View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Ulster University's Research Portal

TITLE PAGE

DO LOYALTY CARD SCHEMES ENCOURAGE PHYSICAL ACTIVITY? A CLUSTER RCT

Ruth F. Hunter[1,2]*, Aisling Gough[1,2], Jennifer M. Murray[1,2], Jianjun Tang[1,2], Sarah F. Brennan[1,2], Oliver J. Chrzanowski-Smith[3], Angela Carlin[4], Chris Patterson[1,2], Alberto Longo[2,5], George Hutchinson[2,5], Lindsay Prior[2], Mark A. Tully[1,2], David P. French[6], Jean Adams[7], Emma McIntosh[8], Yiqiao Xin[8], Frank Kee[1,2]

[1] Centre for Public Health, Queen's University Belfast, Belfast, Northern Ireland

[2] UKCRC Centre of Excellence for Public Health Research (NI), Queen's University Belfast, Belfast, Northern Ireland

[3] Department for Health, University of Bath, Bath, England

[4] School of Sport, Ulster University, Londonderry, Northern Ireland

[5] School of Biological Sciences, Queen's University Belfast, Belfast, Northern Ireland

[6] School of Psychological Sciences, University of Manchester, Manchester, England

[7] Centre for Diet and Activity Research (CEDAR), MRC Epidemiology Unit, University of Cambridge, Cambridge, England

[8] Health Economics and Health Technology Assessment, University of Glasgow, Glasgow, Scotland

***Corresponding author:** Centre for Public Health, Queen's University Belfast, Royal Victoria Hospital, Grosvenor Road, Belfast, BT12 6BJ; E-mail: ruth.hunter@qub.ac.uk; Tel: 02890978944

Keywords: Physical Activity, Intervention, Workplace, Cluster RCT, Behaviour Change Maintenance, Financial Incentives, Absenteeism, Presenteeism, Process Evaluation, Economic Evaluation, Behavioural Economics, Mediation Analyses

Competing Interests: Prof Frank Kee, Prof Emma McIntosh and Prof David French as members of the PHR Funding Board. Prof Frank Kee is also a member of the PHR Prioritisation Group. Dr French reports personal fees from Medical Research Council, personal fees from British Psychological Society, grants from NIHR, grants from MRC, grants from Greater Sport,

grants from European Commission, grants from Prevent Breast Cancer, grants from Natural Environment Research Council, grants from Multiple Sclerosis Society, grants from Cancer Research UK, outside the submitted work. Jean Adams reports my partner is Director of the National Institute of Health Research's Public Health Programme, which funded this research.
Word Count: 30, 313

ABSTRACT

Background: Increasing physical activity in the workplace can provide physical and mental health benefits for employees and economic benefits through reduced absenteeism and increased productivity for the employer. However, there is limited evidence on effective behaviour change interventions in workplace settings that lead to maintained physical activity. This study aimed to address this gap and contribute to the evidence base for effective, and cost-effective, workplace interventions.

Objectives: To determine the effectiveness and cost-effectiveness of the Physical Activity Loyalty (PAL) Scheme, a multi-component intervention based on concepts similar to those that underpin a high-street loyalty card, aimed at encouraging habitual physical activity behaviour and maintained increases in mean steps/day.

Design: A cluster-randomised controlled trial and embedded economic evaluation, behavioural economic experiments, mediation analyses and process evaluation.
Setting: Office-based employees from public sector organisations in Belfast and Lisburn city centres, Northern Ireland.

Participants: 853 participants (mean age 43.6 years (SD 9.6); 71% female) were randomly allocated by cluster to either the Intervention Group or (Waiting-List) Control Group.

Intervention: The six month intervention consisted of financial incentives (retail vouchers), feedback and other evidence-based behaviour change techniques. Sensors situated in the vicinity of the workplaces allowed participants to monitor their accumulated minutes of physical activity.

Main Outcome Measures: The primary outcome was mean steps/day recorded using a sealed pedometer (Yamax Digiwalker CW-701, Japan) worn on the waist for seven consecutive days at six and 12 months post-intervention. Secondary outcomes included health, mental wellbeing, quality of life, work absenteeism and presenteeism and use of healthcare resources.

