a wider range of behaviours including nutrition, relaxation and mindfulness. Suggested improvements were a better user interface, improved ability to synchronise with (i.e. transfer data from) the Fitbit® or other wearable devices, reminders to log in to the app and/or more automated tracking of activity and recording of goals, and fewer, more meaningful notifications:

> *"I just don't find it links very well with the Fitbit app. It doesn't recognise when you do hit targets and it doesn't sync in with statistics about how far you've run and time as well."*
>
> (Police constable, male, age 18-39)

> *"The app is pretty good, but if there was some way for it to remind you, at say six o'clock in the evening or something that you haven't ticked off any of your goals… That would be quite good, so you can be checking and getting streaks going and things like that."*
>
> (Police staff, male, age 18-39)

> *"I wasn't sure of marking the goals. I set myself a secondary goal of going to the gym… and then I thought, well I'm not actually getting any record of doing that, but I think that was my fault, where I should have clicked on when I've been to the gym, click on it and then that would have recorded it."*
>
> (PCSO, male, age 40+)

> *"... loads of things flagging up but nothing really meaningful. I got a few requests off people to challenge me for certain things. But I didn't know what*

was required in terms of accepting one of these notifications as a challenge… Just keep it simple."

(Police sergeant, male, age 18-39)

Interviewees also suggested that a fairer system for earning rewards and for competing against colleagues was needed. Acceptability of social features and rewards and perceived usefulness in relation to impact are discussed in **section 6.4**.

Figure 20 Summary of positive and negative perceptions of the Fitbit® and Bupa Boost app in relation to usability and usefulness

Fitbit®	**Bupa Boost**
+ Good user interface and easy to navigate + Range of features + Small and lightweight + Durable + Easy to charge + Practical to wear with police uniform - Not waterproof - Perceived as sometimes inaccurate and/or not suitable for some types of activity (e.g. horse riding, boxing) - Less useful for night shift workers	+ Flexible and personalised goal-setting + Good for monitoring diet/nutrition, relaxation and mindfulness as well as physical activity - Poor user interface and difficult to find colleagues - Not enough automated tracking of activity - Did not link well with the Fitbit (or other devices) - Too many (and meaningless) notifications - Points/rewards and competitions were sometimes perceived as unfair

6.3.3 Reasons for lack of engagement

Qualitative responses from the surveys and interviews allowed insight into the observed decline in engagement with the intervention over time. The reasons for lack of engagement with the Fitbit® and the Bupa Boost app, as reported by survey respondents at various time points, are detailed in **Table 60** and **Table 61**. Disengagement from the intervention was frequently related to issues of perceived usability and usefulness, as predicted by the Technology Acceptance Model (Davis, 1989, Holden and Karsh, 2010) but other factors also had an influence. There were some changes over time in the nature and frequency of reasons for disengagement.

Usability issues were reported by participants as reasons for disengagement throughout the study, and these became more frequent over time. For the Fitbit®, the main issues were problems with charging, broken devices and other technical problems such as synchronising with phones or computers. Many participants stopped using the Bupa Boost app as they perceived it as not user friendly and/or too complicated. Several reported not understanding how to use the app, despite being given a simple user guide at the beginning of the study (see **Appendix 4e**). Technical problems with Bupa Boost increased over time, and some participants reported that they had stopped using the app because it did not synchronise well with the Fitbit®. A small number of participants reported forgetting to log in to Bupa Boost throughout the study; this suggests that reminders may be needed, either through the app or from the study team. One participant stopped using Bupa Boost due to 'annoying' notifications. Over the longer term, some participants reported that logging goals manually became a chore and so stopped using the app.

Low perceived usefulness was also clearly related to disengagement, again particularly in the longer term. Some participants stopped using the Fitbit® and/or Bupa Boost app due to perceived inaccuracies in recording their activity. Others did not find the technology helpful or motivating; this was a frequent reason for choosing not to use Bupa Boost in particular, and more prevalent at the 8-month follow-up. At this same time point, six participants reported that they stopped using the Bupa Boost app as they did not wish to use the social components to connect with or compete against their colleagues; this included concerns over privacy.

Some participants were using an alternative activity tracker to the Fitbit®, and/or an alternative app to Bupa Boost; this was reported more frequently as the study progressed. The new technological tool(s) were perceived to be of higher usability and/or usefulness. For example:

> *"[The Fitbit] is more basic than my Garmin watch. The Garmin gives me more details on any physical activity I do. I think for someone that's really into their fitness and the friends that I know… the ones that… do a lot of training or that type of stuff, they will have a Garmin watch."*
>
> (Police constable, female, age 18-39)

For Bupa Boost, one of the most frequently reported reasons for disengagement was preferring to use the Fitbit® app alone. Many participants believed there was no need for the two apps, which they perceived to be similar in terms of function (as noted in **Chapter 3**, the Bupa Boost app only contained one behaviour change technique (BCT) that could not be accessed in the Fitbit® or its associated app, although use of social features was encouraged only within Bupa Boost). Perceived duplication in function was reported by interviewees (*"It almost became like doing the same thing twice"*). Participants generally perceived the Fitbit® app as more usable and useful than the Bupa Boost app, and so remained more engaged with the Fitbit® app.

In addition to perceived usability and usefulness, some additional factors influencing engagement with the PAW-Force intervention were identified. These included convenience and accessibility, health status and motivation.

Convenience and accessibility issues were frequently reported as reasons for lack of engagement. The nature of these issues changed over time, initially primarily involving short-term lifestyle and external factors such as lack of time and annual leave, and later based on longer-term compatibility with one's existing routine and habits. For some participants, if wearing the Fitbit® did not fit in with their lifestyle (e.g. at times of home renovation or high work stress), or if wear became burdensome, they stopped using the device. Others adapted use to their lifestyle and preferences, for example only wearing the Fitbit® as an exercise tool at times that were convenient to them.

Health status was reported by a small number of participants as a reason for lack of engagement and was a fairly consistent influencer throughout the study. Factors relating to health status included illness, injury and adverse reactions to wearing the Fitbit® (see **section 6.6**). For example:

> *"I did download the Bupa Boost, started, thought I would do quite well on it. And then I got poorly and it went downhill then. I didn't do any challenges and it just paled into insignificance, really."*
>
> (Police constable, female, age 40+)

Motivation was the final factor influencing engagement. Lack of enjoyment, boredom, and loss of interest increased in frequency over time, particularly for the Bupa Boost app. Some participants simply stated they did not like the app so did not use it. At the 8-month follow-up, one interviewee reported an observed decline in motivation and engagement amongst his colleagues:

"I know a lot of people at work went, 'I'm not using that [the Bupa Boost app] anymore. I can't be bothered to use it.' There was quite a bit of that."

(Police constable, male, age 18-39)

Table 60 Reasons for lack of engagement with the Fitbit® (as reported by survey respondents)

Reason	**Number of participants stating as a reason for not wearing the Fitbit**		
	Week 6 (n = 3)	**Week 12 (n = 5)**	**Month 8 (n = 24)**
On holiday / annual leave / out of force	3	2	0
Illness or injury	0	2	1
Adverse reaction to wear (e.g. skin irritation)	0	2	2
Prefers to use another fitness watch or tracker	0	0	5
Charging issues including lost charger or charging became burdensome	0	0	4
Lifestyle factors meaning not practical to wear or causing distraction (e.g. home renovation, work stress)	0	0	4
Perceived as inaccurate (including time and PA)	0	0	3
Broken Fitbit or strap	0	0	2
Technical problems (e.g. stopped working or syncing with phone or computer)	0	0	2
Became an encumbrance wearing both Fitbit and a wrist watch	0	0	2
Prefers to wear occasionally for sport only	0	0	2
Has been less active so prefers not to wear	0	0	2
Did not find the Fitbit helpful or motivating	0	0	2
Did not find the Fitbit enjoyable to use	0	0	1
Did not feel the need to continue to monitor activity as aware of PA level	0	0	1

Note: n = number of participants. Some participants stated more than one reason

Table 61 Reasons for lack of engagement with the Bupa Boost app (as reported by survey respondents)

Reason	**Number of participants stating as a reason for not using the Bupa Boost app**		
	Week 6 (n = 55)	**Week 12 (n = 60)**	**Month 8 (n = 104)**
Prefers to use the Fitbit app (or another app) / found no need for two apps	14	9	21
Lack of time (including due to work, lifestyle factors or unspecified)	12	14	7
Not user friendly / too complicated	9	13	17
Forgot to use the app or log in to update goals	8	5	6
Lack of technological skills / do not understand the app	7	4	5
On holiday (did not use in previous week)	5	1	0
Technical problems (e.g. unable to download app, not compatible with smartphone)	4	4	6
Did not find it helpful or motivating (including competitions perceived as unfair)	4	10	20
Not interested in using app	2	3	3
Illness or injury	2	4	4
Not using Fitbit therefore did not use Bupa Boost	1	0	9
Perceived as inaccurate for recording activity	1	1	0
App does not sync well with Fitbit	1	3	2
Do not like the app (unspecified)	0	2	2
Logging goals manually became a chore	0	1	3
Notifications were annoying	0	1	1
Lost interest over time / boredom	0	1	5
No desire to connect with or compete against colleagues (including privacy concerns)	0	0	6
Did not feel the need to continue to monitor activity as aware of PA level	0	0	1

Note: n = number of participants. Some participants stated more than one reason

There were large individual differences in levels of engagement with the intervention over time. For many, engagement was consistently high through the 8 months of the study:

> *"I've pretty much had the Fitbit attached to my wrist from the beginning. It's only left my wrist to charge for about an hour or so. It's on there pretty much permanently."*
>
> (Police inspector, male, age 40+)

Others reported fluctuations over time in their engagement with the technology. For example, one police officer stopped using the Bupa Boost app but then experienced a motivational pull to use it again:

> *"I really missed not going on the app, updating and getting my points up. I find it quite a good motivational tool. So I went back to it about two weeks after stopping."*
>
> (Police constable, male, age 18-39)

Another officer interviewed at the 8-month follow-up felt that the Fitbit® and the Bupa Boost app had already promoted awareness of her PA level, improved her motivation and changed her mind set, and so were no longer needed, but reported she would use the technology again in the event of a relapse in behaviour:

> *"It certainly helped in getting me motivated and getting back into being fit again. I've got back into that mind set now. If I slip again, I'd probably put it on and wear it every day again."*
>
> (Police sergeant, female, age 18-39)

6.4 Perceived impact on physical activity and sedentary time (and mechanisms of change)

The interviews confirmed the quantitative findings, provided explanations for the observed impact of the intervention on PA and sedentary time and gave a detailed insight into the mechanisms of behaviour change. In this section the findings from post-intervention and follow-up interviews are presented, together with survey data on preferred intervention phase (i.e. individual versus social). The findings are interpreted with reference to relevant BCTs (focusing on the CALO-RE taxonomy (Michie et al., 2011b)) and the COM-B model of behaviour change (Michie et al., 2011a).

6.4.1 Mechanisms of behaviour change and preferred intervention components

The interviews offered a potential explanation for why significant increases in daily step count were not observed for participants overall, yet self-reported PA did increase. Regardless of their activity level at baseline, many participants reported making changes to their usual activity type, which would not have been reflected in step count data (see also **section 6.3.2**). For example, some individuals had begun boxing or water-based activities (where it was not practical to wear the Fitbit®), and others reported more gym activity and strength training. This reaffirms the importance of including multiple outcomes to capture a wide range of types of PA.

Those with lower activity levels (and daily steps) at baseline perceived that the intervention had a larger impact on PA in both the short and longer term. Interviewees who were highly active at baseline perceived that the technology helped them to maintain, rather than significantly increase, their PA. This confirms the quantitative findings outlined in **Chapter 5** which demonstrated a larger impact on PA for participants who were less active at baseline. These differences are exemplified in the following quotes:

> *"Since I started this trial I have taken to walking a lot more; walk the dogs further and just a bit of a count with my husband, we go out walking the dogs and sometimes we get up to 10,000, which is brilliant. When I'm walking, I don't just saunter along, I try to get my heartbeat up to the maximum beat and then come down again. I've also started Pilates, which I appreciate doesn't really come into fitness, but it's just something else that is making me move. And I recently re-joined the gym again with a view of getting more fitness, more healthy in the future."*
>
> (Police constable, female, age 40+, <10,000 steps/day and 'moderate activity level' at baseline)

> *"Because I was already reasonably active before I started this... so, for me, it maintained, let's put it that way. I wouldn't say it definitely improved, but definitely maintained."*
>
> (PCSO, female, age 18-39, >10,000 steps/day and 'high activity level' at baseline)

According to the post-intervention survey, (see **Table 62**), the majority of participants preferred the 'individual' (56%) phase of the study to the 'social' (7%) phase. The individual components were goal-setting, self-monitoring, feedback, earning virtual rewards (i.e. wellness points and badges within the Bupa Boost app), and the Bupa library/self-help information. The social phase included the individual components in addition to the social feed (social comparison), messaging colleagues (social support) and competing with colleagues individually or as part of a team. Although there were large individual differences, the interviews confirmed that the individual components were generally perceived as more acceptable and most impactful. This is in accordance with the quantitative findings, where the addition of social components was not associated with any significant changes in PA or sedentary time.

Table 62 Preferred phase of study (individual vs. social), n = 91

Preferred phase	n (%) of participants
Individual: goal-setting, self-monitoring, feedback, earning virtual rewards (wellness points and badges), and Bupa library/self-help information	51 (56)
Social: social feed (social comparison), messaging colleagues (social support), competing with colleagues and taking part in company/team challenges	6 (7)
No preference / liked both phases equally	34 (37)

Note: Only participants who had used the Bupa Boost app during both the individual and social phases were asked to report their preference

One of the most powerful mechanisms of behaviour change appeared to be self-monitoring and awareness of PA and sedentary behaviour (SB), in addition to monitoring of behavioural outcomes such as heart rate and weight, in the Fitbit® and Bupa Boost app. Self-monitoring of behaviour and outcomes are two BCTs in the CALO-RE taxonomy (see **Chapter 3**). All of those interviewed commented on the usefulness of knowing their daily step count, PA level and time spent sedentary. For many, particularly those who were less active at baseline, this was sufficient to motivate them to be more active:

> *"Having it [the Fitbit] has been very useful… because it does show you your step count. I always thought I do quite a bit of activity, but my activity levels aren't as great as I thought they were really, so it has given me, dare I use the word 'boost' to do more really. It has certainly done that."*
>
> (Police constable, male, age 40+, increased steps and activity and reduced sedentary time from baseline to post-intervention)

Others reported that self-monitoring was useful for making them aware of times when they were least active (during shift work, for example) and helped them to self-regulate their behaviour:

"I think the Fitbit raises my awareness to the fact that on certain parts of my shift, I don't do very much. I can have days where I'm only doing 4,000 steps, 4,500 steps."

(Police inspector, male, age 40+, increased activity and reduced sedentary time from baseline to follow-up)

"I was more conscious that I wasn't exercising, when I wasn't. When I had exercised, I would always check the Fitbit app and it would be like… you've done how many steps, exercised 45 minutes today. The actual Fitbit along with the app, would motivate me more than the Bupa one. I think it makes me more conscious of what I'm not doing as well."

(Police sergeant, female, age 18-39, increased activity and reduced sedentary time from baseline to post-intervention and follow-up)

"I always walked the dog before, but if I knew I needed to do 10,000 steps, and I was like at eight and a half, I'd just do slightly longer. If you know you're really close to the 10,000, you'd give it more whereas if you have no idea, you'd just stop it when you stopped, wouldn't you."

(Police staff, female, age 18-39, increased steps and activity and reduced sedentary time from baseline to post-intervention and follow-up)

Goal-setting of behaviours and outcomes, and review of goals, within the Fitbit® and Bupa Boost apps, were also important BCTs in the CALO-RE taxonomy, and were reported by interviewees to be useful in increasing steps and PA. Many also used action planning. The importance of small, achievable goals for participants was clear.

"The goal-setting, I think it made me think about what I wanted to do each week. I would often look at it and go, "I need to run one more time this week."

(Police sergeant, female, age 18-39)

"I've definitely got an eye on doing 10,000 steps a day… I'm out doing a bit more walking as well, which I wouldn't have thought of previously. I've actually just gone out a couple of nights and just gone for a walk around the block, which I wouldn't have done before."

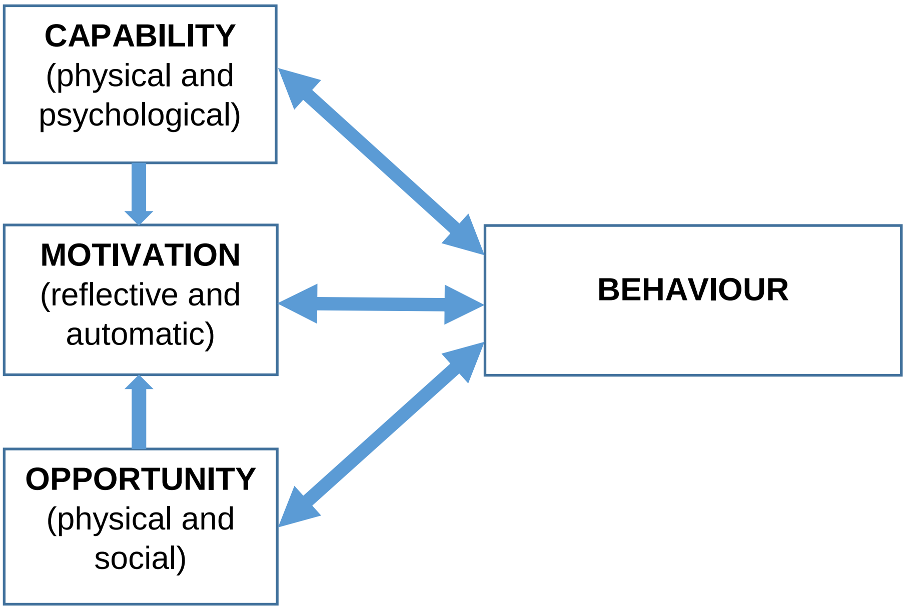
(Police sergeant, male, age 18-39)

There was evidence that feedback on performance, including reviewing behaviour (and outcomes) over the longer term, was also useful. One interviewee reported being motivated by the visual feedback within the Fitbit® app when achieving goals, and also by reviewing past success:

> *"I liked that it all went green when you hit your target. It sounds really silly, but it does make you want to do it because you want to hit it to go green, and it also - when you hit your 10,000 steps, it says, "Congratulations" or "Wow". It's just a little message to say well done. I thought that was good. I set the other targets as well, the amount of miles and the amount of stairs you climb or whatever it is. I set all of them basically. I wanted them all to go green at the end of the day was the idea."*
>
> *"You can look at how well you've done for the last week, but then you can go back a month, three months, a year - we haven't had a year yet - but that's quite good to see over a longer time period. How you have kept it up and how your fitness has changed."*
>
> (Police staff, female, age 18-39)

The COM-B model was proposed by Michie and colleagues as a component of the 'behaviour change wheel', a method for characterising and designing behaviour change interventions (Michie et al., 2011a). The COM-B model forms the centre or 'hub' of the wheel and links with intervention functions, which in turn link with policy categories, but COM-B may be used on its own to understand behaviour and behaviour change. According to this model (**Figure 21**), the three main determinants of behaviour are motivation, capability and opportunity. Capability refers to an individual's physical (e.g. skills) and psychological (e.g. knowledge) ability to perform a behaviour. Opportunity refers to external physical and social factors that influence behaviour. Capability and opportunity may directly influence behaviour or may increase motivation to carry out the behaviour. Motivation may be reflective (i.e. involve conscious decision-making) or automatic (e.g. habits and impulses). Enacting a behaviour may also influence capability, motivation and opportunity (Michie et al., 2011a).

Figure 21 The COM-B Model (adapted from Michie et al., 2011a)



Applying this model to the PAW-Force findings, interviewees felt that increased self-monitoring, awareness and feedback resulted in improvements to their motivation and psychological capability. There was evidence of increased reflective motivation for some interviewees, particularly for those who were less active at baseline:

> *"The job I was doing before was office-based. I was really, really aware that some days I could only do about 3,000 steps and I wasn't feeling very well with it, I was… starting to feel lethargic. That's when I thought, "Right, I'm going to try and make myself do more steps." The Fitbit itself did actually make me more conscious of what I was and wasn't doing. If I didn't have [the Fitbit], I wouldn't have any knowledge of how little I was doing."*
>
> (Police constable, female, age 40+, <10,000 steps/day and 'moderate activity level' at baseline)

> *"It really has made me think about my lifestyle, my activity or lack of, being conscious of, if I've had a really lazy day, I need to do something. I'm not going to get active sat in an armchair or at my desk."*
>
> (Police staff, female, age 40+, <10,000 steps/day and 'low activity level' at baseline)

In contrast, there appeared to be a ceiling effect for participants who were more active at baseline, and already intrinsically motivated:

"I've always done physical activity. So for me, it's been a good recording tool but it's not really made me do any more because I do it anyway and I'll probably always do it."

(PCSO, male, age 40+, <10,000 steps/day and 'high activity level' at baseline)

"It probably maintained my activity level. I didn't wear it to give me a kick up the backside to do my exercise because I'm getting told off for doing too much!"

(Police constable, female, age 18-39, >10,000 steps/day and 'high activity level' at baseline)

Prompts and cues, a further BCT in the CALO-RE taxonomy, were also an important influence, particularly in relation to SB. The 'reminders to move' function of the Fitbit® that encouraged participants to complete 250 steps every hour was perceived as useful by some:

"I did use the notifications. That was useful, especially when we had a meeting. I would stand up and walk around quickly and sit back down [laughs]."

(PCSO, female, age 18-39)

"At work, I would just start my shift, blink, and it's the end of my shift and I found that I was pretty much just sat behind the desk the whole time. Now, with the buzz facility on your wrist, every hour telling you to get up… It is a great reminder because you suddenly think, oh, blimey, it felt like two minutes ago that buzzed, I'd better get moving. So I'll probably just go for a quick two-three minute walk."

(Police inspector, male, age 40+)

For many others, however, prompts and cues were seen as insufficient to significantly reduce sedentary time in the workplace. This was primarily due to perceived pressure of work and organisational culture, which officers and staff did not feel permitted regular breaks in sedentary time during the working day. This illustrates the importance of context in influencing behaviour (see **Chapter 4**), and the need for opportunity (both physical and social), the third main component of the COM-B model:

"Unfortunately with the office role, at least, we don't have the ability to do the movement it requires every hour. It's pretty hard to fit in when I'm actually in the office, so I just turned it off."

(Police staff / special constable, male, age 18-39)

"I don't really have a choice. If I'm sat at a computer, working at a computer, I don't feel that I've got the ability to say... I guess, I do have the ability to say I'm going to leave this now and do something else. But that something else wouldn't necessarily be walking, it would be driving to another police station."

(Police inspector, male, age 40+)

"Well, because of the job I've got, it's quite difficult to-, if I said to my supervisor, "My Fitbit tells me I've got to get up and do 250 paces", I don't know how well that would go down."

(Police staff, female, age 40+)

It was generally perceived as more feasible to reduce sedentary time outside of work rather than within the workplace:

"I tried, but that wasn't particularly easy at work... it's quite difficult in my job. Because if it's busy, you can't just get up. Outside of work, it was easier. You're just sat on the sofa watching telly, and it reminds you to move, and you just get up and make a cup of tea or just walk around a bit, and then sit down again [laughs]. In work it's not always that easy."

(Police staff, female, age 18-39)

Despite these issues, some participants made small changes to reduce sedentary time in the workplace, such as walks during lunch breaks and taking the stairs instead of the lift:

"I used to just go downstairs with some food in my lunch break, but I have started now going into the town on work breaks as well. I've had a walk with some colleagues; there's a few of us that have started doing it. Just me and my colleagues in my staff job, we go out for a walk just to get out of the office."

(Police staff / special constable, male, age 18-39)

"When I've got it [the Fitbit] on, I'm very conscious of how I'm walking. I have been trying to take the stairs rather than the lift even though we're on the fourth floor."

(Police staff, female, age 40+)

Rewards for progress towards or achieving the desired behaviours were two additional BCTs included in the intervention. Several participants did not feel motivated by the virtual rewards (i.e. wellness points and badges) in the Fitbit® and Bupa Boost apps; these were generally perceived as less useful by those who were more active at baseline. This may be explained in terms of motivation; the additional extrinsic motivation did not have an impact on officers and staff who were already intrinsically motivated to be physically active. This is illustrated in the following quote by a sergeant with a 'high activity level' (according to the IPAQ) at baseline:

"When I've logged in to it [Bupa Boost app], it's all these little flashy awards start coming at you. I think that if you haven't got a history of exercise, it might give you a boost. You might think, "Oh, yes, I've achieved this." But that's not for me, really. I don't need lots of flashy stuff."

(Police sergeant, male, age 18-39)

Participants perceived that rewards should be fair and for meaningful achievements, in order to have an impact on their motivation. They also valued transparency in how points and rewards were calculated:

"You can earn badges for having so many friends and bits and pieces on there; I haven't really bothered with those ones so much. It's more the fitness-driven badges that interest me."

(Police constable, male, age 18-39)

"I think challenges for people to drink more water and maybe, do the steps or maybe, sleep better, I think they're all good. Just points for points' sake is... it can make a mockery of the trying to be well and wellness."

(Police inspector, male, age 40+)

"I genuinely don't know how it [Bupa Boost app] works in its points… it doesn't explain it anywhere within the app how it generates points and why it deducts points. That still niggles me to this day because there's some people who will do things just to earn wellness points… not by doing physical activities but

> *having a long list of achievements through the day… They just press the button for the sake of it so there'll be people that eat cheese and they'll say they fell asleep before ten o'clock and yet they pressed the button at five o'clock in the afternoon! There's a bit of that. It's almost like they're cheating in a way. I do find that quite frustrating, that people can get a lot more points and jump ahead of you actually when they don't do any activity or any exercise."*
>
> (Police constable, male, age 18-39)

One interviewee even perceived meaningless rewards within Bupa Boost to be detrimental to motivation:

> *"The levels, every time I went onto the Bupa app, it'd say, "Congratulations, you've moved up to level four", and I'm like, 'Well, I haven't done any exercise; I've not done anything to achieve that level.' I didn't understand the standards. It didn't mean anything, because you hadn't done anything to achieve it. That motivation was gone."*
>
> (Police sergeant, female, age 18-39)

While graded tasks and shaping were included to some extent in the intervention (e.g. graded rewards for greater achievements over time), one interviewee suggested this function could be more automated within the Fitbit® and Bupa Boost:

> *"I suppose the only thing I've set is 10,000 steps, that's my goal. I think if that was upped, so, say, if you're regularly hitting that target, we are going to up you to 12,000, I think that would be a target because people like targets, and then I'd probably go to the 12,000."*
>
> (Police sergeant, male, age 18-39)

General information on consequences of behaviour was the only BCT in the CALO-RE taxonomy that was present in the Bupa Boost app but not the Fitbit® (see **Chapter 3**). As noted previously, this feature (i.e. the Bupa library and self-help information) was accessed by only a small percentage of participants and was not mentioned by interviewees.

Although the 'individual' intervention components were preferred by the majority of participants, a small minority (7%, 6/91) preferred the social components and 37% (34/91) liked the individual and social phases equally (see **Table 62**). The interviews revealed considerable differences between participants, which appeared to be due to personal preferences rather than associated with any identifiable characteristics such as occupation or baseline activity level:

> *"I am not into the social aspect of it. It suits certain people... it certainly doesn't really suit me that much."*
>
> (Police constable, male, age 40+)

> *"I'd say [I prefer] the social stuff. It's okay to set yourself goals, but it's like you don't necessarily have that motivation unless you're one of these individuals that are fortunate to have that. It's fine to have aspirations to complete the goals, but if you're doing it as a group or you're competing against other people, I find… with me it motivates me more anyway."*
>
> (Police constable, male, age 18-39)

On the whole, the quantitative findings in relation to preferred phase supported this observation. For example, Chi-squared tests revealed no significant differences in preference according to age group (18-39 vs. 40+ years; $\chi^2_{(2)} = 1.53$, p = 0.465), mean daily steps at baseline (<10,000 vs. ≥10,000 steps/day; $\chi^2_{(2)} = 1.58$, p = 0.454), occupation (police officers vs. police staff vs. PCSOs and special constables; $\chi^2_{(4)} = 3.31$, p = 0.507) or education (secondary school vs. professional, technical or university; $\chi^2_{(2)} = 3.86$, p = 0.145). There was some suggestion of gender differences; the observation that females were more likely than males to prefer the individual phase, and less likely to prefer the social phase, neared significance (male vs. female; $\chi^2_{(2)} = 5.39$, p = 0.068). However, no gender differences were apparent from the interviews.

Participants were encouraged to link up with their colleagues to use the social components within the Bupa Boost app. These included social support, social comparison and competitions. The social support features focused on 'likes' and in-app messages of encouragement; these incorporated the BCTs of information about others' approval and planning social support / social change:

> *"There's a couple of people, I know them through work and I will send them messages and bits and pieces on the Bupa Boost, you can do the likes and all that stuff."*
>
> (Police constable, male, age 18-39)

The social support features were only used by a small percentage of participants (see **Table 57**), but there was evidence that social support and social opportunities for being physically active had increased ‘offline’ as well as ‘online’. Examples include the aforementioned group walks at lunchtime and “improved camaraderie” in the workplace that was reported by interviewees.

Social comparison and competitions included the social feed, competing against colleagues individually while viewing progress on a virtual leader board, and taking part in team/company challenges. These incorporated a further two BCTs in the CALO-RE taxonomy (provide normative information about others’ behaviour and facilitate social comparison). The participants that used these features perceived them as having a positive impact on their level of motivation and consequently PA, and several talked about social comparison and competition appealing to their competitive nature:

> *"I think future-wise, a lot of people in my organisation would probably be motivated by challenges. I think that would be a big factor in improving people's fitness, where we could clearly see others… If my phone was flashing up saying, "So-and-so has just been for a run", that would motivate me, because that's the way I work."*
>
> (Police sergeant, female, age 18-39)

> *"It's that tracking with other people. I quite like on the Bupa Boost app, you can challenge other people, you can set the challenge. I just press the button and include everybody, so there's about 400 people involved. Not everybody takes you up on it of course. I've set one up which we're halfway through at the moment, which is like a 30-day bike ride to see who can do the most kilometres in 30 days essentially. I'm constantly on that. I'm in second place at the moment. The guy that's ahead of me has got 50 kilometres on top of me. This morning now, I was thinking I really can't be bothered to get out my bike, I'm quite tired. Then I thought I need to get my kilometres up so I've got a chance of beating this guy. It's really motivational in that sense."*
>
> (Police constable, male, age 18-39)

> *"I liked the bit where you joined in with others... a bit of healthy competition! For each section... how many miles they could do in a week, and then we did that for a few weeks. That was a good motivator."*
>
> (Police staff, female, age 18-39)

"That's my nature. I'm very competitive. When I'm at work, I get very competitive [laughs]. When it came to the competing against each other... there was that stage where I pushed myself."

(PCSO, female, age 18-39)

Social accountability was one mechanism for increasing motivation:

"You know that your mates are watching you… you do something just so that people can see that you're doing activities and stuff."

(Police constable, male, age 18-39)

The interviews revealed some important considerations regarding the social intervention components. While there were no differences in preferred phase according to age or baseline activity level (see above), several participants expressed that they wished to compare and compete with colleagues of a similar age and activity level. For example:

"I think a lot of them that do the fitness stuff, they're lot younger than me and are probably a lot more competitive. It's probably a bit of an age thing. I couldn't really be bothered with competing with somebody who's 25, who's done 30,000 steps and you know, who thinks it's really exciting. It just doesn't do anything for me."

(Police constable, female, age 40+)

"I have a couple of friends on the app but they're either a lot less active than me or a lot more active than me… If I had a friend that was at a similar level to me and there was that little bit of, not competition necessarily but a little bit of shared support or, "Have you done your steps today? Do you want to go for a walk?" or whatever, that would be nice. I don't really have that."

(Police staff, female, age 18-39)

This would help to overcome the interpersonal barrier of perceived social pressure (see **Chapter 4**). Some preferred team competitions to individual competitions, as they felt this reduced such pressure:

"It was more when there was a team that's going to work together, that was the good thing. It's a little bit less pressure isn't it?"

(Police staff, female, age 18-39)

Other suggested improvements for the social components included ensuring team challenges were fair, with equal numbers of team members (*“some teams have got massively more staff than others”*), and easier communication within the Bupa Boost app:

> *“It took me quite a while to realise where it was when people would send me messages… I think the social phase could be a bit more fun if something like messaging and some form of blog was there. I think if you could almost have like a WhatsApp chat group among the people in your group. I think if you could communicate within the app a little bit easier it might be a little bit more social."*
>
> (Police inspector, male, age 40+)

Concerns regarding privacy and sharing of information with colleagues were also mentioned:

> *"I got a couple of requests off people. A couple who I worked with, but I didn't feel comfortable not knowing too much about it to start challenging other people, and a few people from Devon and Cornwall who requested as well, but I didn't know them so I just felt a bit uncomfortable with that."*
>
> (Police sergeant, male, age 40+)

6.4.2 Long-term behaviour change

Supporting the quantitative findings, the follow-up interviews provided evidence of longer-term behaviour change. Automatic motivation and habit formation were evident in terms of both wearing the Fitbit® and increasing PA:

> *"I see no reason why I would not wear it [the Fitbit]. It's just a bit of a habit now to always check and see how much I've slept and I like to know what my heart rate is when I'm out running and so on."*
>
> (Police constable, male, age 18-39)

> *"With walking, I wouldn't tend to go out walking but now I find myself after an hour just getting out and going for a walk around and checking my steps quite a lot so I still do it."*
>
> (Police sergeant, male, age 18-39)

After using the technology for eight months, some participants had adopted new behaviours to increase their PA level, such as joining a walking group or finding a personal trainer. This was mainly seen in those with a low baseline activity level. Others had adapted their previous activities, for example to replace habitual walks with runs:

> *"When I take the dog for a walk now... I bought a running belt, with my dog. I run with the dog which makes me run more, because he's got to be walked anyway. I'm doing quite a lot of longer dog walks than I perhaps did before."*
>
> (Police sergeant, female, age 18-39)

For some participants, there was evidence of improved reflective motivation through increased confidence and self-efficacy (i.e. belief in one's ability to perform a behaviour). For example, one police sergeant felt that achieving goals improved his confidence over the longer term:

> *"Because of the goals that I've achieved since I've worn it [the Fitbit], I do feel more confident."*
>
> (Police sergeant, male, age 18-39)

Some interviewees reported increased reflective motivation and changes in mind set that were sustained over the longer term:

> *"It's just made me more mindful of it. It's so easy when you're-, particularly in policing, when there are days when you feel like you're on a treadmill. When you just go to work, you're working flat out, you come home tired, all you want to do is have tea and go to bed. Had I not had this, that was probably my lifestyle before and, had I not had this, I'd probably be still doing that, to be honest. Whereas now, it's a case of, right, get home from work, I'm going to force myself to go out with the dogs for a walk for an hour just to get the steps up. Or I need to go out and go for a half an hour run or something like that. It's that motivator. It's just become normality now, I think."*
>
> (Police inspector, male, age 40+)

> *"It made me more aware of what I wasn't doing. It certainly had an impact on me from that point of view because I personally was getting into a bit of a rut with eating and not doing any physical exercise. For me, once I get into that stage, I tend to ignore it and hope it will go away. So when we got the Fitbits and started using them, it was a daily reminder of, look what you haven't*

done. It certainly helped in getting me motivated and getting back into being fit again. I've got back into that mind set now."

(Police sergeant, female, age 18-39)

For a small number of participants, there was evidence of declining engagement with the intervention over time, and initial increases in PA were not sustained. This appeared to be primarily experienced by those who were highly active at baseline and therefore did not feel that they needed continued extrinsic motivation from the activity monitor and app:

"I think when I started it I went out and tried it to see how far I could get. But as the project continued I probably went back to my normal exercise routine. In my case I don't think I've improved that much. Probably when I first got it I would say I tried to exceed what I was doing."

(PCSO, male, age 40+, <10,000 steps/day and 'high activity level' at baseline)

In contrast, participants who were less active at baseline appeared to show higher sustained engagement and PA in the longer term:

"I would say it's had a really positive impact on me. Funnily enough, just in preparation for this call, I had a look back right to the very start, last year, at my average step count and my average sleep. I think my average step count was around 40,000 a week, something like that, and it's up at around 55-60,000 consistently now, which is a big improvement. It's just got me focused on every step counts."

(Police inspector, male, age 40+, <10,000 steps/day and 'moderate activity level' at baseline)

The perceived impact of the intervention on sedentary time was fairly consistent throughout the study. Where there were initially perceived contextual barriers, or lack of opportunities for breaks in the workplace, these remained unchanged in later interviews. Consequently for some individuals, engagement with the reminders to move function declined over the longer term. They spoke about using this function initially but it later became 'irritating' or 'annoying' so they switched it off. For example:

"I did use it [reminders to move] initially but then I switched it off… Especially for the police, we'll do an activity and then you'll spend a long time doing paperwork. So you're sat down and your watch is buzzing and you can't get

up to move... I did switch it off at times because it became sort of an annoyance if you like. If you're having to do something you don't get that chance to get that break, because it's got to be done there and then."

(PCSO, male, age 40+)

6.4.3 Summary table of mechanisms of change

Table 63 summarises the key COM-B components and associated mechanisms of change that were identified in this section. The long-term behaviour change mechanisms are highlighted in bold.

Table 63 Summary of COM-B concepts and associated behaviour change mechanisms evidenced in the PAW-Force intervention

COM-B concept	**Part of intervention?**	**Examples of associated behaviour change mechanisms / intervention components**
Physical capability	No	
Psychological capability	Yes	Self-monitoring, awareness and feedback (behaviours and outcomes) Goal-setting and action planning Self-regulation
Reflective motivation	Yes	**Change in mind set regarding PA (long-term)** **Improved self-efficacy (long-term)**
Automatic motivation	Yes	Prompts and cues, i.e. reminders to move Virtual rewards (wellness points and badges) **Habit formation (long-term, including Fitbit wear and PA routines)**
Physical opportunity	No	
Social opportunity	Yes	Social support (‘likes’ and in-app messaging) Social comparison and competitions (social feed, leader board and individual/team challenges) ‘Offline’ activities – e.g. lunchtime group walks

6.5 Perceived wider impact

Although the quantitative findings (reported in **Chapter 5**) suggested that there were no significant changes in health and wellbeing outcomes throughout the study (with the exception of mental health-related quality of life at 8 months), the majority of those interviewed reported that they perceived wider benefits of using the intervention and/or increasing their PA level. These matched participants’ initial expectations of the intervention (see **section 6.2**). As detailed in **section 6.3.1**, many participants used the Fitbit® and/or the Bupa Boost app to manage their nutrition, mindfulness and relaxation. Interviewees reported that the intervention had a positive impact on these behaviours in the short (week 12) and longer term (month 8). For example, some used relaxation techniques in the Fitbit®, and/or set ‘mindfulness’ goals within Bupa Boost:

> *"I quite like the relaxation thing, which does make me slow down for a couple of minutes if nothing else. I've been trying to do mindfulness. The relaxation button, it does try to encourage you when to breathe and how to do it, so I think that's also been a help."*
>
> (Police constable, female, age 40+)

This offers one potential explanation for the observed significant improvement in mental health-related quality of life in the longer term, which was independent of increased PA.

Weight loss was a commonly reported outcome (particularly over the longer term), which was perceived as a consequence of both increased PA and goal-setting and self-monitoring of nutrition and weight within the Fitbit® and/or the Bupa Boost app.

> *"I've lost weight and that was one of my targets at the start. Well, from the first of January, I've lost 10 kilograms in about 12 and a half weeks."*
>
> (Police sergeant, male, age 18-39)

Several interviewees reported that the Fitbit® led to greater awareness of poor quality sleep (which they attributed to stress and shift work) throughout the study. Most perceived that their sleep had improved at the 8-month follow-up as a result of self-monitoring, and for some this highlighted the link between increased PA and improved sleep:

> *"The sleep count… that's a good reminder of your quality of sleep, how much sleep you're getting, and it's taught me that I have to get up at a certain time, so I can't influence how long I sleep in. What I can influence is what time I go to bed and that's got earlier as well as a result of the Fitbit and Bupa Boost."*
>
> (Police inspector, male, age 40+)

“Just monitoring the sleep patterns is quite interesting as well. I've noticed that if I ever work out later on in the evening, I sleep better. So that's something that the Fitbit app told me."

(Police sergeant, male, age 18-39)

"I liked that it tracked how much I slept… probably, the more exercise I did, the better I slept. I think that was noticeable.”

(Police staff, female, age 18-39)

For a small number however, monitoring sleep had negative consequences, leading to increased anxiety (see **section 6.6**).

Other perceived wider benefits as a result of increased PA were feeling fitter (including reduced resting heart rate), feeling healthier and having more energy, improved mood, and feeling less stressed. In the follow-up interviews, participants reemphasised the perceived links between PA, sedentary time and physical and mental health (as discussed in **Chapter 4**), and reported that self-monitoring using the intervention had increased their awareness of these relationships.

Several interviewees reported feeling less stressed as a result of increased PA in both the short and longer term. For example:

"I do probably feel healthier. It's hard to monitor your stress levels but I definitely feel healthier and I think that's probably reduced stress that I might have had."

(Police sergeant, male, age 18-39)

For those who felt able to take them, breaks in sedentary time were associated with reduced stress levels as well as feeling more energetic:

"I'll probably just go for a quick two to three minute walk. It's enough just to be able to disengage from all the stress and grief that we’re dealing with.”

(Police inspector, male, age 40+)

"It [the Fitbit] gets you out of the office and you can go for a little walk around and get some fresh air. I think when you get home from work, then you feel just a bit more energetic... I probably just felt more energetic, a bit more energy."

(Police sergeant, male, age 18-39)

This appeared to contrast with the lack of a significant change in overall perceived stress according to the quantitative data. However, as discussed in **section 6.4**, many interviewees experienced a lack of opportunity (or perceived opportunity) to take breaks in the workplace. It is possible that for the majority, the main impact on stress levels was observed outside of the workplace, which might not have been sufficient to improve the overall Perceived Stress Scale (PSS-4) score. This is expressed by the following interviewees:

"I do quite a stressful job anyway. I'm not sure if the Fitbit has made any difference to that. But in my free time, it probably does."