Results: The mean steps/day were significantly lower for the Intervention Group compared to the control group (6,990 [SD 3,078] vs 7,576 [SD 3,345] respectively, adjusted mean difference= -336, 95% CI: -612 to -60, P=0.02) at six months post-baseline, but not significantly lower at 12 months post-baseline. There was a small but significant enhancement of mental wellbeing, in the Intervention Group (difference between groups for the Warwick-Edinburgh Mental Wellbeing Scale [WEMWBS] of 1.34 points, 95% CI: 0.48 to 2.20), but not for the other secondary outcomes. An economic evaluation suggested that, overall, the scheme was not cost-effective compared to no intervention. The intervention was £25.85 (95% CI: -29.89, 81.60) more costly per participant but had no effect on QALYs (incremental QALY= -0.0000891, 95% CI: -0.008, 0.008).

Limitations: Significant re-structuring of participating organisations during the study resulted in lower than anticipated recruitment and retention rates. Technical issues affected intervention fidelity.

Conclusions: Overall, assignment to the Intervention Group resulted in a small but significant **decline** in mean pedometer-measured steps/day at six months relative to baseline, compared to the waiting-list control group. The PAL Scheme was deemed to be not cost-effective compared to no intervention, primarily due to no additional QALY gained through the intervention.

Future Work: Research to better understand the mechanisms of physical activity behaviour change maintenance will help the design of future interventions.

Study Registration: ISRCTN17975376 (Registered 19/09/2014).

Funding Details: Funded by the NIHR PHR Board.

(Word count: 386)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

A&E: Accident and Emergency

ANCOVA: Analysis of Covariance

BCTs: Behaviour Change Techniques

BMI: Body Mass Index

CBA: Cost Benefit Analysis

CD: Coefficient of Determination

CEAC: Cost-Effectiveness Acceptability Curves

CIs: Confidence Intervals

CUA: Cost Utility Analysis

DCE: Discrete Choice Experiment

DBDC: Double-Bound Dichotomous Choices

DV: Dependent Variable

EQ-5D: Euroqol 5 dimension

EQ-5D-5L: Euroqol 5 dimension, 5 level

FG: Focus Group

GLM: Generalised Linear Models

GLS: Generalised Least Squares

GP: General Practice

GPAQ: Global Physical Activity Questionnaire

HPQ: Health Work Performance Questionnaire

HrQOL: Health-related Quality of Life

HTA: Health Technology Assessment

ICC: Intra-Class Correlation Coefficient

ICER: Incremental Cost-Effectiveness Ratio

IV: Independent Variable

MICE: Multiple Imputation of Chained Equations

ML: Maximum Likelihood

MV: Mediating Variable

MVPA: Moderate-to-Vigorous Physical Activity

NHS: National Health Service

NICE: The National Institute for Health and Care Excellence

NIHR: National Institute for Health Research

NIPHRN: Northern Ireland Public Health Research Network

OLS: Ordinary Least Squares Regression

ORECNI: Office of Research Ethics Committees Northern Ireland

PA: Physical Activity

PAL: Physical Activity Loyalty

PDRF: Postdoctoral Research Fellow

PPI: Patient and Public Involvement

PSSRU: Personal Social Service Research Unit

QALY: Quality Adjusted Life Year

RCT: Randomised Controlled Trial

RFID: Radio Frequency Identification

ROI: Return on Investment

SDT: Self-Determination Theory

SE: Standard Errors

SEM: Structural Equation Modelling

SF-8: Short Form 8

SRMR: Standardised Root Mean Square Residual

TRIPPA: TRial of Economic Incentives to Promote Physical Activity

UK: United Kingdom

UKCRC: UK Clinical Research Collaboration

US: United States

VAS: Visual Analogue Scale

WE: Work Environment

WEMWBS: Warwick-Edinburgh Mental Wellbeing Scale

WHO: World Health Organization

WTA: Willing-To-Accept

WTP: Willingness-To-Pay

PLAIN ENGLISH SUMMARY

Inactive lifestyles are bad for your health. We developed the Physical Activity Loyalty (PAL) Scheme to encourage office workers to incorporate physical activity into their working day. Our study was designed to find out if the PAL Scheme would help employees get active during the working day, by incentivising walking breaks and providing an interactive website with personalised feedback on goal setting, and the accumulation of PAL 'points' (minutes of activity recorded with remote sensors) that could be redeemed for modest value retail vouchers.

The PAL Scheme involved employees in four public sector workplaces in two city centres in Northern Ireland, half of whom were randomly chosen to receive the PAL Scheme programme and the other half were the control group. We measured participants' levels of physical activity using a pedometer and used questionnaires to look at how they rated their quality of life, number of hours/days absent from work (through sickness) and their wellbeing.

The PAL scheme lasted for six months, with the same measurements taken after the six months and again at 12 months. We held discussion groups with participants to find out their experiences of taking part.

Despite the vouchers, hints, tips and motivational cues from the website and emails, the group who received the PAL Scheme were slightly less active than those who had not. However, ratings of wellbeing were slightly better in those who received the programme and they had fewer hours absent from work. Feedback on the scheme was positive from participants, retailers and employers and it was likely that the scheme had wider benefits for the employer in terms of productivity and cost savings.

Given that the programme failed to achieve sufficient change in physical activity as intended, new approaches are needed to explore levels of engagement with programmes such as this.

(Word count: 299)

SCIENTIFIC SUMMARY

BACKGROUND

Increasing physical activity (PA) levels in the workplace could have physical and mental health benefits for employees and provide potential economic benefits for employers through reduced absenteeism and increased productivity. Current evidence to support the effectiveness of such interventions is mixed, with previous meta-analyses of workplace PA interventions showing small, positive, short-term effects on levels of PA but little long-term effectiveness. Further, there is scant evidence on the cost-effectiveness of such interventions. Financial and non-financial incentives are increasingly used to promote healthy lifestyles, but we know little about whether they offer effective or sustainable means to promote PA in workplace settings.

AIMS AND OBJECTIVES

We conducted a cluster randomised controlled trial (RCT) in a workplace setting of an incentives based intervention for promoting PA, based on the concept of a "loyalty card", to deliver the following objectives:

1. To investigate the effectiveness of the intervention to increase employees' PA levels;
2. To investigate if any change in PA behaviour is maintained over time;
3. To conduct cost-effectiveness, cost-utility and cost-benefit analyses of the intervention;
4. To investigate how the intervention impacts other health behaviours and outcomes;
5. To investigate wider work-related effects including sickness absenteeism and work presenteeism;
6. To investigate the mediators of (a) uptake and use of the loyalty card, (b) initiation, and (c) maintenance of behaviour change;
7. To conduct a parallel qualitative study to identify the reasons for intervention effects, how and why it worked for participants, and to explore possible mediators of behaviour change;
8. To conduct a Discrete Choice Experiment (DCE) to investigate the effective levels of incentives for such interventions;

9. To conduct a behavioural economic field experiment on inter-temporal and risk preferences to investigate the relationship between PA behavioural change, discounting and financial incentives.

METHODS

The study is a cluster randomised controlled trial (RCT), with a parallel qualitative process evaluation and cost-effectiveness analysis. Trial procedures and protocol have been previously reported and are summarised below. A protocol variation, to limit follow-up to 12 months, was approved by NIHR.

The Intervention (duration six months)

The intervention, known as the Physical Activity Loyalty (PAL) Scheme, is based on similar concepts to high street loyalty cards whereby 'points' are rewarded for repeated behaviour. This six month intervention comprised of financial incentives for PA undertaken during the working day and recorded remotely though sensors in the workplace neighbourhood, with participant access to an intervention website offering tools for planning, goal setting, feedback on PA and redemption of earned points. Those assigned to the waiting-list control group were offered the opportunity to participate in the intervention after the 12 month follow-up period.