(Police staff, female, age 18-39)

"I've noticed no difference [in stress] at work. No, since you're so tied. If it was-, breaks were different and it wasn't so regimented, maybe that would be different."

(Police staff, female, age 40+)

Some interviewees reported improved morale and a sense of camaraderie amongst their colleagues which resulted from the social aspects of the intervention (e.g. social support and competitions). This offers another potential explanation for the improved mental health-related quality of life that was evidenced at the 8-month follow-up. For example:

"I feel there are benefits to having them [Fitbits]... that camaraderie and competitiveness between the team, to outstep each other, do that actual run, or do that extra time of physical activity. I think it's all useful and increases morale."

(Police constable, male, age 18-39)

Provoking discussion around PA and learning more about colleagues were also mentioned as positive social outcomes of the intervention:

"I think it's good that the Fitbit prompts debate and discussion around steps and activity. It definitely raises awareness by virtue of the fact that we're talking about it a lot more."

(Police inspector, male, age 40+)

"It's a surprise when you look at some individuals and think, 'Oh really!'... You start to learn more about others, a bit more than you have."

(PCSO, female, age 18-39)

No interviewees experienced a noticeable impact on productivity as a result of using the intervention. This is in line with the quantitative findings, where there were no significant associations between PA, sedentary time and Health and Work Performance Questionnaire (HPQ) scores. However, some reported experiencing improved mood and resilience in the workplace:

"Just by being more physically active you're just able to deal with things with a smile… whereas before it was probably with a grimace! [laughs] That's just a general effect of being more physically active. Yes, that's certainly noticeable."

(Police inspector, male, age 40+)

6.6 <u>Adverse events and unintended consequences</u>

A consideration of unexpected mechanisms of action and consequences is an important part of process evaluation of complex interventions (Moore et al., 2014). Despite this, few studies have explored the potential negative impacts of mHealth technology. These are included in this section.

Adverse physical and psychological effects of the intervention were reported by a small number of participants. Some experienced adverse reactions to wearing the Fitbit® activity monitor; five participants (approximately 3%) reported skin irritation, eczema, sores or swelling to the wrist. As a result, two of these participants stopped wearing the device before the end of the study, and one withdrew after six weeks.

The negative psychological impact of activity tracking was explored in the follow-up interviews. Although most interviewees did not perceive any negative impact, some reported feeling guilty when they did not meet their PA goals, and frustration or despondence when the Fitbit® prompted them to move from their desk or a meeting but they felt unable to do this due to external factors such as perceived organisational culture and pressure of work. For example:

> *"I can see a point of negative because if you don't do it [meet your goal], you'll feel a bit like, 'Oh, I didn't do it that day.'"*
>
> (Police staff, female, age 18-39)

> *"It made me feel bad because I know I'm lazy and I know I should do more."*
>
> (Police constable, female, age 40+)

> *"During work time it seems hard… well, it's quite depressing when I look at it because I've got such a sedentary job... I'm not moving and I don't have the opportunity to be moving around."*
>
> (Police staff, female, age 40+)

Another potential negative consequence is anxiety associated with heightened awareness of activity levels, heart rate and sleep, that may result for a small number of individuals. Two participants withdrew from the study for reasons of anxiety surrounding PA, heart rate and sleep. Three interviewees showed increased anxiety and cognitive rumination as a result of this heightened awareness of poor sleep:

> *"It makes you a bit more paranoid because you monitor your sleep, don't you? How many hours did I get to sleep last night? Then you see that you only had four hours or three hours or something like that."*
>
> (Police constable, male, age 18-39)

> *"If anything, it's made me a little bit frustrated about a couple of things. My average sleep is something like three and a half hours. It's made me realise that my sleep is absolutely shocking. I can have days where I've not slept more than an hour and a half for maybe five days. I worry a little bit more about my shifts and my health from my shift working."*
>
> (Police inspector, male, age 40+)

> *"The trouble is you could look at it and then you get overly anxious about how bad your sleep is. And then that actually can have quite a negative effect because then you're thinking, 'Oh, God, I'm not going to get much sleep tonight.' Or you look at it and go, 'Oh, I haven't got much sleep, so therefore, I*

feel tired.' I think fitness watches are great, but sometimes it can have quite, I think, a negative impact when you look at your results because you're overthinking it."

(Police constable, female, age 18-39)

6.7 <u>Summary / Discussion</u>

This section includes a summary and discussion of the key findings from this chapter, including previous experiences and expectations of mHealth and fitness technology, findings related to engagement, usability and usefulness, perceived impact of the intervention on PA, sedentary time and health and wellbeing, and the potential negative impact. Comparisons with existing studies are incorporated and the main contributions to knowledge are highlighted.

Prior experience and reasons for using mHealth technology have been shown to influence engagement and experiences of technology use (Carter et al., 2018). However, few studies of mHealth interventions have qualitatively assessed participants' expectations of the intervention prior to delivery. A strength of this study was to explore participants' previous experiences and expectations using both quantitative and qualitative data.

At baseline, participants had various levels of familiarity and previous experiences of mHealth and fitness technology. Reasons for participation were diverse, but typically focused on a desire to increase PA, reduce sedentary time and improve health and wellbeing. Participants were keen to use the Fitbit® activity monitor and Bupa Boost app and had high expectations that the technology would work for them. The components that were expected to be impactful were closely related to the BCTs identified in the CALO-RE taxonomy (Michie et al., 2011b) (see **Chapter 3**). Although there were individual differences, participants generally expected to be motivated by self-monitoring, goal-setting, and social influence (including social accountability and support); these are known to be amongst the most common behavioural techniques in activity monitors (Lyons et al., 2014) and health apps (Edwards et al., 2016). A small number of individuals expected to encounter problems with use of the mHealth intervention, due to a perceived lack of technological skills, but all of the participants with pre-intervention concerns reported having support from family members.

It appeared that prior to the intervention, many participants were at the ‘preparation’ stage of change according to the Transtheoretical Model (Prochaska and DiClemente, 1983), ready and prepared to make changes to their behaviour in the near future. Only a few participants were at the earlier ‘precontemplation’ and ‘contemplation’ stages. However, this is a commonly experienced issue in studies recruiting volunteers for PA interventions. High proportions of participants at the ‘preparation’, ‘action’ and ‘maintenance’ stages have been reported by similar mHealth intervention trials (Holmen et al., 2016). Despite this, a strength of the study is the diversity of the sample in terms of the wide range of views on PA and levels of motivation at baseline; while some participants were already highly intrinsically motivated, others felt they were in need of external motivation which they hoped the intervention would provide.

Overall, the findings showed that the diverse group of police officers and staff who were participating in the study expected the intervention to be impactful and acceptable in this context. Expectations were closely matched with actual experiences; these will now be discussed.

Relatively little is known about longer-term engagement with mHealth technology, and previous studies have been characterised by declining engagement over time (Afshin et al., 2016). A need for mixed methods studies, including qualitative interviews and reports of subjective user experiences, to measure effective engagement, has been recognised (Yardley et al., 2016). In response, this study contributed to filling the gaps in our understanding of engagement with mHealth interventions in the long term. Various aspects of engagement and acceptability were explored, with the use of qualitative interview data to complement the quantitative survey findings and to enable a deeper understanding of these issues, in both the short and longer term.

Engagement with the PAW-Force intervention overall was high. For example, 97% of survey respondents reported wearing the Fitbit® at 12 weeks and 83% at 8 months. This compares favourably with existing studies in this field, where percentages of participants still wearing a wearable device at the end of the study have been reported to range from approximately 50% to 75% (Simblett et al., 2018). Usage time for the Fitbit® was also consistently high in the short and longer term.

This higher than typical engagement level may be partially due to the high expected levels of compliance with device wear in the policing occupation, although there are no similar studies in this group for comparison. Engagement with the Bupa Boost app was comparatively lower and declined more rapidly over time, with only 27% of survey respondents still using the app after 8 months. Survey and interview data revealed that this was mainly due to lower perceived usability and usefulness of Bupa Boost. These findings support the Technology Acceptance Model (Davis, 1989, Holden and Karsh, 2010), which proposes that perceived ease of use and perceived usefulness are the principal determinants of acceptance and use of a new technology. The findings reported here indicate that this model is still relevant and that its principles apply to new mHealth technologies. A recent systematic review also reported that usability issues were the most common reason for dropout in studies of mHealth interventions (Simblett et al., 2018).

Participants' views on positive and negative aspects of the Fitbit® and Bupa Boost app were captured, together with suggested improvements for usability and usefulness. Some of these were general; for example, participants suggested a need for waterproofing of the Fitbit®, and an improved user interface and more automated tracking within Bupa Boost. A need for a more comprehensive user guide and greater technical support for the Bupa Boost app was expressed, particularly by those with lower self-perceived technological skills. Human support has been shown to increase both use and efficacy of digital health interventions (Patrick et al., 2016, Afshin et al., 2016). Some novel findings that may be of particular relevance to the police force also arose; for example, the colour, size and durability of the Fitbit® meant it was practical to wear with the police uniform. On the other hand, the Fitbit® may be less useful for officers (or staff) who work night shifts as the current cut-off point for measuring daily steps is midnight, which may not suit those with unconventional work routines. This demonstrates the importance of considering setting or context in mHealth interventions, and suggests the potential for tailoring for certain groups, such as the police and shift workers.

All of those interviewed felt that the Fitbit® had met or exceeded their expectations. At the end of the study, several interviewees commented that they were either planning to buy a newer model of Fitbit® for themselves, or that they had recommended the device to friends, family or colleagues. This adds to the evidence for high perceived acceptability of the wearable activity monitor component of the intervention. In contrast, participants generally perceived that the Bupa Boost app was more difficult to use, and some questioned its usefulness, referring to the duplication in function of the Fitbit® and Bupa Boost apps. This is perhaps unsurprising given that the Bupa Boost app only contained one BCT (provision of information on consequences of behaviour in general) that could not be accessed in the Fitbit® or its associated app. However, participants recognised some advantages of using Bupa Boost including more flexible and personalised goal-setting and its applicability to a wider range of health behaviours (i.e. nutrition, relaxation and mindfulness).

A qualitative exploration of the reasons for disengagement with the PAW-Force intervention confirmed and extended previous study findings in relation to mHealth technology. The findings closely matched the key factors influencing engagement with mHealth technology that were recognised in a recent systematic review and content analysis – usability, perceived usefulness (utility/value), convenience and accessibility, health status, and motivation (Simblett et al., 2018). All of these factors were clearly evidenced in the interviews. The novel contribution of the present study was considering changes over time, and which factors may moderate both initial and longer-term engagement. For example, while health status was a consistent influencer throughout the study, usability and usefulness issues increased over time (particularly for the Bupa Boost app), and motivation declined over time (again particularly for Bupa Boost). The nature of convenience and accessibility issues changed in the longer term; at first external circumstances (such as time and annual leave) were more important, and later participants considered whether the technology would fit in with their lifestyle and preferences and adapted its use accordingly.

Other studies have suggested individual and socio-demographic differences in levels of engagement, including differences by age (Patel et al., 2017), ethnicity (Widmer et al., 2016), gender, education, and income (Kohl et al., 2013). While individual differences in engagement were not explored in depth in the current study, this has clear relevance for the tailoring and personalisation of interventions and is worthy of future research.

Some participants showed fluctuations in mHealth technology use over time. This should not necessarily be seen negatively in all cases. This issue is recognised in the existing literature as a 'normal' part of technology use, in which engagement is seen as a dynamic process involving a cycle of engagement, disengagement and reengagement (O'Brien and Toms, 2008, Yardley et al., 2016). 'Effective engagement', i.e. sufficient engagement to achieve desired outcomes, may be more important than continued high use (Yardley et al., 2016). The present study showed that several participants (who appeared to disengage with the intervention) had opted to use an alternative mHealth technology 'tool' which better suited their needs. Others perceived that constant engagement with the intervention was not needed in the longer term as it had served its purpose in promoting awareness and motivation and changing their mind set. This is an important contribution to knowledge.

The study explored the perceived impact of the intervention on PA, sedentary time and secondary health and wellbeing outcomes, including mechanisms of behaviour change. The interviews provided support for, and helped to explain, the quantitative findings which suggested a positive impact of the intervention on PA, particularly for those who were less active at baseline. The interviews also elucidated the reasons for lack of an observed impact on sedentary time. The CALO-RE taxonomy (Michie et al., 2011b), concepts of intrinsic and extrinsic motivation, and the COM-B model (Michie et al., 2011a) were used to interpret the interview findings and enable understanding of the mechanisms of behaviour change. The COM-B model was selected for its logical and simple structure and clear concepts that were able to be identified through the interview data. The recognition of contextual and external influences on behaviour (i.e. opportunity) is a particular strength of the COM-B model, unlike many early psychological theories and models (such as the Health Belief Model - Janz and Becker, 1984) which limited the focus on behavioural determinants to the individual. The important influence of context on PA and SB in the police force was recognised in **Chapter 4**. Furthermore, the model had wide applicability as it had been successfully used in the development and evaluation of health interventions (including eHealth and mHealth) in different clinical and occupational populations (Handley et al., 2016, Keyworth et al., 2017, Herber et al., 2018).

In terms of the three core components of the COM-B model (capability, opportunity and motivation), the intervention seemed to have a large impact on motivation, and some impact on capability and opportunity. Both reflective and automatic motivation were enhanced by the intervention, mainly for those who were less active at baseline. In contrast, there appeared to be a ceiling effect for participants who were already highly intrinsically motivated, and who consequently maintained rather than increased their PA levels. The qualitative comparison of individuals with different levels of motivation was a unique contribution of this study. The intervention appeared to enhance psychological capability (with most pronounced effects for the less active participants) through mechanisms including goal-setting, self-monitoring, awareness, feedback and self-regulation. The intervention did not target physical capability to be more active (for example with skills training). Social opportunities to be physically active were provided by the intervention, both online (through mechanisms including in-app social support, social comparison and competition) and offline (for example, the intervention led some participants to initiate group walks during lunch breaks). However, the interviews highlighted the importance of context (see **Chapter 4**) and the need for additional opportunities to reduce sedentary time in the workplace. Perceived pressure of work and organisational culture appeared to be the most prominent barriers to reducing sedentary time; these issues are common amongst desk-based workers (Cole et al., 2015). Although some participants found the ‘reminders to move’ prompts delivered by the Fitbit® helpful, many did not feel they were able to leave their desks to take breaks. This shows the importance of providing physical opportunities and of challenging perceived social norms. Different strategies may be needed for PA and SB; an ecological approach that targets contextual and external factors has been recommended for interventions that aim to reduce SB (Owen et al., 2011, Spence et al., 2016). The inclusion of environmental restructuring may be an effective strategy for reducing sedentary time (Gardner et al., 2016).

Quantitative and qualitative data provided important insights into acceptability of the various intervention components. The majority of participants (56%) preferred the ‘individual’ to the ‘social’ features, with only 7% of participants preferring the social phase of the study. This is an interesting finding that aligns with the observed lack of an additional impact of the social phase on PA levels (see **Chapter 5**). This finding may be context-dependent in the police force, as previous studies have suggested that social components of mHealth can enhance engagement and/or behaviour change (Hamari and Koivisto, 2015, King et al., 2016). However, the interviews revealed important individual differences in preferences for, and perceived impact of, individual versus social components. This variation seemed to be more associated with individual personality differences than any identifiable characteristics such as occupation or baseline activity level. For example, social competitions appealed to those with a competitive nature. There was some suggestion from quantitative data that females were more likely than males to prefer the individual phase, and less likely to prefer the social phase. Research from video gaming has suggested similar gender differences, for example males tend to prefer competition whereas females may show less motivation to play in social situations (Lucas and Sherry, 2004). Although the interview findings in the present study did not show any clear gender differences, this may require further exploration in future studies. Nevertheless, such individual differences suggest the importance of tailored, personalised mHealth interventions. Personalisation and tailoring has previously been recognised as a key theme and such interventions have been recommended by mHealth researchers (Carter et al., 2018).

The perceived acceptability of some specific components, such as rewards and social competitions, was variable. Rewards were generally perceived as more useful by those who were less active at baseline and lacking in intrinsic motivation. Importantly, participants noted that the rewards system should be transparent and clearly calculated, which they felt was lacking in the Bupa Boost app. The high value placed by smartphone users on transparency of reward systems has been previously noted (Middelweerd et al., 2015). The present study additionally suggested that rewards should be fair and for meaningful achievements; meaningless rewards did not enhance, and in some cases, reduced, motivation. Participants wished to compete against those of a similar age and PA level to themselves, and desired easier communication with their colleagues within the Bupa Boost app. The issue of counterproductive competition, whereby competing against those of a much higher activity level can reduce motivation, has previously been reported in a qualitative evaluation of activity tracker use (Kanstrup et al., 2018). The PAW-Force study findings reiterate the importance of parity in activity monitor and app-based competitions.

Issues of security and privacy in relation to mHealth interventions for PA have previously received little attention in the literature (Carter et al., 2018), although a qualitative study of wearable devices in a workplace setting (focus groups with professional truck drivers) reported that employees had concerns surrounding monitoring of their health and lifestyle data by their employer (Greenfield et al., 2016). Privacy concerns were noted in the present study, and extended to sharing data with colleagues. The findings suggested that social competitions with colleagues should be optional, or perhaps sharing of information should be limited to colleagues within the same section or office. These factors might have contributed to the relatively lower uptake and perceived impact of the social features compared with the individual features.

In line with the quantitative findings, the interviews provided evidence of long-term behaviour change, i.e. sustained increases in PA after using the intervention for eight months. The main mechanisms were increased automatic motivation or habit formation (Fitbit® wear and increased PA), and increased reflective motivation through improved confidence, self-efficacy, and sustained changes in mind set relating to PA. Generally, the participants that were less active at baseline perceived the highest usefulness of the intervention over the longer term. They also showed greater sustained engagement and PA levels at follow-up; this has been reported in earlier eHealth studies. For example, Robroek and colleagues found that the least motivated and least active employees were more likely to sustain participation in an internet-delivered workplace health promotion programme (Robroek et al., 2012). These are promising findings for future interventions.

While the intervention was highly acceptable overall, particularly for the less active participants, there may be room for improvement to the number of BCTs and type of motivational strategies included. This may help to improve both engagement and the likelihood of behaviour change in future mHealth interventions. Although 20 of a possible 40 BCTs from the CALO-RE taxonomy were included in the Fitbit® and Bupa Boost app (see **Chapter 3**) and all were used by participants (to a greater or lesser extent), there was a desire for more graded tasks and shaping. Further techniques to promote self-efficacy, such as prompting use of imagery, generalisation of target behaviours and motivational interviewing, were not identified in the PAW-Force intervention, and have previously been reported as lacking in wearable activity monitors (Mercer et al., 2016). Similarly, BCTs related to detailed planning of the behaviour, instruction, modelling, problem solving, anticipated regret and fear arousal tend to be infrequent or lacking in common activity trackers (Mercer et al., 2016) and were not identified here. This may have important implications as these may be effective techniques to promote PA, with problem solving and modelling particularly effective in older adults (French et al., 2014).

Future interventions should aim to include these BCTs and address the missing components of the COM-B model. For example, future interventions could aim to enhance physical capability by including elements such as skills training or modelling, either as part of the mHealth tool or as a separate ‘offline’ component. Future workplace interventions seeking to reduce SB will need to increase physical opportunities and take a more socio-ecological approach, aiming to address barriers at multiple levels – individual, interpersonal, organisational and environmental.

Interviewees perceived a number of wider benefits to their health and wellbeing, which matched their initial expectations. These were partly due to use of additional features in the Fitbit® and Bupa Boost app (e.g. managing nutrition, mindfulness and relaxation) and partly as a consequence of increased PA. While the quantitative findings (see **Chapter 5**) showed no significant changes in health, wellbeing, or stress (with the exception of an improvement in mental health-related quality of life at 8-month follow-up), the qualitative findings suggested that there were various benefits that would not have been recorded by the assessment tools, such as improved fitness, improved sleep, and weight loss. Although this was beyond the scope of the present study, it suggests a need for consideration of other outcomes (including physical and physiological outcomes such as weight or Body Mass Index (BMI), heart rate and sleep) in future effectiveness trials. Future trials should further investigate the use of mHealth to improve sleep in the police force, or in shift workers in general. While the interviews revealed that some participants experienced reduced stress levels as a result of increased PA and breaks in sedentary time, it appeared that this was mainly experienced outside of the workplace, which might not have been sufficient to produce detectable changes in perceived stress (PSS-4) scores. This again emphasises the importance of context and external factors in influencing PA, sedentary time and stress within the workplace.

The interviews offered potential explanations for the long-term improvement in mental health-related quality of life (which was not associated with changes in PA). These included the use of relaxation and mindfulness features within the intervention, and/or increased morale and camaraderie in the workplace that resulted from social support and competition. Improved mood and resilience were also reported in follow-up interviews. Some interviewees reported having learned more about their colleagues as a result of participating in the study and commented that the use of Fitbits® had prompted discussion and awareness of PA in the workplace.

Consistent with the quantitative findings (see **Chapter 5**), there were no perceived changes in productivity in the short or longer term. It is possible that any changes were small and unnoticed by participants. Future trials of the impact of mHealth interventions should therefore carefully consider how productivity is assessed, whether through quantitative or qualitative self-report, objective absenteeism or presenteeism measures, or assessments of work performance. More objective measures and/or more sensitive, standardised and valid questionnaires may be needed. There may also be a need for longer follow-up (i.e. greater than 8 months) for changes in productivity to become apparent.

The study findings contribute to knowledge of the impact of digital interventions on health and wellbeing outcomes in a workplace setting, a field which is still in its infancy (Howarth et al., 2018). The use of mixed methods is a strength of the research; most existing studies have relied on quantitative surveys alone to explore the wider impact of digital health interventions (Howarth et al., 2018, Buckingham et al., 2019). Although the interview findings initially appeared to contrast with the quantitative results, they clarified the mechanisms through which positive changes to health and wellbeing occurred and suggested additional outcomes for exploration in future studies.

A final contribution of this study was to explore the potential negative impacts of mHealth technology. Adverse physical and psychological consequences of using the Fitbit® and Bupa Boost app were experienced by a minority of participants. A small number (5/180) experienced localised skin irritation as a result of wearing the

Fitbit®, and for three of these this was severe enough to discontinue wear. This has previously been reported in activity monitor interventions (Brakenridge et al., 2016b) and is an issue that should be taken into account by developers and manufacturers of wearable activity monitors.

The negative psychological impact of activity tracking has received limited attention in the literature, although awareness of this issue is increasing. Some of the study participants reported feelings of failure and guilt when they did not meet their PA goals. Others experienced anxiety associated with heightened awareness of their activity level, heart rate and sleep, which led to a negative cycle of cognitive rumination and a detrimental impact on behaviour. These findings mirror those of a mixed methods study of activity tracker use in young adults with depression and anxiety, which reported the negative impacts of feelings of guilt and increased anxiety (Kanstrup et al., 2018). A qualitative interview study also found evidence of 'unhealthy preoccupation' and 'obsession' relating to health and fitness app use in college students (Gowin et al., 2015). The present study shows that similar negative psychological effects may be experienced in an employee population. Police officers and staff also reported frustration resulting from the Fitbit®-delivered prompts to move during the working day but feeling unable to reduce their sedentary time due to lack of perceived opportunity for breaks.

While overall experiences were positive for the majority of study participants, the negative impact of activity tracking, and of mHealth use in general, requires further attention. It is important to consider the implications of such findings in future trials, particularly when recruiting those who already suffer from anxiety or are experiencing high stress levels. Participants should be informed of all potential adverse effects, both physical and psychological, at the outset of the trial in order to make a fully informed decision regarding participation. Some may require additional support from occupational health or an appropriate medical professional.

In summary, the results indicated that in general the PAW-Force intervention was highly acceptable to police officers and staff, although the usability issues with the Bupa Boost app would need to be resolved before it is rolled out more widely. The study improves our understanding of the factors influencing engagement with mHealth technology over the short and longer term and improves knowledge of how such technology works to facilitate behaviour change. Higher engagement with, and perceived impact of, the intervention was experienced by those with lower baseline activity levels. There was evidence that use of the intervention led to both short- and longer-term behaviour change. Context and the potential for personalised, tailored interventions were recurring themes and should be emphasised. These findings will aid designers of mHealth technology and intervention developers to produce more engaging, usable, useful and impactful mHealth tools in the future.

CHAPTER 7 FEASIBILITY FOR THE WIDER WORKFORCE AND REFLECTION ON STUDY METHODS

7.1 Introduction and overview

Key aims of this pilot study were to assess the feasibility and acceptability of the PAW-Force intervention for the wider workforce, to explore the feasibility of implementation and delivery of the intervention, and to assess the feasibility of the study methods. These aspects are addressed in this chapter.

In **section 7.2** the survey findings of perceptions of managers, commissioners and occupational health staff are reported, to assess whether the intervention was seen as feasible, acceptable and cost-effective at all levels of the Devon and Cornwall and Dorset Police forces. In **section 7.3**, an overview of implementation and delivery of the intervention is presented, followed by an exploration of multiple aspects of the feasibility and acceptability of the study methods in **section 7.4**. The chapter concludes with an integrated summary and discussion (**section 7.5**).

7.2 Feasibility and acceptability for the wider workforce: Survey with managers, commissioners and occupational health staff

Twenty-three managers, commissioners and occupational health staff were invited to complete the survey to assess perceived feasibility and acceptability of the intervention for the wider workforce (methods are described in **Chapter 3**). Of these, there were 10 respondents (shown in **Table 64**). Six of these were participants in the PAW-Force study. Seven respondents were employed by Devon and Cornwall Police, two were employed by Dorset Police and one represented the Devon and Cornwall Office of the Police and Crime Commissioner. The respondents included sergeants in a range of roles (all with supervisory responsibilities), a local policing inspector, a wellbeing manager, an exercise facilities manager, and a manager in criminal justice, partnerships and commissioning.

Table 64 Respondents in the managers, commissioners and occupational health staff survey

Respondent number	Organisation	Work role	Study participant?
1	Devon & Cornwall Police	Detective Sergeant	Yes
2	Devon & Cornwall Office of the Police & Crime Commissioner	Criminal Justice, Partnerships and Commissioning Manager	No
3	Devon & Cornwall Police	Response Sergeant	Yes
4	Dorset Police	Wellbeing Manager	No
5	Devon & Cornwall Police	Police Sergeant	Yes
6	Devon & Cornwall Police	Inspector - Local Policing	No
7	Devon & Cornwall Police	Detective Sergeant	Yes
8	Devon & Cornwall Police	Custody Sergeant	Yes
9	Dorset Police	Neighbourhood Policing Team Sergeant	Yes
10	Devon & Cornwall Police	Exercise Facilities Manager	No

The responses to each question will now be presented in turn.

Q1. Based on your knowledge and experience, do you think that mobile health (mHealth) and fitness technology (i.e. Fitbit® wearable activity monitor and 'Bupa Boost' app) is a useful intervention within your workplace? Please explain your answer.

Nine respondents perceived mHealth and fitness technology as a useful intervention within their workplace. Reasons for perceived usefulness were classed into three categories:

1. Improved awareness of physical activity (PA) and other indicators of health and wellbeing

   Information, knowledge and awareness of PA, fitness, sleep, resting heart rate and general wellbeing were seen as important benefits of mHealth technology. For example:

   > *"… it assists those trying to reduce their resting heart rate to understand what it does at various points during the day."*

   > *"Any item that makes you aware of your daily exercise, health, sleep quality and wellbeing must be a benefit."*

2. Increased motivation to keep active

   Several respondents commented on the perceived impact of the Fitbit® and Bupa Boost app on PA levels, through increased motivation. Self-monitoring and social influence from other officers and staff were both seen as important contributors:

> *"It has promoted the need to keep active and through the thirst to maintain a level of steps. This has been influenced by peer pressure and self-motivation when in possession of stats."*

3. Reduced sedentary behaviour in the workplace

   Three respondents reported the main perceived benefit to be a reduction in sedentary behaviour (SB) in the workplace. Information, awareness and prompts for those who are desk-based were perceived to have a positive impact on behaviour:

> *"It is very easy to sit at the desk for hours without moving – they [activity monitor and app] prompt you to be more active."*
>
> *"It motivates those who sit at a desk to get up and move around."*
>
> *"It provides a visible guide to how sedentary you are during a shift and prompts you to walk more."*

One respondent believed that mHealth and fitness technology was a potentially useful intervention but should be optional rather than compulsory for all officers/staff:

> *"It would be a useful offer for those interested in health and wellbeing but should be optional."*

Data protection and privacy concerns were also raised:

> *"There would need to be a clear agreement of what any related data could/should be used for."*

Q2. What do you perceive to be the main benefits (if any) of the use of mHealth and fitness technology for individual officers and staff?

Perceived benefits of mHealth technology for individual officers and staff were improved physical and mental health and wellbeing (n = 6), improved fitness levels (n = 4), reduced stress (n = 3) and reduced fatigue (n = 1). Three respondents stated that an increased awareness of health and wellbeing was an important benefit for individuals; the technology was seen as encouraging staff to *"start thinking about wellbeing and health"*, *"focus on health and wellbeing"*, and *"getting staff to think* *about health through the device and making sure they are getting up regularly, completing steps for the day, completing exercise challenges"*.

Two respondents referred to the perceived benefits of social components and competing against colleagues using the activity monitor and app. These respondents believed the technology enhanced *"natural competitiveness between officers"* and promoted *"healthy competition".*

Additional benefits, each mentioned by one respondent, were supporting a "*positive team culture*", and officers and staff *"viewing their employer as caring and being interested in their health and wellbeing"*.

Q3. What do you perceive to be the main benefits (if any) for the organisation?

All respondents perceived that the use of mHealth and fitness technology would have some benefits for the organisation; these were generally perceived to be a result of improved health and fitness of officers and staff. The most commonly stated organisational benefit was reduced sickness absence (n = 8). For example, as one respondent indicated, *"Regular movement and exercise during the day are likely over the long term to impact positively on sickness absence"*. Increased motivation, morale or happiness of staff were identified as potential organisational benefits by seven of the 10 respondents. Improved capability, efficiency and/or productivity were stated by five respondents as possible benefits of mHealth technology for the organisation. These were perceived as being due to improved health and fitness (e.g. *"fitter more capable staff"*) or as a result of breaks in sedentary time during the working day: *"Taking breaks and exercising are likely to improve productivity"* and *"I find that when I step away from the desk for five minutes I am more focused and productive when I return"*.

Additional perceived organisational benefits, each reported by one respondent, included reduced presenteeism, improved team focus, more resilient staff with greater ability to cope with stress, and *"the perception that the organisation cares about its staff"*. The latter was reported as a potential benefit for both individual staff and the organisation overall.

One respondent believed there were important potential organisational benefits, but echoed the views of another respondent that the use of mHealth technology in the workplace should be optional:

> *"People who wear them [Fitbits] have to want to wear them, otherwise there will be no benefits."*

Q4. Would you consider commissioning wearable activity monitors for individual staff within your department or organisation at a cost of approximately £80 per device?

All 10 respondents stated that they would consider commissioning wearable activity monitors at a cost of approximately £80 per device, for their team or organisation. This was the approximate cost of the Fitbit® Charge 2 device used in the PAW-Force study.

Q5. Do you have any other recommendations for policies or strategies to encourage staff to become more physically active / less sedentary?

Managers, commissioners and occupational health and wellbeing staff recognised the importance of the workplace as a setting for promoting a healthy lifestyle. For example:

> *"We are wanting to look after their health and wellbeing 24/7, therefore enabling work to be part of an overall lifestyle of looking after yourself is very important."*

The recommendations made by survey respondents for promoting PA and reducing SB were closely related to the suggestions of study participants (as detailed in **Chapter 4**). As emphasised by participants in interviews, recommendations from the survey included scheduled time during working hours for PA, fitness and/or wellbeing (n = 3) and organised fitness classes for officers and staff (n = 2). However, there were some differences in opinion surrounding whether exercise should be compulsory or optional, and when it should take place. For example, one respondent believed fitness classes should be mandatory during the work day, while another recommended optional fitness classes during lunch breaks:

> *"Bring in mandatory fitness sessions into shift patterns, to encourage team working, morale, fitness development and wellness… potentially circuit training which can be done at individual speeds."*
>
> *"Any strategy which encourages staff to take lunch breaks is likely to assist… and the opportunity for arranged fitness classes."*

Two respondents advocated the continuation of subsidised gym membership. One respondent suggested that wider use of mHealth and fitness technology (in particular the Fitbit® activity monitor) should be part of the strategy to promote PA and reduce SB in the police force.

Other key recommendations included support from senior management in reducing sedentary time, and monitoring progress against the Health and Wellbeing Strategy implemented by the Devon and Cornwall Police. Two respondents emphasised the importance of communication, encouragement and support from senior management to encourage staff to take lunch breaks and other breaks from sedentary time. This was also one of the suggestions made by study participants (see **Chapter 4**). A change in organisational culture and staff perceptions of the legitimacy of work breaks may be needed:

> *"Encouragement from senior management for staff to take meal breaks away from their desk – staff still feel reluctant to do this for some reason."*
>
> *"There needs to be communication from our senior management teams that breaks are encouraged, and in our lunch break, it is acceptable to exercise. There is still an uneasy feeling about going to the gym during lunch, almost like you are doing something wrong and 'sliding away secretly'. It shouldn't feel like this."*

A further two respondents recognised the importance of the 'healthy behaviours' component of the Devon and Cornwall Police Force Health and Wellbeing Strategy, which includes diet and nutrition and sleep quality in addition to PA. These respondents believed that the strategy worked well but should continue to be carefully monitored and updated. As one respondent stated:

> *"We are already evolved around the wellness strategy. However I think we need to remain receptive to new and innovative ideas and continue to respond to new ideas."*

7.3 Implementation of the intervention

While implementation is a core component of process evaluation of complex interventions (Moore et al., 2014), it is generally infrequently and poorly reported in mHealth intervention studies (Blackman et al., 2013). Implementation includes an assessment of what is actually delivered and how delivery is achieved (Moore et al., 2014). This section provides a brief overview of the main aspects of implementation of the PAW-Force intervention, including intervention fidelity, cost, adaptation and co-interventions.

Assessment criteria for fidelity of PA interventions include design fidelity, training, delivery, receipt and enactment (Lambert et al., 2017). Design fidelity was less relevant as the Fitbit® and Bupa Boost app were a prescribed intervention, rather than designed by the research team (see **Chapter 3** for a full description of the intervention). Training fidelity was also not applicable, as the intervention was self-facilitated with minimal human support except for e-mailed instructions and guidance from the researcher sent at regular time points (see **Chapter 3**). The researcher answered technical queries where possible and referred participants to online Fitbit® help guides and the Bupa Boost support team for unresolved queries. Reported technical issues included faulty devices, failure to synchronise data, problems downloading step data, difficulties logging in to the Bupa Boost app and incompatible devices. Additional issues in implementation of the intervention included loss of the Fitbit® or charger and broken or poorly fitting straps. A supply of spare Fitbits®, chargers and straps was kept at a central administrative location within the Plymouth Basic Command Unit (BCU) and replacements were issued as needed. Fidelity of delivery (i.e. use/engagement with the intervention), receipt (e.g. demonstration of knowledge, skills and self-efficacy) and enactment (e.g. performance of intervention skills) were assessed using quantitative and qualitative methods and are described in **Chapter 6**. Overall, intervention fidelity was high.

The cost of the intervention was approximately £80 per Fitbit® device, and the devices were provided to study participants at no cost, while they were employed by the participating police forces. The Bupa Boost app was provided to the police forces for free as part of an agreement with Bupa. As reported in **section 7.2**, the intervention was perceived by managers, commissioners and occupational health staff as cost-effective.

The intervention was delivered according to the protocol, with no adaptations made during the eight month study. Co-interventions were minimal; at the 8-month follow-up survey only six of 143 (4%) participants reported having taken part in a new PA programme in their workplace since the beginning of the study. This included becoming a new member of the local police force gym (n = 2), joining a running group with colleagues (n = 2) and starting a new exercise programme with the support of a police force trainer (n = 2).

7.4 Reflection on study methods

The following measures of the feasibility of study methods have been recommended by the National Institute for Health Research (National Institute for Health Research, 2019) and reported in previous studies of mHealth interventions in workplace settings (Buckingham et al., 2019):

- reach and recruitment;
- adherence and attrition;
- comparison of study completers and non-completers;
- feasibility and acceptability of data collection procedures and outcomes; and
- acceptability of study participation as perceived by participants

In the PAW-Force study, these aspects were assessed using a combination of quantitative and qualitative methods, and the findings are presented in this section (characteristics of interviewees are described in **Chapter 4**).

7.4.1 Reach and recruitment

Study participants were volunteers, recruited using a wide range of methods as outlined in **Chapter 3**. The precise number of officers and staff reached was unknown. However, due to the diversity of recruitment methods used it was expected that the majority of those employed at the Plymouth BCU and North Dorset territorial area would have been approached to participate. Several of those recruited reported they were aware of the study from multiple sources, and snowballing/word of mouth from colleagues was also an effective recruitment method. Many were enthusiastic and excited about taking part:

> *"We had one of the fitness teams come in to work and did our BMI and all different measurements. Somebody mentioned it then, that there was going to be the trial and I just kept looking out for the e-mail all the time thinking I want to take part in this."*
>
> (Police staff, female, age 40+)

It was apparent that the incentive of a 'free' Fitbit® was an important factor in encouraging recruitment. One interviewee reported that he would otherwise not have taken part:

> *"It's nice to get something free as part of a trial as well... Well, if it wasn't for the free Fitbit... no, I don't think I would have [signed up]. I would probably have just read it and ignored it."*
>
> (Police sergeant, male, age 18-39)

Participants felt that there was good reach and representation across the various work streams:

> *"It [the study], as far as I could see, stretched across a fair number of the work streams as well. Everybody had the opportunity to take part in it."*
>
> (Police sergeant, female, age 18-39)

The recruitment target of approximately 150 was quickly met and exceeded. During the two month recruitment period (April-May 2017), 241 officers and staff from the two study sites expressed an interest in taking part. Due to resource limitations, 35 of these individuals were added to a waiting list for future research studies and/or receipt of a Fitbit® (sampling methods are described in **Chapter 3**). An additional 16 potential participants did not complete the consent form. This left 190 officers and staff who consented to take part in the study, giving an approximate recruitment rate of 79% (the number completing consent relative to those expressing an interest in the study). A further 8 individuals were excluded post-consent (see **Chapter 4**).

Overall, the study participants were representative of the wider police force populations at the Plymouth BCU and North Dorset sites in terms of occupation, gender and ethnicity (see **Table 65**; note complete data for Devon and Cornwall Police and Dorset Police as a whole were not available). Approximately 10% (128/1268) of those employed at Plymouth BCU were recruited to the study, and the numbers were reasonably representative of the various occupational roles, although police officers were slightly overrepresented and police staff slightly underrepresented in the study. Data on average age for Plymouth BCU officers and staff were not available, although the modal age group for police officers employed at this site was 36-45 years, which was close to the mean age of 39 years in the study sample. The majority of individuals working at the four North Dorset sites were recruited to the study, i.e. 52 of approximately 73 officers and staff (71%). The sample was representative of occupation, gender and ethnicity of the police population in North Dorset. The average age of police officers employed in North Dorset was 41 years, also close to the mean age of 39 in the study sample.

Table 65 Comparison of study participants from the Plymouth BCU and North Dorset sites with the wider Plymouth BCU and North Dorset police populations: numbers in occupational roles, gender and ethnicity

	Plymouth BCU		**North Dorset**	
	n (%) in study sample	**n (%) in Plymouth BCU**	**n (%) in study sample**	**n (%) in North Dorset**
Occupation				
Police officers	81 (63)	654 (52)	33 (63)	45 (62)
Police staff	33 (26)	488 (38)	3 (6)	7 (10)
PCSOs and special constables	14 (11)	126 (10)	16 (31)	21 (29)
Gender				
Male	73 (57)	691 (54)	34 (65)	48 (66)
Female	55 (43)	577 (46)	18 (35)	25 (34)
Ethnicity				
White	125 (98)	976 (77)	52 (100)	70 (96)
Other ethnicity	2 (2)	14 (1)	0 (0)	2 (3)
Undisclosed	1 (1)	278 (22)	0 (0)	1 (1)
Total	*128 (100)*	*1268 (100)*	*52 (100)*	*73 (100)*

Note: Data provided by Performance & Analysis teams, Devon & Cornwall Police (2017) and Dorset Police (2017)
Percentages may not total 100 due to rounding.

7.4.2 Adherence and attrition

Attrition was defined as the percentage of participants who consented to participate and began the intervention but failed to provide any data at 8 month follow-up. The attrition rate was 21% (39/182). Conversely, the retention rate was 79% (143/182).

Of the 39 participants who did not complete the study, seven officially withdrew between completing the baseline questionnaire and the end of the study at 8 months. The following reasons were given (see also **Chapter 4** for participant flow diagram):

- Physical adverse effects of the intervention (i.e. skin irritation resulting from Fitbit® wear) (n = 1)
- Psychological adverse effects - anxiety surrounding PA and sleep as a result of heightened awareness from using the Fitbit® and Bupa Boost app (n = 2)
- Leaving the police force before the end of the study (n = 2)
- Lack of time to increase PA, particularly during the work day (n = 1)
- Did not wish to continue with study (unspecified) (n = 1)

A further proportion failed to complete questionnaires at each data collection point, with the number of non-completers increasing slightly as the study progressed, i.e. 19/178 (11%) at week 6, 25/176 (14%) at week 12, and 30/173 (17%) at month 8 (see **Chapter 4**).

Adherence in relation to engagement with the intervention is discussed in **Chapter 6**.