Sample

Participants were healthy adults working in office-based occupations in public sector organisations in Lisburn and Belfast city centres, Northern Ireland. To be eligible, participants had to be based at the worksite for at least four hours/day on three days/week, have a current contract lasting the duration of the study, and have no recent history of conditions that would restrict their ability to take part in PA. Clusters were defined as the smallest organisational unit, for example, specific buildings or departments participating within the trial. Of the 1,209 assessed for eligibility, 853 participants from 37 clusters were randomised into two groups (n=457 Intervention Group; n=396 Control Group). Further, 71 were lost to follow-up either from being uncontactable or having moved workplaces, and 154 withdrew during trial follow-up. After omitting those who did not supply valid pedometer data (primary outcome measure), pedometer data for the primary outcome at six months was analysed for n= 485 (Intervention Group n=249, Control Group n=236).

Outcomes and Measures

The primary outcome was PA behaviour change at six months (mean steps/day). Pedometer data was considered valid if the participant provided ≥250 steps/day for three or more days at each data collection period. Participants completed the Global Physical Activity Questionnaire (GPAQ) to elucidate the context of PA undertaken.

Secondary outcomes included measures of health and well-being using the Short Form-8 (SF-8), the Quality of Life (EQ-5D-5 L), and the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS).

Work-related impacts, including absenteeism and presenteeism were measured using the World Health Organization (WHO) Health Work Performance Questionnaire (HPQ). Putative mediators of behaviour change initiation and maintenance were measured at baseline, four weeks and six months. A range of variables for use in the economic evaluation were measured; and, in subsets of participants, variables used in contingent valuation and economic experiments were measured.

Data Collection and Analyses

Data were collected (i) at baseline, including demographic characteristics, PA behaviour, health and wellbeing, work-related impacts, behavioural mediators and moderators, double-bound dichotomous choices (DBDC) and a Discrete Choice Experiment (DCE) questionnaire; (ii) at four weeks (putative mediators) (iii) six months (PA behaviour, behavioural mediators, health and wellbeing, work-related impacts) and (iv) 12 months (PA behaviour). All self-report measures were collected online via Qualtrics (www.Qualtrics.com). Outcomes were compared between groups using ANCOVA adjusting for baseline values, randomisation stratum and season, with standard errors corrected for clustering. Indirect effects of hypothesised mediators on six and 12 month PA were examined using the structural equation modelling (SEM) based product-of-coefficients approach with confidence intervals formed using the bias-corrected bootstrap procedure.

Process Evaluation

A qualitative process evaluation was conducted alongside the trial in order to provide in-depth qualitative data on both the implementation and outcomes of the intervention. Focus groups were conducted to solicit participants' views of the PAL scheme post-intervention (six months

post-baseline). Semi-structured interviews and focus groups were also undertaken with employers and retailers (who offered retail vouchers as incentives as part of the intervention) and again with participants at 12 months post-intervention to describe facets of and reasons for any maintained behaviour change. All qualitative data were analysed using a thematic framework.

Health Economics

The primary economic evaluation took the form of a within trial Cost Utility Analysis (CUA), adopting a public sector perspective. Costs included the intervention costs (apportioned per participant) and health-care resource use. Health outcomes were expressed in terms of quality-adjusted life-years (QALYs) accrued over the six month follow-up period in the CUA.

The primary analysis used an incremental cost-effectiveness ratio (ICER) estimated by dividing the adjusted difference in mean costs between groups by the adjusted difference in mean QALYs between groups. ICER estimates were compared with a £20,000 - £30,000 per QALY threshold applied by NICE (National Institute for Health and Care Excellence). A supplementary Cost Benefit Analysis (CBA) was undertaken from an employer's perspective by using a 'net-cost model', incorporating not only the intervention costs but also the avoided costs of absenteeism and productivity loss due to sick days. All analyses were undertaken according to the principle of intention-to-treat.

Behavioural Economics

To determine the financial incentive level that might stimulate behaviour change, two stated preference methods were employed, namely: (i) Contingent Valuation, used to measure participants' willingness-to-accept (WTA) financial incentives for increasing their PA; and (ii) DCE which examined the monetary value required for increasing different types and levels of PA.

RESULTS ("*INTENTION TO TREAT*" ANALYSIS)

Primary Outcomes: Immediate Post-intervention (Six Months Post-baseline)

At six months post-intervention, mean steps/day were significantly lower for the Intervention Group compared with the control group (6,990 [SD 3,078] vs 7,576 [SD 3,345] respectively, mean difference= -336, 95% CI: -612 to -60, P=0.02) after adjustment for baseline values,

randomisation stratum and season, and correction for cluster effects. There was also a significant difference between the Intervention Group compared to the control group for minutes/week of self-reported work-related PA (mean difference=-33.3, 95% CI: -65.44 to -1.24, P=0.04) but not for minutes/week self-report moderate to vigorous PA (MVPA) (mean difference=4.12, 95% CI: -47.07 to 55.31, P=0.88).

Effects at 12 Months Post-baseline

There was a non-significant difference between the Intervention and Control Groups in steps/day (7,790 [SD 3,462] vs 8,203 [SD 3,401], adjusted mean difference =-570, 95% CI: -1,267 to 127, P=0.11) after adjustment for baseline values, randomisation stratum and season, and correction for cluster effects. There were also non-significant differences between groups for minutes/week of self-reported work-related PA (mean difference=7.0, 95% CI: -12.6 to 26.6, P=0.48) and minutes/week self-report MVPA (mean difference=77.0, 95% CI: -7.9 to 162.0, P=0.08).

Secondary Outcomes

There was a significant difference between groups for the WEMWBS, in favour of the Intervention Group (adjusted mean difference=1.34, 95% CI: 0.48 to 2.20, P<0.01), but not for the other secondary outcomes.

Sensitivity Analysis

Given the loss to follow-up and missing mean steps/day data, a further analysis was carried out following imputation by chained equations. In this case the difference in mean steps/day was -526 (95% CI: -948 to -104, p=0.02), in the same direction as the primary analysis.

Process Evaluation

Feedback from participants on the PAL scheme was positive. A number of themes emerged from the focus group discussions on the benefits participants received from their participation in the intervention. Benefits identified by participants included increased levels of PA, health benefits, social benefits and increased productivity. Participants highlighted how the intervention had instigated changes in their usual routine, leading to increased PA across the working day. However, barriers and facilitators to the PAL scheme were also identified with regard to work demands, time and the weather.

Cost Effectiveness Analysis

The intervention was £25.85 (95% CI: -29.89, 81.60) more costly per participant but had no effect on QALYs (incremental QALY= -0.0000891, 95% CI: -0.008, 0.008). The bootstrapped cost and QALY pairs spread over the four quadrants of the cost-effectiveness plane indicating large uncertainty especially surrounding the effect on QALY. Overall, the findings suggest the scheme was not cost-effective. A CBA demonstrated fewer hours absent from work through sickness in the Intervention Group (2.97 hours over a four week period; p=0.62), which pro-rata equates to 17.82 hours for the six-month intervention period and could result in savings ranging from £66 to £735 on average depending on the wage rate employed at current intervention costs (=£55.68). The estimated cost saving is associated with great uncertainty; due to the statistically insignificant effect of the absenteeism, the probability that the PAL scheme is cost saving (net cost < £0) for employers ranged from 57% to 64%.

Behavioural Economics

On average, participants' WTA financial incentives was £1.38 (95% CI: £1.16, £1.61) for increasing PA for 30 minutes/week and was £2.80 (95% CI: £2.32, £3.27) for 60 minutes additional PA/week. The average money required by participants for increasing walking or cycling to and from places, moderate-intensity and vigorous-intensity recreational PA was £2.88/hour (95% CI: £2.33; £3.43), £1.02/hour (95% CI: £0.68; £1.37), and £3.29/hour (95% CI: £2.72; £3.86), respectively. The minimum monetary incentives necessary for increasing PA (60 mins of moderate PA) differed significantly for inactive (£3.24 [95% CI: £2.30, £4.17]) and active (£0.92 [95% CI: £0.15, £1.68]) participants at baseline, with inactive participants requiring significantly more monetary incentive.