7.4.3 Comparison of study completers and non-completers

Study completers (i.e. participants who provided data for one or more outcomes at the 8-month follow-up) and non-completers (participants providing no data at 8-month follow-up) did not differ significantly on baseline PA level (objective or self-reported), baseline sedentary time, or baseline measures of any secondary outcomes (physical and mental health-related quality of life, perceived stress and productivity) (see **Table 66** and **Table 67**). As shown in **Table 68**, the two groups were also similar in terms of socio-demographics, with the exception of presence of a self-reported health condition. A significantly higher proportion of non-completers

(18/36, 50%) than completers (42/142, 30%) reported having a health condition at baseline ($\chi^2 = 5.36$, p = 0.021). Compared with study completers, non-completers were significantly more likely to have arthritis (osteo or rheumatoid) ($\chi^2 = 4.90$, p = 0.027), depression or anxiety ($\chi^2 = 10.77$, p = 0.001) or type 1 diabetes ($\chi^2 = 7.82$, p = 0.005).

Table 66 Baseline activity level of study completers and non-completers

Outcome	**Non-completers**		**Completers**		**t value [1]**	**p-value for difference**
	Mean (SD)	**95% CI**	**Mean (SD)**	**95% CI**		
Mean daily step count	9,764 (3,234) n = 29	8,534 to 10,994	10,721 (3,251) n = 138	10,174 to 11,268	-1.44	0.151
Total physical activity (minutes/week)	156.5 (129.1) n = 37	113.4 to 199.5	174.0 (99.6) n = 143	157.5 to 190.5	-0.89	0.373
Total physical activity (MET-minutes/week)	2,968.5 (3,065.8) n = 37	1,946.4 to 3,990.7	3,237.4 (2,378.4) n = 143	2,844.2 to 3,630.5	-0.58	0.566
Moderate-to-vigorous physical activity (MET-minutes/week)	1,620.0 (2,246.6) n = 37	870.9 to 2,369.1	1,744.1 (1,713.6) n = 143	1,460.8 to 2,027.3	-0.37	0.714
Sedentary time (hours/typical weekday)	6.97 (2.97) n = 37	5.98 to 7.96	6.27 (2.93) n = 143	5.79 to 6.75	1.29	0.198

Note: [1] Independent samples t-test completers vs. non-completers
Only participants providing baseline data are included. Study completers includes participants who provided any data at 8-month follow-up.
SD = Standard Deviation; n = number of observations; 95% CI = 95% Confidence Interval
p-values where significant (i.e. <0.05) are highlighted in bold

Table 67 Physical and mental health-related quality of life, perceived stress and productivity of study completers and non-completers

Outcome	Non-completers		Completers		t value [1]	p-value for difference
	Mean (SD)	95% CI	Mean (SD)	95% CI		
SF-12 physical component score[2]	54.71 (6.12) n = 31	52.47 to 56.96	54.15 (6.93) n = 141	53.00 to 55.31	0.42	0.678
SF-12 mental component score[2]	46.51 (10.94) n = 31	42.50 to 50.53	47.79 (8.87) n = 141	46.32 to 49.27	-0.70	0.487
Perceived stress (PSS-4 score)[3]	5.55 (3.79) n = 31	4.20 to 6.90	4.70 (3.07) n = 141	4.19 to 5.21	1.34	0.182
Combined relative absenteeism and relative presenteeism (HPQ score)[4]	1.00 (0.31) n = 31	0.89 to 1.12	1.09 (0.37) n = 143	1.03 to 1.15	-1.24	0.216

Note: [1] Independent samples t-test completers vs. non-completers
[2] Higher scores indicate higher physical and mental quality of life
[3] Higher scores indicate higher perceived stress
[4] Higher scores indicate higher productivity
Only participants providing baseline data are included. Study completers includes participants who provided any data at 8 month follow-up.
SD = Standard Deviation; n = number of observations; 95% CI = 95% Confidence Interval
p-values where significant (i.e. <0.05) are highlighted in bold

Table 68 Socio-demographic and occupational characteristics of study completers and non-completers

Variable	Non-completers	Completers	Test statistic[1]	p-value for difference
Age, mean (SD)	37.4 (10.5) n = 37	39.8 (9.3) n = 143	t = -1.38	0.170
Gender, n (%)				
Male	21 (57)	86 (60)	$\chi^2 = 0.14$	0.709
Female	16 (43)	57 (40)		
Years of police service, mean (SD)	10.7 (9.1) n = 37	12.4 (7.6) n = 143	t = -1.18	0.238
Police force, n (%)				
Devon & Cornwall Police	25 (68)	103 (72)	$\chi^2 = 0.28$	0.594
Dorset Police	12 (32)	40 (28)		
Occupation, n (%)				
Police officers	19 (51)	95 (66)	$\chi^2 = 3.40$	0.183
PCSOs and special constables	7 (19)	23 (16)		
Police staff	11 (30)	25 (18)		
Marital status, n (%)				
Married or civil partnership	21 (58)	91 (64)	$\chi^2 = 0.41$	0.523
Single, divorced, separated or widowed	15 (42)	51 (36)		
Health condition at baseline, n (%)				
Yes	18 (50)	42 (30)	$\chi^2 = 5.36$	**0.021**
No	18 (50)	100 (70)		
Smoking, n (%)				
Current or ex-smoker	13 (35)	44 (31)	$\chi^2 = 0.26$	0.611
Never smoked	24 (65)	99 (69)		
Alcohol consumption, n (%)				
Once a week or more	19 (51)	78 (55)	$\chi^2 = 0.12$	0.728
Less than once a week	18 (49)	65 (46)		
Education, n (%)				
Lower or upper secondary school	20 (54)	62 (43)	$\chi^2 = 3.84$	0.146
Professional or technical qualification	4 (11)	37 (26)		
University	13 (35)	44 (31)		
Previous activity monitor use, n (%)				
Yes	17 (46)	49 (34)	$\chi^2 = 1.73$	0.189
No	20 (54)	94 (66)		

Note: [1] Independent samples t-test used for continuous variables, Chi-square for categorical variables.
Only participants providing baseline data are included. Study completers includes participants who provided any data at 8 month follow-up.
SD = Standard Deviation; n = number of observations
p-values where significant (i.e. <0.05) are highlighted in bold
Percentages may not total 100 due to rounding.

7.4.4 Feasibility and acceptability of data collection procedures and outcomes

Table 69 shows the proportions of participants providing primary outcome data (i.e. Fitbit®-recorded daily step count) compared with self-reported PA data at each time point. Completeness of step count data was lower than that of self-reported outcome measures throughout the study. For example, 83% of remaining participants (143/173) completed the 8-month follow-up questionnaire compared with only 50% (87/173) that provided step count data at this time point.

Table 69 Completeness of outcome data: step count vs. questionnaires

Time point	Number (%)[1] of participants providing step count data	Number (%) of participants completing questionnaire[2]
Baseline (week 0)	167/180 (93)	180/180 (100)
Mid-intervention (week 6)	118/178 (66)	159/178 (89)
Post-intervention (week 12)	114/176 (65)	151/176 (86)
Follow-up (month 8)	87/173 (50)	143/173 (83)

1 Total number of participants excludes those who had withdrawn from the study
2 Partial questionnaire completers included

Qualitative data showed that many participants experienced no difficulty in downloading their step data from the Fitbit® website and uploading it to the online questionnaire or e-mailing it as an attachment to the researcher. They found the written instructions easy to follow:

> *"I found getting the step data off the Fitbit website and sending it through, that was relatively easy. It did need the crib sheet each time, but the crib sheet, it was 'Janet and John', which was very helpful indeed. I'm not particularly good at computing."*
>
> (Police inspector, male, age 40+)

Others reported problems obtaining or transferring their step data, either experienced by themselves or their colleagues (*"I know there were a few issues with people getting the data from the Fitbit site and downloading"*). This appeared to be the main reason for lower completeness of step data. Interviewees recommended having a designated 'workplace champion' to assist with this task in any future trials:

> *"It might be useful if... the likes of *** [PCSO] and *** [Police Sergeant]… could be champions, you could almost delegate it to them, to make sure that*

everybody gets it done... for five minutes just to guide them through it, that wouldn't be too much of an onerous task."

(Police inspector, male, age 40+)

Some participants wished to download data and complete questionnaire surveys on their mobile phone but experienced some technical problems with obtaining step data and completing questionnaires using this method:

"I would have liked to be able to do the questionnaire on my phone but I couldn't do that… I don't know if some people could do it on their phones, whether it's just me having an old phone… Because you're using an app, you're using the Fitbit app, you're using the Bupa Boost app. If you could have done the questionnaires somehow through the app that would have made it easier rather than thinking I've got to do that, I've got to get the laptop out and work that out."

(Police staff, female, age 18-39)

Interview data suggested that acceptability of the questionnaires was high and perceived burden of completion was low. Participants reported that questionnaires were easy to complete and that there was sufficient time to complete questionnaires and take part in interviews:

"I thought it was really easy. I didn't think any of the questionnaires that you sent out were challenging or difficult or time consuming."

(Police inspector, male, age 40+)

"It wasn't intrusive, it was fine. I've done a couple of interviews on the phone and they're no more than we would expect really."

(Police constable, female, age 40+)

Participants also appreciated the use of the International Physical Activity Questionnaire (IPAQ) to capture data on a wider range of physical activities in addition to step count. It was perceived by some as very important that activities such as boxing, swimming and strength training were recognised (see also **Chapter 6**):

"My steps of 12,000 a day, I set myself. But it's different because… I think if you're a runner you would hit those really easily, where it's actually quite a

disadvantage to people like me because I do CrossFit and I'm doing a lot of Olympic lifting as well, which exerts your body but that can't be measured in steps."

(Police constable, female, age 18-39)

Completeness and quality issues were experienced in relation to secondary data. The lack of availability of sickness absence data prevented analysis of this outcome for the North Dorset study site. For Plymouth BCU, some information was available but it was subject to data quality issues (see **Chapter 5**).

7.4.5 Acceptability of study participation

At the 8-month follow-up, all interviewees reported a positive experience of participating in the study. Frequently used terms were 'enjoyable', 'something different', 'fun' and 'motivational'. For example:

"I've found that it's been really enjoyable in the sense that you've been able to socialise and compete with people at work. I found it hugely motivational."

(Police constable, male, age 18-39)

"I suppose it's been a bit different and quite exciting in that respect. It just takes something small sometimes to just focus people and get people motivated. I think it's worked. I'm grateful for the Fitbit and being part of it really."

(Police sergeant, male, age 18-39)

Some showed a clear willingness to take part in future similar studies or trials:

"If there's any more [studies] coming on I'll do it again!"

(PCSO, male, age 40+)

Interviewees appreciated the opportunity to focus on their physical and psychological health, and felt that the study demonstrated that the organisation cared about their health and wellbeing:

"Overall, I think it's a really useful bit of work… It's a massive step in the right direction… We're all having to do more, having to work longer, under more

stressful conditions. Health and wellbeing's now getting the spotlight it really does need."

(Police inspector, male, age 40+)

Others suggested that the organisation, particularly higher level managers, should have greater involvement in such PA and wellness programmes. The following quote also reemphasises the importance of the wider context and the need for more opportunities for PA in the workplace:

"It's just having the organisation getting involved a bit more. We were utilised as the guinea pigs so to speak, but the hierarchy could have been involved and actually given us times when we could go out and actually do stuff or put events on within the force whilst the trial was going on… Hopefully it'll have a positive effect where we work. I would like to think that our senior management would look at it and maybe alter a few things within the organisation."

(Police constable, male, age 18-39)

7.5 Summary / Discussion

Overall, the PAW-Force intervention was perceived as feasible and acceptable for the wider workforce. All respondents of the survey with managers, commissioners and occupational health staff were supportive of the intervention, and believed that mHealth technology (i.e. the Fitbit® and Bupa Boost app) was potentially useful within their organisation. Benefits were perceived for individual officers and staff (such as improved fitness, health and wellbeing) in addition to organisational benefits such as increased resilience, improved morale and productivity. These were similar to the expected benefits reported by interviewees in **Chapter 6**. The intervention was low cost, and importantly perceived as cost-effective by the managers, commissioners and occupational health staff surveyed. Cost-effectiveness is seen as a major advantage of mHealth technology in general (Direito et al., 2017, Stephenson et al., 2017, Sullivan and Lachman, 2016), and there is evidence that the use of wearable fitness technology as part of a workplace wellness programme can lead to substantial healthcare cost savings for employers (Daniels et al., 2016).

Whole-system approaches to workplace wellness programmes, involving support from all levels of the organisation, are associated with successful behaviour change and improved health outcomes (Hendriksen et al., 2016b, Brand et al., 2017). The finding of support from managers, commissioners and occupational health staff is therefore encouraging. It was clear that those surveyed perceived the workplace as an important setting for promoting PA, health and wellbeing, and the organisations' focus on health and wellbeing was appreciated by officers and staff who were interviewed. It was apparent that mHealth was seen as one component of a wider health promoting culture within the police forces, and the consensus was that the use of this technology should be optional rather than compulsory for all officers and staff.

Managers also recognised the influence of contextual factors on PA and SB, a recurring theme throughout the thesis. Suggestions for promoting PA and reducing sedentary time in the workplace closely matched those of interviewees (see **Chapter 4**), and included scheduled time for PA or wellness, organised fitness classes, and a change in organisational culture regarding breaks from sedentary time. More visible support and active involvement of the most senior managers was recognised as a need by both the participants interviewed and managers surveyed. This should be addressed in future wellness programmes within the police force.

A consideration of implementation of the intervention is a strength of the research, as this has been underreported in previous mHealth intervention studies (Blackman et al., 2013). The intervention was feasible to implement and fidelity of delivery, receipt and enactment were high, i.e. the intervention was delivered as planned, associated with high engagement and well received by participants in both the short and longer term. As reported in **Chapter 6**, the Fitbit® was perceived by the majority as more usable and useful than the Bupa Boost app. While designed as a self-directed intervention, many participants required technical support during the study, and some needed assistance with downloading their step data. Due to limited resources, support with such queries was provided by a single researcher. The need for technical support in trials of wearable activity monitors has previously been reported (Harrison et al., 2014) and should not be underestimated when planning new mHealth interventions. Some participants suggested having a workplace champion (or champions) to assist with data downloads and oversee the PA challenges within Bupa Boost. The use of one or more individuals to perform these roles should be considered for any future effectiveness trials.

The desire for a workplace champion to facilitate mHealth interventions for PA has similarly been reported by other employee populations including nurses (Torquati et al., 2018). Although the present study was designed as a standalone mHealth intervention, there is evidence that workplace health promotion interventions that include both ‘online’ components and ‘offline’ elements such as human support are more likely to be successful (Nuffield Health, 2018, Schoeppe et al., 2016). Future trials would need to consider training fidelity to ensure consistent input and support across sites by workplace champions.

Due to technical issues and loss of equipment (another common issue with this type of intervention - Harrison et al., 2014), it was also necessary to have a supply of spare Fitbit® devices, straps and chargers and to ensure participants were aware of how to obtain these. For many, this promoted continued engagement with the intervention and study.

A reflection based on quantitative and qualitative data indicated that study methods were feasible and acceptable for participants. The findings suggested some possible ways in which future similar studies or a larger trial within the police force could be enhanced. Recruitment methods were effective and efficient; the target was quickly met with a waiting list of potential participants for future studies. The incentive of a free Fitbit® was an important contributing factor; incentives have been associated with higher participation in previous workplace health promotion programmes (Robroek et al., 2009). Reach was high due to the wide range of recruitment methods used.

Overall, study participants had a similar socio-demographic profile to the officer and staff populations. However, despite the use of maximal variation sampling to aim for representation of the various occupational groups, proportions of staff recruited were slightly lower than expected relative to the wider population at the Plymouth BCU site. Future recruitment methods should therefore aim to target police staff. All participants were volunteers (i.e. self-selected rather than randomly selected or systematically referred from the target population); this may be seen as contributing to selection bias (Thomas et al., 2004). However, recruitment of volunteers is the typical method used in workplace interventions, including those with an mHealth component (Buckingham et al., 2019). The reporting of reach and representativeness is another strength of this research as these aspects tend to be underreported (Blackman et al., 2013, Buckingham et al., 2019). Complementing the quantitative findings, interviewees perceived good reach and representativeness across the two police forces.

Attrition was reasonably low, particularly considering the relatively long duration of the study in comparison to existing studies in this field (Buckingham et al., 2019, Afshin et al., 2016). The overall attrition rate (i.e. the percentage of enrolled participants who failed to provide any data at the 8-month follow-up) was 21%, lower than many previous studies of mHealth interventions to promote PA and reduce SB in the workplace, which have reported attrition rates of up to 74% (Buckingham et al., 2019, Torquati et al., 2018). Adverse consequences of the intervention, both physical and psychological, were the reason for the dropout of a small number of individuals (see **Chapter 6**); it is important to consider these and inform participants of potential adverse outcomes during enrolment in future trials.

A comparison of study completers and non-completers found little evidence of bias in dropout. The two groups were similar in baseline PA level (objective and self-reported) and self-reported sedentary time. Study completers and non-completers were also similar in terms of a range of socio-demographic and occupational characteristics, and there were no differences in likelihood of completing the study between the two participating police forces. This contrasts with a previous questionnaire study of obesity in police officers, which reported that older and more experienced officers were less likely to adhere to the research protocol by fully completing the questionnaire (Can and Hendy, 2014).

The only significant difference between completers and non-completers in the present study was that a higher proportion of non-completers reported a health condition at baseline. The qualitative findings also noted that health status was a barrier to both PA (see **Chapter 4**) and engagement with the intervention (see **Chapter 6**). It is conceivable that those with arthritis were less likely to adopt and/or maintain increased PA levels to the end of the study and so did not wish to provide final outcome data. Those with anxiety and depression were also less likely to complete the study; this may be for the same reason as these conditions are associated with lower sports participation and general activity (Hiles et al., 2017).

Two participants withdrew for reasons of increased anxiety associated with the intervention (see **Chapter 6**). It is important that all participants, especially those who are most vulnerable, are made fully aware of potential negative psychological consequences at the outset of any trial. However, those with chronic health conditions (both physical and psychological) may receive the greatest benefits from increased PA (Penedo and Dahn, 2005) and from mHealth technology use (Phillips et al., 2018), and should therefore not be dissuaded from participating in such trials.

The results indicated that data collection procedures were feasible. Overall outcome completion was high, with 83% of police officers and staff remaining in the study at 8 months completing all or part of the final follow-up questionnaire. According to the interviews, participants were satisfied with the questionnaires and reported low burden of completion. Participants valued the use of the IPAQ to capture a range of activities in addition to steps. One possible suggested improvement was for the questionnaire surveys to be more easily completed using a smartphone. A key finding was that completeness of step count data was lower than self-reported PA and secondary outcomes. Future studies could overcome this issue by using a different primary outcome or by exploring alternative ways of obtaining data from the Fitbit®. Methods reported by previous studies have included use of the Application Programming Interface (API) for the wearable device (Harrison et al., 2014), management of individual Fitbit® accounts by the researcher (Yeung et al., 2017) or the use of social app functions to track participants' PA (Jones, 2016). Workplace champions could also provide valuable support with data downloads. A lack of complete and accurate data on sickness absence was another issue that should be resolved if this outcome is to be captured in future trials.

In summary, the intervention was feasible to implement and deliver, and perceived as highly acceptable by officers and staff at various levels of the police forces involved. Study methods were generally feasible and acceptable. Participants reported a positive experience of study participation, which was seen as novel, enjoyable and motivational. Wider implementation and upscaling of the PAW-Force intervention and research methods should be carefully considered given that this was an uncontrolled pilot study in a fairly specific setting. Nevertheless, the study has produced some important implications and recommendations for the design and conduct of future mHealth intervention studies in workplace settings; these are discussed further in **Chapter 8**.

CHAPTER 8 INTEGRATION OF FINDINGS, DISCUSSION AND CONCLUSIONS

8.1 Introduction and overview

The aim of this chapter is to summarise the findings of the systematic review and PAW-Force study, to explore the separate and combined contributions of the quantitative and qualitative methods in the PAW-Force study, and to consider the implications of the findings for academia (i.e. what the research adds to existing theory and evidence), public service (policy and practice in the police force) and future workplace wellness programmes. An overview of the key findings is given rather than a comprehensive discussion, which has been included in the preceding chapters (**Chapters 4 to 7**).

A brief discussion of the systematic review findings is first presented in **section 8.2**. A revised logic model for the PAW-Force study is presented in **section 8.3**, then the answers to the research questions introduced in **Chapter 1** are summarised in terms of quantitative findings, qualitative findings and contributions to knowledge (**section 8.4**). Strengths and limitations of the PAW-Force study are discussed in **section 8.5**. This is followed by recommendations and impact for the police force (**section 8.6**) and directions for future research, including recommendations for mHealth interventions in workplace settings (**section 8.7**). The overall knowledge contributions of the PhD are presented in a summary table in **section 8.8**, and final conclusions are given in **section 8.9**.

8.2 Summary of systematic review findings

A systematic review of mobile health (mHealth) interventions to promote physical activity (PA) and reduce sedentary behaviour (SB) in workplace settings was conducted alongside the PAW-Force study. The purpose of the review was threefold: to summarise the existing evidence base and provide a background for the research; to identify gaps in the literature and subsequently inform the specific research questions in the PAW-Force study; and to use the findings to guide the study design.

The review was the first known published review of studies of mHealth interventions conducted specifically in workplace settings. The limitations of previous reviews had included poor reporting of review procedures and included studies, the omission of grey literature, and a lack of a comprehensive description of interventions (Bardus et al., 2016). These issues had resulted in low confidence in findings and conclusions. The review overcame many of these limitations with clear and comprehensive reporting of procedures, study designs and outcomes, the inclusion of grey literature, and systematic coding of the theoretical components of interventions in the included studies. Compared with previous reviews, this review was more comprehensive by considering multiple aspects of mHealth – feasibility, acceptability and engagement in addition to effectiveness.

Despite heterogeneity of the included studies (which precluded meta-analysis), there was reasonable evidence that mHealth interventions may be useful in workplace settings for promoting PA, at least in the short term. This was in line with the findings of reviews of studies in non-workplace contexts (Fanning et al., 2012, Bort-Roig et al., 2014, Muntaner et al., 2016, Schoeppe et al., 2016, Direito et al., 2017). The longer-term impact was less clear as interventions and studies tended to be short in duration; other reviews had similarly reported typical study durations ranging from a few weeks to six months (Fanning et al., 2012, Bort-Roig et al., 2014, Afshin et al., 2016). This highlighted the importance of exploring both the short- and long-term impact, and the need for a study with longer than six months' follow-up (taking into account the constraints of the PhD timescale). Relatively few studies had used qualitative methods to examine feasibility and acceptability of mHealth interventions; this gap was an important one to fill. Additional gaps identified in the review included the unclear impact of mHealth on SB and wider outcomes such as health, wellbeing and productivity; these also led to research questions in the PAW-Force study.

The limitations of the included studies were used to guide the design of the PAW-Force study. In many previous studies, interventions and outcomes had not been clearly defined or reported. In the PAW-Force study, the intervention (including the component behaviour change techniques (BCTs)) and outcomes were clearly reported. Some of the reviewed studies had relied on outcome measures with poor or unknown validity; it was ensured that the primary outcome in the PAW-Force study was objective, valid and reliable. Inadequate control and/or reporting of confounders was an issue in several of the reviewed studies; this emphasised the need to consider, capture and report confounders (i.e. factors influencing PA levels) in the PAW-Force study.

8.3 Revised logic model

Figure 22 shows the logic model produced using the quantitative and qualitative findings of the PAW-Force study. This is an updated version of the pre-study logic model introduced in **Chapter 3** (**Figure 3**). The main changes and additions are as follows:

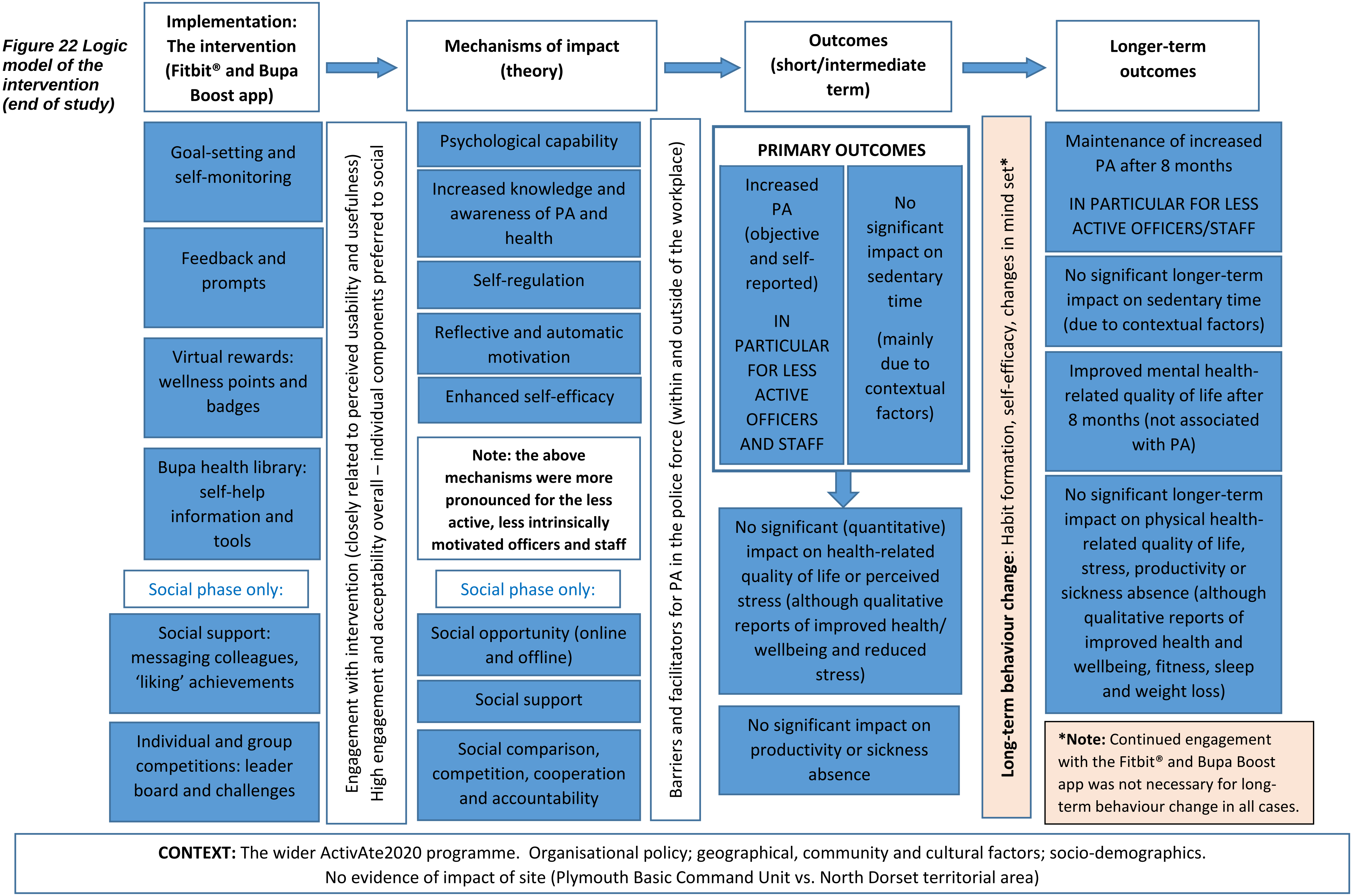
- Engagement with the intervention was high overall, and closely related to perceived usability and usefulness (see **Chapter 6**)
- The findings showed that continued engagement with the Fitbit® and Bupa Boost app was not always necessary for long-term behaviour change (discussed in **Chapter 6**)
- Individual intervention components were generally preferred to social components (see **Chapter 6**)
- Some additional mechanisms of impact were observed, including improved psychological capability, reflective and automatic motivation, and social opportunity according to the COM-B model (Michie et al., 2011a) (as discussed in **Chapter 6**)
- There was no evidence for social facilitation and co-action, but social accountability was observed as a behaviour change mechanism (see **Chapter 6**)
- The observed long-term behaviour change mechanisms included changes in mind set in addition to the predicted habit formation and improved self-efficacy (see **Chapter 6**)
- The observed mechanisms were most pronounced for officers and staff who were less active at the beginning of the study. Increases in PA were also greatest for this group. (Discussed in **Chapters 5 and 6**)
- Some expected outcomes were not observed, including reduced sedentary time, improved physical health-related quality of life, reduced stress, improved productivity and reduced sickness absence. However, there were qualitative

reports of improved health and wellbeing, fitness, sleep and weight loss. (See **Chapters 5 and 6**)

- Lack of an expected impact on sedentary time was mainly due to contextual factors (discussed in **Chapter 6**)
- Several of the expected contextual factors (e.g. organisational policy and culture) appeared to influence implementation, mechanisms and outcomes, but there was no evidence for a differential impact of site (Plymouth Basic Command Unit (BCU) vs. North Dorset territorial area) (see **Chapters 5 to 7**)

The revised logic model should be referred back to throughout **section 8.4**, where the main findings are summarised and discussed.

Figure 22 Logic model of the intervention (end of study)



8.4 Integrated summary of findings and contributions to theory and evidence

8.4.1 Q1. What is the context (prevalence, opportunities, barriers, facilitators) of physical activity (PA) and sedentary behaviour (SB) in the police force?

Quantitative findings: Survey data indicated the sedentary nature of the policing role. The majority (58%) of study participants reported their role as mainly sedentary, with only 17% reporting having a mainly active role. Mean self-reported weekday sedentary time as measured by the International Physical Activity Questionnaire (IPAQ) was 6.41 hours. Almost half (46%) of participants had a mean baseline daily step count less than 10,000 and 13% were classed as 'low activity level' according to the IPAQ. There were some subgroup differences in baseline activity levels, based on gender, age, length of service and marital status.

Qualitative findings: The qualitative findings confirmed the need for an intervention, by adding support to, as well as expanding on, the quantitative results. Overall, officers and staff reported sedentary roles and compensated with higher activity outside of work hours. Interviews revealed further differences in patterns of activity based on occupation and work stream, indicating for example that PCSOs and special constables, and those employed as response officers, tended to be more active during work hours. The interviews also provided in-depth information on opportunities, barriers and facilitators for PA and SB in the police force. While there were several existing opportunities for PA in the workplace (gyms, exercise classes and sports teams), the extent to which these were used varied greatly between individuals. The main barriers and facilitators to PA (both within and outside of the workplace) were classed into four categories according to the Socio-Ecological Model - individual (e.g. motivation), interpersonal (e.g. social support), organisational (e.g. shift work) and community/environmental (e.g. season) level factors. Recommendations for promoting PA and reducing sedentary time in the police force were made by interviewees.

Contribution to knowledge: The findings showed a clear need for interventions to promote PA and reduce SB in the police force, supporting previous findings of the sedentary nature of the policing role (Ramey et al., 2014, Lagestad and Van Den Tillaar, 2014). The results suggested that the most inactive subgroups were females, younger officers and staff, and those with fewer than 10 years of police force service. These findings may be used to inform future targeting of interventions.

To the author's knowledge, this was the first in-depth qualitative study of barriers and facilitators to PA for the police force using a Socio-Ecological approach. Some of the identified barriers may be particularly pertinent in the policing role (or similar occupations), such as shift work, lack of availability or access to workplace exercise facilities (despite fitness testing being a requirement of the role), and perceived social pressure from more active colleagues.

As included in the logic model under 'context' (see **Figure 22**), the findings highlight the need to consider socio-demographic factors and wider determinants of behaviour (factors outside of the individual). These may affect the success of an intervention by influencing implementation, mechanisms of impact and outcomes (Moore et al., 2014) and should therefore be taken into account in any workplace PA intervention. Recommendations for promoting PA and reducing sedentary time in the Devon and Cornwall and Dorset Police forces (which are likely to be relevant at a national or international level) are described fully in **Chapter 4**. Some of these recommendations had an impact on policy and practice (see **section 8.6**).

8.4.2 Q2. Is mHealth and fitness technology (the Fitbit® and Bupa Boost app) a feasible and acceptable intervention in the police force?

Quantitative findings: Survey findings suggested that the intervention was acceptable to police officers and staff. Engagement according to self-reported usage data was high (for example, 83% of survey respondents were still wearing the Fitbit® at 8 months), but engagement with the Bupa Boost app was lower and declined more rapidly over time (only 27% of respondents were still using the app at 8 months). Usability and usefulness ratings showed that the Fitbit® was perceived as more user friendly and useful in promoting PA than the Bupa Boost app.

Qualitative findings: Qualitative survey and interview findings confirmed, explained and expanded on the quantitative results. Interviews confirmed high overall perceived acceptability of the intervention, with higher acceptability of the Fitbit® compared with the Bupa Boost app. The interviews provided in-depth views on positive and negative aspects of the two intervention components, and suggestions to improve the technology were made by participants. Some of these were general, for example the Bupa Boost app was viewed as difficult to navigate, while other comments were context-specific. As an example, the Fitbit® was seen as practical to wear with the police uniform, but it was suggested that its algorithm for capturing activity data could be adapted for night shift workers. Some participants perceived duplication in function of the Fitbit® and Bupa Boost app, with many preferring to use the Fitbit® on its own.

As shown in the logic model (**Figure 22**), there was a clear link between engagement with the intervention and perceived usability and usefulness. Low perceived usability and/or usefulness were the main reasons for lack of engagement, which determined the ability of the intervention to change behaviour, i.e. exposure to mechanisms of impact. Engagement was also influenced by convenience and accessibility, health status and motivation.

Although experiences of the intervention were positive overall, the qualitative data highlighted some potential negative consequences of mHealth and fitness technology use for a small number of individuals. These included skin irritation as a result of Fitbit® wear, and negative psychological effects including feelings of failure and guilt when not meeting goals, and anxiety and cognitive rumination resulting from tracking activity and sleep.

The intervention was perceived as feasible, acceptable and cost-effective by managers, commissioners and occupational health staff. It was feasible to implement and deliver, although some participants needed additional technical support for issues such as lost or malfunctioning devices.

Contribution to knowledge: A major contribution of this study was to explore engagement and acceptability, which are of central importance yet understudied and underreported in the mHealth literature (McCallum et al., 2018). The quantitative and qualitative findings together indicated high feasibility and acceptability of the intervention at all levels of the police force, for officers and staff as well as for managers, commissioners and occupational health staff. This was the first known study to explore the acceptability of mHealth and fitness technology in the policing context. The detailed qualitative feedback from study participants will be helpful for others looking to develop or evaluate similar mHealth interventions, and the context-specific feedback will be a valuable guide for those planning interventions for the police force or more generally for the emergency services or shift workers.

The study used qualitative methods and investigated engagement with the intervention in the longer term, topics that have been underdeveloped in existing studies (Yardley et al., 2016, McCallum et al., 2018). The findings underlined the importance of perceived usability and usefulness as key determinants of engagement, which provided support for the Technology Acceptance Model (Davis, 1989, Holden and Karsh, 2010) (see **Chapter 6**). Additional determinants of engagement – convenience and accessibility, health status, and motivation – matched those identified in a recent systematic review by Simblett and colleagues (Simblett et al., 2018). The present study produced some novel findings in relation to changes over time, and the differential impact of these factors in the short and longer term. For example, while health status consistently influenced engagement, usability and usefulness issues became a more frequent cause of disengagement as the study progressed. The results also suggested that constant engagement with mHealth technology may not be necessary for sustaining behaviour change in the longer term, a much debated issue in the mHealth literature (Yardley et al., 2016, Perski et al., 2017).

A further contribution was the identification of potential negative physical and psychological consequences of mHealth technology use. The present study is one of only a few to explore these qualitatively. The findings add support to this limited evidence base (Brakenridge et al., 2016b, Kanstrup et al., 2018, Gowin et al., 2015) and highlight the importance of considering and preparing for such outcomes.

While the intervention was feasible to implement and deliver in the police forces involved, the findings suggested a need for input from all organisational levels and additional support from a workplace champion or technical specialists. This is consistent with the findings of existing 'real-world' mHealth studies (Torquati et al., 2018, Harrison et al., 2014). This knowledge will assist future researchers, intervention developers and designers of workplace wellbeing programmes.

8.4.3 Q3. Does the intervention assist police officers and staff in increasing physical activity and reducing sedentary time? If so, who is likely to benefit most?

Quantitative findings: There was evidence of increased self-reported PA in both the short term (including total weekly PA and moderate-to-vigorous PA (MVPA) at 6 weeks and 12 weeks) and longer term (increased MVPA at 8 months), but there appeared to be no overall impact on objective daily step count. The results of subgroup analysis and exploratory regression indicated that the less active officers and staff (those achieving an average daily step count of less than 10,000 at baseline) appeared to show the greatest benefits from mHealth technology use. This group showed large and significant increases in both objectively measured PA (i.e. Fitbit®-recorded steps) and self-reported PA, in the short and longer term. For example, the less active officers and staff showed a mean increase in total PA of approximately 42 minutes per week from baseline to 8-month follow-up. Logistic regression also indicated that older officers and staff (those aged over 40 years) and those with fewer than 10 years of police force service showed greater increases in steps (controlling for baseline PA level). Although self-reported sedentary time showed a small reduction from baseline to 8 months (mean decrease of 0.24 hours per day), this was not statistically significant.

Qualitative findings: Interview data supported the quantitative findings, adding strength to the conclusions that the intervention did result in increased PA, and particularly for those who were less active at baseline. There was evidence of perceived increases in PA over the short and longer term. The interviews helped to explain why for some participants self-reported PA increased but there was no significant change in steps, with those participants instead focusing on alternative activities to walking and running, where it had not been practical to wear the Fitbit® (e.g. swimming, boxing) or where activity might not have been recorded as steps (e.g. gym activity, strength training).

Qualitative interviews were also valuable in explaining how the intervention worked (see 'mechanisms of impact' in **Figure 22**). Using the COM-B model as a framework (Michie et al., 2011a) (see **Chapter 6**), the Fitbit® and Bupa Boost app enhanced psychological capability, reflective and automatic motivation, and social opportunity. Several BCTs were identified from the CALO-RE taxonomy (Michie et al., 2011b) (introduced in **Chapter 3**) as having played a role, and key behavioural mechanisms identified in interviews included goal-setting, self-monitoring and awareness, self-regulation, social support and social comparison (see **Chapter 6** for a full discussion). There was also evidence of longer-term behaviour change including habit formation (Fitbit® wear and PA routines), enhanced self-efficacy and changes in mind set regarding PA. Behaviour change was most pronounced for the less active officers and staff. There was evidence of a ceiling effect for those who were already active and motivated at baseline; many of these interviewees perceived that the intervention had led them to maintain rather than increase their existing PA levels.

Confirming the quantitative findings, the majority of those interviewed perceived little impact of the intervention on sedentary time. Qualitative data explained that this was mainly due to perceived pressure of work and organisational culture/social norms where breaks were not perceived as appropriate.

Contribution to knowledge: The findings support those of our systematic review presented in **Chapter 2**, which indicated that mHealth technology is a potentially impactful means for promoting PA in a workplace setting (Buckingham et al., 2019). The work extends knowledge by providing evidence that this applies to a novel occupational setting (the police force) and that mHealth may lead to both short- and long-term behaviour change in this context. Many existing studies have been short in duration and/or have shown little evidence of longer-term behaviour change (Fanning et al., 2012, Bort-Roig et al., 2014, Afshin et al., 2016, Stephenson et al., 2017), so the finding that an mHealth intervention has the potential to produce relatively long-term behaviour change is important to the research field. The findings supported the COM-B model of behaviour change (Michie et al., 2011a) and the CALO-RE taxonomy (Michie et al., 2011b) in explaining mechanisms of impact. This is an important contribution as our understanding of how mHealth technology works to promote behaviour change is limited due to the lack of a theoretical basis of many mHealth tools (activity monitors and apps) and poor reporting of intervention components (Bort-Roig et al., 2014, Direito et al., 2017).

Although precise estimates of effect sizes could not be obtained in this uncontrolled study, there was evidence of statistically and potentially clinically significant increases in PA. Bringing the quantitative and qualitative findings together strongly suggested that the less active officers and staff may benefit most, a similar finding to existing mHealth studies in other populations (Schrager et al., 2017, Xian et al., 2017). A unique contribution of the present study was to qualitatively explore differences in levels of motivation and to relate these to behaviour change, comparing individuals with low and high intrinsic motivation. Behaviour change was most pronounced for those who were less active and less motivated at baseline. Increases in steps in the less active subgroup were similar in magnitude to those reported by other workplace mHealth interventions for PA (Brakenridge et al., 2016b, Poirier et al., 2016, Gremaud et al., 2018). Increasing daily steps has previously been associated with many health benefits, including a lower risk of cardiovascular events (Yates et al., 2014), improved insulin sensitivity and reduced adiposity (Dwyer et al., 2011) and lower all-cause mortality (Dwyer et al., 2015). Increasing PA has similarly been associated with positive health outcomes, with the greatest improvements observed amongst those who change from performing no activity to some activity (Blair and Connelly, 1996, Arem et al., 2015, Warburton and Bredin, 2017). The increases in both total weekly PA and MVPA are therefore likely to be clinically important for all participants in the PAW-Force study, but the greatest health benefits are likely to be gained by those who were less active at the beginning of the study. The novel finding that older officers and staff and those with fewer years of policing service may show greater benefit (i.e. greater increases in PA levels) should also be explored further. The findings will inform which subgroups should be targeted in future interventions.

The intervention was less useful for reducing sedentary time; this echoes the findings of some similar studies of activity monitors and smartphone apps in workplace settings (Slootmaker et al., 2009, Torquati et al., 2018). The qualitative findings supported the Socio-Ecological Model, reiterating the importance of context and external factors in changing behaviour. The COM-B model was also supported in highlighting the need for more opportunities (both physical and social), including 'offline' opportunities to be more active in the workplace. There was also a perceived need for greater support from managers in taking breaks during the working day.

8.4.4 Q4. Which intervention components are preferred and potentially most likely to result in behaviour change?

Quantitative findings: The study found no differential impact of the 'individual' and 'social' phases on steps, self-reported PA or sedentary time, in either the short or longer term. The majority of participants (56%) preferred the individual components of the Fitbit® and Bupa Boost app (e.g. goal-setting, self-monitoring) compared with only 7% who preferred the social features. There were no statistical differences in preferred phase by baseline activity level, age or occupation. There was some evidence that females were more likely to prefer the individual phase while males were more likely to prefer the social phase.