Mediation Analysis

Random-effects regression analyses showed that there were significant increases at four weeks post-baseline in intentions (b=0.29, p=0.02), identified regulation (b=0.14, p=0.01), integrated regulation (b=0.23, p<0.01), intrinsic motivation (b=0.18, p<0.01), and social norms (b=0.23, p<0.01), for intervention compared to control participants. Where b represents the coefficient for group assignment (intervention vs. control) in random-effects regressions. None of the changes in these variables were significantly related to six month changes in mean steps/day, controlling for group assignment.

However, there were significant and positive, indirect effects of group assignment on change in mean steps/day at six months through changes in several of the mediators of maintenance (measured at baseline and six months). At six months, there were significant increases in identified regulation (b=0.11, p=0.02), integrated regulation (b=0.26, p<0.01), intrinsic motivation (b=0.17, p<0.01) and habit (b=0.48, p<0.01), for intervention compared to control participants. Where b represents the indirect effect (i.e. the multiplication of the coefficient of the path relating group assignment to the mediator and the coefficient of the path relating the mediator to PA). These changes were shown in random-effects regression analyses to be related to changes in pedometer steps/day at six months, controlling for group assignment.

CONCLUSIONS

The trial demonstrated that assignment to the Intervention Group resulted in a small but significant decline in the primary outcome (i.e. mean pedometer-measured steps/day at six months) relative to baseline, compared to the waiting-list control group. Self-reported minutes/week of workplace PA also declined (compared to control) but there was **no** significant change for total MVPA. At 12 months, those in the Intervention Group still had a lower mean step/day total than did the control group, but the difference was not significant.

Feedback on the scheme was generally positive from participants, retailers and employers. In addition to the marginal though statistically discernible from zero gain in mental wellbeing for participants, the scheme had wider benefits for the employer in terms of productivity, as well as enabling networks and partnerships to be built between businesses and retailers in relation to the rewards element of the scheme.

Whilst the results showed that intervention was not likely to be cost-effective from an NHS perspective over the six month time horizon, results highlighted that the Intervention Group consumed less health care resources compared to the control group, though this difference was not statistically significant. However, there was a net cost saving for the employers for intervention participants arising due to reduced absenteeism, ranging from £66 to £735 depending on the wage rate employed. Whilst the decline in absenteeism hours was not statistically significant they are arguably economically significant. Hence, from this perspective and valuing absenteeism using a human capital approach, the intervention could be deemed to be cost-beneficial.

Whilst assignment to the Intervention Group led to increases in some putative mediators of initiation, these increases were not related to PA behaviour change at six months. Two potential reasons are: (i) changes in these mediators of initiation do not induce change in PA behaviour; (ii) changes in mediators of initiation are not carried through to PA behaviour change at six months.

Though the primary outcomes are not positive in relation to increased and maintained PA, there were some positive aspects which merit further attention and which should be examined in future PA intervention studies, such as: the use of self-regulation techniques, with social and environmental prompts, to promote habit formation as new PA behaviours becomes more automatic. Monitoring of behavioural outcomes is an additional self-regulation technique which should be explored, as it can potentially encourage participants to focus on their satisfaction with behaviour change.

Trial Registration: ISRCTN17975376 (Registered 19/09/2014).
Funding: Funded by NIHR PHR Board.
(Word count: 2384)

1 INTRODUCTION

1.1 RATIONALE FOR CURRENT STUDY

Previously the authors completed a pilot study of the feasibility of conducting a randomised controlled trial (RCT) investigating the use of financial incentives for physical activity (PA) in the workplace[1] which: assessed the recruitment process and retention of office-based employees to a trial; tested outcome measures and data collection processes; and assessed the feasibility of programme implementation in a public sector organisation. In that pilot a high uptake rate (63% of those invited took part) indicated that our recruitment strategy was successful and that the intervention was acceptable to the target population. Our data showed that over 50% of participants recruited at baseline were categorised as having low PA levels, and that the intervention was acceptable to a wide range of individuals, including those currently physically inactive, with the potential for significant reach in the population. Furthermore, this strategy was successful in recruiting a representative sample, in terms of gender and age from office-based public sector organisations. High retention rates at six months (85%) showed that our electronic method of data collection was acceptable, and that the outcome measures were feasible. Although there was no significant difference in PA levels (ascertained by the GPAQ) between the Intervention and Control Groups in the pilot, the study was not powered to demonstrate an effect size of Cohen's d~0.21 (suggested by the literature) which initially informed our trial design. Finally, over 90% (n=331) of participants were satisfied with taking part in the scheme with 89% (n=322) stating that the PAL card was 'very helpful' in encouraging them to undertake more PA.