Qualitative findings: The interviews expanded on the above findings and revealed individual differences in preferences and perceived impact of the different intervention components. Factors such as personality were important; for example, social features appealed to those with a competitive nature. There was no qualitative support for the apparent gender differences in preferred intervention components, although this was not explored in depth. The interviews provided further insight into which intervention components/app features were preferred and potentially most impactful. For example, participants wanted rewards to be transparent, fair and for meaningful achievements, and preferred to compete against others of a similar age and fitness level to themselves.

Contribution to knowledge: The findings of no apparent difference in impact of the individual and social intervention components, and general preference for individual over social features, failed to provide support for social theories of behaviour change and contradicted the findings of some previous mHealth studies. For example, social app features have been associated with higher activity levels than analytic and affective components (King et al., 2016). Potential explanations for the observed lack of a difference may include low engagement with the Bupa Boost app and the fact that social components may be more important for longer-term engagement and behaviour change (Hamari and Koivisto, 2015, Du et al., 2016). Alternatively the preference for individual intervention components may be a unique finding in the policing population or workplace context, which may require further exploration. The finding does need to be interpreted with caution, however, as the aim of this study was to explore the potential impact and acceptability of individual and social components rather than to establish their definitive effectiveness.

Individual differences in preferred and perceived impactful features again suggest the importance of tailoring and personalised interventions. Potential gender differences suggested by the quantitative data could be explored further in future qualitative or mixed methods studies.

Findings regarding perceptions of virtual rewards and competitions will be of value to future app designers and intervention developers. The results supported previous findings that rewards in smartphone apps should be transparent (Middelweerd et al., 2015), and also fair and for meaningful achievements in order to increase motivation. Competitions between users should ideally be for those of comparable age and fitness or activity level.

8.4.5 Q5. Are there any wider benefits in terms of improved health and wellbeing, reduced stress, improved productivity and reduced sickness absence?

Quantitative findings: Although at baseline there were correlations between objective and self-reported PA, physical and mental health-related quality of life, and stress, there was no change in any secondary outcomes (health and wellbeing, stress or productivity) in the short or longer term (see **Figure 22**). The only exception was mental health-related quality of life which showed a small but significant improvement from baseline to 8-month follow-up, although this was not correlated with PA. It was also unclear whether this change was large enough to be clinically important. Due to data quality issues, it was not possible to draw any definitive conclusions on sickness absence, although there was no significant change in the number of sickness episodes within the Plymouth BCU, and the number of duty days and total days lost due to sickness appeared to increase from pre- to post-intervention.

Qualitative findings: The interviews confirmed the perceived links between PA and health and wellbeing, including sleep and stress. In apparent contrast with the quantitative survey findings, several participants reported feeling fitter and healthier, less stressed and more energetic as a result of increased PA. Potential explanations for this discrepancy suggested by interview data were that additional benefits such as weight loss, improved fitness and sleep were not captured as outcomes, and that the reported reductions in stress occurred mainly outside of the workplace. The interviews also provided possible explanations for the survey-assessed long-term improvement in mental health-related quality of life, such as the use of relaxation and mindfulness features within the Fitbit® and Bupa Boost app, and improved morale and camaraderie resulting from social support and competitions.

Contribution to knowledge: A major contribution of this study was the use of mixed methods to explore the potential wider impact of mHealth technology in a workplace intervention. Previously, few studies had employed a mixed methods approach, and there was a particular gap in the use of qualitative methods to explore subjective views and perceptions (Howarth et al., 2018, Buckingham et al., 2019). The qualitative findings were informative in explaining the observed longer-term improvement in mental health-related quality of life, and suggested other potential benefits that were not captured in the quantitative surveys. The work highlighted some remaining gaps in evidence that will require investigation in future mHealth trials, such as the impact on body mass index (BMI) and physiological outcomes (e.g. sleep and physiological indicators of stress).

As our systematic review revealed, the impact of workplace mHealth and PA programmes on work-related outcomes such as productivity and sickness absence had been particularly under-researched (Buckingham et al., 2019). While the impact on these outcomes remains unclear (conclusions are tentative due to data quality issues), the study findings of no impact on productivity and sickness absence closely match those of existing studies (Reijonsaari et al., 2012, Brakenridge et al., 2016b).

8.4.6 Q6. Will a future larger scale effectiveness trial be feasible and acceptable in this context?

Quantitative findings: Recruitment was rapid and efficient with the target of 150 participants exceeded within the two-month recruitment period. Study participants were representative of the wider police force sites in terms of age, gender and occupation, although the proportion of police staff in the study sample was slightly lower than expected relative to the Plymouth BCU site. The attrition rate was reasonably low at 21%, and study completers and non-completers had a similar socio-demographic profile, although participants with a health condition at baseline were statistically more likely to drop out. Outcome completion was high, with 83% of police officers and staff who remained in the study at month 8 completing the final follow-up questionnaire, although provision of step data was lower than that of self-reported PA data. While sickness absence was included as an outcome, there were issues with the availability, completeness and accuracy of data from staff records.

Qualitative findings: The interview findings complemented, expanded on and explained the quantitative results relating to feasibility of the study methods. The incentive of a free Fitbit® was revealed as an important factor in promoting recruitment, and participants perceived high reach, i.e. recruitment of a range of occupational groups. Qualitative data offered possible explanations for higher dropout amongst those with a health condition, with health status as a barrier to PA and engagement with the intervention. Interviews added further support for high acceptability of study outcomes; high satisfaction with questionnaires and low perceived burden of completion were reported. The interviews also explained why lower numbers of participants supplied step data; this was due to technical difficulties in obtaining and transferring these data. Overall participants expressed very high satisfaction with their experience of study participation.

Contribution to knowledge: The reflection on study methods based on both quantitative and qualitative data provided strong support for feasibility and acceptability of a larger scale effectiveness trial of this intervention (or similar future mHealth interventions) in the police force. This component of the work was comprehensive in comparison with many existing studies, and several of the aspects considered, such as reach and representativeness, have been underreported in the literature (Blackman et al., 2013, Buckingham et al., 2019). The qualitative findings suggested some possible ways in which the conduct of future mHealth studies in a similar context could be further improved. For example, the needs of those with chronic health conditions should be carefully considered to encourage engagement and retention, and alternative methods for capturing step data and/or support of a workplace champion may be required.

A need for improved reporting of sickness absence data in the Devon and Cornwall and Dorset Police forces was identified; this would greatly facilitate future research. It is not known whether the data quality issues encountered are more widespread across other police forces.

8.5 Strengths and limitations

The findings should be interpreted in the light of some key strengths and limitations relating to the design, conduct and analysis of the PAW-Force study (strengths and limitations of the systematic review are discussed in **Chapter 2** and **section 8.2**). The use of mixed methods, as recommended by the Medical Research Council (MRC) for the evaluation of complex interventions (Moore et al., 2014), is a major strength of this work. Mixed methods were used for multiple purposes; in addition to the quantitative findings being used to inform sampling for interviews and to aid analysis at an individual participant level, the qualitative findings from surveys and interviews were integrated during the interpretation phase to expand on, explain, and triangulate with (confirm and complement) the quantitative results. These are all important functions of mixed methods research (Greene et al., 1989). As a result of integrating qualitative data with quantitative results, a richer, more in-depth understanding was achieved, with the resulting knowledge greater than the sum of its quantitative and qualitative parts (Barbour, 1999). Both comprehensiveness and validity of the findings were improved through the use of mixed methods.

The inclusion of quantitative and qualitative feasibility outcomes (such as engagement and acceptability) was another key strength of the research, as these outcomes have been understudied in the fields of digital health and mHealth (Huang et al., 2019, McCallum et al., 2018). Aspects such as implementation and delivery of the intervention, feasibility for the wider workforce, and feasibility of study methods were also assessed to give a comprehensive overview of feasibility and acceptability. The key components of process evaluation – implementation, mechanisms of impact and context (Moore et al., 2014) – were integrated throughout the thesis.

An additional strength and contribution was the relatively long duration of the study (8 months) which facilitated an analysis of short- and longer-term engagement and behaviour change. In contrast, most previous studies of mHealth interventions for PA have been characterised by short-term interventions and follow-up (generally less than 6 months) (Fanning et al., 2012, Bort-Roig et al., 2014, Afshin et al., 2016).

As this was an uncontrolled pilot study, the findings (particularly those in relation to impact) should be interpreted with caution. This methodology was selected for practical reasons, which are described fully in **Chapter 3**. The potential confounding of changes in PA levels by temporal or seasonal factors should not be ignored, as seasonal factors have been shown to influence levels of PA and SB, with individuals tending to be more active in spring and summer and more sedentary in winter (O'Connell et al., 2014). The fact that activity levels increased in the present study despite the baseline measures being taken in summer, the 12-week follow-up in autumn and the 8-month follow-up in winter adds support to suggest a positive impact of the intervention.

Confidence in the conclusions is supported by several strengths of the study, including the capture of potential confounders (i.e. self-reported data on factors influencing PA levels), the incorporation of qualitative data, and use of multiple outcomes and multiple time points. These are all recommended ways to improve the rigour of non-randomised studies (Craig et al., 2012, Moore et al., 2014).

Some potential limitations surrounding data collection and analysis should be considered. Firstly, these tasks were performed by a single researcher, and the possibility of researcher bias should be mentioned. This was unavoidable due to resource limitations and the constraints of a PhD, although the potential for bias in survey data was minimal due to absence of any in-person contact when collecting outcomes. Subjectivity in interpretation of interview data is a common issue in qualitative analysis (Tong et al., 2007) and it is possible that a different researcher would have interpreted the findings differently. However, the potential influence of the researcher was considered reflexively (see **Chapter 3**) and the themes identified were checked by an independent researcher. These are important ways to improve the rigour and quality of qualitative studies (Tong et al., 2007).

Another limitation was the use of multiple hypothesis testing in the quantitative analysis. There is therefore a potential for false positives or type 1 errors (Sedgwick, 2014). Future effectiveness studies with larger samples should consider correcting for this, for example using the Bonferroni method (Bland and Altman, 1995). Again, the use of mixed methods, multiple outcomes and time points added strength to the findings.

The decision to use Fitbit®-captured step counts as the primary outcome measure may be questioned, particularly as accelerometers are widely used in scientific studies and have high validity and reliability for capturing data on PA and sedentary time (Esliger et al., 2011, Kozey-Keadle et al., 2011). Consultations with the Devon and Cornwall Police Health and Wellbeing Board indicated that wearing an accelerometer in addition to the Fitbit® would be considered burdensome, particularly for officers who already have to wear a uniform and carry weighty equipment. This would have also posed logistical issues given the geographical dispersion of the study participants. Fitbit® data on step count is known to be valid and reliable compared with research-grade accelerometers in different settings (Evenson et al., 2015, Feehan et al., 2018). Although accelerometers would have provided an objective measure of sedentary time in the PAW-Force study, this was not felt to be essential as reducing SB was a secondary aim. Nevertheless, future intervention studies should use objective, valid and reliable measures of PA and SB wherever possible.

The diversity and representativeness of the overall participant sample (which included both urban and rural police sites), and the relatively large sample size (n = 180) are strengths of this research. Recruitment of volunteers may be seen as a weakness, as many participants were already moderately active, motivated and at the later 'stages of change' of the Transtheoretical Model (Prochaska and DiClemente, 1983). However, opportunistic sampling is a commonly used method in workplace mHealth interventions (Buckingham et al., 2019) and it is a common experience that workplace PA programmes tend to attract those with higher levels of motivation and activity (Bravata et al., 2007). Nevertheless, participants with a range of baseline activity levels were recruited in the present study. A benefit of this was an improved understanding of how the intervention worked by enabling a comparison between the less active/motivated and sufficiently active/motivated participants. As a result of purposive sampling, interview participants were also diverse and representative of the study sample.

The extent to which the findings can be applied more widely within the police force nationally or internationally must be considered. It is recognised that the setting of the study (police forces in South West England) is specific. In line with the pragmatist approach taken in this research, it is argued that the findings are neither entirely context-specific nor generalisable, but transferable to other settings. It is important to maximise the use of knowledge gained and make the most appropriate use of that knowledge in other circumstances (Morgan, 2007). The limitations of the the PAW-Force study resulting from its uncontrolled pilot design and data collection procedures are recognised. However, the detailed assessment of how the intervention was implemented, how it worked to change behaviour, and contextual factors, will be valuable in improving external validity and transferability of the findings to future interventions in similar settings (O'Cathain et al., 2013). Many of the contextual issues identified are likely to be commonly experienced within the policing occupation and other service occupations. For example, shift work is likely to be a common barrier to PA, and lack of time, pressure of work and lack of access to exercise facilities have been previously identified as barriers to PA in the police force (Soroka and Sawicki, 2014, Lagestad and Van Den Tillaar, 2014). The high ecological validity of this study, conducted in a real-world setting, is an important strength that should be highlighted.

8.6 Recommendations and impact

Research impact is defined by Research England as "an effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life, beyond academia" (Research England, 2018, p83). In addition to the contributions to academic knowledge (outlined in **section 8.4** and summarised in **section 8.8**), the work produced a number of recommendations for the police force (see **Chapter 4, section 4.9**), which in turn had an impact on policy and practice in the Devon and Cornwall and Dorset Police.

Engagement with stakeholders is an important aspect of impactful research (Fast Track Impact, 2019). The police forces involved in the PAW-Force study were consulted throughout the research process. Prior to the intervention being delivered, Health and Wellbeing Board meetings were attended by the researcher to develop a clear understanding of context and the needs of police officers, staff and the wider organisations. The study protocol was presented as a central component of the wellness programme at the launch of ActivAte 2020; the audience at this event comprised over 50 stakeholders including the Chief Constables for Devon and Cornwall and Dorset Police, health and wellbeing leads, representatives of the College of Policing, and champions of the Bupa Boost app. At the end of the study, the findings were disseminated (see **Appendix 7** for summary findings report) and used to inform discussions and policy decisions at Health and Wellbeing Board meetings in Plymouth, Exeter and Dorset.

Following the recommendations of this work and ongoing consultations within the organisations, the Health and Wellbeing Strategic Coordination Group for Devon and

Cornwall and Dorset Police introduced discretionary wellness sessions for police officers and staff. These sessions may be used for PA or other activities to support health and wellbeing. It is anticipated that this new policy may contribute to a change in organisational culture, and may positively influence attitudes and perceptions relating to PA and the appropriateness of breaks in sedentary time in the workplace. Attitudinal and cultural changes are two important types of research impact (Fast Track Impact, 2019).

As a result of the research, the Chief Medical Officer for Devon and Cornwall Police made the decision to include Fitbits® (but not the Bupa Boost app) in the future police force wellbeing programme. The results informed which groups of individuals should be targeted (i.e. the less active and most sedentary). The Chief Medical Officer perceived that the intervention had a positive impact at an individual and organisational level:

> *"We are now planning to use Fitbits in our lifestyle programme. The intention is that they will be offered to people who are essentially sedentary and wanting to be more active, probably linked to a weight loss programme and healthy eating. This is part of our implementation plan for 2019/20. However, I don't think we will be promoting Bupa Boost in the future. The overall impact on the participants and the BCU has been positive in terms of individual wellbeing and organisational morale and culture."*
>
> Professor John Harrison, Chief Medical Officer, Devon & Cornwall Police

There was also evidence of impact outside of the Devon and Cornwall and Dorset Police forces. Participants' views on usability and usefulness of the Bupa Boost app were fed back to Bupa. These will be used to guide future development of this app and any similar apps designed by Bupa. In addition, as a result of this work, there has been interest from other police forces (such as Avon and Somerset Police) in implementing similar mHealth programmes, with a particular interest in the use of Fitbits® to promote PA.

8.7 Directions for future research

Rigorous experimental studies are important for all research. However, there is an increasing movement towards alternative methods to the randomised controlled trial

(RCT) approach for future evaluations of mHealth interventions. Technology-based interventions have several unique features which pose problems with RCT studies, including their rapidly changing nature, complexity with multiple interacting components, and infeasibility of including a control group or blinding participants (Arigo et al., 2019). Alternatives including n-of-1 studies, factorial trials and the Multiphase Optimisation Strategy (MOST) have been proposed as more appropriate and efficient for such interventions (Dallery et al., 2013, Pham et al., 2016, McCallum et al., 2018, Arigo et al., 2019). These types of designs also allow participants to act as their own control which minimises between-individual variation and allows efficient assessment of the relative impact and acceptability of various intervention components. Future studies should use robust methods such as factorial and crossover designs, to assess the definitive effectiveness of the different components (e.g. individual versus social components or specific BCTs). This was not possible in the present study due to practical considerations and its feasibility nature.

Following the PAW-Force study as an example, future studies evaluating mHealth interventions should clearly identify and report the component BCTs. There may be a need to develop new behaviour change taxonomies (or adapt existing taxonomies) that can readily be applied to eHealth and mHealth, which differ from 'traditional' interventions in both the nature of the intervention and the way in which they are delivered. The theoretical basis should be clear to enable understanding, replication and development of useful interventions that promote effective behaviour change.

It is recommended that future studies of mHealth interventions for PA should aim to recruit less active participants, as the present study showed that these individuals were most in need of intervention, and also benefited most and showed the greatest behaviour change. There is clear evidence that individuals who change from being inactive to at least somewhat active also experience the greatest improvements to health, including reduced morbidity and mortality (Blair and Connelly, 1996, Arem et al., 2015, Warburton and Bredin, 2017). Further subgroup differences suggested by the present study (such as impact according to age and years of policing service) should be explored in more depth. The relative impact and acceptability of individual and social components should be examined in different populations and contexts. The potential for negative consequences of mHealth technology use should not be overlooked, and the needs of workers with chronic health conditions (both physical and psychological) should be a key consideration.

Although the present study included a wide range of officers and staff working in urban and rural settings, future studies in the policing context should recruit from a wider range of sites and police forces. For example, metropolitan police forces outside of Devon, Cornwall and Dorset are quite different in terms of the nature of the role and socio-demographics of police officers and staff (e.g. greater ethnic diversity). There is also a need to extend and explore the use of mHealth technology in other workplace settings, including the emergency services and shift workers.

The impact of mHealth interventions and PA programmes on secondary outcomes such as health, wellbeing and stress, and on work-related outcomes such as productivity and sickness absence, should be investigated further as this is still unclear. The systematic review found limited evidence for an impact on these outcomes, and in the PAW-Force pilot study, quantitative evidence of the health, wellbeing or productivity benefits of increasing PA and/or using the technology did not become apparent over 8 months. In this study, the ability to draw definitive conclusions on sickness absence was hindered by a lack of complete and accurate data from the organisations involved. The outcome of productivity (incorporating presenteeism) was captured using the absenteeism and presenteeism questions of the Health and Work Performance Questionnaire (HPQ); this was selected for its widespread applicability and known validity, reliability and sensitivity to change (Kessler et al., 2003, Kessler et al., 2004). However, the reliance on self-report, which may be subject to various biases such as recall and social desirability bias, is a limitation that should be addressed by future studies. A need for a standardised definition and standardised measures of presenteeism has been recognised (Cancelliere et al., 2011, Brown et al., 2011). Future studies could use more objective measures of productivity such as targets met or assessments of work performance. Where possible, standardised and objective measures of primary and secondary outcomes should be used in preference to self-report. Future studies could also include physiological and clinical health outcomes such as cholesterol, blood pressure and BMI. Studies with larger samples and longer-term follow-up will be necessary to elucidate the effects on secondary outcomes.

While the findings of the PAW-Force study supported the use of mHealth technology as a standalone intervention for the promotion of PA in the police force, it was clear that wider contextual factors were of vital importance, particularly for SB. The importance of external influences on SB has previously been recognised in the literature, with an ecological approach recommended for interventions (Owen et al., 2011, Spence et al., 2016). It is therefore recommended that future studies aiming to reduce SB in the workplace should use multi-component interventions, with 'offline' components (such as standing desks or walking meetings) to complement the 'online' or technological aspects. Interventions that are specifically designed to target SB (Prince et al., 2014), and that incorporate environmental cues and restructuring (Bond et al., 2014, Gardner et al., 2016) are likely to be successful. Multi-component interventions may also help to enhance the impact of mHealth technology on PA (Schoeppe et al., 2016, Nuffield Health, 2018, Buckingham et al., 2019). Multi-level interventions that target multiple tiers of an organisation are recommended. The involvement of internal workplace champions to promote engagement and to assist in the day-to-day running and oversight of such interventions should also be encouraged.

Based on the findings of the systematic review and PAW-Force study, a summary of recommendations for the design and conduct and future mHealth intervention studies in workplace settings is given in **Figure 23**. Overall, there is a need for further high quality mixed methods studies that use alternatives to RCT designs, in a wider range of workplace settings, in order to address the remaining knowledge gaps surrounding the usefulness of mHealth interventions in promoting PA, reducing SB and improving health and wellbeing. A focus on longer-term engagement, acceptability and impact of mHealth interventions should be central in future research.

Figure 23 Recommendations for the design and conduct of future mHealth intervention studies in workplace settings

Intervention Design	• Clearly identify and report the theoretical basis of interventions (e.g. coded BCTs) • Multi-level interventions with the support of individuals at various organisational 'tiers' are recommended • Workplace wellness champions may help to promote engagement, assist with data collection etc. • Ensure adequate technical support (staff, time, resources) is available • Multi-component interventions that take a socio-ecological approach may be more successful
Research Methods	• Use robust study designs, not necessarily RCTs (e.g. factorial trials, crossover trials, MOST) • Use mixed methods to explore impact/effectiveness, feasibility, acceptability and engagement • Duration of interventions and follow-up should be maximised – aim for ≥6 months • Recruit a diverse range of staff, in a range of occupations, with targeted recruitment of those most in need of intervention (e.g. less physically active) • Consider behaviour both within and outside the workplace • Explore the wider impact on health and wellbeing, productivity and sickness absence • Objective outcome measures should be used where possible for primary and secondary outcomes. Where this is not possible, standardised questionnaires should be used. • Consider potential negative physical and/or psychological consequences of technology use, and record any adverse effects

8.8 Summary of contributions of the PhD

In bringing together the fields of workplace wellness, mHealth and behaviour change, this PhD project has made several important contributions. The main contributions to knowledge are summarised in **Table 70**. The study methods and resulting findings will be a useful template for future intervention designers, mHealth developers and researchers. For example, unlike many previous mHealth studies, the intervention was fully described with clear coding of BCTs, analysed using a mixed methods approach, and evaluated using relevant theories and MRC guidance. The findings are also of practical benefit to the police force, and have already had a direct impact on policy and practice (outlined in **section 8.6**). Finally, the findings are of more general relevance to employers who are considering including an mHealth component in future workplace wellness programmes.

Table 70 Summary of main contributions to knowledge of the PhD

What is already known	What this study adds
Mobile health (mHealth) technology is a popular, low cost and accessible tool which offers the potential to promote PA and improve public health.	Despite the low methodological quality of many studies, there is reasonable evidence that mHealth interventions are potentially effective, feasible and acceptable for promoting PA in workplace settings.
Policing is an increasingly sedentary and highly stressful occupation. There is a need for interventions to increase PA and reduce SB in officers and staff.	An mHealth intervention (Fitbit® activity monitor and Bupa Boost app) in the police force was potentially impactful for increasing PA, feasible to deliver and acceptable at all organisational levels. Officers and staff who are less active may benefit most. There was higher observed and perceived impact and greater perceived acceptability in this group. Context and external factors are of central importance, particularly for reducing SB. Opportunities to take breaks in the workplace and support from managers should be given. Multi-component interventions (with online and offline components) that target barriers at different levels of the Socio-Ecological Model may be more successful than standalone mHealth interventions.
mHealth interventions are associated with declining engagement over time and may be most acceptable and impactful in the short term. Factors influencing engagement include usability, perceived usefulness, convenience and accessibility, health status, and motivation.	mHealth technology may be used successfully to promote longer-term behaviour change (up to 8 months after beginning the intervention), including habit formation, changes in mind set and improved self-efficacy. Perceived usability and usefulness are the main factors that influence engagement with mHealth technology. Convenience and accessibility, health status and motivation are also influential. The nature and relative importance of these factors may change over time. Continued use of mHealth technology is not necessarily required to sustain healthy behaviours.
Social features of mHealth technology (e.g. social support and competitions) may be associated with greater engagement and/or behaviour change than individual components.	Individual components (e.g. self-monitoring and goal-setting) may be preferred to, and potentially as impactful as, social features in certain contexts or populations (such as the police force). However, there are large individual differences in preferences.

8.9 Conclusions

This thesis adds to the growing body of evidence that supports the use of mHealth technology to promote PA in the workplace. A systematic review found that mHealth interventions in workplace settings are potentially effective, acceptable and feasible for promoting PA, at least in the short term. A mixed methods pilot study of an mHealth intervention in a previously unstudied setting, the police force, showed that the use of a Fitbit® activity monitor and 'Bupa Boost' app was a feasible, acceptable and potentially impactful way to increase PA within this context.

This research contributes to filling a gap in evidence surrounding the longer-term acceptability and impact of mHealth technology; for the majority of officers and staff, engagement with the technology and associated behaviour change was sustained in the longer term. A further important finding was that those in greatest need of intervention, the less active officers and staff, experienced the greatest benefits. This standalone mHealth intervention was less useful for reducing sedentary time; the work highlights the importance of wider socio-ecological factors, and the need for multi-component interventions to reduce sedentary time in the workplace.

The impact on secondary outcomes, including health-related quality of life, stress, productivity and sickness absence, remains unclear. There is a need for further high quality, mixed methods studies in a range of workplace settings to explore the use of mHealth technology in the reduction of SB and the wider impact on health, wellbeing and productivity. In addition to the contributions to academic knowledge, the findings of this study were used to inform health and wellbeing policy and practice in the police forces involved.

APPENDICES

Appendix 1: Systematic review documentation

1a. Systematic review protocol

The systematic review protocol is registered with PROSPERO at: http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD420170588566

Title

Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: a systematic review

Project team

Sarah Buckingham
Dr. Andrew James Williams
Dr. Karyn Morrissey
Prof. John Harrison
Dr. Lisa Price

Background

Physical inactivity is considered one of the biggest public health problems of the 21st century (Blair, 2009). Failing to meet the recommended guidelines is associated with an increased risk of morbidity due to cardiovascular disease, cancer and metabolic syndrome (Wannamethee and Shaper, 2001, Liu et al., 2016, Metzger et al., 2010) and general mortality (Arem et al., 2015, Dwyer et al., 2015). There is now also substantial evidence that sedentary behaviour is an independent predictor of poor health and mortality (Owen et al., 2010, Dunstan et al., 2011, Wilmot et al., 2012).

Interventions to increase physical activity (PA) levels and reduce sedentary behaviour (SB) are clearly vital. The workplace is viewed as an important setting for health promotion and disease prevention (Institute of Occupational Safety and Health, 2015). Around half of weekday sitting time is work-related (Miller and Brown, 2004, Kazi et al., 2014) and up to 71% of working hours in office workers are spent sedentary (Clemes et al., 2014). Occupational sedentary time is predicted to further increase in future with rises in automation and information technology use (Hendriksen et al., 2016a). Promotion of PA in the workplace has many potential benefits, including improved health and wellbeing of employees and economic benefits for employers (Hendriksen et al., 2016b).

Mobile health (mHealth) technology has rapidly gained popularity in the general population. mHealth technology includes wearable PA monitors or trackers and smartphone applications (apps) designed to help people to manage their own health and wellbeing. The potential value of mHealth in health promotion lies in its widespread appeal, accessibility and ability to reach large populations at a low cost (Sullivan and Lachman, 2016). It also offers the potential for tailoring of interventions to the needs of individuals or specific groups.

Studies have investigated the use of mHealth to promote PA in a range of settings, including the workplace (Sullivan and Lachman, 2016). Whilst the results of clinical and general population studies suggest that mHealth may be a feasible and cost-effective way to promote PA (Direito et al., 2017), the findings of existing reviews have been inconclusive. Some reviews have reported nonsignificant effects of mobile technology on PA levels (Flores Mateo et al., 2015), and where beneficial effects are reported, effect sizes have generally been small (Fanning et al., 2012, Bort-Roig et al., 2014, Muntaner et al., 2016, Direito et al., 2017). Additional limitations of previous reviews are the inclusion of studies where mHealth devices were used as a data collection tool rather than as an intervention in their own right (O'Reilly and Spruijt-Metz, 2013, Bort-Roig et al., 2014), and a lack of a comprehensive description of interventions and study procedures (Fanning et al., 2012). Furthermore, apart from two recent exceptions (Schoeppe et al., 2016, Direito et al., 2017), few reviews of mHealth interventions have assessed both PA and SB outcomes.

Identification of behaviour change techniques (BCTs) using standardised taxonomies is important for recognition of effective and acceptable components, to allow replication and comparison of interventions, and to facilitate further development and testing of theories (Abraham and Michie, 2008). There is also evidence that including established BCTs is associated with greater intervention effectiveness (Greaves et al., 2011). Despite this, previous reviews have concluded that many mHealth interventions lack an explicit theoretical basis (Fanning et al., 2012, Bort-Roig et al., 2014) and it remains unclear which components are most effective and accepted (Sullivan and Lachman, 2016). Identification or coding of included BCTs, and identifying the theoretical basis of existing studies are therefore important gaps to address.

As mHealth is such a rapidly progressing field due to advances in technology, studies have increased exponentially in a short space of time. Early reviews predominantly comprised studies of text messaging (SMS) interventions but the emergence of new technologies (e.g. tablets, commercial wearable activity monitors, and ‘exergaming’) means the evidence should be frequently reviewed in order to accurately reflect the current status. Furthermore, the use and effectiveness of mHealth interventions in specific population groups remains unclear (Schoeppe et al., 2016). It is important to consider setting or context in the evaluation of mHealth interventions as due to their complex nature, various components may produce different outcomes for different individuals in different settings (McCallum et al., 2018). Workplace mHealth interventions may differ from general interventions in terms of both intervention content and timing of effectiveness (Stephenson et al., 2017). To the authors' knowledge, there has been no previous systematic review of mHealth technology for promoting PA and reducing SB in workplace settings. A recent review of general digital health interventions in the workplace concluded that the evaluation of smartphone apps in this context is an important 'next step' for future research (Howarth et al., 2018).

Employee populations potentially have much to gain from mHealth interventions for PA and SB, yet little is known about the impact of this technology in a workplace context. Feasibility and acceptability are important aspects to consider but remain understudied and underreported (McCallum et al., 2018). This review will therefore aim to provide a comprehensive synthesis of current evidence in relation to the effectiveness, feasibility and acceptability of mHealth interventions in the promotion of PA and reduction of SB in the workplace. This will include a description of intervention content in terms of common BCTs using an established behaviour change taxonomy, and a consideration of subgroup differences and the wider impact of interventions on health and related outcomes.

Review questions

- Are mobile health (mHealth) interventions effective at increasing physical activity levels and reducing sedentary behaviour in a workplace setting?
- Are such interventions feasible and acceptable in a workplace setting?
- What are the most commonly used behaviour change techniques and theories?
- Is there any evidence for subgroup differences (e.g. age, gender, shift workers, different occupations) in the effectiveness (or acceptability) of workplace mHealth interventions?
- Are any other related outcomes (e.g. health, wellbeing, productivity) improved after mHealth interventions for physical activity promotion?
- Where are the gaps in current knowledge surrounding the use of mHealth technology in physical activity promotion, and what are the implications for future research?

Search strategy

The following resources will be used:

- The electronic databases MEDLINE, SPORTDiscus, Scopus, EMBASE, PsycINFO and Web of Science using both free text searching and controlled vocabulary
- The Cochrane library including:
    - Cochrane Database of Systematic Reviews (CDSR)
    - Cochrane Central Register of Controlled Trials (CENTRAL)
    - Database of Abstracts of Reviews of Effect (DARE)
    - Health Technology Assessment (HTA)
- Grey literature

- Dissertations and theses (ProQuest Dissertations and Theses Global)
  - mHealth Evidence
  - ‘Fitabase’ research library (studies using the Fitbit® activity tracker)
  - Other reputable grey literature
- Reference lists of included articles and other published related systematic reviews (backward citation searching)
- Forward citation searching using Google Scholar and Web of Science

The search will include studies with a publication date between 2007 and 2017 (February). No language restriction will be applied.

Types of study to be included

A range of study designs will be considered, including experimental (e.g. randomised controlled trials), quasi-experimental (e.g. pre-post/uncontrolled trials) and observational (e.g. cohort trials). Pilot and feasibility trials may be included if they meet the inclusion criteria.

Types of study to be excluded

- Review articles, narrative reviews, editorials, opinions and letters, and reports published as conference abstracts or proceedings only
- Studies that report only qualitative data
- Studies with insufficient methodological details to allow critical appraisal or replication
- Non-human studies (i.e. animals, computer models)
- Clinical population studies and studies of school, college or university based students
- Studies which do not include a valid measure of physical activity or sedentary behaviour

Condition/domain being studied

Physical activity and sedentary behaviour

Participants/population and context

Free-living, working adults (18+ years). Participants should be either recruited in the workplace and/or the intervention should be delivered in the workplace (although may be designed as a general lifestyle intervention).

Intervention/exposure

Interventions/programmes with mHealth as a major component of the intervention (mobile phone, smartphone apps, personal digital assistants (PDAs), tablets, wearable activity monitors/trackers (commercial or research grade)). Smartphone apps for monitoring of physical activity alone or with other behaviours (e.g. diet, weight) may be included. Interventions of any duration will be included.

A mobile device should form part of the intervention, not be used only as a data collection tool. The intervention does not need to be completely automated and may include a human-delivered or facilitated component, such as meetings or telephone contacts.

Web-only interventions, and interventions involving text messaging (SMS) or traditional pedometers (i.e. not able to transmit data to a consumer interface) alone will be excluded.

Comparator/control

For experimental studies there should be a control group or comparator (i.e. those not exposed to the intervention, usual care, wait-list group).

For quasi-experimental and observational studies, as a minimum, pre- and post-exposure data should be available.

Outcomes

Primary outcomes (essential for inclusion):

Any measure of physical activity and/or sedentary behaviour (e.g. MET-minutes, weekly minutes of moderate-to-vigorous physical activity (MVPA), step count, sedentary time or patterns of sedentary behaviour).

The outcome may be objective or self-reported, and may comprise occupational physical activity alone or in combination with non-occupational activity.

Secondary outcomes

Any other related outcomes, such as:

- Anthropometrics and cardiometabolic risk factors, for example body mass index, weight, waist circumference, waist-hip ratio, body composition, biomarkers (e.g. cholesterol), blood pressure, fitness level
- Self-reported or objective physical or psychological health measures, wellbeing assessments or quality of life
- Health service usage
- Productivity measures
- Sickness absence

Study selection procedures

Study selection

The results of all electronic searches will be imported into EndNote reference management software. All duplicate papers will be carefully identified and excluded. Two independent reviewers (SB and AJW) will screen the titles and abstracts of all retrieved articles for relevance, using the pre-specified inclusion and exclusion criteria. Full paper or electronic copies of all potentially relevant studies will be obtained. Any disagreements regarding inclusion will be resolved through discussion and involving a third reviewer (KM) where necessary.

Data extraction and quality assessment

Standardised data abstraction forms, including quality assessments, will be completed by SB and verified by AJW. Any disagreements will be resolved by consensus and a third reviewer (KM) will be involved as necessary.

The review team will use a suitable quality assessment tool (i.e. the Effective Public Health Practice Project (EPHPP) quality assessment tool for quantitative studies) to evaluate study quality and assess the risk of bias. The general principles of the NHS Centre for Reviews and Dissemination and the Cochrane Collaboration will be applied to the quality assessment.

Data synthesis and analysis

Data collected in extraction forms will be summarised in evidence tables, including study characteristics and results. The following study characteristics will be reported: authors; year of publication; country; participants (number and characteristics); setting; study design; the intervention (type of mHealth/tool, intervention components, motivational strategies/behaviour change techniques, duration and frequency); comparison groups; length of follow-up; primary and secondary outcome measures. The results table will include analysis methods used and key findings for each study, including any subgroup findings and comments on engagement and attrition.

Effects of the intervention will be assessed using meta-analysis of pooled effect sizes for physical activity and sedentary behaviour outcomes where possible (e.g. weighted mean difference using both fixed and random effects models). Data synthesis and meta-analysis will be performed in Stata (version 14.0) and RevMan (version 5.3). Forest plots will be produced for primary outcomes. If meta-analysis is not possible, narrative review will be used instead.

Heterogeneity of study populations, interventions, methods and outcomes will be explored descriptively and statistically using the Chi-square test for homogeneity and the I^2 statistic. Small study effects (including publication bias) will be assessed visually using funnel plots and quantified using Egger's statistic.

Analysis of subgroups or subsets

If sufficient comparable data are available, findings will be synthesised according to subgroups of interest such as gender, age, type of setting/occupation, shift- vs. non-shift work, type of mHealth intervention (e.g. wearable physical activity monitors/trackers, smartphone apps) and objectively measured vs. self-reported outcomes. Subgroup data will be reported in summary tables and forest plots.

Conduct and reporting

The systematic review will be conducted and reported in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement.

Dissemination plans

Findings will be reported in a peer-reviewed journal, and at national or international conferences.

1b. Additional files 2 and 3 of published paper

Additional File 2: Master search strategy

Intervention: mHealth OR m-health OR "mobile health" OR (mobile adj3 (program* or intervention* or app* or device)) OR "mobile phone" OR (cell* phone) OR cellphone OR smartphone OR "smart phone" OR PDA OR "personal digital assistant" OR tablet OR (smart* adj2 (sensor or app*)) OR (wearable adj3 (monitor or tracker or device)) OR (activity adj2 (monitor or tracker or device)) OR (digital adj3 (monitor or tracker or device)) OR (mobile gam*) OR exergam*

AND

Outcome: "Physical activity" OR PA OR exercise OR fitness OR inactiv* OR sedentar* OR sitting OR "screen time" OR "leisure activity" OR walk* OR step* OR sport*

AND

Population/Setting: Worksite* OR work-base* OR workplace* OR worker* OR occupation* OR employ* OR profession* OR organi?ation* OR office*

Additional File 3: Example search strategy for Ovid MEDLINE®

1. exp Telemedicine/
2. mHealth.ti,ab.
3. exp Cell Phones/
4. m-health.ti,ab.
5. "mobile health".ti,ab.
6. exp Mobile Applications/
7. (mobile adj3 (program* or intervention* or app* or device)).ti,ab.
8. "mobile phone".ti,ab.
9. (cell* phone or cellphone).ti,ab.
10. exp Computers, Handheld/ or exp Smartphone/
11. (smartphone or smart phone).ti,ab.
12. PDA.ti,ab.
13. "personal digital assistant".ti,ab.
14. tablet.ti,ab.
15. (smart* adj2 (sensor or app*)).ti,ab.
16. (wearable adj3 (monitor or tracker or device)).ti,ab.
17. (activity adj2 (monitor or tracker or device)).ti,ab.
18. (digital adj3 (monitor or tracker or device)).ti,ab.
19. exp Fitness Trackers/
20. mobile gam*.ti,ab.
21. exp Video Games/
22. exergam*.ti,ab.
23. exp Exercise/
24. "physical activity".ti,ab.
25. PA.ti,ab.
26. exercise.ti,ab.
27. exp Physical Fitness/
28. fitness.ti,ab.
29. inactiv*.ti,ab.

30. exp Sedentary Lifestyle/

31. sedentar*.ti,ab.

32. sitting.ti,ab.

33. "screen time".ti,ab.

34. "leisure activity".ti,ab.

35. exp Walking/ or exp Walk Test/

36. walk*.ti,ab.

37. step*.ti,ab.

38. exp Sports/

39. sport*.ti,ab.

40. exp Work/

41. (worksite* or work-base* or workplace* or worker*).ti,ab.

42. exp Occupations/

43. occupation*.ti,ab.

44. exp Occupational Health/

45. employ*.ti,ab.

46. organi?ation.ti,ab.

47. office*.ti,ab.

48. profession*.ti,ab.

49. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22

50. 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39

51. 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48

52. 49 and 50 and 51

53. limit 52 to (humans and yr="2007 -Current")

1c. Search and screening log

Database	Host	Limits	Date	Hits
Medline in-process & other non-indexed citations, 1946-2017	Ovid SP	Publication date 2007 to Feb 2017 Humans	21/2/17	502
Embase 1974-2017	Ovid SP	Publication date 2007 to Feb 2017 Humans Including articles in press and InProcess status. Excluding conference abstracts.	21/2/17	230
PsycINFO 1806-2017	Ovid SP	Publication date 2007 to Feb 2017 Humans	21/2/17	233
Web of Science Core Collection (Science Citation Index Expanded, Social Sciences Citation Index, Emerging Sources Citation Index) Inception-Feb 2017	Thomson Reuters	Publication date 2007 to Feb 2017	21/2/17	701
SPORTDiscus Inception-Feb 2017	EBSCO host	Publication date 2007 to Feb 2017. Publication type = academic journal, book, electronic resource, monograph or government document, periodical, primary source document, proceeding, report, serial publication, thesis or dissertation, trade publication, URL	27/2/17	279
Scopus Inception-Feb 2017	Elsevier	Publication date 2007 to Feb 2017. Subject area = Medicine, Health Professions, Social Sciences, Psychology, Multidisciplinary Source type = journals, books, book series, undefined	27/2/17	719
Cochrane Library (CDSR, CENTRAL, DARE and HTA) Inception-Feb 2017	Cochrane	Publication date 2007 to Feb 2017	27/2/17	151
			Total	2815 (1892 after removal of 923 duplicates)

Five additional records were identified through other sources (one from Proquest, one from Fitabase, one from citation searching using Google Scholar, one from reference lists of systematic reviews – Kirk et al. 2016, one from reference list of included articles – Finkelstein et al. 2016).

1897 records in total (from databases and grey literature/other sources) after removal of duplicates.

Title and Abstract Screening

Reviewer 1 (SAB):

1824 excluded at title and abstract

73 included for full text

Reviewer 2 (AJW):

1827 excluded at title and abstract

70 included for full text

Combined:

Disagreed on **47**

Agreed on **1850**

Discussed and agreed to obtain full text for **71** papers

Full Text Screening

Full text articles excluded:

Not a mobile health intervention / text messaging only (n = 18)

Protocol studies and conference proceedings (n = 7)

Non-workplace setting (e.g. clinical, community) (n = 21)

PA / sedentary behaviour not a primary outcome (n = 5)

Review articles or not a primary research study (n = 2)

Articles for inclusion: n = 18 (15 studies)

Updated Searches

Searches re-run in February 2018 (publication date February 2017 to February 2018):

Ovid Medline

Ovid Embase

Ovid PsycINFO

Web of Science

SPORTDiscus

Scopus

Cochrane Library

5 new articles (4 studies) met the criteria for inclusion in the review (Gilson et al. 2016; Gilson et al. 2017; Losina et al. 2017; Neil-Sztramko 2017; Yeung et al. 2017).

Searches re-run in December 2018 (publication date March 2018 to December 2018):

Ovid Medline

Ovid Embase

Ovid PsycINFO

Web of Science

SPORTDiscus

Scopus

Cochrane Library

7 new articles (6 studies) met the criteria for inclusion in the review (Gremaud et al. 2018; Olsen et al. 2018; Patel et al. 2018; Reed et al. 2018; Simons et al. 2018a; Simons et al. 2018b; Torquati et al. 2018).

1d. Data extraction form and quality assessment tool

Data extraction form

Name of reviewer: Date:

Bibliographic details

Authors:

Title:

Year (publication):

Citation:

Study

Country: Setting:

Study design:

Funding source:

Data collection year(s):

Notes:

Population (baseline characteristics)

	Intervention	**Control / comparator**	**Overall**
No. of participants			
Age in years (mean and SD and/or range)			
% male			
% female			
Ethnicity			

Other baseline characteristics (e.g. overweight)			

Main exclusion criteria:

Control

Was there a control or comparison group?

Yes ☐ No ☐

If yes, how were the control group treated?

Intervention

Type of mHealth technology/tool (tick all that apply):

Smartphone app ☐

Tablet or PDA ☐

Wearable activity monitor or tracker ☐

Other please specify:

Further description of intervention (including type of app or device):

Was it a standalone mHealth intervention or part of a multi-component programme?

Standalone mHealth ☐ Multi-component ☐

If multi-component, please give details:

Was the intervention based on a named behaviour change theory?

Yes ☐ No ☐

If yes, please state theory:

Which motivational strategies/behaviour change techniques were included in the intervention?

Goal-setting ☐ Self-monitoring ☐

Prompts/cues ☐ Feedback ☐

Rewards (virtual or real) ☐ Social support ☐

Social comparison/competition ☐

Other please specify:

Duration of intervention:

Frequency of intervention (if applicable):

<u>Outcomes:</u>

Primary:

Secondary:

Method of assessment:

Objective ☐ Self-reported ☐

Both ☐

Details of assessment method(s) (e.g. surveys, interviews, accelerometers):

Duration of follow-up:

Outcome assessment points:

Analysis (e.g. inferential statistics, ITT, per protocol, clustering):

Results

	Intervention	**Control**	**Total**
No. randomised/ beginning study			
No. lost to follow-up			
Attrition rate (%)			

Any notes on engagement/attrition:

Main results (effect size with 95% CI, p-values):

Outcome	**Within-group findings**	**Between-group findings**

Subgroup findings (if applicable):

Additional findings (e.g. qualitative):

Additional notes (e.g. strengths, limitations, recommendations for future research):

Quality assessment tool for quantitative studies (EPHPP)

STUDY:

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

1. Very likely
2. Somewhat likely
3. Not likely
4. Can't tell

(Q2) What percentage of selected individuals agreed to participate?

1. 80 - 100% agreement
2. 60 – 79% agreement
3. less than 60% agreement
4. Not applicable
5. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

1. Randomized controlled trial
2. Controlled clinical trial
3. Cohort analytic (two group pre + post)
4. Case-control
5. Cohort (one group pre + post (before and after))
6. Interrupted time series
7. Other specify ______________________________
8. Can't tell

Was the study described as randomized? If NO, go to Component C.

No   Yes

If Yes, was the method of randomization described? (See dictionary)

No   Yes

If Yes, was the method appropriate? (See dictionary)

No   Yes

RATE THIS SECTION	**STRONG**	**MODERATE**	**WEAK**
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

1. Yes
2. No
3. Can't tell

The following are examples of confounders:

1. Race
2. Sex
3. Marital status/family
4. Age
5. SES (income or class)
6. Education
7. Health status
8. Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1. 80 – 100% (most)
2. 60 – 79% (some)
3. Less than 60% (few or none)

4. Can’t Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?

1. Yes
2. No
3. Can’t tell

(Q2) Were the study participants aware of the research question?

1. Yes
2. No
3. Can’t tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1. Yes
2. No
3. Can’t tell

(Q2) Were data collection tools shown to be reliable?

1. Yes
2. No
3. Can’t tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1. Yes
2. No
3. Can't tell
4. Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell
5. Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell

(Q2) Was the consistency of the intervention measured?

1. Yes
2. No
3. Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

1. Yes
2. No
3. Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?

1. Yes
2. No
3. Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

1. Yes
2. No
3. Can't tell

GLOBAL RATING

COMPONENT RATINGS

See dictionary for guidance on how to rate this section, available from: http://www.ephpp.ca/PDF/QADictionary_dec2009.pdf

A	**SELECTION BIAS**	STRONG 1	MODERATE 2	WEAK 3	
B	**STUDY DESIGN**	STRONG 1	MODERATE 2	WEAK 3	
C	**CONFOUNDERS**	STRONG 1	MODERATE 2	WEAK 3	
D	**BLINDING**	STRONG 1	MODERATE 2	WEAK 3	
E	**DATA COLLECTION METHOD**	STRONG 1	MODERATE 2	WEAK 3	
F	**WITHDRAWALS AND DROPOUTS**	STRONG 1	MODERATE 2	WEAK 3	Not applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1 STRONG (no WEAK ratings)

2 MODERATE (one WEAK rating)

3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No Yes

If yes, indicate the reason for the discrepancy

1. Oversight

2. Differences in interpretation of criteria

3. Differences in interpretation of study

Final decision of both reviewers (circle one):

1 STRONG

2 MODERATE

3 WEAK

Appendix 2: Ethics committee approval letter

Our Ref: RG/CB/17/02/116

27 March 2017

Sarah Buckingham
PhD Student
University of Exeter Medical School
European Centre for Environment & Human Health
Room F37b
Knowledge Spa
Royal Cornwall Hospital
Truro
TR1 3HD

Please reply to:
Ruth Garside, PhD
Chair, UEMS Research Ethics Committee
University of Exeter Medical School
c/o Carol Barkle
Administrator to UEMS REC
Knowledge Spa
Royal Cornwall Hospital
TRURO
Cornwall
TR1 3HD
Tel: 01872 256460
Email: **c.barkle@exeter.ac.uk**

Dear Sarah

Application Number: 17/02/116
Project Title: The PAW-Force (Physical Activity Wearables in the Police Force) Trial

I am writing to confirm that I am now happy that you have addressed all the points made by the UEMS Research Ethics Committee relating to the above project. I have approved this project under Chair's Action with immediate effect and have pleasure in enclosing your Certificate of Approval.

Approval of this study will be formally ratified by the University of Exeter Medical School Research Ethics Committee at its next meeting on the 27th April 2017.

Good luck with your study.

Yours sincerely

RGarside.

Ruth Garside, PhD
Chair
University of Exeter Medical School Research Ethics Committee

University of Exeter Medical School Knowledge Spa Royal Cornwall Hospital Truro Cornwall TR1 3HD UK
Tel +44 (0)1872 256460 Email : c.barkle@exeter.ac.uk

Pro Vice Chancellor and Dean **Professor Clive Ballard**

Appendix 3: Sample size calculations

All calculations were performed in Stata version 14.0, using the paired test comparing two correlated means. Based on 80% power and 5% significance level (p = 0.05).

Primary Outcome Variable: Step Count

Large variation in effect sizes reported in existing trials – e.g. a systematic review of internet and mobile interventions for improving diet, physical activity, obesity, tobacco and alcohol use found that the increase in step count from baseline to follow-up ranged from 900 to 4500 steps/day (Afshin et al., 2016)

(Wang et al., 2016) - Fitbit mobile app use – change in daily step count from baseline to 6 week follow-up: 6228 (SD = 468) to 6773 (SD = 481) (p = 0.04, n = 22)

(Poirier et al., 2016) - activity tracker and internet-based adaptive walking programme. Intervention group (n = 107) change in daily step count from baseline to 6 week follow-up: 5102 (SD = 1901) to 5411 (SD = 2277). Sedentary group (n = 58) change in daily step count from baseline to follow-up: 3769 (SD = 970) to 4363 (SD = 1517)

(Fukuoka et al., 2010) - mobile phone PA intervention for sedentary women (n = 42). Increase in total steps from 5394 (95% CI 4563–6224) to 6210 (95% CI 5379–7041) over three weeks.

(Bort-Roig et al., 2014) - systematic review: three pre-post studies reported effects on step count. Mean physical activity increases ranged from 800 to 1,104 steps per day.

(Ganesan et al., 2016) - international mobile health intervention (n = 36,562). Mean increase in step count of 3,519 steps (95% confidence interval 3,484 to 3,553 steps).

There is a large variation in effect sizes in studies using the outcome of step count, but mean increase in daily step count ranges from 300 to 1000 or higher, with standard deviations (SD) of approximately 500 to 2000.

Table 1: Estimated sample size based on common standard deviation of 1000

Mean difference in daily step count	**Correlation**		
	0.3	**0.5**	**0.8**
300	125	90	37
500	46	34	15
800	20	15	8
1000	14	10	10

Table 2: Estimated sample size based on common standard deviation of 2000

Mean difference in daily step count	Correlation		
	0.3	**0.5**	**0.8**
300	491	351	142
500	178	128	53
800	71	52	22
1000	46	34	15

Assuming a correlation of 0.3 for pre- and post-intervention outcomes, a sample size of 125 would be needed to detect a mean difference in daily step count of 300 (with a standard deviation, SD, of 1000), or assuming a correlation of 0.5, a sample size of 128 would be needed to detect a mean difference of 500 steps (with SD of 2000).

IPAQ Outcomes: Weekly minutes of moderate-to-vigorous physical activity (MVPA)

As with step count, the magnitude of effect size is variable; a systematic review found change in total duration of physical activity ranged from 1.5 to 153 minutes/week (Afshin et al., 2016).

(Thompson et al., 2016) - mean weekly minutes of MVPA in community-dwelling men = 852 mins (SD = 386) (accelerometry/commercial activity monitor-derived). The authors recommended aiming for a post-intervention target of around 1000 minutes.

Table 3: Estimated sample size based on weekly minutes of moderate-to-vigorous physical activity (MVPA) (based on Thompson et al., 2016)

Mean weekly minutes of MVPA – pre-intervention	Mean weekly minutes of MVPA – post-intervention	Standard deviation (pre-intervention)	Correlation		
			0.3	**0.5**	**0.8**
852	1000	386	77	56	24

SF-12 Health Survey

The SF-12 is based on a national norm (18+ years) score of 50 (where 0 = lowest level of health and 100 = highest level) and SD of 10.0. The clinically important difference for SF-12 is one standard deviation (i.e. 10-point difference) (Shulman et al., 2010, Clement et al., 2014). However, a smaller change (e.g. 3 points) may be expected in the study population.

Table 4: Estimated sample size based on mean change in SF-12 of 3 points

Mean change in SF-12 score	Standard deviation	Correlation		
		0.3	0.5	0.8
3	10	125	90	37

Based on the above, a sample size of 128 will be sufficient to detect:

- A change in mean daily step count of 300 (SD = 1000) assuming a correlation of 0.3 for pre- and post-intervention outcomes
- A change in mean daily step count of 500 (SD = 2000) assuming a correlation of 0.5 for pre- and post-intervention outcomes
- An approximate mean increase in weekly minutes of moderate-to-vigorous physical activity (MVPA) of 150 (SD = 386) assuming a correlation of 0.3 for pre- and post-intervention outcomes
- A mean increase in SF-12 score of 3 points (SD = 10) assuming a correlation of 0.3 for pre- and post-intervention outcomes

To allow for around 15% attrition over time, it was decided to increase the sample size to 150.

(Due to higher than expected demand for participation and an ample supply of Fitbit® devices, sample size was later increased to 180 to allow for a higher attrition rate of approximately 25-30%.)

Sample size calculations were not performed for the Perceived Stress Scale (PSS-4) and WHO Health and Work Performance Questionnaire (HPQ) (absenteeism and presenteeism) as these were secondary outcomes and there were few comparative studies. This was particularly the case for the short version (absenteeism and presenteeism) of the HPQ as it is a relatively new outcome measure.

Appendix 4: Participant documents

4a. Main study: Participant information sheet

The PAW-Force (Physical Activity Wearables in the Police Force) Trial

INFORMATION SHEET FOR PARTICIPANTS
VERSION NUMBER [2]: DATE [07/03/17]

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate.

What is the aim of the project?

This project is being undertaken as part of a PhD in Medical Studies with the University of Exeter. The main aim is to determine the impact of introducing wearable fitness technology (personal activity monitors linked to a smartphone app) as a motivator for increasing physical activity within the police force. We will look at which features of the app are most effective and who they work best for. The project will also look at the wider effects of physical activity on health, wellbeing and stress, and absence and personal perceived productivity in the workplace.

Description of participants required

We are looking for male and female participants of all ages, who are currently employed at either the Plymouth Basic Command Unit of the Devon & Cornwall Police or at the North Dorset site of the Dorset Police.

It doesn't matter if you have or haven't used a physical activity monitor before, as long as you have a smartphone or tablet (Apple or Android 4.0.3 or higher), with Bluetooth and internet access, that you are able to use during the study*. We are looking for participants with various levels of fitness, from a range of occupations (including police officers, PCSOs, special constables and other police staff).

What will participants be asked to do?

Should you agree to take part in this project, you will be given an activity monitor (Fitbit Charge 2™) which you will be asked to wear on its own for a week, before linking it to the Bupa Boost app on your smartphone or tablet (you will be given instructions on how to register the activity monitor and link it to the app*). We will ask you to set yourself some specific goals, including improving your daily step count, and record these on the app. You will then be asked to continue to wear the Fitbit for the next 12 weeks while monitoring your activity levels and using various features of the app (we will tell you which features you can use and when). After the

12 week period, you will be free to continue to use the activity monitor and app as you please for the next 5 months.

Week 0	**Weeks 1-6**	**Weeks 7-12**	**Week 13 to end of study (8 months from beginning)**
			
Beginning of study: Wear Fitbit on its own for a week	Wear Fitbit and use ‘Bupa Boost’ app	Wear Fitbit and use ‘Bupa Boost’ app with colleagues	Continue to wear Fitbit and use ‘Bupa Boost’ app as desired

You will be asked to complete an online questionnaire and send data on your step count (collected by the Fitbit) at four time points: **the beginning of the study,** after **6 weeks**, after **12 weeks**, and after **8 months**. Again, you will be given clear instructions on how to do this.

A small number of participants will be followed up by the researcher to take part in up to three short interviews, to find out more about their experience of using the fitness technology. Interviews will take place at the **beginning of the study,** after **12 weeks** and after **8 months**. You can choose whether you want to be on the list for possible interviews at the beginning of the study. Interviews will take place in your workplace or via phone or Skype at a time that is convenient to you.

If you choose to give permission, we will also look at your individual staff sickness and absence records within your organisation. Your personal details will be kept anonymous and will only be reported combined with the other participants.

**The Bupa Boost app is available to download on iPhone 4S+, iPad (2nd Gen)+ and iPad Mini (1st Gen)+ devices running iOS 7.0 or later. It is supported by most Android devices – you can check compatibility on Google Play. The Fitbit app works with most iOS and Android devices – you can check compatibility here: https://www.fitbit.com/uk/devices.*

Time commitment

Taking part in the research project will involve giving up some time but we hope your involvement will be enjoyable and not too time-consuming overall. The project will be carried out over 8 months.

You will be asked to wear the Fitbit continuously (during waking hours and overnight if you are able to) for a total of 13 weeks initially, and over the following five months as you desire.

Each of the four questionnaires should take no more than 15-20 minutes to complete (including sending your step count data). Interviews are expected to last around 30 minutes, and will be no longer than an hour.

Will there be any cost to me in taking part?

You will be expected to use your own smartphone or tablet to access the Fitbit and Bupa Boost apps, and you will also need access to a computer with internet access for completing the questionnaires and sending your step count data. This will involve data usage which may incur a charge from your mobile network or internet service provider.

Will I be able to keep the Fitbit at the end of the study?

The Fitbits are property of the Devon & Cornwall Police. The Force has agreed that participants may continue to use them after the study has finished, as long as they remain employed by either the Devon & Cornwall Police or Dorset Police.

Can participants change their mind and withdraw from the project?

You may withdraw from participation in the project at any time without any disadvantage to yourself of any kind. Any data you have provided up to this point will be removed from our records at your request. If you withdraw while using one of the wearable activity monitors this will need to be returned to your workplace.

What data or information will be collected and what use will be made of it?

What information are you collecting?

We will collect information on your physical activity level at various time points. This will include your step count, the time you have spent being inactive (sedentary time), and the time you have spent engaging in light, moderate and vigorous physical activity over the previous week.

The questionnaires will collect some information about your background (e.g. ethnicity, marital status), your occupation, your general health status, stress, absence and personal perceived productivity at work. We would also like to know how long you have spent using the activity monitor and app and how useful you found it.

If you agree to take part we will keep some personal details such as your name, date of birth and work address for identification purposes during the study only. All information you give will be kept completely confidential.

What will be collected from the interviews?

In the interviews we will ask you more about your experiences of using the wearable fitness technology and the impact it has had on your activity levels, health and wellbeing.

<u>Please note:</u>

The project involves an open-questioning technique where the precise nature of the questions asked have not been determined in advance, but will depend on the way in which the interview develops. Consequently, although the School Research Ethics Committee is aware of the general areas to be explored in the interview, the Committee has not been able to review the precise questions to be used.

In the event that the line of questioning does develop in such a way that you feel hesitant or uncomfortable, you are reminded of your right to decline to answer any particular question(s) and also that you may withdraw from the project at any stage without any disadvantage to yourself of any kind.

What will my data be used for?

Your individual data will be securely stored and accessed only by the research team and not passed to any external organisations. Some data collected by Bupa Boost (goals set and achieved, activity, badges and wellness points) will be available to the police force, but this will all be anonymous and in a summary for the whole workforce (not you as an individual). Bupa do not intend to collect any data for their own use.

The overall findings may be used to inform a larger trial of wearable fitness technology within the Devon & Cornwall Police and Dorset Police. Results of this project may be published but any data included will not be individually identifiable.

What data will be shared with my colleagues?

In weeks 7 to 12, we will ask you to link up with your colleagues in the Bupa Boost app. If you wish, you will be able to share data on goals set and achieved, step count, badges and wellness points, and take part in competitions and group challenges. You will have full control over which goals and information you choose to share and which you want to keep to yourself (we will tell you how to do this), and which colleagues you share these with. It will not be possible to share private data such as date of birth and weight.

Can I see a summary of the results?

Participants will be given a summary of the findings at the end of the study, and provided with a copy of the final report on request.

What if participants have any questions?

If you have any questions about our project, either now or in the future, please feel free to contact either:-

Sarah Buckingham (student) or Dr. Karyn Morrissey (supervisor)
European Centre for Environment and Human Health (ECEHH),
Knowledge Spa, Truro, TR1 3HD
Tel: (01872) 255179 or (01872) 258158
Email: sab246@exeter.ac.uk or K.Morrissey@exeter.ac.uk

Complaints

If you have any complaints about the way in which this study has been carried out please contact the Co-chairs of the University of Exeter Medical School Research Ethics Committee:-
Ruth Garside, PhD *or* Rob Anderson, PhD
Co-chairs of the UEMS Research Ethics Committee
Email:uemsethics@exeter.ac.uk

This project has been reviewed and approved by the University of Exeter Medical School Research Ethics Committee (Reference Number: 17/02/116)

4b. Main study: Screening and consent forms

UNIVERSITY OF EXETER | MEDICAL SCHOOL

The PAW-Force (Physical Activity Wearables in the Police Force) Trial

CONSENT FORM FOR PARTICIPANTS : MAIN TRIAL
VERSION NUMBER [2]: DATE [07/03/17]

I have read the Information Sheet Version Number [2] Dated [07/03/17] concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I agree to:

1. Participate in the trial (involving wearing of a Fitbit physical activity monitor, use of the Bupa Boost app and completing online questionnaires);

Yes/No

2. Use my own smartphone or tablet to access the Fitbit and Bupa Boost apps and my own PC to complete online questionnaires (which may incur data usage charges from my mobile network or internet service provider)

Yes/No

I know that:

1. My participation in the project is entirely voluntary;

Yes/No

2. I am free to withdraw from the project at any time without any disadvantage;

Yes/No

3. My personal details and individual questionnaire data will be retained in secure storage and accessed only by the research team;

Yes/No

4. I have full control over the data (i.e. goals set and achieved, step count, badges and wellness points) I choose to share in the Bupa Boost app and which colleagues I share this with;

Yes/No

5.	Data on my physical activity goals may be shared with the police force via Bupa Boost, but this will not allow identification of me as an individual;

Yes/No

6.	The results of the project may be published but my anonymity will be preserved.

Yes/No

I agree to take part in this project.

This project has been reviewed and approved by the University of Exeter Medical School Research Ethics Committee **UEMS REC REFERENCE NUMBER: 17/02/116**

The PAW-Force (Physical Activity Wearables in the Police Force) Trial

CONSENT FORM FOR PARTICIPANTS : INTERVIEWS

VERSION NUMBER [2]: DATE [07/03/17]

I have read the Information Sheet Version Number [2] Dated [07/03/17] concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I agree to:

1. Take part in up to three interviews, which will involve open-questioning about my expectations and experiences of using the wearable fitness technology and physical activity

Yes/No

I know that:

1. My participation in the interviews is entirely voluntary;

Yes/No

2. I am free to withdraw from taking part in the interviews at any time without any disadvantage;

Yes/No

3. The audio recordings and interview transcripts will be retained in secure storage and accessed only by the research team;

Yes/No

4. The results of the project may be published but my anonymity will be preserved.

Yes/No

I agree to take part in this project.

This project has been reviewed and approved by the University of Exeter Medical School Research Ethics Committee **UEMS REC REFERENCE NUMBER: 17/02/116**

The PAW-Force (Physical Activity Wearables in the Police Force) Trial

CONSENT FORM FOR PARTICIPANTS : ATTENDANCE RECORDS
VERSION NUMBER [2]: DATE [07/03/17]

I have read the Information Sheet Version Number [2] Dated [07/03/17] concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I agree to:

1. Allow my staff attendance record including sickness and absence data to be accessed by the research team;

Yes/No

2. Allow data to be collected on absence time and reasons for sickness absence just prior to and for the duration of the study.

Yes/No

I know that:

1. My participation in this aspect of the study is entirely voluntary;

Yes/No

2. I am free to withdraw my data at any time without any disadvantage;

Yes/No

3. The data collected will be retained in secure storage and accessed only by the research team;

Yes/No

4. My personal details will be kept anonymous and all data collected will only be reported combined with the other participants.

Yes/No

I agree to take part in this project.

This project has been reviewed and approved by the University of Exeter Medical School Research Ethics Committee **UEMS REC REFERENCE NUMBER: 17/02/116**

4c. Survey with managers, commissioners and occupational health staff: Participant information sheet

The PAW-Force (Physical Activity Wearables in the Police Force) Trial

INFORMATION SHEET FOR MANAGERS/COMMISSIONERS SURVEY
VERSION NUMBER [1]: DATE [24/01/17]

Thank you for showing an interest in this project. Please read this information sheet carefully before deciding whether or not to participate.

What is the aim of the project?

This project is being undertaken as part of a PhD in Medical Studies with the University of Exeter. The main aim is to determine the impact of introducing wearable fitness technology (personal activity monitors linked to a smartphone application) as a motivator for increasing physical activity within the police force. We are assessing which features of the app are most effective and who they work best for. The project is also looking at the wider effects of physical activity on health, wellbeing and stress, and absence and perceived productivity in the workplace.

In addition to interviews with participants taking part in the trial regarding their experiences of using the technology and its impact on their physical activity levels, we would like to find out the opinions of some managers, commissioners and occupational health staff within the police force. We would like your views on feasibility and acceptability of using such technology within your organisation, and effective and cost-effective ways to improve activity levels of police officers and staff.

Description of participants required

We are looking for managers, commissioners and occupational health staff who are currently employed by either the Devon & Cornwall Police or the Dorset Police Force.

What will participants be asked to do?

You will be asked to complete a brief online questionnaire.

Time commitment

The questionnaire should take around 10-15 minutes to complete.

Can participants change their mind and withdraw from the project?

You may withdraw from participation in the project at any time without any disadvantage to yourself of any kind.

What data or information will be collected and what use will be made of it?

What information are you collecting?

We will collect information on your occupation, recommended policies or strategies for improving physical activity levels within the workplace, your views on wearable fitness technology, and whether you think this might be an effective, cost-effective, acceptable and feasible intervention.

What will my data be used for?

The information you give will be securely stored, accessed only by the research team and not passed to any external organisations. This study will test the research methods we are using and the findings will be used to inform a larger trial of wearable fitness technology within the Devon & Cornwall Police and Dorset Police. Results of this project may be published but any data included will not be individually identifiable.

Can I see a summary of the results?

Participants will be given a summary of the findings of the trial at the end of the study, and provided with a copy of the final report on request.

What if participants have any questions?

If you have any questions about our project, either now or in the future, please feel free to contact either:-

Sarah Buckingham (student)
Room F37b
European Centre for Environment and Human Health (ECEHH)
Knowledge Spa
TR1 3HD

University Tel. No:- (01872) 255179
Email address: sab246@exeter.ac.uk

or

Dr. Karyn Morrissey (supervisor)
European Centre for Environment and Human Health (ECEHH)
Knowledge Spa
TR1 3HD

University Tel. No:- (01872) 258158
Email address: K.Morrissey@exeter.ac.uk

Complaints

If you have any complaints about the way in which this study has been carried out please contact the Co-chairs of the University of Exeter Medical School Research Ethics Committee:-

Ruth Garside, PhD *or* Rob Anderson, PhD
Co-chairs of the UEMS Research Ethics Committee
Email: uemsethics@exeter.ac.uk

This project has been reviewed and approved by the University of Exeter Medical School Research Ethics Committee (Reference Number: 17/02/116)

4d. Survey with managers, commissioners and occupational health staff: Screening and consent form

UNIVERSITY OF EXETER | MEDICAL SCHOOL

The PAW-Force (Physical Activity Wearables in the Police Force) Trial

CONSENT FORM FOR MANAGERS/COMMISSIONERS
VERSION NUMBER [1]: DATE [24/01/17]

I have read the Information Sheet Version Number [1] Dated [24/01/17] concerning this project and understand what it is about. All my questions have been answered to my satisfaction. I understand that I am free to request further information at any stage.

I agree to:

1. participate in this study and am happy to complete a questionnaire about the use of wearable fitness technology within the police force **Yes/No**

I know that:

1. my participation in the project is entirely voluntary; **Yes/No**
2. I am free to withdraw from the project at any time without any disadvantage; **Yes/No**
3. the data will be retained in secure storage; **Yes/No**
4. the results of the project may be published but my anonymity will be preserved. **Yes/No**

This project has been reviewed and approved by the University of Exeter Medical School Research Ethics Committee UEMS REC REFERENCE NUMBER 17/02/116.

4e. Participant instructions/user guide for Fitbit® and Bupa Boost app

Introduction

Thank you for registering to take part in the Physical Activity Wearables in the police force (PAW-Force) study. I am pleased to inform you that you are eligible to take part. You will be given a Fitbit Charge 2 activity monitor which you will be asked to wear continuously (during waking hours and overnight if possible) for 13 weeks from Thursday 15th June.

- **Details of when/where to collect the Fitbit (North Dorset – Inspector's office, Blandford Police Station and Plymouth – Admin Services Centre, Crownhill). Deadline for collection 14th June.**

When you have collected the Fitbit you will be sent the first instructions via e-mail. You should not set up or start to wear the Fitbit until you receive these instructions. Please do not remove the sticker from the screen!

After Collecting the Fitbit: The First Week (sent to all participants at the same time).

You should now have collected your Fitbit. If not please collect it as soon as possible; Fitbits will be available for collection until **Wednesday 14th June.** You should register and begin to wear the Fitbit **on Thursday 15th June**.

Each Fitbit has a unique sticker to cover the screen – it is important that you do not remove this sticker until after you have worn the Fitbit for at least 8 days!

What to do now:

Registering the Fitbit:

- First set up the Fitbit on your smartphone or tablet (see www.fitbit.com/setup). You will need to download the Fitbit app (from the App Store (Apple) or Google Play (Android)) and follow the on-screen instructions to set up the device. See the following page for more help with setting up the Fitbit: http://help.fitbit.com/articles/en_US/Help_article/1873/?l=en_US&fs=Search&pn=1
- Register the Fitbit as prompted with your e-mail address and unique password. Make sure you have read the terms and conditions and privacy policy carefully, including the medical information (see box below).

It is important to read the 'Consult your doctor before using the Fitbit service' section of the Terms and Conditions. If you have a medical or heart condition, you should consult your doctor before using the Fitbit or engaging in any exercise programme. You should also consult your doctor prior to use if you have epilepsy, are sensitive to flashing lights or are taking any photosensitive medicine, if you have reduced circulation or bruise easily, or have tendonitis, carpal tunnel syndrome or other musculoskeletal disorders.

Some people may experience skin irritation when wearing the Fitbit. To reduce the risk of irritation: 1. Keep it clean 2. Keep it dry 3. Don't wear it too tight and 4. Give your wrist a rest by removing the band for an hour after extended wear. If you do notice any irritation, remove the device and consult your occupational health service or doctor if symptoms persist.

- Fill in details of your height, weight, age and gender as requested (this is used to calculate your Base Metabolic Rate (BMR) and personalised fitness statistics).
- Select 'I agree' when you have read all terms and policies, then follow the on-screen instructions to pair the Fitbit with the app. If 'update available' screen appears in the app, select 'Next'. Keep the Fitbit close to your smartphone or tablet during setup – this may take up to 10 minutes!
- Select 'Next' and follow the on-screen tutorial.
- Select which wrist you will be wearing the Fitbit on when prompted – this should be your non-dominant wrist.
- When you get to 'Answer 3 questions so we can help you set the right goals for you', **CLOSE THIS SCREEN** (select the 'x' on the top left of the screen). **Don't set any goals yet as this will be done when you start using Bupa Boost.**
- Go to the 'General settings' screen within the app (click on the Fitbit icon to get to this screen) and do the following:
    - Set 'handedness' and 'wrist placement' (this will be right handedness and left wrist placement for right-handed people, vice-versa for left-handed people)
    - Turn 'Reminders to Move' off
    - Leave all other settings as default. 'All-Day Sync' should be left on.
- Verify your e-mail address by clicking on the link in the e-mail sent to you after registration. This will allow you to log in to the Fitbit website later.

Charging the Fitbit:

- The Fitbit should be charged at least every 5 days by connecting the charger to your computer. It is a good idea to charge it before the first time you start to wear it. See the following page for instructions: http://help.fitbit.com/articles/en_US/Help_article/1799/?l=en_US&c=Topics%3ABattery&fs=Search&pn=1

Wearing the Fitbit:

- Wear the Fitbit continuously for 8 days **from Thursday 15th June to Friday 23rd June while continuing with your normal day-to-day life and maintaining your usual activity levels**. This first week is just to give us an idea of your normal activity levels so we can see how this changes through the study.
- You should not use the Fitbit app to view your data, or view the data on the Fitbit screen during this time. It is important that you do not remove the sticker from the screen! Please also ignore any e-mails and mobile notifications from the app.
- After wearing the Fitbit for 8 days you will receive an e-mail with a link to the first questionnaire of the study. Look out for this e-mail as the first questionnaire will only be available for a few days!

The full manual for the Fitbit Charge 2 is available via the following link: https://staticcs.fitbit.com/content/assets/help/manuals/manual_charge_2_en_US.pdf

Please note the Fitbit is not waterproof! You should remove it for swimming, showering etc.

End of First Week

Please follow the instructions below **on Friday 23rd June (or for the late starters, after you have worn the Fitbit for 8 consecutive days!)**.

What to do now:

- Remove the sticker from the screen of the Fitbit!
- Sync the Fitbit with the app – this should be done automatically as long as ‘All-Day Sync’ is on. For more instructions see: http://help.fitbit.com/articles/en_US/Help_article/1877/?l=en_US&c=Topics%3ASyncing&fs=Search&pn=1
- Log in to the Fitbit website on your computer using the username and password you used to register: https://www.fitbit.com/login. The ‘Dashboard’ page should open. **Please note you will need a PC or laptop for this step – it is not possible within the Fitbit app.**

- To obtain your step data for the last 7 days:
    - Go to the ‘View Settings’ icon (top right of screen) on the Fitbit dashboard and select ‘Settings’
    - Select ‘Data Export’ from the Settings menu
    - On the ‘Export your data’ screen enter the Custom time period from a week ago to yesterday, e.g. if today’s date is 23/6/17, select from 16/6/17 to 22/6/17.
    - Make sure the ‘Activities’ box is checked under ‘Data’.
    - Keep the default file format of CSV and click ‘Download’.
    - Open and save your data file using either the default file name or change it to one of your own. You will need to access this file to upload it later.
    - (If your file doesn’t show complete step count data for the last 7 days, make sure it is up to date by checking you have recently synched the Fitbit)
- Complete the first questionnaire (see link in separate e-mail). Upload your Fitbit data file within the questionnaire as requested.

Next 6 Weeks: Individual Goal-Setting

Please do not read this section until after you have completed the first questionnaire! (Remember the questionnaire can be completed after wearing the Fitbit for 7 full days)

What to do after completing the first questionnaire:

- Turn back on ‘Reminders to Move’ under ‘General settings’ within the Fitbit app if you wish.

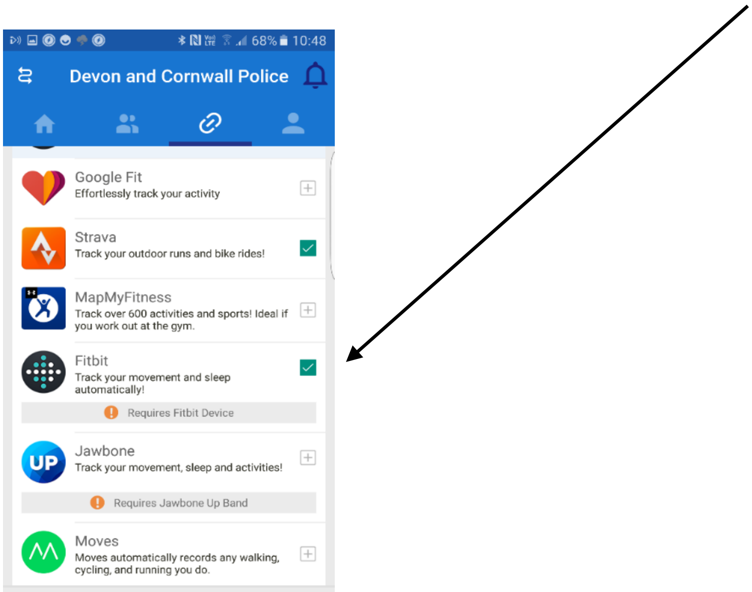
Register with Bupa Boost:

- Download the Bupa Boost app from either the App Store (Apple) or Google Play (Android)*
- Open the app and follow the on-screen instructions to register:
    - Enter the company PIN – this is ‘********’ (you will automatically be part of the ‘DCD Police Alliance’ group)
    - Sign up with your name, e-mail address (we recommend you use the same address that you used to register with Fitbit) and a unique password. Don’t forget to verify your e-mail address.

*The Bupa Boost app is available to download on iPhone 4S+, iPad (2nd Gen)+ and iPad Mini (1st Gen)+ devices running iOS 7.0 or later. It is supported by most Android devices – you can check compatibility on Google Play.

Connect to Fitbit:

- Connect to the Fitbit app when prompted – tick the 'Fitbit' box:



- Log in to the Fitbit account when prompted (with your e-mail and password)
- Tick the boxes for the data that you would like to be linked – make sure 'Activity and Exercise' is checked. On the next screen you can choose whether to hide your sleep data.

Set your first fitness goal(s):

- Set yourself at least one specific fitness goal. **One of these <u>must</u> be to improve your daily step count.** Walking 10,000 steps per day is recommended but if you are sedentary you might want to start with somewhere between 5,000 and 10,000 steps per day. (This will be a custom / 'Create your own' goal):

- Enter the goal description, e.g. ‘walk 10,000 steps per day’
- Select goal type = fitness
- Choose whether to make your goal private in ‘goal privacy’
- Set frequency as ‘every day’
- Set a reminder if you wish
- You are welcome to set yourself any additional fitness goals (either suggested by the app or custom goals) or goals from any of the other areas (mindfulness, nutrition and relaxation) as well. Remember that goals should be SMART:

SMART Goals

Specific: target a specific area for improvement, e.g. cycle to work three days per week

Measurable: easy to determine whether it has been achieved (e.g. 10,000 steps per day)

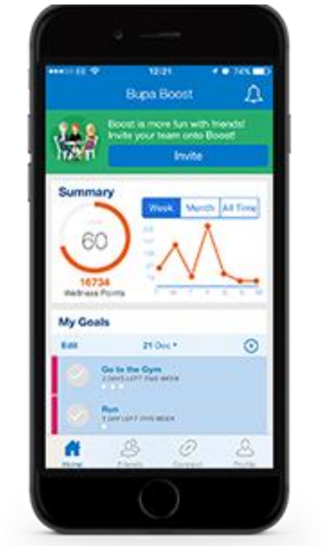
Achievable: is it attainable and realistic?

Relevant: is it important to you?

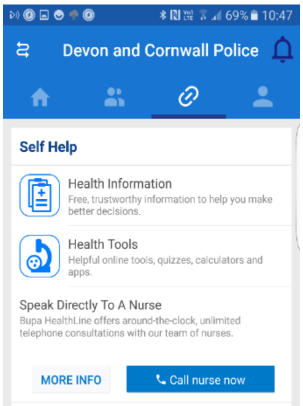
Time-bound: make sure you have a specific timeframe for completion (e.g. 12 weeks’ time)

- For the next 6 weeks you should use the individual goal-setting features of the app – monitoring goals set and achieved, earning wellness points, badges and working your way through the various levels. Mark off (tick) your goals in the app as you achieve them. Each user receives a set amount of wellness points for completing a goal. More points are awarded if a user hits their weekly target or “goal streak”. Points for steps are awarded based on the number of steps you have done on a given day.

- You will mainly be using the ‘Summary’ screen (home icon) for this part of the study, to view a summary of your activity, goals and points:



- You may also use the third screen (paperclip icon) to access Bupa’s advice and information. The app provides you with access to a bank of A to Z health information guides, health tools such as BMI calculator, brainy app, nutrition quiz, diabetes risk score etc. and also the ability to speak directly to a BUPA nurse 24/7.



- The fourth screen (single person icon) gives an additional summary of the points and badges you have earned.
- **Don’t use the app to find friends, team members or to join in with any fitness challenges. This is important - at this stage we just want you to concentrate on your own fitness goals.** (If you have already connected with others within the Fitbit or Bupa Boost app, please refrain from sending messages and taking part in any social challenges for the next 6 weeks).
- During weeks 5-6 you will be sent a link via e-mail to complete the next questionnaire and to upload your Fitbit data.

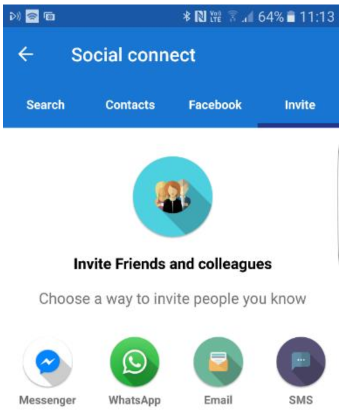
Next 6 Weeks: Social Features of the Bupa Boost App

Only read this section after you have completed the second (mid-intervention) questionnaire!

What to do now:

Find your colleagues:

- For this part of the study you may now link up with colleagues within the app:
    - Go to the second screen in Bupa Boost (two people icon) and select 'Invite friends'.
    - From the next screen (see below) you can find friends and colleagues who are also using the app via 'Search' (other users in your workplace), 'Contacts' (your phone contacts), 'Facebook', or 'Invite' new users (but you should not share your PIN with anyone outside of your organisation).



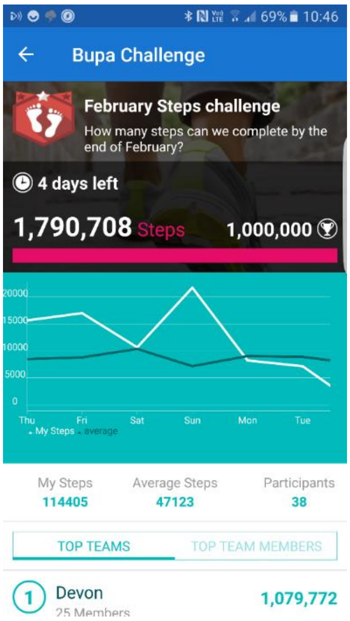
- Now you can stay in the loop with social updates (the social feed) to see how colleagues and friends are getting on. You can celebrate, encourage and reward your peers through the Bupa Boost messaging platform (or you can also jeer and taunt them!!)

Take part in competitions/challenges:

You can compete against your colleagues either on your own (group challenges/leader board) or as part of a team (force-wide/team challenges).

- If you wish, you can set your own challenges and invite your friends to participate (e.g. to see who can walk the most steps!).
- You can now sign up for any force-wide challenges. Just choose your team and then add your weight to getting your team to the top and winning the challenge (no prizes, just kudos and bragging rights!)

- Challenges may be steps, running, cycling, swimming or wellness point based, e.g. walk a million steps, cycle 100k miles, swim the length of the English Channel.
- **Your original goal of improving your daily step count is still important for this part of the study! A steps challenge has been set for August – please join in with this!**



Teams:

- You will automatically be part of the ‘DCD Police Alliance’ group but you can create your own teams, (for example your station or department), invite friends to your team and see where you are on your own leader board.

- *During weeks 11-12 you will be sent a link via e-mail to complete the next questionnaire and to upload your Fitbit data.*

<u>Months 4 to 8</u>

For the remainder of the study you should continue to use the Fitbit and the Bupa Boost app together as you wish. We hope you will still find them useful. At the end of month 8 you will receive a link to the final questionnaire, where you will also be able to upload your Fitbit data.

Appendix 5: Online surveys

5a. Baseline (week 0) questionnaire

Thank you for agreeing to complete the first questionnaire for this study. Please answer all of the questions on the following pages. The questionnaire should take around **10-15 minutes** to complete.

Your individual answers will be kept secure and confidential and accessed only by the research team. Your data will not be shared in an identifiable way with the police force or any external organisations.

If during completion you have any concerns about your work, health or any related issues, you may wish to contact your line manager, your occupational health department or your GP.

You may request to withdraw your data at any time by e-mailing the researcher (Sarah Buckingham) at: sab246@exeter.ac.uk.

This study has been approved by the University of Exeter Medical School Research Ethics Committee.

About You

Please enter your age in years *

What is your gender? *

Please choose **only one** of the following:

- ○ Male
- ○ Female
- ○ Other

What is your ethnic group? *

Please choose **only one** of the following:

- ○ White: English / Welsh / Scottish / Northern Irish / British
- ○ White: Irish
- ○ White: Gypsy or Irish Traveller
- ○ Any other White background
- ○ Mixed: White and Black Caribbean
- ○ Mixed: White and Black African
- ○ Mixed: White and Asian
- ○ Any other Mixed / Multiple ethnic background
- ○ Asian: Indian

- ◯ Asian: Pakistani
- ◯ Asian: Bangladeshi
- ◯ Asian: Chinese
- ◯ Any other Asian background
- ◯ Black African
- ◯ Black Caribbean
- ◯ Any other Black / African / Caribbean background
- ◯ Arab
- ◯ Any other ethnic group
- ◯ Prefer not to say

Choose one option that best describes your ethnic group or background.

If other ethnicity, please describe... *

What is your current legal marital or same-sex civil partnership status? *

Please choose **only one** of the following:

- ◯ Single / Never married or civil partnered
- ◯ Married or civil partnership
- ◯ Separated (still legally married or in a civil partnership)
- ◯ Divorced (or in a legally dissolved civil partnership)
- ◯ Widowed or surviving civil partner
- ◯ Prefer not to say

Do you live alone in your current home? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No
- ◯ Prefer not to say

(If no to living alone) Not including yourself, how many people aged 18 years or over live with you in your current home?

Please write your answer here:

-

Enter '0' if you don't live with any other adults aged 18 years or over.

(If no to living alone) How many children aged under 18 years live with you in your current home?

Please write your answer here:

-

Enter '0' if you don't live with any children under 18 years.

Are you a dog owner? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

What is the postcode of your current main residence?

Please write your answer here:

Do you consider your residence to be... *

Please choose **only one** of the following:

- ◯ Urban - city
- ◯ Urban - town
- ◯ Suburban
- ◯ Rural village
- ◯ Hamlet or isolated rural dwelling

What is the highest level of education you have completed? *

Please choose **only one** of the following:

- ◯ No formal qualifications
- ◯ Lower secondary school (e.g. GCSE / CSE / O-level / Standard Grade / Intermediates)
- ◯ Upper secondary school (e.g. AS or A-level / Scottish Highers)
- ◯ Professional or technical qualification below degree level (e.g. HNC / HND / City & Guilds advanced certificate / Level 3 Diploma in Policing)
- ◯ University / college degree
- ◯ Postgraduate (masters / PhD)

About Your Work

Which organisation do you work for? *

Please choose **only one** of the following:

- ◯ Devon and Cornwall Police

- ○ Dorset Police

On which site are you based? *

Please choose **only one** of the following:

- ○ Plymouth - Crownhill
- ○ Plymouth - Charles Cross
- ○ Plymouth - Devonport
- ○ Plymouth - Plympton or Plymstock
- ○ Plymouth - Exmouth Road
- ○ North Dorset - Blandford
- ○ North Dorset - Gillingham
- ○ North Dorset - Shaftesbury
- ○ North Dorset - Sturminster Newton
- ○ Other

If you work on more than one site, please select the site where you work most frequently.

Which of the following teams are you a part of? *

Please choose **only one** of the following:

- ○ Response policing
- ○ Local Policing Partnerships / Neighbourhood
- ○ Local investigation (e.g. SODAIT)
- ○ Other

What is your job title? *

Please write your answer here:

If you have more than one occupation within the police force, please state your main role.

What is your job category? *

Please choose **only one** of the following:

- ○ Police Officer
- ○ Police Community Support Officer
- ○ Special Constable
- ○ Police Staff

If you are a police officer, what is your rank? *

Please choose **only one** of the following:

- ○ Constable
- ○ Sergeant
- ○ Inspector
- ○ Chief Inspector
- ○ Superintendent
- ○ Chief Officers Group

Would you consider your role to be mainly active or sedentary (e.g. desk-based)? *

Please choose **only one** of the following:

- ○ Mainly active (standing, walking or manual work)
- ○ Mainly sedentary (sitting at a desk or driving)
- ○ Equally active and sedentary

In total, how many years have you worked for the police force (including the current and any previous organisations)? *

Please write your answer here:

Please state the number of complete years of service. This may be approximate if you are not sure.

How many hours per week do you normally work? (Including contracted hours and overtime) *

Please choose **only one** of the following:

- ○ Less than 15
- ○ 15 to 30
- ○ 30 to 40
- ○ More than 40

Do you work shifts? *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If yes, do you work... *

Please choose **all** that apply:

- ☐ Morning (early) shift
- ☐ Afternoon (late) shift

- ☐ Night shift
- ☐ Rotating shift (e.g. 2+2+2 duty scheme)

What is your annual salary? *

Please choose **only one** of the following:

- ○ Less than £10,000
- ○ £10,000 to £19,999
- ○ £20,000 to £29,999
- ○ £30,000 to £39,999
- ○ £40,000 to £49,999
- ○ £50,000 to £59,999
- ○ £60,000 to £69,999
- ○ £70,000 or higher
- ○ Prefer not to say

About Your Health and Physical Activity

Are you... *

Please choose **only one** of the following:

- ○ A current smoker
- ○ An ex-smoker
- ○ Never smoked

If you are a current smoker...

Do you currently smoke at least once a day?

Please choose **only one** of the following:

- ○ Yes
- ○ No

If yes, how many cigarettes do you smoke on a typical day? *

Please write your answer here:

-

Enter '0' if none.

If you are a current smoker, how many cigars do you smoke on a typical day? *

Please write your answer here:

-

Enter '0' if none.

If you are a current smoker, how much tobacco (roll-ups) do you smoke in a typical week?

Please write your answer(s) here:

- Weight in ounces (approx)
- Weight in grams (approx)

Please complete one of the above boxes. Enter '0' if none.

If you are a current smoker, how much pipe tobacco do you smoke in a typical week?

Please write your answer(s) here:

- Weight in ounces (approx)
- Weight in grams (approx)

Please complete one of the above boxes. Enter '0' if none.

How often do you drink alcohol? (i.e. beer, lager, cider, wine, sherry, vermouth or spirits) *

Please choose **only one** of the following:

- ○ More than once a week
- ○ Once a week
- ○ Less than once a week, but once a month or more
- ○ Less than once a month / a few times a year
- ○ Never

Do you have any of the following doctor-diagnosed health conditions? *

Please choose **all** that apply:

- ☐ Asthma
- ☐ COPD, chronic bronchitis or emphysema
- ☐ Arthritis (osteo or rheumatoid)
- ☐ Chronic back pain
- ☐ Other connective tissue or musculoskeletal disorder
- ☐ Depression or anxiety
- ☐ Other mental health condition
- ☐ Heart disease or angina
- ☐ Hypertension (high blood pressure)
- ☐ Peripheral vascular disease
- ☐ Type 1 (autoimmune) diabetes

- ☐ Type 2 diabetes
- ☐ Chronic kidney disease or failure
- ☐ Liver disease
- ☐ Visual impairment
- ☐ None of the above
- ☐ Prefer not to say
- ☐Other please specify:

Prior to taking part in this study, had you ever used a wearable activity monitor or tracker (e.g. Fitbit, Jawbone, Withings, Garmin)? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

If yes, please give the brand name and model name if known (e.g. Fitbit Flex)

Please write your answer here:

Have you previously used a health or fitness app on your smartphone or tablet? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

If yes, which app or apps have you used? (Select all that apply) *

Please choose **all** that apply:

- ☐ Fitbit
- ☐ Apple Health
- ☐ Strava
- ☐ MapMyFitness
- ☐ One You Couch to 5K
- ☐ Google Fit
- ☐ Jawbone
- ☐Other please specify (if known):

On how many days have you worn the Fitbit over the last 7 days (not including today)? *

Please choose **only one** of the following:

- ◯ 0

- ◯ 1
- ◯ 2
- ◯ 3
- ◯ 4
- ◯ 5
- ◯ 6
- ◯ 7

If 0, please state why... *

Please write your answer here:

Over the last 7 days, for how many hours per day did you typically wear the Fitbit? *

Please write your answer here:

-

Enter a number between 0 and 24.

Have there been any events that you feel have affected your physical activity levels over the last 7 days? (E.g. illness, annual leave) *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

If yes, please give details... *

Please write your answer here:

International Physical Activity Questionnaire

These questions will ask you about the time you spent being physically active in the last 7 days. Please answer each question even if you do not consider yourself to be an active person. Please think about the activities you do at work, as part of your house and yard work, to get from place to place, and in your spare time for recreation, exercise or sport.

The following information sheet will help you to decide which activities can be classed as vigorous and moderate:

/upload/surveys/715495/files/Moderate%20and%20vigorous%20activities.pdf

Think about all the vigorous activities that you did in the last 7 days. Vigorous physical activities refers to activities that take hard physical effort and make you breathe much harder than normal. Think *only* about those physical activities that you did for at least 10 minutes at a time.

1. During the last 7 days, on how many days did you do vigorous physical activities like heavy lifting, digging, aerobics, or fast bicycling? *

Please choose **only one** of the following:

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

2. How much time did you usually spend doing vigorous physical activities on one of those days? *

Please write your answer(s) here:

- Hours per day:
- Minutes per day:

Please estimate the time in hours and minutes per day (e.g. 1 hour 30 minutes)

Think about all the moderate activities that you did in the last 7 days. Moderate activities refer to activities that take moderate physical effort and make you breathe somewhat harder than normal. Think only about those physical activities that you did for at least 10 minutes at a time.

3. During the last 7 days, on how many days did you do moderate physical activities like carrying light loads, bicycling at a regular pace, or doubles tennis? Do not include walking. *

Please choose **only one** of the following:

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

4. How much time did you usually spend doing moderate physical activities on one of those days? *

Please write your answer(s) here:

- Hours per day:
- Minutes per day:

Please estimate the time in hours and minutes per day (e.g. 1 hour 30 minutes)

Think about the time you spent walking in the last 7 days. This includes at work and at home, walking to travel from place to place, and any other walking that you have done solely for recreation, sport, exercise, or leisure.

5. During the last 7 days, on how many days did you walk for at least 10 minutes at a time? *

Please choose **only one** of the following:

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

6. How much time did you usually spend walking on one of those days? *

Please write your answer(s) here:

- Hours per day:
- Minutes per day:

Please estimate the time in hours and minutes per day (e.g. 1 hour 30 minutes)

The last question is about the time you spent sitting on weekdays during the last 7 days. Include time spent at work, at home, while doing course work and during leisure time. This may include time spent sitting at a desk, visiting friends, reading, or sitting or lying down to watch television.

7. During the last 7 days, how much time did you spend sitting on a week day? *

Please write your answer(s) here:

- Hours per day:
- Minutes per day:

Please estimate the time spent sitting per day in hours and minutes.

Step Count Upload

Please upload your step count data for the last 7 days. For instructions on how to do this click here:

/upload/surveys/715495/files/How%20to%20Obtain%20and%20Upload%20Your%20Step%20Count(1).pdf

Please upload at most one file

Kindly attach the aforementioned documents along with the survey

Click 'Upload files' above, then 'Select File' to find the step count file you have saved to your computer. This should be a CSV file. Select the file and click 'Open' to upload it.

SF-12 Health Survey

This survey asks for your views about your health. This information will help keep track of how you feel and how well you are able to do your usual activities.

For each of the following questions, please tick the one box that best describes your answer.

1. In general, would you say your health is:

Please choose **only one** of the following:

- ○ Excellent
- ○ Very good
- ○ Good
- ○ Fair
- ○ Poor

2. The following questions are about activities you might do during a typical day. Does your health now limit you in these activities? If so, how much?

Please choose the appropriate response for each item:

	Yes, limited a lot	Yes, limited a little	No, not limited at all
MODERATE ACTIVITIES, such as moving a table, pushing a vacuum	○	○	○

	Yes, limited a lot	Yes, limited a little	No, not limited at all
cleaner, bowling, or playing golf			
Climbing SEVERAL flights of stairs	○	○	○

3. During the <u>past 4 weeks</u>, how much of the time have you had any of the following problems with your work or other regular daily activities <u>as a result of your physical health</u>?

Please choose the appropriate response for each item:

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
ACCOMPLISHED LESS than you would like	○	○	○	○	○
Were limited in the KIND of work or other activities	○	○	○	○	○

4. During the <u>past 4 weeks</u>, how much of the time have you had any of the following problems with your work or other regular daily activities <u>as a result of any emotional problems</u> (such as feeling depressed or anxious)?

Please choose the appropriate response for each item:

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
ACCOMPLISHED LESS than you would like	○	○	○	○	○
Did work or other activities LESS CAREFULLY THAN USUAL	○	○	○	○	○

5. During the <u>past 4 weeks</u>, how much did pain interfere with your normal work (including both work outside the home and housework)?

Please choose **only one** of the following:

- ○ Not at all
- ○ A little bit
- ○ Moderately
- ○ Quite a bit

- ○ Extremely

6. These questions are about how you feel and how things have been with you <u>during the past 4 weeks</u>. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the <u>past 4 weeks</u>...

Please choose the appropriate response for each item:

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
Have you felt calm and peaceful?	○	○	○	○	○
Did you have a lot of energy?	○	○	○	○	○
Have you felt downhearted and low?	○	○	○	○	○

7. During the <u>past 4 weeks</u>, how much of the time has your <u>physical health or emotional problems</u> interfered with your social activities (like visiting with friends, relatives, etc.)?

Please choose **only one** of the following:

- ○ All of the time
- ○ Most of the time
- ○ Some of the time
- ○ A little of the time
- ○ None of the time

Perceived Stress Scale

The questions in this scale ask you about your feelings and thoughts during the last month. In each case, you will be asked to indicate by selecting <u>how often</u> you felt or thought a certain way.

In the last month, how often have you...

Please choose the appropriate response for each item:

	Never	Almost Never	Sometimes	Fairly Often	Very Often
Felt that you were unable to control the important things in your life?	○	○	○	○	○
Felt confident about your ability to handle your personal problems?	○	○	○	○	○
Felt that things were going your way?	○	○	○	○	○
Felt difficulties were piling up so high that you could not overcome them?	○	○	○	○	○

Health and Work Performance Questionnaire: Absenteeism and Presenteeism

1. About how many hours altogether did you work <u>in the past 7 days</u>?

Please write your answer here:

-

Enter a number between 0 and 97. (If more than 97, enter 97.)

2. How many hours does your employer expect you to work in a typical 7-day week? (If it varies, estimate the average)

Please write your answer here:

-

Please enter a number between 0 and 97. (If more than 97, enter 97.)

3. Now please think of your work experiences over the <u>past 4 weeks</u> (28 days). Select the number of days you spent in each of the following work situations.

In the <u>past 4 weeks</u> (28 days), how many days did you...

	Number of days (0 to 28)
Miss an ENTIRE work day because of problems with your physical or mental health? (Please include only days missed for your OWN health, not someone else's health)	
Miss an ENTIRE work day for any other reason (including vacation)?	
Miss PART of a work day because of problems with your physical or mental health? (Please include only days missed for your OWN health, not someone else's)	

Number of days (0 to 28)

Miss PART of a work day for any other reason (including vacation)?

Come in early, go home late, or work on your day off?

About how many hours altogether did you work in the past 4 weeks (28 days)? (See examples below)

Your answer must be between 0 and 400

Please write your answer here:

-

Examples for Calculating Hours Worked in the Past 4 Weeks

40 hours per week for 4 weeks = 160 hours

35 hours per week for 4 weeks = 140 hours

40 hours per week for 4 weeks with two 8-hour days missed = 144 hours

40 hours per week for 4 weeks with three 4-hour partial days missed = 148 hours

35 hours per week for 4 weeks with two 8-hour days missed and three 4-hour partial days missed = 112 hours

5. On a scale from 0 to 10 where 0 is the WORST job performance anyone could have at your job and 10 is the performance of a TOP worker...

Please choose the appropriate response for each item:

	0	1	2	3	4	5	6	7	8	9	10
How would you rate the usual performance of MOST workers in a job similar to yours?	○	○	○	○	○	○	○	○	○	○	○
How would you rate YOUR USUAL job performance over the	○	○	○	○	○	○	○	○	○	○	○

	0	1	2	3	4	5	6	7	8	9	10
PAST YEAR OR TWO?											
How would you rate your OVERALL job performance on the days you worked during the PAST 4 WEEKS?	○	○	○	○	○	○	○	○	○	○	○

Thank you very much for completing the questionnaire. Your responses have been recorded.

If you have any further questions about the study please contact Sarah Buckingham at sab246@exeter.ac.uk.

5b. Mid-intervention (week 6) questionnaire

Thank you for agreeing to complete the second questionnaire for this study. Please answer all of the questions on the following pages. The questionnaire should take around **10-15 minutes** to complete.

Your individual answers will be kept secure and confidential and accessed only by the research team. Your data will not be shared in an identifiable way with the police force or any external organisations.

If during completion you have any concerns about your work, health or any related issues, you may wish to contact your line manager, your occupational health department or your GP.

You may request to withdraw your data at any time by e-mailing the researcher (Sarah Buckingham) at: sab246@exeter.ac.uk.

This study has been approved by the University of Exeter Medical School Research Ethics Committee.

The Fitbit, Bupa Boost and your Physical Activity

Now that you have been wearing the Fitbit and using the 'Bupa Boost' app, we would like to know a bit more about what you thought of them.

On how many days have you worn the Fitbit over the last 7 days (not including today)? *

Please choose **only one** of the following:

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

If 0, please state why... *

Please write your answer here:

Over the last 7 days, for how many hours per day did you typically wear the Fitbit? *

Please write your answer here:

-

Enter a number between 0 and 24.

On a scale of 1-5, where 5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree, 1 = strongly disagree, how much do you agree with the following statements...? *

Please choose the appropriate response for each item:

	5	4	3	2	1	0 - I did not use the Fitbit
The Fitbit tracker was easy to use	○	○	○	○	○	○
The Fitbit tracker helped me to be more physically active	○	○	○	○	○	○

Have you used the 'Bupa Boost' app over the last 7 days (not including today)? *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If no, please state why... *

Please write your answer here:

Over the last 7 days, on how many days did you use the 'Bupa Boost' app? *

Please choose **only one** of the following:

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

Over the last 7 days, how long did you spend using the 'Bupa Boost' app on a typical day? *

Please choose **only one** of the following:

- ○ Less than one minute

- ◯ 1-5 minutes
- ◯ 5-15 minutes
- ◯ 15-30 minutes
- ◯ 30-45 minutes
- ◯ 45-60 minutes
- ◯ More than an hour

Which features of the app did you use? *

Please choose **all** that apply:

- ☐ Setting my own goals and earning wellness points and badges
- ☐ Bupa's health library / self-help information and tools

Did you set yourself any other fitness goals in addition to the recommended goal of improving your daily step count? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

If yes, please state the goal(s) you set yourself... *

Please write your answer here:

For example, cycle to work three days per week.

Which of the following areas of your health and wellbeing did you manage using the app? *

Please choose **all** that apply:

- ☐ Fitness
- ☐ Nutritional
- ☐ Mindfulness
- ☐ Relaxation

On a scale of 1-5, where 5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree, 1 = strongly disagree, how much do you agree with the following statements...? *

Please choose the appropriate response for each item:

	5	4	3	2	1	0 - I did not use the 'Bupa Boost' app
The 'Bupa Boost' app was easy to use	◯	◯	◯	◯	◯	◯

	5	4	3	2	1	0 - I did not use the 'Bupa Boost' app
The 'Bupa Boost' app helped me to be more physically active	○	○	○	○	○	○

Have there been any events that you feel have affected your physical activity levels over the last 7 days? (E.g. illness, annual leave) *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If yes, please give details... *

Please write your answer here:

International Physical Activity Questionnaire (as baseline questionnaire)

Step Count Upload (as baseline questionnaire)

SF-12 Health Survey (as baseline questionnaire)

Perceived Stress Scale (as baseline questionnaire)

Health and Work Performance Questionnaire: Absenteeism and Presenteeism (as baseline questionnaire)

Thank you very much for completing the questionnaire. Your responses have been recorded.

If you have any further questions about the study please contact Sarah Buckingham at sab246@exeter.ac.uk.

5c. Post-intervention (week 12) questionnaire

Thank you for agreeing to complete the third questionnaire for this study. Please answer all of the questions on the following pages. The questionnaire should take around **10-15 minutes** to complete.

Your individual answers will be kept secure and confidential and accessed only by the research team. Your data will not be shared in an identifiable way with the police force or any external organisations.

If during completion you have any concerns about your work, health or any related issues, you may wish to contact your line manager, your occupational health department or your GP.

You may request to withdraw your data at any time by e-mailing the researcher (Sarah Buckingham) at: sab246@exeter.ac.uk.

This study has been approved by the University of Exeter Medical School Research Ethics Committee.

The Fitbit, Bupa Boost and your Physical Activity

Now that you have been wearing the Fitbit and using the 'Bupa Boost' app, we would like to know a bit more about what you thought of them.

On how many days have you worn the Fitbit over the last 7 days (not including today)? *

Please choose **only one** of the following:

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

If 0, please state why... *

Please write your answer here:

Over the last 7 days, for how many hours per day did you typically wear the Fitbit? *

Please write your answer here:

-

Enter a number between 0 and 24.

On a scale of 1-5, where 5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree, 1 = strongly disagree, how much do you agree with the following statements...? *

Please choose the appropriate response for each item:

	5	4	3	2	1	0 - I did not use the Fitbit
The Fitbit tracker was easy to use	○	○	○	○	○	○
The Fitbit tracker helped me to be more physically active	○	○	○	○	○	○

Have you used the 'Bupa Boost' app over the last 7 days (not including today)? *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If no, please state why... *

Please write your answer here:

Over the last 7 days, on how many days did you use the 'Bupa Boost' app? *

Please choose **only one** of the following:

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

Over the last 7 days, how long did you spend using the 'Bupa Boost' app on a typical day? *

Please choose **only one** of the following:

- ○ Less than one minute
- ○ 1-5 minutes
- ○ 5-15 minutes

- ◯ 15-30 minutes
- ◯ 30-45 minutes
- ◯ 45-60 minutes
- ◯ More than an hour

Which features of the app did you use? *

Please choose **all** that apply:

- ☐ Setting my own goals and earning wellness points and badges
- ☐ Bupa's health library / self-help information and tools
- ☐ Group challenges / leader board to compete with colleagues on my own
- ☐ Company or team challenges (competing with colleagues as a group)
- ☐ Social feed (to see how colleagues were getting on)
- ☐ Messages to colleagues (e.g. motivation and encouragement)

Which of the following did you prefer? *

Please choose **only one** of the following:

- ◯ Individual goal-setting and earning wellness points and badges
- ◯ Social elements - messaging colleagues, competing with colleagues and taking part in company/team challenges
- ◯ No preference / liked both equally

Did you set yourself any other fitness goals in addition to the recommended goal of improving your daily step count? *

Please choose **only one** of the following:

- ◯ Yes
- ◯ No

If yes, please state the goal(s) you set yourself... *

Please write your answer here:

For example, cycle to work three days per week.

Which of the following areas of your health and wellbeing did you manage using the app? *

Please choose **all** that apply:

- ☐ Fitness
- ☐ Nutritional
- ☐ Mindfulness
- ☐ Relaxation

On a scale of 1-5, where 5 = strongly agree, 4 = agree, 3 = neutral, 2 = disagree, 1 = strongly disagree, how much do you agree with the following statements...? *

Please choose the appropriate response for each item:

	5	4	3	2	1	0 - I did not use the 'Bupa Boost' app
The 'Bupa Boost' app was easy to use	○	○	○	○	○	○
The 'Bupa Boost' app helped me to be more physically active	○	○	○	○	○	○

Have there been any events that you feel have affected your physical activity levels over the last 7 days? (E.g. illness, annual leave) *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If yes, please give details... *

Please write your answer here:

International Physical Activity Questionnaire (as baseline questionnaire)

Step Count Upload (as baseline questionnaire)

SF-12 Health Survey (as baseline questionnaire)

Perceived Stress Scale (as baseline questionnaire)

Health and Work Performance Questionnaire: Absenteeism and Presenteeism (as baseline questionnaire)

Thank you very much for completing the questionnaire. Your responses have been recorded. If you have any further questions about the study please contact Sarah Buckingham at sab246@exeter.ac.uk.

5d. Follow-up (month 8) questionnaire

Thank you for agreeing to complete the final questionnaire for this study. Please answer all of the questions on the following pages. The questionnaire should take around **10-15 minutes** to complete.

Your individual answers will be kept secure and confidential and accessed only by the research team. Your data will not be shared in an identifiable way with the police force or any external organisations.

If during completion you have any concerns about your work, health or any related issues, you may wish to contact your line manager, your occupational health department or your GP.

You may request to withdraw your data at any time by e-mailing the researcher (Sarah Buckingham) at: sab246@exeter.ac.uk.

This study has been approved by the University of Exeter Medical School Research Ethics Committee.

The Fitbit, Bupa Boost and your Physical Activity

Please answer the following questions about if and how you are still using the Fitbit and Bupa Boost app.

Are you still using the Fitbit? *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If no, please explain why...

Please write your answer here:

On how many days have you worn the Fitbit over the last 7 days (not including today)? *

Please choose **only one** of the following:

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

Over the last 7 days, for how many hours per day did you typically wear the Fitbit? *

Your answer must be between 0 and 24

Please write your answer here:

-

Are you still using the 'Bupa Boost' app? *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If no, please explain why...

Please write your answer here:

Over the last 7 days, on how many days did you use the 'Bupa Boost' app? *

Please choose **only one** of the following:

- ○ 0
- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ 6
- ○ 7

If you are still using the 'Bupa Boost' app, over the last 7 days, how long did you spend using the app on a typical day? *

Please choose **only one** of the following:

- ○ Less than one minute
- ○ 1-5 minutes
- ○ 5-15 minutes
- ○ 15-30 minutes
- ○ 30-45 minutes
- ○ 45-60 minutes
- ○ More than an hour
- ○ I did not use the app

Which features of the app have you used during the last five months? *

Please choose **all** that apply:

- ☐ Setting my own goals and earning wellness points and badges
- ☐ Bupa's health library / self-help information and tools
- ☐ Group challenges / leader board to compete with colleagues on my own
- ☐ Company or team challenges (competing with colleagues as a group)
- ☐ Social feed (to see how colleagues were getting on)
- ☐ Messages to colleagues (e.g. motivation and encouragement)
- ☐ None of the above

Which of the following areas of your health and wellbeing did you manage using the app? *

Please choose **all** that apply:

- ☐ Fitness
- ☐ Nutritional
- ☐ Mindfulness
- ☐ Relaxation

Have there been any events that you feel have affected your physical activity levels over the last 7 days? (E.g. illness, annual leave) *

Please choose **only one** of the following:

- ○ Yes
- ○ No

If yes, please give details... *

Please write your answer here:

Apart from participating in the PAW-Force trial, have you taken part in any new physical activity schemes or activities run by the police force since June 2017?

For example, the Love2Ride cycling scheme or an organised lunchtime running club. This may be part of ActivAte2020 or otherwise. Don't include anything you were already doing before June 2017.

Please choose **only one** of the following:

- ○ Yes
- ○ No

If yes, please give details. *

Please write your answer here:

International Physical Activity Questionnaire (as baseline questionnaire)

Step Count Upload (as baseline questionnaire)

SF-12 Health Survey (as baseline questionnaire)

Perceived Stress Scale (as baseline questionnaire)

Health and Work Performance Questionnaire: Absenteeism and Presenteeism (as baseline questionnaire)

Thank you very much for completing the questionnaire. Your responses have been recorded.

If you have any further questions about the study please contact Sarah Buckingham at sab246@exeter.ac.uk.

5e. Survey with managers, commissioners and occupational health staff

Thank you for taking an interest in this research project. Please read the information sheet and answer the questions on the following page if you agree to take part.

Q1. Based on your knowledge and experience, do you think that wearable fitness technology (i.e. Fitbit wearable activity monitor and 'Bupa Boost' app) is a useful intervention within your workplace?
Please explain your answer.

Q2. What do you perceive to be the main benefits (if any) of the use of wearable fitness technology for individual staff?
(For example, reduced stress, improved health and wellbeing)

Q3. What do you perceive to be the main benefits (if any) for the organisation?
(For example, improved staff productivity, morale, reduced sickness absence)

Q4. Would you consider commissioning wearable activity monitors for individual staff within your department or organisation at a cost of approximately £80 per device?
Yes / No

Q5. Do you have any other recommendations for policies or strategies to encourage staff to become more physically active / less sedentary?

Thank you very much for completing this survey. If you have any further questions please contact Sarah Buckingham at: sab246@exeter.ac.uk.

Appendix 6: Topic guides for interviews (pre-intervention, post-intervention and follow-up)

Participants' expectations of mobile health (mHealth) and fitness technology will be assessed at baseline (pre-intervention). At around 12 weeks (immediately post-intervention), engagement with, experience of and response to the activity monitor (Fitbit Charge 2®) and Bupa Boost app will be characterised and linked to prior expectations where possible. At eight months (five months post-intervention) overall use and benefit and maintenance of physical activity behaviours will be explored. Any differences between participants will be noted.

Researcher says:

My name is Sarah and I am a PhD student working on the PAW-Force study (Physical Activity Wearables in the Police Force).

Thank you for agreeing to take part in this interview; we really appreciate your time and help with the study. We would like to audio record the interview to help make sense of what you are saying about your experiences of taking part. Are you happy for our discussion to be recorded? The recording will be stored securely by the University of Exeter.

We want to find out how the mobile health and fitness technology works or doesn't work for you and what could be done better. There are no right or wrong answers; we're really interested to hear about your thoughts and experiences.

Researcher Notes

A selection of the following questions will be used and the order will be at the discretion of the researcher depending on what the participant says. Questions may be modified to suit the participant, and the researcher will be able to explore other topics raised by the participant or as a result of feedback from other sources (e.g. questionnaires). The interview may be stopped at the participant's request or at the discretion of the researcher.

The interviewer should use 'reflective listening' statements and brief summaries to check their understanding of what the participant has said during the interview. This will also help to build empathy and encourage further disclosure.

Leading questions should be avoided; however prompts or scenarios based on other participants' experiences may be used to encourage openness. The following general prompts may be used:

"Tell me more about…"

Reflect back… "Have I got this right?" "You said…" "It sounds like…" etc.

Give vocal encouragement and use terms such as, "Interesting thought", "Go on".

Allow silences to be used to allow the participant to think about his/her answer.

Use attentive body language such as leaning forward (if interviews are face-to-face).

<u>Pre-Intervention (Baseline) Interview</u>

Topic	Questions	Prompts	Theoretical basis
Current role	Tell me about your current role in the police force	Desk-based or active?	Context and baseline activity level.
Existing physical activity levels	Tell me a bit about how you keep physically active at the moment...	Home or work Different activities (sports, walking, cycling to work)	As above. May also help to determine existing attitudes to PA and existing intrinsic motivation.
Existing initiatives to improve physical activity in the workplace	Are there any existing physical activity/fitness programmes available in your workplace? Do you use them?	E.g. police force gyms, Love2Ride scheme (encouraging cycling for commuting), taster sessions with sports clubs, ActivAte2020 programme	Context. Opportunity component of the COM-B model. Organisational layer of the Socio-Ecological Model (SEM).
Prior experience of mHealth/ fitness technology	Have you ever used a wearable activity monitor or tracker? If yes, tell me more about it….	Which one? (e.g. Fitbit, Jawbone, Withings, Garmin) When? For how long? How did you find it? Was it helpful to you?	Individual-level factors in the SEM – knowledge, skills and attitude in relation to mHealth and fitness technology.
	Have you ever used a health or fitness app on your smartphone or tablet? If yes, tell me more about it…	Which one? (e.g. Fitbit, Apple Health, Strava, MapMyFitness, Google Fit, Jawbone) What did you use it for? When? For how long? How did you find it? Was it helpful to you?	As above.

Expectations of mHealth/fitness technology and the intervention (including perceived barriers)	Why did you choose to take part in this study? What are your expectations of taking part/using the wearable device and app?	What does the participant think the intervention sets out to do? How do they think it will work?	Process evaluation – assess reach and engagement with research. Assess expectations (and expected barriers) of the intervention.
	What do you hope to achieve? Do you have any specific fitness goals (in addition to improving your step count)?	Explore priorities and goals – general (e.g. improving health and fitness, reducing stress) and specific (e.g. walking 10,000 steps/day, 150 mins/week of moderate intensity activity such as fast walking or cycling, strength training two days/week)	Contemplation and preparation phases of transtheoretical model / stages of change. Establish existing level of (intrinsic) motivation.
	Do you expect to encounter any problems or difficulties in taking part in the study or using the technology to monitor your activity levels?	E.g. lack of technical skills, time, lack of support from managers or colleagues	Expected barriers to behaviour change (technology or study participation) at the various levels of the SEM – individual, interpersonal, organisational.
	Are there any other factors that you expect might prevent you from getting more active or achieving your goals?	E.g. time, health problems, lack of support from family or colleagues	Expected barriers to behaviour change (PA) at various levels of the SEM – individual, interpersonal, organisational.

Post-Intervention (Week 12) Interview

Topic	Questions	Prompts	Theoretical basis
Engagement with mHealth/fitness technology	How did you use the Fitbit and Bupa Boost app?	Wear time Usage of app When/for how long?	Acceptability and engagement with intervention.
Response to and experience of mHealth/fitness technology	Did the technology help you to improve your physical activity levels? If so, how? If not, why?	E.g. setting myself goals, monitoring and tracking progress, motivation from colleagues	Action phase of transtheoretical model/stages of change. Self-regulation. Social support and influence. Self-efficacy. Self-determination (intrinsic and extrinsic motivation). Capability, motivation and (social) opportunity components of the COM-B model. Technology acceptance model (perceived usefulness).
	Did it match the expectations you had at the beginning of the study? If not, why?	Refer to first interview for expectations.	Compare experiences with prior expectations.
	Did it help you to achieve your fitness goals (general and specific)?	Refer to first interview for specific goals.	Self-regulation (and as above).
	If the Fitbit and Bupa Boost app helped you to be more active, did you feel any other benefits of this?	E.g. improved general health/wellbeing, reduced stress, more productive at work	Assess wider impact on health, wellbeing, stress and productivity (gaps in current evidence).
	Did you find the Fitbit and Bupa Boost app easy to use? If not, why? Could they be improved in any way?	User friendly? Any problems? Recommendations for future users.	Technology acceptance model (usability).
	Which features of the app did you use? What did you think of them?	Fitness only or mindfulness, relaxation and nutrition?	Identify preferred (and perceived effective) behaviour change components (behaviour

		Specific features - Setting my own goals, group challenges/leader board, company/team challenges, social feed, messages to colleagues, Bupa's health library/tools	change taxonomies – e.g. CALORE). Explore intrinsic and extrinsic motivation in relation to components (rewards, competition). Social support and influence theories. Individual and interpersonal layers of the Socio-Ecological Model (SEM).
	Did you find the individual goal setting or social/competitive/ collaborative elements more useful? Why?	Competing against colleagues, social support. Individual preferences.	As above. Individual differences.
Current physical activity level and context	Could you briefly summarise what you currently do to keep physically active…	Home or work Different activities (sports, walking, cycling to work)	Improved understanding of context. Compare with baseline interview.
	Do you now use any physical activity or fitness programmes available in your workplace? If yes, is this as a result of taking part in this study?	E.g. police force gyms, Love2Ride scheme, taster sessions with sports clubs, ActivAte2020 programme	Context. Expansion in mixed methods – add to quantitative findings. Opportunity component of the COM-B model. Organisational layer of the SEM.
Barriers and facilitators to use of technology and physical activity	Did you have any problems or difficulties in taking part in the study or using the technology to monitor your activity levels?	E.g. lack of technical skills, lack of time, lack of support from managers or colleagues	Barriers to engagement with the intervention.
	Were there any other factors that prevented you from getting more active or achieving your goals? (i.e. what gets in the way?)	E.g. time, health problems, lack of support from family or colleagues	Barriers to PA at all levels of the SEM – focus on individual, interpersonal and organisational.
	What might help you to keep up your	E.g. continued use of	Facilitators to PA at all levels of the SEM – focus

	physical activity levels in the future? Is there anything else that you feel your employer could do to help?	mHealth/fitness technology, other initiatives/ incentives in the workplace	on individual, interpersonal and organisational.

<u>Follow-Up (Month 8) Interview</u>

Topic	Questions	Prompts	Theoretical basis
Continued engagement with mHealth/fitness technology (including barriers and facilitators)	Are you still using the Fitbit and Bupa Boost app? If yes, how? If no, when did you stop using it and why? What would have encouraged you to use it for longer?	Device wear time and usage of app – is this lower than five months earlier?	Assess longer-term acceptability and engagement with intervention. Barriers and facilitators to technology use.
	If no, are you still tracking your activity levels?	Other methods of tracking activity, e.g. use of different device or app, paper or mental record?	Elucidate whether continued technology use is necessary to sustain behaviour change. Habit formation (activity tracking).
	Do you plan to use the Fitbit and Bupa Boost app after the study ends?		Longer-term acceptability. Perceived necessity of use to sustain behaviour change.
Longer-term experience of mHealth/fitness technology and maintenance of physical activity levels	Do you feel that the technology helped you to keep up your physical activity levels? If so, how? If not, why?	E.g. setting myself goals, monitoring and tracking progress, motivation from colleagues	Maintenance phase of transtheoretical model/stages of change (after 6 months). Self-efficacy. Self-determination (intrinsic motivation). Habit formation (PA behaviours). Motivation and (social) opportunity components of the COM-B model.

	Are you still achieving the goals you set yourself at the beginning of the study? Did using the technology change the way you feel about being physically active? If the Fitbit and Bupa Boost app have helped you to maintain or improve your physical activity/fitness level, have you felt any other benefits of this? Which features of the app did you carry on using? Why did you find these most useful?	Refer to earlier interviews for specific goals. Confidence about being physically active. Ability, skills, attitudes, knowledge. E.g. improved general health/wellbeing, reduced stress, more productive at work Fitness only or mindfulness, relaxation and nutrition? Specific features - Setting my own goals, group challenges/leader board, company/team challenges, social feed, messages to colleagues, Bupa's health library/tools	Habit formation (goal-setting and PA). Self-efficacy. Self-determination. Capability component of the COM-B model. Assess longer-term impact on health, wellbeing, stress and productivity (gaps in current evidence). Potential pathways of change, e.g. reduced stress due to setting of relaxation goals or increased PA? Identify preferred (and perceived effective) behaviour change components in the longer term (behaviour change taxonomies – e.g. CALORE). Explore intrinsic and extrinsic motivation – e.g. whether rewards and competitions were still effective. Social support and influence theories. Individual differences.
Current physical activity level and context	Could you briefly summarise what you currently do to keep physically active? Do you now take part in any physical activity or fitness programmes in	Home or work Different activities (sports, walking, cycling to work) E.g. police force gyms, Love2Ride scheme, taster sessions with	Context. Compare with earlier interviews. Context. Expansion in mixed methods – add to quantitative findings. Opportunity component

	your workplace? If yes, is this as a result of taking part in this study?	sports clubs, ActivAte2020.	of the COM-B model. Organisational layer of the Socio-Ecological Model.
Barriers and facilitators to maintaining physical activity	Over the last few months, were there any other factors that prevented you from getting more active or achieving your goals? (i.e. what got in the way?) What would have encouraged you to stay more active?	E.g. lack of time, health problems, lack of support from family or colleagues E.g. reminders from technology/ employer/ colleagues, allocated time in work for exercise/ lunchtime walks	Long-term barriers to PA at all levels of the SEM – particularly individual, interpersonal and organisational. Long-term facilitators to PA at all levels of the SEM – focus on individual, interpersonal and organisational.
Experience of study participation	Overall, how did you find taking part in the study? Is there any way that your experience could have been improved?	Experience of taking part in the study Thoughts about the research process – e.g. data collection, frequency of contacts, workplace support.	Process evaluation. Assess acceptability of the research process for participants.

At the end of each interview:

Summarise the discussion and ask if there is anything else they would like to add.

Thank the participant for taking part in the discussion and for their help with the research.

Appendix 7: Summary of findings report for the police force

The PAW-Force Study: Summary Findings November 2018

Sarah Buckingham, PhD student, University of Exeter

This report presents a summary of findings of the Physical Activity Wearables in the police force (PAW-Force) study, from June 2017 to February/March 2018. The report includes and extends the findings of the interim summary report that was circulated in June 2018.

Why was the study needed?

This study was part of the three year ActivAte2020 programme which aims to improve the health and wellbeing of police officers and staff, including physical activity, diet and nutrition and sleep quality. The study mainly focused on physical activity.

Wearable or mobile fitness technology (such as Fitbit® activity monitors and smartphone apps) is an accessible, personalised, cost-effective and fun way to help people to be more active. As policing is associated with a high risk of lifestyle-related health conditions (including cardiovascular disease and obesity[1]), is becoming a more sedentary occupation, and no previous studies had looked at the use of wearable fitness technology in the police force, the study was much needed.

What was the aim of the study?

We wanted to know whether wearable fitness technology (a Fitbit activity monitor used with the Bupa Boost app) would be a feasible, acceptable and potentially effective way to help police officers and staff to be active, in both the short and longer term.

What did taking part involve?

All participants were asked to wear the Fitbit and use the Bupa Boost app for an initial 12 weeks (using different features of Bupa Boost), followed by a 5 month ‘maintenance phase’:

Week 0	**Weeks 1-6**	**Weeks 7-12**	**Week 13 to end of study (8 months from beginning)**
			
Beginning of study: Wear Fitbit on its own for a week	Wear Fitbit and use ‘Bupa Boost’ app: individual goal-setting, wellness points and badges	Wear Fitbit and use ‘Bupa Boost’ app with colleagues: social support, competitions and challenges	Continue to wear Fitbit and use ‘Bupa Boost’ app as desired

How was the data collected?

Participants sent their step data and completed surveys at four different time points – week 0, week 6, week 12 and month 8. Some participants were selected to take part in more in-depth interviews about their expectations and experiences (32 interviews in total).

Who was recruited?

- 180 participants were recruited: 128 from Devon & Cornwall Police (Plymouth Basic Command Unit) and 52 from Dorset Police (North Dorset site).
- 63% were police officers (ranging from constable to superintendent), 17% were PCSOs or special constables, and 20% were police staff.
- 59% male, 41% female
- Average age was 39 years, and the average length of service was 12 years.
- 80% were shift workers
- 39% of participants had never used a wearable activity monitor or health or fitness app before taking part.

How active were participants at the start of the study?

- Average daily step count at the beginning of the study ranged from around 3,800 in some participants to 20,800 in others. The average across all participants was 10,555 steps per day.

- Participants had different levels of motivation and views on being physically active:

> *"I can't get my head round people enjoying sports and gyms and things like that. I've tried it but didn't enjoy it."*
> **Police staff, female, age 40+**

> *"I've always been interested in sport. I love it and I have always loved it."*
> **PCSO, male, age 40+**

- 58% reported their role was mainly sedentary (e.g. spent sitting at a desk). Participants reported spending around six and a half hours sitting on a typical weekday.

> *"It is completely office-based. You sit at a desk with your phone. You haven't got much movement at all."*
> **Police staff, male, age 18-39**

> *"In my day-to-day job I am sat behind a desk all day."*
> **Police sergeant, male, age 18-39**

- Male officers and staff were more active than females, with higher daily steps, total physical activity and moderate and vigorous activity.
- Police officers were more active than PCSOs and special constables overall, but less active during the working day.

<u>What were participants' expectations at the beginning of the study?</u>

Participants wanted to take part for the following reasons:

Improving physical activity level or trying a new type of exercise
Comparing their activity level with colleagues
Reducing sedentary time at work
Improving fitness
Maintaining or improving their health and wellbeing
Wanting to lose weight
Reducing stress
Improving sleep
Improving confidence and enjoyment of life
Desire to learn about fitness technology

> *"I wanted to do this trial because somebody came in [to work] and did a fitness check. After this health check I thought, no, I need to do something."*
> **Police staff, female, age 40+**

Many thought the technology would help them to be more aware of their activity level, set small, achievable goals, and motivate them to be more active.

What did participants think of the Fitbit and Bupa Boost app?

On a scale of 1 to 5 (1 = lowest rating, 5 = highest rating):

- The Fitbit was given an average rating of 4.7 for being easy to use and 3.9 for its usefulness in helping people to be more active
- The Bupa Boost app was rated 3.6 for being easy to use and 3.2 for its usefulness

The main positives and negatives were:

Fitbit	Bupa Boost
+ Good user interface + Range of features + Small and lightweight + Durable + Easy to charge - Not waterproof - Not always accurate for monitoring heart rate - Not suitable for some types of activity (e.g. horse riding)	+ Flexible and personalised goal-setting + Good for monitoring diet/nutrition, relaxation and mindfulness as well as physical activity - Poor user interface and difficult to find colleagues - Not enough automated tracking of activity - Did not link well with the Fitbit (or other devices) - Too many (and meaningless) notifications - Points and competitions were sometimes unfair

At the end of the study, several participants talked about wanting to buy a Fitbit for their partner or family members.

How much did participants actually use the technology?

After **12 weeks…**

97% of participants were wearing the Fitbit, for an average of **6.6 days a week** and for **22 hours a day**	**60% of participants** were using Bupa Boost, for an average of **5.1 days a week**. Most of these logged in to the app for **1 to 5 minutes each day**.

After **8 months…**

83% of participants were still wearing the Fitbit, for an average of **6.5 days a week** and for **21 hours a day**	Only **27% of participants** were still using Bupa Boost, for an average of **5.1 days a week**, and for **1 to 5 minutes each day**

Why did so many participants stop using Bupa Boost?

The most common reasons for not using the Bupa Boost app after 8 months were:

- Preferring to use the Fitbit app or another app / finding no need for two apps
- Not finding it helpful or motivating

- It was seen as too complicated / not user friendly

<u>Was there any impact on physical activity levels (up to 12 weeks)?</u>

- Overall, there was no increase in average daily step count between the beginning of the study and 12 weeks…
- However, there was an increase in:
    - Total physical activity (by an average of **23 minutes/week** or **465 MET-minutes/week***)
    - Moderate and vigorous physical activity (by an average of **403 MET-minutes/week***)
- The least active participants (those with an average of less than 10,000 steps per day at the beginning of the study) seemed to benefit most:
    - Increase in steps by **1,028 steps/day** on average
    - Increase in total physical activity by an average of **50 minutes/week** or **1,094 MET-minutes/week***
    - Increase in moderate and vigorous physical activity by an average of **796 MET-minutes/week***

"The Fitbit itself did actually make me more conscious of what I was and wasn't doing. If I didn't have [the Fitbit], I wouldn't have any knowledge of how little I was doing."
Police constable, female, age 40+

"Generally, I'm trying to get 10,000 [steps] a day. I'm out doing a bit more walking as well, which I wouldn't have thought of previously. I've actually just gone out a couple of nights and just gone for a walk around the block, which I wouldn't have done before."
Police sergeant, male, age 18-39

"I liked the fact that it recorded it weekly as well, so I could look at last week or the week before and say, 'I didn't do quite as many steps, I need to do that this week' or, 'I beat the steps last week.' That was good to compare it to something rather than just looking at it daily."
Police staff, female, age 18-39

Was there any longer term impact on physical activity level?

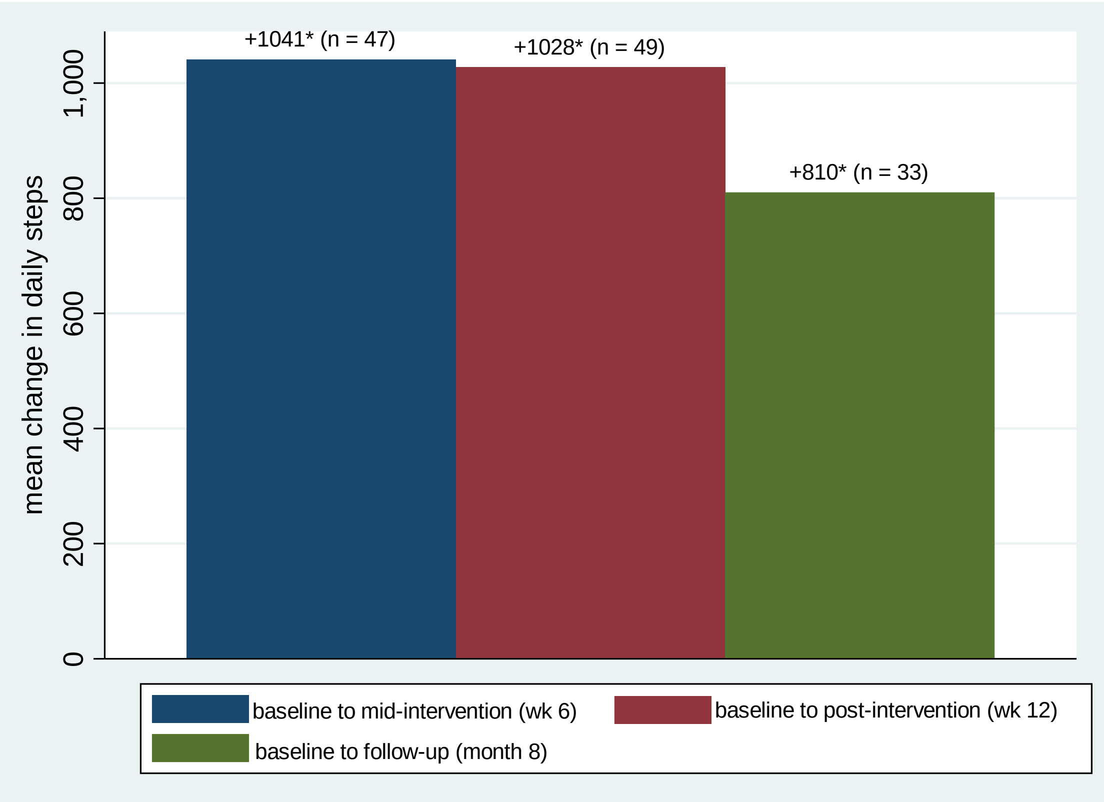
- Overall, there was no increase in average daily step count between the beginning of the study and 8 months…
- However, there was an increase in:
  - Total physical activity (by an average of **19 minutes/week** or **318 MET-minutes/week***)
  - Moderate and vigorous physical activity (by an average of **421 MET-minutes/week***)
- Again, the least active participants (those with an average of less than 10,000 steps per day at the beginning of the study) seemed to benefit most:
  - Increase in steps by **810 steps/day** on average
  - Increase in total physical activity by an average of **42 minutes/week** or **1,068 MET-minutes/week***
  - Increase in moderate and vigorous physical activity by an average of **840 MET-minutes/week***

"I wouldn't tend to go out walking but now I find myself after an hour just getting out and going for a walk around and checking my steps quite a lot, so I still do it."
Police sergeant, male, age 18-39

"I would say it's had a really positive impact on me… I had a look back right to the very start, last year… I think my average step count was around 40,000 a week, something like that, and it's up at around 55-60,000 consistently now, which is a big improvement. It's just got me focused on every step counts."
Police inspector, male, age 18-39

"It has made me think. It really has made me think about my lifestyle, my activity or lack of, being conscious of, if I've had a really lazy day, I need to do something. I'm not going to get active sat in an armchair or at my desk. You see it in black and white that you've been lazy today and you need to do more. "
Police staff, female, age 40+

Changes in average daily step count from baseline (i.e. beginning of study) to week 6, week 12 and month 8: Participants with step count less than 10,000/day at baseline



n = number of participants

Was there an impact on any other outcomes?

There was no overall significant change in:

- Sedentary time
- Perceived stress
- Physical health-related quality of life
- Perceived productivity at work

There was a small improvement in mental health-related quality of life from the beginning of the study to 8 months.

Some participants reported feeling fitter and healthier, losing weight, having more energy and feeling less stressed. Some noticed improved morale amongst their colleagues.

We are still waiting for some data on sickness absences, but early indications are that there was no impact. It may be difficult to draw any definite conclusions though as sickness absence is affected by many different factors.

Did participants prefer the individual or social phase of the study?

- **56% preferred the individual phase** (self-monitoring, goal-setting and earning wellness points and badges, weeks 1-6)
- **Only 7% preferred the social phase** (social support, competing and comparing with colleagues, weeks 7-12)
- **37% liked both the individual and social phases equally**

> *"I'm probably not one for the social side of it, really. I just prefer to monitor it myself rather than having to challenge other people."* **Police sergeant, male, age 18-39**

> *"It's fine to have aspirations to complete the goals, but if you're doing it as a group or you're competing against other people… with me it motivates me more anyway."* **Police constable, male, age 18-39**

Interestingly, there was no difference in physical activity level between the two phases.

What were the views of managers, commissioners and occupational health staff?

A separate survey with 10 managers, commissioners and occupational health staff found…

- Wearable fitness technology was seen to be useful for the police force, but those surveyed believed it should be optional.
- There were perceived benefits for officers and staff – improved fitness, health and wellbeing, reduced stress and fatigue, promoting "healthy competition" and supporting a "positive team culture".

- Benefits were also expected for the organisation - reduced sickness absence and presenteeism, a more resilient workforce, and improved morale and productivity.
- All those surveyed would consider commissioning wearable activity monitors at a cost of around £80 per device.

Were there any additional findings?

- Sleep and stress were key themes; participants recognised that being physically active helped to improve these. For many, the Fitbit helped with this awareness by monitoring activity, sleep and heart rate.
- The main barriers to physical activity were shift work, time pressures, home and family commitments, health problems and the weather.
- Suggestions for improving physical activity (and reducing sedentary time) in the police force included:
    - Designated time for exercise/improving wellness for all work streams
    - More health and wellbeing champions to organise group activities and workplace competitions
    - Greater individual support and feedback for annual fitness tests
    - More exercise facilities for small and rural police stations
    - Encouragement from managers and supervisors to take breaks away from the desk

Is there a need for any further research?

Yes. The results of this study show that wearable fitness technology is a feasible and acceptable way to increase physical activity in the police force. The findings suggest that it is likely to have a positive impact on activity levels in the short and longer term; however, we cannot draw any definitive conclusions on this as there was no control or comparison group and it is possible that changes might have occurred anyway. A larger, controlled trial would improve our confidence in the results. Further studies with other police forces in the UK will also tell us whether these findings apply more generally.

[1] Zimmerman, F.H. Cardiovascular disease and risk factors in law enforcement personnel: a comprehensive review. *Cardiology in review* 2012; 20(4):159-66.

*****A MET (or Metabolic Equivalent of Task) is a measure of intensity of physical activity based on oxygen consumption, compared to when at rest. A MET-minute refers to the intensity of activity in one minute.

Bibliography

ABDIN, S., WELCH, R. K., BYRON-DANIEL, J. & MEYRICK, J. 2018. The effectiveness of physical activity interventions in improving well-being across office-based workplace settings: a systematic review. *Public Health,* 160**,** 70-76.

ABRAHAM, C. & MICHIE, S. 2008. A taxonomy of behavior change techniques used in interventions. *Health Psychol,* 27**,** 379-387.

AFSHIN, A., BABALOLA, D., MCLEAN, M., YU, Z., MA, W., CHEN, C. Y., ARABI, M. & MOZAFFARIAN, D. 2016. Information Technology and Lifestyle: A Systematic Evaluation of Internet and Mobile Interventions for Improving Diet, Physical Activity, Obesity, Tobacco, and Alcohol Use. *J Am Heart Assoc,* 5**,** e003058.

AINSWORTH, B. E., HASKELL, W. L., HERRMANN, S. D., MECKES, N., BASSETT, D. R., JR., TUDOR-LOCKE, C., GREER, J. L., VEZINA, J., WHITT-GLOVER, M. C. & LEON, A. S. 2011. 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc,* 43**,** 1575-1581.

AIRAKSINEN, J., JOKELA, M., VIRTANEN, M., OKSANEN, T., KOSKENVUO, M., PENTTI, J., VAHTERA, J. & KIVIMAKI, M. 2018. Prediction of long-term absence due to sickness in employees: development and validation of a multifactorial risk score in two cohort studies. *Scand J Work Environ Health,* 44**,** 274-282.

ALTHOFF, T., SOSIČ, R., HICKS, J. L., KING, A. C., DELP, S. L. & LESKOVEC, J. 2017. Large-scale physical activity data reveal worldwide activity inequality. *Nature*, 547, 336-339.

ANDERSON, A. A., YOO, H. & FRANKE, W. D. 2016. Associations of Physical Activity and Obesity With the Risk of Developing the Metabolic Syndrome in Law Enforcement Officers. *J Occup Environ Med,* 58**,** 946-951.

APPFIGURES. 2019. *Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 1st quarter 2019 [Graph]* [Online]. Statista. Available: https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/ [Accessed 02.08.19].

AREM, H., MOORE, S. C., PATEL, A., HARTGE, P., BERRINGTON DE GONZALEZ, A., VISVANATHAN, K., CAMPBELL, P. T., FREEDMAN, M., WEIDERPASS, E., ADAMI, H. O., LINET, M. S., LEE, I. M. & MATTHEWS, C. E. 2015. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. *JAMA Intern Med,* 175**,** 959-967.

ARIGO, D., JAKE-SCHOFFMAN, D. E., WOLIN, K., BECKJORD, E., HEKLER, E. B. & PAGOTO, S. L. 2019. The history and future of digital health in the field of behavioral medicine. *J Behav Med,* 42**,** 67-83.

ARMIJO-OLIVO, S., STILES, C. R., HAGEN, N. A., BIONDO, P. D. & CUMMINGS, G. G. 2012. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract,* 18**,** 12-18.

ASTHMA UK. 2019. *Asthma facts and statistics* [Online]. Available: https://www.asthma.org.uk/about/media/facts-and-statistics/ [Accessed 29.01.19].

AZJEN, I. & FISHBEIN, M. 1980. *Understanding attitudes and predicting social behaviour,* London, Pearson.
BAICKER, K., CUTLER, D. & SONG, Z. 2010. Workplace wellness programs can generate savings. *Health Aff (Millwood),* 29**,** 304-311.
BANDURA, A. 1997. *Self-Efficacy: The exercise of control,* New York, Freeman.
BARBOUR, R. S. 1999. The case for combining qualitative and quantitative approaches in health services research. *J Health Serv Res Policy,* 4**,** 39-43.
BARDUS, M., SMITH, J. R., SAMAHA, L. & ABRAHAM, C. 2016. Mobile and Web 2.0 interventions for weight management: an overview of review evidence and its methodological quality. *Eur J Public Health,* 26**,** 602-610.
BATORSKY, B., VAN STOLK, C. & LIU, H. 2016. Is More Always Better in Designing Workplace Wellness Programs?: A Comparison of Wellness Program Components Versus Outcomes. *J Occup Environ Med,* 58**,** 987-993.
BILLINGHAM, S. A., WHITEHEAD, A. L. & JULIOUS, S. A. 2013. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. *BMC Med Res Methodol,* 13**,** 104.
BLACKMAN, K. C., ZOELLNER, J., BERREY, L. M., ALEXANDER, R., FANNING, J., HILL, J. L. & ESTABROOKS, P. A. 2013. Assessing the internal and external validity of mobile health physical activity promotion interventions: a systematic literature review using the RE-AIM framework. *J Med Internet Res,* 15**,** e224.
BLAIR, S. N. 2009. Physical inactivity: the biggest public health problem of the 21st century. *Br J Sports Med,* 43**,** 1-2.
BLAIR, S. N. & CONNELLY, J. C. 1996. How much physical activity should we do? The case for moderate amounts and intensities of physical activity. *Res Q Exercise Sport,* 67**,** 193-205.
BLAND, J. M. & ALTMAN, D. G. 1995. Multiple significance tests: the Bonferroni method. *BMJ,* 310**,** 170.
BOND, D. S., THOMAS, J. G., RAYNOR, H. A., MOON, J., SIELING, J., TRAUTVETTER, J., LEBLOND, T. & WING, R. R. 2014. B-MOBILE--a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: a within-subjects experimental trial. *PLoS One,* 9**,** e100821.
BORT-ROIG, J., GILSON, N. D., PUIG-RIBERA, A., CONTRERAS, R. S. & TROST, S. G. 2014. Measuring and influencing physical activity with smartphone technology: a systematic review. *Sports Med,* 44**,** 671-686.
BOWEN, D. J., KREUTER, M., SPRING, B., COFTA-WOERPEL, L., LINNAN, L., WEINER, D., BAKKEN, S., KAPLAN, C. P., SQUIERS, L., FABRIZIO, C. & FERNANDEZ, M. 2009. How we design feasibility studies. *Am J Prev Med,* 36**,** 452-457.
BRAKENRIDGE, C. L., FJELDSOE, B. S., YOUNG, D. C., WINKLER, E. A., DUNSTAN, D. W., STRAKER, L. M., BRAKENRIDGE, C. J. & HEALY, G. N. 2016a. Organizational-Level Strategies With or Without an Activity Tracker to Reduce Office Workers' Sitting Time: Rationale and Study Design of a Pilot Cluster-Randomized Trial. *JMIR Res Protoc,* 5**,** e73.
BRAKENRIDGE, C. L., FJELDSOE, B. S., YOUNG, D. C., WINKLER, E. A., DUNSTAN, D. W., STRAKER, L. M. & HEALY, G. N. 2016b. Evaluating the effectiveness of organisational-level strategies with or without an activity tracker to reduce office workers' sitting time: a cluster-randomised trial. *Int J Behav Nutr Phys Act,* 13**,** 115.

BRAND, S. L., THOMPSON COON, J., FLEMING, L. E., CARROLL, L., BETHEL, A. & WYATT, K. 2017. Whole-system approaches to improving the health and wellbeing of healthcare workers: A systematic review. *PLoS One,* 12**,** e0188418.

BRAUN, V. & CLARKE, V. 2006. Using thematic analysis in psychology. *Qual Res Psychol,* 3**,** 77-101.

BRAVATA, D. M., SMITH-SPANGLER, C., SUNDARAM, V., GIENGER, A. L., LIN, N., LEWIS, R., STAVE, C. D., OLKIN, I. & SIRARD, J. R. 2007. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA,* 298**,** 2296-2304.

BRONFENBRENNER, U. 1989. Ecological systems theory. *In:* VASTA, R. (ed.) *Annals of Child Development: Vol. 6.* London, UK: Jessica Kingsley Publishers.

BROWN, H. E., GILSON, N. D., BURTON, N. W. & BROWN, W. J. 2011. Does physical activity impact on presenteeism and other indicators of workplace well-being? *Sports Med,* 41**,** 249-62.

BRYMAN, A. 2006. Integrating quantitative and qualitative research: how is it done? *Qual Res,* 6**,** 97-113.

BUCKINGHAM, S. A., WILLIAMS, A. J., MORRISSEY, K., PRICE, L. & HARRISON, J. 2019. Mobile health interventions to promote physical activity and reduce sedentary behaviour in the workplace: A systematic review. *Digit Health,* 5**,** 2055207619839883.

BUCKLEY, J. P., HEDGE, A., YATES, T., COPELAND, R. J., LOOSEMORE, M., HAMER, M., BRADLEY, G. & DUNSTAN, D. W. 2015. The sedentary office: an expert statement on the growing case for change towards better health and productivity. *Br J Sports Med,* 49**,** 1357-1362.

BUIJINK, A. W., VISSER, B. J. & MARSHALL, L. 2013. Medical apps for smartphones: lack of evidence undermines quality and safety. *Evid Based Med,* 18**,** 90-92.

CAN, S. H. & HENDY, H. M. 2014. Behavioral Variables Associated with Obesity in Police Officers. *Ind Health,* 52**,** 240-247.

CANCELLIERE, C., CASSIDY, J. D., AMMENDOLIA, C. & CÔTÉ, P. 2011. Are workplace health promotion programs effective at improving presenteeism in workers? A systematic review and best evidence synthesis of the literature. *BMC Public Health,* 11**,** 395.

CANEPPELE, S. & AEBI, M. F. 2017. Crime Drop or Police Recording Flop? On the Relationship between the Decrease of Offline Crime and the Increase of Online and Hybrid Crimes. *Policing,* 13**,** 66-79.

CARTER, D. D., ROBINSON, K., FORBES, J. & HAYES, S. 2018. Experiences of mobile health in promoting physical activity: A qualitative systematic review and meta-ethnography. *PLoS One,* 13**,** e0208759.

CHENG, H. L. 2016. *A simple, easy-to-use spreadsheet for automatic scoring of the International Physical Activity Questionnaire (IPAQ) Short Form* [Online]. ResearchGate. Available: https://www.researchgate.net/profile/Hoi_Lun_helen_Cheng/publication/310953872_A_simple_easy-to-use_spreadsheet_for_automatic_scoring_of_the_International_Physical_Activity_Questionnaire_IPAQ_Short_Form/data/583bbee208ae3a74b4a06f27/IPAQ-SF-Scoring.xlsx [Accessed 16.07.19].

CHINAPAW, M. J., SLOOTMAKER, S. M., SCHUIT, A. J., VAN ZUIDAM, M. & VAN MECHELEN, W. 2009. Reliability and validity of the Activity Questionnaire for Adults and Adolescents (AQuAA). *BMC Med Res Methodol,* 9**,** 58.
CIALDINI, R. 2001. *Influence, science and practice,* Boston, Allen & Bacon.
CLELAND, V., HUGHES, C., THORNTON, L., VENN, A., SQUIBB, K. & BALL, K. 2015. A Qualitative Study of Environmental Factors Important for Physical Activity in Rural Adults. *PLoS One,* 10**,** e0140659.
CLEMENT, N. D., MACDONALD, D. & SIMPSON, A. H. 2014. The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc,* 22**,** 1933-1939.
CLEMES, S. A., O'CONNELL, S. E. & EDWARDSON, C. L. 2014. Office workers' objectively measured sedentary behavior and physical activity during and outside working hours. *J Occup Environ Med,* 56**,** 298-303.
COHEN, J. 1988. *Statistical power analysis for the behavioral sciences (2nd ed.),* Hillsdale, New Jersey, Erlbaum.
COHEN, S., KAMARCK, T. & MERMELSTEIN, R. 1983. A global measure of perceived stress. *J Health Soc Behav,* 24**,** 385-396.
COHEN, S. & WILLIAMSON, G. 1988. Perceived stress in a probability sample of the United States. *In:* SPACAPAN, S. & OSKAMP, S. (eds.) *The Claremont Symposium on Applied Social Psychology. The social psychology of health.* Thousand Oaks, CA, US: Sage Publications, Inc.
COLE, J. A., TULLY, M. A. & CUPPLES, M. E. 2015. "They should stay at their desk until the work's done": a qualitative study examining perceptions of sedentary behaviour in a desk-based occupational setting. *BMC Res Notes,* 8**,** 683.
CONROY, D., YANG, C. & MAHER, J. 2014. Behavior Change Techniques in Top-Ranked Mobile Apps for Physical Activity. *Am J Prev Med,* 46**,** 649-652.
CRAIG, C. L., MARSHALL, A. L., SJOSTROM, M., BAUMAN, A. E., BOOTH, M. L., AINSWORTH, B. E., PRATT, M., EKELUND, U., YNGVE, A., SALLIS, J. F. & OJA, P. 2003. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc,* 35**,** 1381-1395.
CRAIG, P., COOPER, C., GUNNELL, D., HAW, S., LAWSON, K., MACINTYRE, S., OGILVIE, D., PETTICREW, M., REEVES, B., SUTTON, M. & THOMPSON, S. 2012. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health,* 66**,** 1182-1186.
CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ,* 337.
CRUTZEN, R. & GÖRITZ, A. S. 2011. Does social desirability compromise self-reports of physical activity in web-based research? *Int J Behav Nutr Phys Act,* 8**,** 31.
DALL, P. M., ELLIS, S. L. H., ELLIS, B. M., GRANT, P. M., COLYER, A., GEE, N. R., GRANAT, M. H. & MILLS, D. S. 2017. The influence of dog ownership on objective measures of free-living physical activity and sedentary behaviour in community-dwelling older adults: a longitudinal case-controlled study. *BMC Public Health,* 17**,** 496.
DALLERY, J., CASSIDY, R. N. & RAIFF, B. R. 2013. Single-case experimental designs to evaluate novel technology-based health interventions. *J Med Internet Res,* 15**,** e22.

DANIELS, P., REASEN, R., WIENEKE, G. & BITTLES, K. 2016. *Wearable technology: Unlocking ROI of workplace wellness. An employer case study in health care cost management.* [Online]. Springbuk. Available: https://www.springbuk.com/our-insights/wearable-technology-unlocking-roi-of-workplace-wellness/ [Accessed 14.03.19].
DAVIS, F. D. 1989. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *Mis Quarterly,* 13**,** 319-340.
DE LEON, E., FUENTES, L. W. & COHEN, J. E. 2014. Characterizing periodic messaging interventions across health behaviors and media: systematic review. *J Med Internet Res,* 16**,** e93.
DECI, E. & RYAN, R. 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry,* 11**,** 227-268.
DEVON AND CORNWALL POLICE. 2016. *Health and wellbeing: Business as usual* [Online]. Available: http://www.insidegovernment.co.uk/uploads/2016/03/johnharrisson.pdf [Accessed 28.10.19].
DIAZ-ARRIBAS, M. J., FERNANDEZ-SERRANO, M., ROYUELA, A., KOVACS, F. M., GALLEGO-IZQUIERDO, T., RAMOS-SANCHEZ, M., LLORCA-PALOMERA, R., PARDO-HERVAS, P. & MARTIN-PARIENTE, O. S. 2017. Minimal Clinically Important Difference in Quality of Life for Patients With Low Back Pain. *Spine (Phila Pa 1976),* 42**,** 1908-1916.
DICICCO-BLOOM, B. & CRABTREE, B. F. 2006. The qualitative research interview. *Med Educ,* 40**,** 314-321.
DIREITO, A., CARRAÇA, E., RAWSTORN, J., WHITTAKER, R. & MADDISON, R. 2017. mHealth Technologies to Influence Physical Activity and Sedentary Behaviors: Behavior Change Techniques, Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Ann Behav Med,* 51**,** 226-239.
DOMINICK, G. M., WINFREE, K. N., POHLIG, R. T. & PAPAS, M. A. 2016. Physical Activity Assessment Between Consumer- and Research-Grade Accelerometers: A Comparative Study in Free-Living Conditions. *JMIR Mhealth Uhealth,* 4**,** e110.
DONTJE, M. L., DE GROOT, M., LENGTON, R. R., VAN DER SCHANS, C. P. & KRIJNEN, W. P. 2015. Measuring steps with the Fitbit activity tracker: an inter-device reliability study. *J Med Eng Technol,* 39**,** 286-290.
DU, H., VENKATAKRISHNAN, A., YOUNGBLOOD, G. M., RAM, A. & PIROLLI, P. 2016. A Group-Based Mobile Application to Increase Adherence in Exercise and Nutrition Programs: A Factorial Design Feasibility Study. *JMIR Mhealth Uhealth,* 4**,** e4.
DUNSTAN, D. W., THORP, A. A. & HEALY, G. N. 2011. Prolonged sitting: is it a distinct coronary heart disease risk factor? *Curr Opin Cardiol,* 26**,** 412-419.
DWYER, T., PEZIC, A., SUN, C., COCHRANE, J., VENN, A., SRIKANTH, V., JONES, G., SHOOK, R. P., SHOOK, R., SUI, X., ORTAGLIA, A., BLAIR, S. & PONSONBY, A. L. 2015. Objectively Measured Daily Steps and Subsequent Long Term All-Cause Mortality: The Tasped Prospective Cohort Study. *PLoS One,* 10**,** e0141274.
DWYER, T., PONSONBY, A. L., UKOUMUNNE, O. C., PEZIC, A., VENN, A., DUNSTAN, D., BARR, E., BLAIR, S., COCHRANE, J., ZIMMET, P. & SHAW, J. 2011. Association of change in daily step count over five years with insulin sensitivity and adiposity: population based cohort study. *BMJ,* 342**,** c7249.

EDWARDS, E. A., LUMSDEN, J., RIVAS, C., STEED, L., EDWARDS, L. A., THIYAGARAJAN, A., SOHANPAL, R., CATON, H., GRIFFITHS, C. J., MUNAFÒ, M. R., TAYLOR, S. & WALTON, R. T. 2016. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. *BMJ Open,* 6**,** e012447.

ELDRIDGE, S. M., LANCASTER, G. A., CAMPBELL, M. J., THABANE, L., HOPEWELL, S., COLEMAN, C. L. & BOND, C. M. 2016. Defining Feasibility and Pilot Studies in Preparation for Randomised Controlled Trials: Development of a Conceptual Framework. *PLoS One,* 11**,** e0150205.

EMERSON, N. D., MERRILL, D. A., SHEDD, K., BILDER, R. M. & SIDDARTH, P. 2017. Effects of an employee exercise programme on mental health. *Occup Med (Lond),* 67**,** 128-134.

ESLIGER, D. W., ROWLANDS, A. V., HURST, T. L., CATT, M., MURRAY, P. & ESTON, R. G. 2011. Validation of the GENEA Accelerometer. *Med Sci Sports Exerc,* 43**,** 1085-1093.

ESSIET, I. A., BAHAROM, A., SHAHAR, H. K. & UZOCHUKWU, B. 2017. Application of the Socio-Ecological Model to predict physical activity behaviour among Nigerian University students. *Pan Afr Med J,* 26**,** 110.

EVENSON, K. R., GOTO, M. M. & FURBERG, R. D. 2015. Systematic review of the validity and reliability of consumer-wearable activity trackers. *Int J Behav Nutr Phys Act,* 12**,** 159.

EYSENBACH, G. 2005. The law of attrition. *J Med Internet Res,* 7**,** e11.

FANNING, J., MULLEN, S. P. & MCAULEY, E. 2012. Increasing physical activity with mobile devices: a meta-analysis. *J Med Internet Res,* 14**,** e161.

FAST TRACK IMPACT. 2019. *Fast Track Impact blog* [Online]. Available: https://www.fasttrackimpact.com/blog [Accessed 28.05.19].

FEEHAN, L. M., GELDMAN, J., SAYRE, E. C., PARK, C., EZZAT, A. M., YOO, J. Y., HAMILTON, C. B. & LI, L. C. 2018. Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. *JMIR Mhealth Uhealth,* 6**,** e10527.

FINKELSTEIN, E. A., HAALAND, B. A., BILGER, M., SAHASRANAMAN, A., SLOAN, R. A., NANG, E. E. K. & EVENSON, K. R. 2016. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. *Lancet Diabetes Endocrinol,* 4**,** 983-995.

FINKELSTEIN, E. A., SAHASRANAMAN, A., JOHN, G., HAALAND, B. A., BILGER, M., SLOAN, R. A., NANG, E. E. & EVENSON, K. R. 2015. Design and baseline characteristics of participants in the TRial of Economic Incentives to Promote Physical Activity (TRIPPA): a randomized controlled trial of a six month pedometer program with financial incentives. *Contemp Clin Trials,* 41**,** 238-247.

FITZGERALD, M. & MCCLELLAND, T. 2017. What makes a mobile app successful in supporting health behaviour change? *Health Educ J,* 76**,** 373-381.

FLORES MATEO, G., GRANADO-FONT, E., FERRÉ-GRAU, C. & MONTAÑA-CARRERAS, X. 2015. Mobile Phone Apps to Promote Weight Loss and Increase Physical Activity: A Systematic Review and Meta-Analysis. *J Med Internet Res,* 17**,** e253.

FRANKE, W. D., RAMEY, S. L. & SHELLEY, M. C. 2002. Relationship between cardiovascular disease morbidity, risk factors, and stress in a law enforcement cohort. *J Occup Environ Med,* 44**,** 1182-1189.

FREAK-POLI, R., WOLFE, R., BRAND, M., DE COURTEN, M. & PEETERS, A. 2013a. Eight-month postprogram completion: change in risk factors for chronic disease amongst participants in a 4-month pedometer-based workplace health program. *Obesity (Silver Spring),* 21**,** E360-368.

FREAK-POLI, R. L., CUMPSTON, M., PEETERS, A. & CLEMES, S. A. 2013b. Workplace pedometer interventions for increasing physical activity. *Cochrane Database Syst Rev***,** Cd009209.

FRENCH, D. P., OLANDER, E. K., CHISHOLM, A. & MC SHARRY, J. 2014. Which behaviour change techniques are most effective at increasing older adults' self-efficacy and physical activity behaviour? A systematic review. *Ann Behav Med,* 48**,** 225-234.

FUKUOKA, Y., VITTINGHOFF, E., JONG, S. S. & HASKELL, W. 2010. Innovation to motivation--pilot study of a mobile phone intervention to increase physical activity among sedentary women. *Prev Med,* 51**,** 287-289.

GALE, N. K., HEATH, G., CAMERON, E., RASHID, S. & REDWOOD, S. 2013. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol,* 13**,** 117.

GAN, Y., YANG, C., TONG, X., SUN, H., CONG, Y., YIN, X., LI, L., CAO, S., DONG, X., GONG, Y., SHI, O., DENG, J., BI, H. & LU, Z. 2015. Shift work and diabetes mellitus: a meta-analysis of observational studies. *Occup Environ Med,* 72**,** 72-78.

GANESAN, A. N., LOUISE, J., HORSFALL, M., BILSBOROUGH, S. A., HENDRIKS, J., MCGAVIGAN, A. D., SELVANAYAGAM, J. B. & CHEW, D. P. 2016. International Mobile-Health Intervention on Physical Activity, Sitting, and Weight: The Stepathlon Cardiovascular Health Study. *J Am Coll Cardiol,* 67**,** 2453-2463.

GARDNER, B., SMITH, L., LORENCATTO, F., HAMER, M. & BIDDLE, S. J. 2016. How to reduce sitting time? A review of behaviour change strategies used in sedentary behaviour reduction interventions among adults. *Health Psychol Rev,* 10**,** 89-112.

GERSHON, R. R., LIN, S. & LI, X. 2002. Work stress in aging police officers. *J Occup Environ Med,* 44**,** 160-167.

GIBSON, R., ERIKSEN, R., SINGH, D., VERGNAUD, A. C., HEARD, A., CHAN, Q., ELLIOTT, P. & FROST, G. 2018. A cross-sectional investigation into the occupational and socio-demographic characteristics of British police force employees reporting a dietary pattern associated with cardiometabolic risk: findings from the Airwave Health Monitoring Study. *Eur J Nutr,* 57**,** 2913-2926.

GILSON, N. D., PAVEY, T. G., VANDELANOTTE, C., DUNCAN, M. J., GOMERSALL, S. R., TROST, S. G. & BROWN, W. J. 2016. Chronic disease risks and use of a smartphone application during a physical activity and dietary intervention in Australian truck drivers. *Aust N Z J Public Health,* 40**,** 91-93.

GILSON, N. D., PAVEY, T. G., WRIGHT, O. R., VANDELANOTTE, C., DUNCAN, M. J., GOMERSALL, S., TROST, S. G. & BROWN, W. J. 2017. The impact of an m-Health financial incentives program on the physical activity and diet of Australian truck drivers. *BMC Public Health,* 17**,** 467.

GOOGLE PLAY. 2019. *Bupa Boost* [Online]. Available: https://play.google.com/store/apps/details?id=com.bupa.boost&hl=en_GB [Accessed 02.08.19].

GOWIN, M., CHENEY, M., GWIN, S. & WANN, T. F. 2015. Health and Fitness App Use in College Students: A Qualitative Study. *Am J Health Educ,* 46**,** 223-230.
GRAM, B., HOLTERMANN, A., BULTMANN, U., SJOGAARD, G. & SOGAARD, K. 2012. Does an exercise intervention improving aerobic capacity among construction workers also improve musculoskeletal pain, work ability, productivity, perceived physical exertion, and sick leave?: a randomized controlled trial. *J Occup Environ Med,* 54**,** 1520-1526.
GREAVES, C. J., SHEPPARD, K. E., ABRAHAM, C., HARDEMAN, W., RODEN, M., EVANS, P. H., SCHWARZ, P. & GROUP, I. S. 2011. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health,* 11**,** 119.
GREENE, J. C., CARACELLI, V. J. & GRAHAM, W. F. 1989. Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educ Eval Policy Anal,* 11**,** 255-274.
GREENFIELD, R., BUSINK, E., WONG, C. P., RIBOLI-SASCO, E., GREENFIELD, G., MAJEED, A., CAR, J. & WARK, P. A. 2016. Truck drivers' perceptions on wearable devices and health promotion: a qualitative study. *BMC Public Health,* 16**,** 677.
GREMAUD, A. L., CARR, L. J., SIMMERING, J. E., EVANS, N. J., CREMER, J. F., SEGRE, A. M., POLGREEN, L. A. & POLGREEN, P. M. 2018. Gamifying Accelerometer Use Increases Physical Activity Levels of Sedentary Office Workers. *J Am Heart Assoc,* 7, e007735.
GUO, X., COBERLEY, C., POPE, J. E. & WELLS, A. 2015. The Value of a Well-Being Improvement Strategy: Longitudinal Success across Subjective and Objective Measures Observed in a Firm Adopting a Consumer-Driven Health Plan. *J Occup Environ Med,* 57**,** 1055-1062.
HAMARI, J. & KOIVISTO, J. 2015. "Working out for likes": An empirical study on social influence in exercise gamification. *Comput Hum Behav,* 50**,** 333-347.
HANDLEY, M. A., HARLEMAN, E., GONZALEZ-MENDEZ, E., STOTLAND, N. E., ALTHAVALE, P., FISHER, L., MARTINEZ, D., KO, J., SAUSJORD, I. & RIOS, C. 2016. Applying the COM-B model to creation of an IT-enabled health coaching and resource linkage program for low-income Latina moms with recent gestational diabetes: the STAR MAMA program. *Implement Sci,* 11**,** 73.
HARRIES, T., ESLAMBOLCHILAR, P., RETTIE, R., STRIDE, C., WALTON, S. & VAN WOERDEN, H. C. 2016. Effectiveness of a smartphone app in increasing physical activity amongst male adults: a randomised controlled trial. *BMC Public Health,* 16**,** 925.
HARRISON, D., MARSHALL, P., BERTHOUZE, N. & BIRD, J. 2014. Tracking physical activity: problems related to running longitudinal studies with commercial devices. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication.* Seattle, Washington: ACM.
HARTLEY, T. A., BURCHFIEL, C. M., FEKEDULEGN, D., ANDREW, M. E. & VIOLANTI, J. M. 2011. Health disparities in police officers: comparisons to the U.S. general population. *Int J Emerg Ment Health,* 13**,** 211-220.
HEATH, G., CAMERON, E., CUMMINS, C., GREENFIELD, S., PATTISON, H., KELLY, D. & REDWOOD, S. 2012. Paediatric 'care closer to home': stake-holder views and barriers to implementation. *Health Place,* 18**,** 1068-1073.

HENDRIKSEN, I. J., BERNAARDS, C. M., STEIJN, W. M. & HILDEBRANDT, V. H. 2016a. Longitudinal Relationship Between Sitting Time on a Working Day and Vitality, Work Performance, Presenteeism, and Sickness Absence. *J Occup Environ Med,* 58**,** 784-789.

HENDRIKSEN, I. J., SNOIJER, M., DE KOK, B. P., VAN VILSTEREN, J. & HOFSTETTER, H. 2016b. Effectiveness of a Multilevel Workplace Health Promotion Program on Vitality, Health, and Work-Related Outcomes. *J Occup Environ Med,* 58**,** 575-583.

HERBER, O. R., ATKINS, L., STORK, S. & WILM, S. 2018. Enhancing self-care adherence in patients with heart failure: a study protocol for developing a theory-based behaviour change intervention using the COM-B behaviour model (ACHIEVE study). *BMJ Open,* 8**,** e025907.

HIGGINS, J. P. T. & GREEN, S. 2011 (eds). Cochrane handbook for systematic reviews of interventions. Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available: www.handbook.cochrane.org [Accessed 20.02.18].

HILES, S. A., LAMERS, F., MILANESCHI, Y. & PENNINX, B. W. J. H. 2017. Sit, step, sweat: longitudinal associations between physical activity patterns, anxiety and depression. *Psychol Med,* 47**,** 1466-1477.

HOLDEN, R. J. & KARSH, B. T. 2010. The technology acceptance model: its past and its future in health care. *J Biomed Inform,* 43**,** 159-172.

HOLMEN, H., WAHL, A., TORBJORNSEN, A., JENUM, A. K., SMASTUEN, M. C. & RIBU, L. 2016. Stages of change for physical activity and dietary habits in persons with type 2 diabetes included in a mobile health intervention: the Norwegian study in RENEWING HEALTH. *BMJ Open Diabetes Res Care,* 4**,** e000193.

HOWARTH, A., QUESADA, J., SILVA, J., JUDYCKI, S. & MILLS, P. R. 2018. The impact of digital health interventions on health-related outcomes in the workplace: A systematic review. *Digit Health,* 4**,** 2055207618770861.

HUANG, Y., BENFORD, S. & BLAKE, H. 2019. Digital Interventions to Reduce Sedentary Behaviors of Office Workers: Scoping Review. *J Med Internet Res,* 21**,** e11079.

INSTITUTE OF OCCUPATIONAL SAFETY AND HEALTH. 2015. *Working well: Guidance on promoting health and wellbeing at work* [Online]. Available: www.iosh.co.uk/workingwell [Accessed 01.10.17].

INTERNATIONAL DATA CORPORATION. 2019. *Forecast wearables unit shipments worldwide from 2014 to 2023 (in millions) [Graph]* [Online]. Statista. Available: https://www.statista.com/statistics/437871/wearables-worldwide-shipments/ [Accessed 02.08.19].

IRVINE, A. B., PHILIPS, L., SEELEY, J., WYANT, S., DUNCAN, S. & MOORE, R. W. 2011. Get moving: a web site that increases physical activity of sedentary employees. *Am J Health Promot,* 25**,** 199-206.

JANZ, N. K. & BECKER, M. H. 1984. The Health Belief Model: a decade later. *Health Educ Q,* 11**,** 1-47.

JETTE, M., SIDNEY, K. & BLUMCHEN, G. 1990. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. *Clin Cardiol,* 13**,** 555-565.

JONES, C. A. 2016. *Examining the efficacy and feasibility of digital activity monitors and shared active desks to reduce employee sedentary behavior.* 10119815 Ph.D., The University of North Carolina at Chapel Hill.

KANSTRUP, A. M., BERTELSEN, P. & JENSEN, M. B. 2018. Contradictions in digital health engagement: An activity tracker's ambiguous influence on vulnerable young adults' engagement in own health. *Digit Health,* 4**,** 2055207618775192.
KAZI, A., DUNCAN, M., CLEMES, S. & HASLAM, C. 2014. A survey of sitting time among UK employees. *Occup Med (Lond),* 64**,** 497-502.
KESSLER, R. C., AMES, M., HYMEL, P. A., LOEPPKE, R., MCKENAS, D. K., RICHLING, D. E., STANG, P. E. & USTUN, T. B. 2004. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. *J Occup Environ Med,* 46**,** S23-37.
KESSLER, R. C., BARBER, C., BECK, A., BERGLUND, P., CLEARY, P. D., MCKENAS, D., PRONK, N., SIMON, G., STANG, P., USTUN, T. B. & WANG, P. 2003. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med,* 45**,** 156-174.
KEYWORTH, C., HART, J., THOONG, H., FERGUSON, J. & TULLY, M. 2017. A Technological Innovation to Reduce Prescribing Errors Based on Implementation Intentions: The Acceptability and Feasibility of MyPrescribe. *JMIR Hum Factors,* 4**,** e17.
KING, A. C., HEKLER, E. B., GRIECO, L. A., WINTER, S. J., SHEATS, J. L., BUMAN, M. P., BANERJEE, B., ROBINSON, T. N. & CIRIMELE, J. 2016. Effects of Three Motivationally Targeted Mobile Device Applications on Initial Physical Activity and Sedentary Behavior Change in Midlife and Older Adults: A Randomized Trial. *PLoS One,* 11**,** e0156370.
KNIGHT, E., STUCKEY, M. I., PRAPAVESSIS, H. & PETRELLA, R. J. 2015. Public health guidelines for physical activity: is there an app for that? A review of android and apple app stores. *JMIR Mhealth Uhealth,* 3**,** e43.
KNUTH, A. G. & HALLAL, P. C. 2009. Temporal trends in physical activity: a systematic review. *J Phys Act Health,* 6**,** 548-559.
KOHL, L. F., CRUTZEN, R. & DE VRIES, N. K. 2013. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. *J Med Internet Res,* 15**,** e146.
KORLIN, J., ALEXANDERSON, K. & SVEDBERG, P. 2009. Sickness absence among women and men in the police: A systematic literature review. *Scand J Public Health,* 37**,** 310-319.
KOYANAGI, A., STUBBS, B. & VANCAMPFORT, D. 2018. Correlates of low physical activity across 46 low- and middle-income countries: A cross-sectional analysis of community-based data. *Prev Med,* 106**,** 107-113.
KOYLE, A. E. 2013. *The value of infusing self-efficacy theory with smartphone technology to sustain walking for exercise in a worksite population.* 3595640 Ph.D., The University of Utah.
KOZEY-KEADLE, S., LIBERTINE, A., LYDEN, K., STAUDENMAYER, J. & FREEDSON, P. S. 2011. Validation of wearable monitors for assessing sedentary behavior. *Med Sci Sports Exerc,* 43**,** 1561-1567.
KURTI, A. N., LOGAN, H., MANINI, T. & DALLERY, J. 2015. Physical Activity Behavior, Barriers to Activity, and Opinions About a Smartphone-Based Physical Activity Intervention Among Rural Residents. *Telemed E-Health,* 21**,** 16-23.
LAGESTAD, P. & VAN DEN TILLAAR, R. 2014. Longitudinal changes in the physical activity patterns of police officers. *Int J Police Sci Manag,* 16**,** 76-86.

LAMBERT, J. D., GREAVES, C. J., FARRAND, P., CROSS, R., HAASE, A. M. & TAYLOR, A. H. 2017. Assessment of fidelity in individual level behaviour change interventions promoting physical activity among adults: a systematic review. *BMC Public Health,* 17**,** 765.

LEAR, S. A., HU, W., RANGARAJAN, S., GASEVIC, D., LEONG, D., IQBAL, R., CASANOVA, A., SWAMINATHAN, S., ANJANA, R. M., KUMAR, R., ROSENGREN, A., WEI, L., YANG, W., CHUANGSHI, W., HUAXING, L., NAIR, S., DIAZ, R., SWIDON, H., GUPTA, R., MOHAMMADIFARD, N., LOPEZ-JARAMILLO, P., OGUZ, A., ZATONSKA, K., SERON, P., AVEZUM, A., POIRIER, P., TEO, K. & YUSUF, S. 2017. The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study. *Lancet*, 390, 2643-2654.

LEARY, M. R. 1992. Self-Presentational Processes in Exercise and Sport. *J Sport Exercise Psy,* 14**,** 339-351.

LEE, B. C., BENDIXSEN, C., LIEBMAN, A. K. & GALLAGHER, S. S. 2017. Using the Socio-Ecological Model to Frame Agricultural Safety and Health Interventions. *J Agromedicine,* 22**,** 298-303.

LEISCHIK, R., FOSHAG, P., STRAUß, M., LITTWITZ, H., GARG, P., DWORRAK, B. & HORLITZ, M. 2015. Aerobic Capacity, Physical Activity and Metabolic Risk Factors in Firefighters Compared with Police Officers and Sedentary Clerks. *PLoS One,* 10**,** e0133113.

LIAO, Y., CHOU, C. P., HUH, J., LEVENTHAL, A. & DUNTON, G. 2017. Examining acute bi-directional relationships between affect, physical feeling states, and physical activity in free-living situations using electronic ecological momentary assessment. *J Behav Med,* 40**,** 445-457.

LIMESURVEY GMBH 2003. LimeSurvey: An Open Source survey tool. Hamburg, Germany: LimeSurvey GmbH.

LIU, L., SHI, Y., LI, T., QIN, Q., YIN, J., PANG, S., NIE, S. & WEI, S. 2016. Leisure time physical activity and cancer risk: evaluation of the WHO's recommendation based on 126 high-quality epidemiological studies. *Br J Sports Med,* 50**,** 372-378.

LOSINA, E., SMITH, S. R., USISKIN, I. M., KLARA, K. M., MICHL, G. L., DESHPANDE, B. R., YANG, H. Y., SMITH, K. C., COLLINS, J. E. & KATZ, J. N. 2017. Implementation of a workplace intervention using financial rewards to promote adherence to physical activity guidelines: a feasibility study. *BMC Public Health,* 17**,** 921.

LUCAS, K. & SHERRY, J. 2004. Sex differences in video game play: A communication-based explanation. *Commun Res,* 31**,** 499-523.

LYONS, E. J., LEWIS, Z. H., MAYRSOHN, B. G. & ROWLAND, J. L. 2014. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. *J Med Internet Res,* 16**,** e192.

MAINSBRIDGE, C. P., COOLEY, P. D., FRASER, S. P. & PEDERSEN, S. J. 2014. The effect of an e-health intervention designed to reduce prolonged occupational sitting on mean arterial pressure. *J Occup Environ Med,* 56**,** 1189-1194.

MATEVEY, C., ROGERS, L. Q., DAWSON, E. & TUDOR-LOCKE, C. 2006. Lack of Reactivity During Pedometer Self-Monitoring in Adults. *Meas Phys Educ Exerc Sci,* 10**,** 1-11.

MCCALLUM, C., ROOKSBY, J. & GRAY, C. M. 2018. Evaluating the Impact of Physical Activity Apps and Wearables: Interdisciplinary Review. *JMIR Mhealth Uhealth,* 6**,** e58.

MERCER, K., LI, M., GIANGREGORIO, L., BURNS, C. & GRINDROD, K. 2016. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. *JMIR Mhealth Uhealth,* 4**,** e40.

METZGER, J., CATELLIER, D., EVENSON, K., TREUTH, M., ROSAMOND, W. & SIEGA-RIZ, A. 2010. Associations between patterns of objectively measured physical activity and risk factors for the metabolic syndrome. *Am J Health Promot,* 24, 161-169.

MICHIE, S., ASHFORD, S., SNIEHOTTA, F. F., DOMBROWSKI, S. U., BISHOP, A. & FRENCH, D. P. 2011b. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. *Psychol Health,* 26**,** 1479-1498.

MICHIE, S., RICHARDSON, M., JOHNSTON, M., ABRAHAM, C., FRANCIS, J., HARDEMAN, W., ECCLES, M. P., CANE, J. & WOOD, C. E. 2013. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med,* 46**,** 81-95.

MICHIE, S., VAN STRALEN, M. & WEST, R. 2011a. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implement Sci,* 6, 42.

MIDDELWEERD, A., VAN DER LAAN, D. M., VAN STRALEN, M. M., MOLLEE, J. S., STUIJ, M., TE VELDE, S. J. & BRUG, J. 2015. What features do Dutch university students prefer in a smartphone application for promotion of physical activity? A qualitative approach. *Int J Behav Nutr Phys Act,* 12**,** 31.

MILLER, R. & BROWN, W. 2004. Steps and sitting in a working population. *Int J Behav Med,* 11**,** 219-224.

MOHER, D., LIBERATI, A., TETZLAFF, J., ALTMAN, D. G. & GROUP, P. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol,* 62**,** 1006-1012.

MOORE, G., AUDREY, S., BARKER, M., BOND, L., BONELL, C., HARDEMAN, W., MOORE, L., O'CATHAIN, A., TINATI, T., WIGHT, D. & BAIRD, J. 2014. Process evaluation of complex interventions: UK Medical Research Council (MRC) guidance. London: MRC Population Health Science Research Network.

MOORE, G. F., AUDREY, S., BARKER, M., BOND, L., BONELL, C., HARDEMAN, W., MOORE, L., O'CATHAIN, A., TINATI, T., WIGHT, D. & BAIRD, J. 2015. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ,* 350**,** h1258.

MORGAN, D. L. 2007. Paradigms Lost and Pragmatism Regained:Methodological Implications of Combining Qualitative and Quantitative Methods. *J Mix Methods Res,* 1**,** 48-76.

MORRIS, A., MURPHY, R., SHEPHERD, S. & GRAVES, L. 2018. Multi-Stakeholder Perspectives of Factors That Influence Contact Centre Call Agents' Workplace Physical Activity and Sedentary Behaviour. *Int J Environ Res Public Health,* 15, 1484.

MUNTANER, A., VIDAL-CONTI, J. & PALOU, P. 2016. Increasing physical activity through mobile device interventions: A systematic review. *Health Inform J,* 22**,** 451-469.

NATIONAL INSTITUTE FOR HEALTH AND CARE EXCELLENCE. 2007. *Behaviour change: General approaches. Public health guideline.* [Online]. Manchester: NICE. Available: http://nice.org.uk/guidance/ph6 [Accessed 05.07.19].
NATIONAL INSTITUTE FOR HEALTH RESEARCH. 2019. *Feasibility and pilot studies* [Online]. Available: https://www.nihr.ac.uk/funding-and-support/documents/funding-for-research-studies/research-programmes/PGfAR/CCF-PGfAR-Feasibility-and-Pilot-studies.pdf [Accessed 18.04.19].
NEIL-SZTRAMKO, S. E., GOTAY, C. C., SABISTON, C. M., DEMERS, P. A. & CAMPBELL, K. C. 2017. Feasibility of a telephone and web-based physical activity intervention for women shift workers. *Transl Behav Med,* 7**,** 268-276.
NHS CENTRE FOR REVIEWS AND DISSEMINATION 2009. *Systematic reviews: CRD's guidance for undertaking reviews in health care,* York, Centre for Reviews and Dissemination, University of York.
NUFFIELD HEALTH. 2018. *A Healthier Workplace: How employers can reduce physical inactivity* [Online]. Available: https://www.sportspartnershiphw.co.uk/uploads/a-healthier-workplace-nuffield-health-whitepaper.pdf [Accessed 13.12.18].
O'BRIEN, H. L. & TOMS, E. G. 2008. What is user engagement? A conceptual framework for defining user engagement with technology. *J Am Soc Inf Sci Tec,* 59**,** 938-955.
O'CATHAIN, A., MURPHY, E. & NICHOLL, J. 2010. Three techniques for integrating data in mixed methods studies. *BMJ,* 341**,** c4587.
O'CATHAIN, A., THOMAS, K. J., DRABBLE, S. J., RUDOLPH, A. & HEWISON, J. 2013. What can qualitative research do for randomised controlled trials? A systematic mapping review. *BMJ Open,* 3, e002889.
O'CONNELL, S. E., GRIFFITHS, P. L. & CLEMES, S. A. 2014. Seasonal variation in physical activity, sedentary behaviour and sleep in a sample of UK adults. *Ann Hum Biol,* 41**,** 1-8.
O'DONOGHUE, G., PERCHOUX, C., MENSAH, K., LAKERVELD, J., VAN DER PLOEG, H., BERNAARDS, C., CHASTIN, S. F., SIMON, C., O'GORMAN, D., NAZARE, J. A. & CONSORTIUM, D. 2016. A systematic review of correlates of sedentary behaviour in adults aged 18-65 years: a socio-ecological approach. *BMC Public Health,* 16**,** 163.
O'REILLY, G. A. & SPRUIJT-METZ, D. 2013. Current mHealth technologies for physical activity assessment and promotion. *Am J Prev Med,* 45**,** 501-507.
OKONKWO, O. C., ROTH, D. L., PULLEY, L. & HOWARD, G. 2010. Confirmatory factor analysis of the validity of the SF-12 for persons with and without a history of stroke. *Qual Life Res,* 19**,** 1323-1331.
OLSEN, H. M., BROWN, W. J., KOLBE-ALEXANDER, T. & BURTON, N. W. 2018. A Brief Self-Directed Intervention to Reduce Office Employees' Sedentary Behavior in a Flexible Workplace. *J Occup Environ Med,* 60**,** 954-959.
OWEN, N., HEALY, G., MATTHEWS, C. & DUNSTAN, D. 2010. Too much sitting: the population health science of sedentary behavior. *Exerc Sport Sci Rev,* 38, 105-113.
OWEN, N., SUGIYAMA, T., EAKIN, E. E., GARDINER, P. A., TREMBLAY, M. S. & SALLIS, J. F. 2011. Adults' sedentary behavior determinants and interventions. *Am J Prev Med,* 41**,** 189-196.
PATEL, A. V., BERNSTEIN, L., DEKA, A., FEIGELSON, H. S., CAMPBELL, P. T., GAPSTUR, S. M., COLDITZ, G. A. & THUN, M. J. 2010. Leisure time spent

sitting in relation to total mortality in a prospective cohort of US adults. *Am J Epidemiol,* 172**,** 419-29.

PATEL, M. S., ASCH, D. A., ROSIN, R., SMALL, D. S., BELLAMY, S. L., EBERBACH, K., WALTERS, K. J., HAFF, N., LEE, S. M., WESBY, L., HOFFER, K., SHUTTLEWORTH, D., TAYLOR, D. H., HILBERT, V., ZHU, J. S., YANG, L., WANG, X. M. & VOLPP, K. G. 2016b. Individual Versus Team-Based Financial Incentives to Increase Physical Activity: A Randomized, Controlled Trial. *J Gen Intern Med,* 31**,** 746-754.

PATEL, M. S., ASCH, D. A., ROSIN, R., SMALL, D. S., BELLAMY, S. L., HEUER, J., SPROAT, S., HYSON, C., HAFF, N., LEE, S. M., WESBY, L., HOFFER, K., SHUTTLEWORTH, D., TAYLOR, D. H., HILBERT, V., ZHU, J. S., YANG, L., WANG, X. M. & VOLPP, K. G. 2016a. Framing Financial Incentives to Increase Physical Activity Among Overweight and Obese Adults A Randomized, Controlled Trial. *Ann Intern Med,* 164**,** 385-394.

PATEL, M. S., FOSCHINI, L., KURTZMAN, G. W., ZHU, J., WANG, W., RARESHIDE, C. A. L. & ZBIKOWSKI, S. M. 2017. Using Wearable Devices and Smartphones to Track Physical Activity: Initial Activation, Sustained Use, and Step Counts Across Sociodemographic Characteristics in a National Sample. *Ann Intern Med,* 167**,** 755-757.

PATEL, M. S., VOLPP, K. G., ROSIN, R., BELLAMY, S. L., SMALL, D. S., HEUER, J., SPROAT, S., HYSON, C., HAFF, N., LEE, S. M., WESBY, L., HOFFER, K., SHUTTLEWORTH, D., TAYLOR, D. H., HILBERT, V., ZHU, J., YANG, L., WANG, X. & ASCH, D. A. 2018. A Randomized, Controlled Trial of Lottery-Based Financial Incentives to Increase Physical Activity Among Overweight and Obese Adults. *Am J Health Promot,* 32**,** 1568-1575.

PATRICK, K., HEKLER, E. B., ESTRIN, D., MOHR, D. C., RIPER, H., CRANE, D., GODINO, J. & RILEY, W. T. 2016. The Pace of Technologic Change: Implications for Digital Health Behavior Intervention Research. *Am J Prev Med,* 51**,** 816-824.

PEDERSEN, S. J., COOLEY, P. D. & MAINSBRIDGE, C. 2014. An e-health intervention designed to increase workday energy expenditure by reducing prolonged occupational sitting habits. *Work,* 49**,** 289-95.

PEEL, E., DOUGLAS, M., PARRY, O. & LAWTON, J. 2010. Type 2 diabetes and dog walking: patients' longitudinal perspectives about implementing and sustaining physical activity. *Br J Gen Pract,* 60**,** 570-577.

PENEDO, F. J. & DAHN, J. R. 2005. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Curr Opin Psychiatry,* 18**,** 189-193.

PERNEGER, T. V. 1998. What's wrong with Bonferroni adjustments. *BMJ,* 316**,** 1236-1238.

PERSKI, O., BLANDFORD, A., WEST, R. & MICHIE, S. 2017. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. *Transl Behav Med,* 7, 254-267.

PHAM, Q., WILJER, D. & CAFAZZO, J. A. 2016. Beyond the Randomized Controlled Trial: A Review of Alternatives in mHealth Clinical Trial Methods. *JMIR Mhealth Uhealth,* 4**,** e107.

PHILLIPS, S. M., CADMUS-BERTRAM, L., ROSENBERG, D., BUMAN, M. P. & LYNCH, B. M. 2018. Wearable Technology and Physical Activity in Chronic Disease: Opportunities and Challenges. *Am J Prev Med,* 54**,** 144-150.

POIRIER, J., BENNETT, W. L., JEROME, G. J., SHAH, N. G., LAZO, M., YEH, H. C., CLARK, J. M. & COBB, N. K. 2016. Effectiveness of an Activity Tracker- and Internet-Based Adaptive Walking Program for Adults: A Randomized Controlled Trial. *J Med Internet Res,* 18, e34.

PRINCE, S. A., SAUNDERS, T. J., GRESTY, K. & REID, R. D. 2014. A comparison of the effectiveness of physical activity and sedentary behaviour interventions in reducing sedentary time in adults: a systematic review and meta-analysis of controlled trials. *Obes Rev,* 15**,** 905-919.

PROCHASKA, J. O. & DICLEMENTE, C. C. 1983. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol,* 51**,** 390-395.

PRONK, N. P., MARTINSON, B., KESSLER, R. C., BECK, A. L., SIMON, G. E. & WANG, P. 2004. The association between work performance and physical activity, cardiorespiratory fitness, and obesity. *J Occup Environ Med,* 46**,** 19-25.

PROPER, K. I., VAN DEN HEUVEL, S. G., DE VROOME, E. M., HILDEBRANDT, V. H. & VAN DER BEEK, A. J. 2006. Dose-response relation between physical activity and sick leave. *Br J Sports Med,* 40**,** 173-178.

PUBLIC HEALTH ENGLAND. 2018. *Public Health Profiles* [Online]. Public Health England. Available: https://fingertips.phe.org.uk/search/public%20health%20profiles#page/0/gid/1/pat/6/par/E12000004/ati/102/are/E06000015 [Accessed 14.01.19].

PUTTONEN, S., HÄRMÄ, M. & HUBLIN, C. 2010. Shift work and cardiovascular disease - pathways from circadian stress to morbidity. *Scand J Work Environ Health,* 36**,** 96-108.

QSR INTERNATIONAL 2015. NVivo Qualitative Data Analysis Software: Version 11. QSR International Pty Ltd.

RAMEY, S. L., PERKHOUNKOVA, Y., HEIN, M., CHUNG, S., FRANKE, W. D. & ANDERSON, A. A. 2016. Building Resilience in an Urban Police Department. *J Occup Environ Med,* 58**,** 796-804.

RAMEY, S. L., PERKHOUNKOVA, Y., MOON, M., TSENG, H. C., WILSON, A., HEIN, M., HOOD, K. & FRANKE, W. D. 2014. Physical Activity in Police Beyond Self-Report. *J Occup Environ Med,* 56**,** 338-343.

REED, J. L., COLE, C. A., ZISS, M. C., TULLOCH, H. E., BRUNET, J., SHERRARD, H., REID, R. D. & PIPE, A. L. 2018. The Impact of Web-Based Feedback on Physical Activity and Cardiovascular Health of Nurses Working in a Cardiovascular Setting: A Randomized Trial. *Front Physiol,* 9**,** 142.

REID, R. E., INSOGNA, J. A., CARVER, T. E., COMPTOUR, A. M., BEWSKI, N. A., SCIORTINO, C. & ANDERSEN, R. E. 2017. Validity and reliability of Fitbit activity monitors compared to ActiGraph GT3X+ with female adults in a free-living environment. *J Sci Med Sport,* 20**,** 578-582.

REIJONSAARI, K., VEHTARI, A., KAHILAKOSKI, O. P., VAN MECHELEN, W., ARO, T. & TAIMELA, S. 2012. The effectiveness of physical activity monitoring and distance counseling in an occupational setting - Results from a randomized controlled trial (CoAct). *BMC Public Health,* 12**,** 344.

REIJONSAARI, K., VEHTARI, A., VAN MECHELEN, W., ARO, T. & TAIMELA, S. 2009. The effectiveness of physical activity monitoring and distance counselling in an occupational health setting--a research protocol for a randomised controlled trial (CoAct). *BMC Public Health,* 9**,** 494.

RESEARCH ENGLAND. 2018. *REF 2021: Draft guidance on submissions* [Online]. Available: https://www.ref.ac.uk/media/1016/draft-guidance-on-submissions-ref-2018_1.pdf [Accessed 24.05.19].

RILEY, W. T., RIVERA, D. E., ATIENZA, A. A., NILSEN, W., ALLISON, S. M. & MERMELSTEIN, R. 2011. Health behavior models in the age of mobile interventions: are our theories up to the task? *Transl Behav Med,* 1**,** 53-71.

RITCHIE, J. & SPENCER, L. 1994. Qualitative data analysis for applied policy research. *In:* BRYMAN, A. & BURGESS, R. G. (eds.) *Analysing Qualitative Data.* London: Routledge.

ROBINSON, T. 2008. Applying the socio-ecological model to improving fruit and vegetable intake among low-income African Americans. *J Community Health,* 33**,** 395-406.

ROBROEK, S. J., VAN LENTHE, F. J., VAN EMPELEN, P. & BURDORF, A. 2009. Determinants of participation in worksite health promotion programmes: a systematic review. *Int J Behav Nutr Phys Act,* 6**,** 26.

ROBROEK, S. J. W., LINDEBOOM, D. E. M. & BURDORF, A. 2012. Initial and Sustained Participation in an Internet-delivered Long-term Worksite Health Promotion Program on Physical Activity and Nutrition. *J Medl Internet Res,* 14, e43.

ROLAND BERGER. 2016. *Global digital health market from 2015 to 2020, by major segment (in billion U.S. dollars) [Graph]* [Online]. Statista. Available: https://www.statista.com/statistics/387867/value-of-worldwide-digital-health-market-forecast-by-segment/ [Accessed 02.08.19].

ROWE-ROBERTS, D., CERCOS, R. & MUELLER, F. 2014. Preliminary results from a study of the impact of digital activity trackers on health risk status. *Stud Health Techn Inform,* 204**,** 143-148.

RUIZ, J. & MORROW, E. 2005. Retiring the Old Centurion: Life after a Career in Policing—An Exploratory Study. *Int J Publ Admin,* 28**,** 1151-1186.

SCHOEPPE, S., ALLEY, S., VAN LIPPEVELDE, W., BRAY, N. A., WILLIAMS, S. L., DUNCAN, M. J. & VANDELANOTTE, C. 2016. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. *Int J Behav Nutr Phys Act,* 13**,** 127.

SCHRAGER, J. D., SHAYNE, P., WOLF, S., DAS, S., PATZER, R. E., WHITE, M. & HERON, S. 2017. Assessing the Influence of a Fitbit Physical Activity Monitor on the Exercise Practices of Emergency Medicine Residents: A Pilot Study. *JMIR Mhealth Uhealth,* 5**,** e2.

SCHWARZER, R., LIPPKE, S. & LUSZCZYNSKA, A. 2011. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). *Rehabil Psychol,* 56**,** 161-70.

SEDENTARY BEHAVIOUR RESEARCH NETWORK 2012. Letter to the Editor: Standardized use of the terms “sedentary” and “sedentary behaviours”. *Appl Physiol Nutr Metab,* 37**,** 540-542.

SEDGWICK, P. 2014. Pitfalls of statistical hypothesis testing: multiple testing. *BMJ,* 349**,** g5310.

SHULMAN, L. M., GRUBER-BALDINI, A. L., ANDERSON, K. E., FISHMAN, P. S., REICH, S. G. & WEINER, W. J. 2010. The clinically important difference on the unified Parkinson's disease rating scale. *Arch Neurol,* 67**,** 64-70.

SIMBLETT, S., GREER, B., MATCHAM, F., CURTIS, H., POLHEMUS, A., FERRAO, J., GAMBLE, P. & WYKES, T. 2018. Barriers to and Facilitators of Engagement With Remote Measurement Technology for Managing Health:

Systematic Review and Content Analysis of Findings. *J Med Internet Res,* 20**,** e10480.
SIMONS, D., DE BOURDEAUDHUIJ, I., CLARYS, P., DE COCKER, K., VANDELANOTTE, C. & DEFORCHE, B. 2018a. A Smartphone App to Promote an Active Lifestyle in Lower-Educated Working Young Adults: Development, Usability, Acceptability, and Feasibility Study. *JMIR Mhealth Uhealth,* 6**,** e44.
SIMONS, D., DE BOURDEAUDHUIJ, I., CLARYS, P., DE COCKER, K., VANDELANOTTE, C. & DEFORCHE, B. 2018b. Effect and Process Evaluation of a Smartphone App to Promote an Active Lifestyle in Lower Educated Working Young Adults: Cluster Randomized Controlled Trial. *JMIR Mhealth Uhealth,* 6**,** e10003.
SKOGSTAD, M., LUNDE, L. K., SKARE, O., MAMEN, A., ALFONSO, J. H., OVSTEBO, R. & ULVESTAD, B. 2016. Physical activity initiated by employer and its health effects; an eight week follow-up study. *BMC Public Health,* 16**,** 377.
SLOOTMAKER, S. M., CHINAPAW, M. J. M., SCHUIT, A. J., SEIDELL, J. C. & VAN MECHELEN, W. 2009. Feasibility and Effectiveness of Online Physical Activity Advice Based on a Personal Activity Monitor: Randomized Controlled Trial. *J Med Internet Res,* 11**,** e27.
SMALL, M. L. 2011. How to Conduct a Mixed Methods Study: Recent Trends in a Rapidly Growing Literature. *Annu Rev Sociol,* 37, 57-86.
SOROKA, A. & SAWICKI, B. 2014. Physical activity levels as a quantifier in police officers and cadets. *Int J Occup Med Environ Health,* 27, 498-505.
SPENCE, J. C., RHODES, R. E. & CARSON, V. 2016. Challenging the Dual-Hinge Approach to Intervening on Sedentary Behavior. *Am J Prev Med,* 52, 403-406.
STATACORP 2015. Stata Statistical Software: Release 14. College Station, TX: StataCorp LP.
STATACORP 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC.
STEINHARDT, M., GREENHOW, L. & STEWART, J. 1991. The relationship of physical activity and cardiovascular fitness to absenteeism and medical care claims among law enforcement officers. *Am J Health Promot,* 5**,** 455-460.
STEPHENSON, A., MCDONOUGH, S. M., MURPHY, M. H., NUGENT, C. D. & MAIR, J. L. 2017. Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. *Int J Behav Nutr Phys Act,* 14**,** 105.
STERNE, J. A., WHITE, I. R., CARLIN, J. B., SPRATT, M., ROYSTON, P., KENWARD, M. G., WOOD, A. M. & CARPENTER, J. R. 2009. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ,* 338**,** b2393.
STOKOLS, D. 1996. Translating social ecological theory into guidelines for community health promotion. *Am J Health Promot,* 10**,** 282-298.
STRATH, S. J., KAMINSKY, L. A., AINSWORTH, B. E., EKELUND, U., FREEDSON, P. S., GARY, R. A., RICHARDSON, C. R., SMITH, D. T., SWARTZ, A. M., AMERICAN HEART ASSOCIATION PHYSICAL ACTIVITY COMMITTEE OF THE COUNCIL ON, L., CARDIOMETABOLIC, H., CARDIOVASCULAR, E. C. R., PREVENTION COMMITTEE OF THE COUNCIL ON CLINICAL, C. & COUNCIL 2013. Guide to the assessment of physical activity: Clinical and

research applications: a scientific statement from the American Heart Association. *Circulation,* 128**,** 2259-2279.
STURGES, J. E. & HANRAHAN, K. J. 2004. Comparing telephone and face-to-face qualitative interviewing: a research note. *Qual Res,* 4**,** 107-118.
SULLIVAN, A. N. & LACHMAN, M. E. 2016. Behavior Change with Fitness Technology in Sedentary Adults: A Review of the Evidence for Increasing Physical Activity. *Front Public Health,* 4**,** 289.
SUN, M., FENG, W., WANG, F., LI, P., LI, Z., LI, M., TSE, G., VLAANDEREN, J., VERMEULEN, R. & TSE, L. A. 2018. Meta-analysis on shift work and risks of specific obesity types. *Obes Rev,* 19**,** 28-40.
TAKAYANAGI, N., KITAMURA, K., YAMAUCHI, T. & TOKIMITSU, I. 2018. Effects of promoting daily physical activity on physical and mental health in older individuals. *J Phys Ther Sci,* 30**,** 1315-1322.
TASHAKKORI, A. & CRESWELL, J. W. 2007. Editorial: Exploring the nature of research questions in mixed methods research. *J Mix Methods Res,* 1**,** 207-211.
TASHAKKORI, A. & TEDDLIE, C. 2010. Overview of contemporary issues in mixed methods research. *In:* TASHAKKORI, A. & TEDDLIE, C. (eds.) *Sage Handbook of Mixed Methods in Social and Behavioural Research.* London: SAGE.
TAYLOR, A. W., WILSON, D. H. & WAKEFIELD, M. 1998. Differences in health estimates using telephone and door-to-door survey methods--a hypothetical exercise. *Aust N Z J Public Health,* 22**,** 223-236.
TEYHEN, D. S., ALDAG, M., EDINBOROUGH, E., GHANNADIAN, J. D., HAUGHT, A., KINN, J., KUNKLER, K. J., LEVINE, B., MCCLAIN, J., NEAL, D., STEWART, T., THORNDIKE, F. P., TRABOSH, V., WESENSTEN, N. & PARRAMORE, D. J. 2014. Leveraging technology: creating and sustaining changes for health. *Telemed J E Health,* 20**,** 835-849.
THOMAS, B. H., CILISKA, D., DOBBINS, M. & MICUCCI, S. 2004. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs,* 1**,** 176-184.
THOMPSON, D., BATTERHAM, A. M., PEACOCK, O. J., WESTERN, M. J. & BOOSO, R. 2016. Feedback from physical activity monitors is not compatible with current recommendations: A recalibration study. *Prev Med,* 91**,** 389-394.
THORNDIKE, A. N., MILLS, S., SONNENBERG, L., PALAKSHAPPA, D., GAO, T., PAU, C. T. & REGAN, S. 2014. Activity monitor intervention to promote physical activity of physicians-in-training: randomized controlled trial. *PLoS One,* 9**,** e100251.
TONG, A., SAINSBURY, P. & CRAIG, J. 2007. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care,* 19**,** 349-357.
TORQUATI, L., KOLBE-ALEXANDER, T., PAVEY, T. & LEVERITT, M. 2018. Changing Diet and Physical Activity in Nurses: A Pilot Study and Process Evaluation Highlighting Challenges in Workplace Health Promotion. *J Nutr Educ Behav,* 50**,** 1015-1025.
TUDOR-LOCKE, C., BASSETT, D. R., SHIPE, M. F. & MCCLAIN, J. J. 2011a. Pedometry methods for assessing free-living adults. *J Phys Act Health,* 8**,** 445-453.
UTAH DEPARTMENT OF HEALTH 2001. Interpreting the SF-12: Utah Health Status Survey. Utah: Utah Department of Health.

VAN DANTZIG, S., GELEIJNSE, G. & VAN HALTEREN, A. T. 2013. Toward a persuasive mobile application to reduce sedentary behavior. *Pers Ubiquit Comput,* 17**,** 1237-1246.

VAN DEN HEUVEL, S. G., BOSHUIZEN, H. C., HILDEBRANDT, V. H., BLATTER, B. M., ARIËNS, G. A. & BONGERS, P. M. 2005. Effect of sporting activity on absenteeism in a working population. *Br J Sports Med,* 39**,** e15.

VAN UFFELEN, J. G., WONG, J., CHAU, J. Y., VAN DER PLOEG, H. P., RIPHAGEN, I., GILSON, N. D., BURTON, N. W., HEALY, G. N., THORP, A. A., CLARK, B. K., GARDINER, P. A., DUNSTAN, D. W., BAUMAN, A., OWEN, N. & BROWN, W. J. 2010. Occupational sitting and health risks: a systematic review. *Am J Prev Med,* 39**,** 379-388.

WALSH, S. M., MEYER, M. R., GAMBLE, A., PATTERSON, M. S. & MOORE, J. B. 2017. A Systematic Review of Rural, Theory-based Physical Activity Interventions. *Am J Health Behav,* 41**,** 248-258.

WANG, J. B., CATALDO, J. K., AYALA, G. X., NATARAJAN, L., CADMUS-BERTRAM, L. A., WHITE, M. M., MADANAT, H., NICHOLS, J. F. & PIERCE, J. P. 2016. Mobile and Wearable Device Features that Matter in Promoting Physical Activity. *J Mob Technol Med,* 5**,** 2-11.

WANNAMETHEE, S. G. & SHAPER, A. G. 2001. Physical activity in the prevention of cardiovascular disease: an epidemiological perspective. *Sports Med,* 31**,** 101-114.

WARBURTON, D. E. R. & BREDIN, S. S. D. 2017. Health benefits of physical activity: a systematic review of current systematic reviews. *Curr Opin Cardiol,* 32**,** 541-556.

WARE, J., JR., KOSINSKI, M. & KELLER, S. D. 1996. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care,* 34**,** 220-233.

WARTTIG, S. L., FORSHAW, M. J., SOUTH, J. & WHITE, A. K. 2013. New, normative, English-sample data for the Short Form Perceived Stress Scale (PSS-4). *J Health Psychol,* 18**,** 1617-1628.

WELCH, W. A., SPRING, B., PHILLIPS, S. M. & SIDDIQUE, J. 2018. Moderating Effects of Weather-Related Factors on a Physical Activity Intervention. *Am J Prev Med*, 54, e83-e89.

WIDMER, R. J., ALLISON, T. G., KEANE, B., DALLAS, A., BAILEY, K. R., LERMAN, L. O. & LERMAN, A. 2016. Workplace Digital Health Is Associated with Improved Cardiovascular Risk Factors in a Frequency-Dependent Fashion: A Large Prospective Observational Cohort Study. *PLoS One,* 11**,** e0152657.

WILMOT, E., EDWARDSON, C., ACHANA, F., DAVIES, M., GORELY, T., GRAY, L., KHUNTI, K., YATES, T. & BIDDLE, S. 2012. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. *Diabetologia,* 55**,** 2895-2905.

WORLD HEALTH ORGANISATION. 2011. *mHealth: New horizons for health through mobile technologies: second global survey on eHealth* [Online]. Switzerland: WHO. Available: https://www.who.int/goe/publications/goe_mhealth_web.pdf [Accessed 31.07.19].

WORLD HEALTH ORGANISATION. 2019. *Global strategy on diet, physical activity and health: Physical activity* [Online]. WHO. Available: https://www.who.int/dietphysicalactivity/pa/en/ [Accessed 30.07.19].

WYATT, K. M., BRAND, S., ASHBY-PEPPER, J., ABRAHAM, J. & FLEMING, L. E. 2015. Understanding how healthy workplaces are created: implications for developing a National Health Service healthy workplace program. *Int J Health Serv,* 45**,** 161-185.

XIAN, Y., XU, H., XU, H., LIANG, L., HERNANDEZ, A. F., WANG, T. Y. & PETERSON, E. D. 2017. An Initial Evaluation of the Impact of Pokemon GO on Physical Activity. *J Am Heart Assoc,* 6, e005341.

YARDLEY, L., SPRING, B. J., RIPER, H., MORRISON, L. G., CRANE, D. H., CURTIS, K., MERCHANT, G. C., NAUGHTON, F. & BLANDFORD, A. 2016. Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. *Am J Prev Med,* 51**,** 833-842.

YATES, T., HAFFNER, S. M., SCHULTE, P. J., THOMAS, L., HUFFMAN, K. M., BALES, C. W., CALIFF, R. M., HOLMAN, R. R., MCMURRAY, J. J., BETHEL, M. A., TUOMILEHTO, J., DAVIES, M. J. & KRAUS, W. E. 2014. Association between change in daily ambulatory activity and cardiovascular events in people with impaired glucose tolerance (NAVIGATOR trial): a cohort analysis. *Lancet,* 383**,** 1059-1066.

YEUNG, J., MAZLOOMDOOST, D., CRISP, C. C., KLEEMAN, S. & PAULS, R. N. 2017. Impact of Electronic Feedback and Peer Comparisons on Residents' Physical Activity Level. *Journal of Graduate Medical Education,* 9**,** 527-530.

YU, J., ABRAHAM, J. M., DOWD, B., HIGUERA, L. F. & NYMAN, J. A. 2017. Impact of a workplace physical activity tracking program on biometric health outcomes. *Prev Med,* 105**,** 135-141.

ZIMMERMAN, F. H. 2012. Cardiovascular disease and risk factors in law enforcement personnel: a comprehensive review. *Cardiol Rev,* 20**,** 159-166.