1.2 EXISTING RESEARCH

1.2.1 Physical Activity Levels

Physical inactivity is estimated to cause 6-10% of deaths worldwide from non-communicable diseases such as coronary heart disease, type 2 diabetes and certain cancers.[2] On a global scale, estimates have highlighted that physical inactivity cost $ (INT$) 53·8 billion worldwide in 2013.[3] The impact of increasingly inactive lifestyles is thought to cost the NHS over £1billion annually.[4] Within the UK, adults are recommended to undertake at least 150 minutes/week of moderate to vigorous PA (MVPA), for example, cycling, brisk walking or running.[5] The current guidelines also include recommendations to reduce sedentary behaviour across the day, by taking regular breaks at work and incorporating active travel.[6] Despite government

guidelines highlighting the health benefits of regular PA,[4] still over 80% of adults in the UK[6] and over 60% of adults in Northern Ireland[7,8] are not meeting current recommended levels.

Within the most recent NICE (The National Institute for Health and Care Excellence) guidelines,[9] employers are encouraged to promote PA in the work place by identifying workplace 'Champions' or by implementing workplace walking and cycling programmes. As a high percentage of the adult population worldwide is in employment, and most of the waking day is spent at work, workplace-based interventions are practical settings for PA promotion and have the potential to significantly contribute to habitual PA and other beneficial health behaviours. With increasing numbers of inactive office-based occupations, improvements in PA levels may also contribute positively to wellbeing, mental health and productivity.[10] It has been suggested that there is a strong business case for investment in the health and wellbeing of the workforce,[11] and in a report published by Price Waterhouse Coopers it was estimated that for every £1 invested in workplace health and wellbeing, there is a potential return of over £4 as a consequence of improvements in absenteeism and productivity,[12] which are thought to cost in the region of £30 billion annually.[10]

1.2.2 Workplace-based PA Interventions

Current evidence to support the effectiveness of workplace PA interventions is mixed,[13] with reviews,[14–16] calling for more robust research on workplace interventions and well-designed RCTs. Previous meta-analyses of workplace PA interventions have shown small, positive, short-term effects,[15,16] on levels of walking but little long-term effectiveness is evident.[9] Thus there is a recognised need to develop workplace interventions purposefully designed to encourage PA behaviour change maintenance.

More recently, a systematic review and meta-analysis of 26 intervention studies published in 2016 identified three main intervention strategies to reduce sedentary time and increase PA within the workplace: educational or behavioural change programmes (for example, goal setting, motivational interviewing), environmental changes (for example, sit-stand workstations) and multicomponent interventions (for example, environmental changes coupled with behaviour change techniques).[17] The review found that workplace-based interventions, especially multi-component interventions (implementing both an educational and environmental aspect to the intervention), effectively reduced workplace sitting time.

However, most of these studies demonstrated behaviour change in the short-term, with the longest follow-up from the included studies recorded at 12 months.

1.2.3 Financial Incentives

Financial incentives have been proposed as a method of promoting healthy lifestyles by the UK government[18–20] but evidence in this field is relatively limited to date and considered controversial by some. Financial incentives have been shown to be effective in the short-term in smoking cessation studies[21] and have also proved effective in encouraging discrete health behaviours such as attendance at vaccination programmes.[20]

In a systematic review published in 2015,[22] which examined the effectiveness of financial incentives in changing health-related behaviours in general, it was reported that behaviour change was maintained until 18 months post-baseline and three months after incentive removal. Effectiveness measured at more than 6-12 months post-baseline was moderated by the participants' socioeconomic circumstances. In a recent trial, rebates were offered to participants when they purchased fruit and vegetables and this led to improvements in Healthy Eating Index-2010 scores.[23] The follow-up in this trial was conducted at four-six months and nine-11 months after implementation, so only relatively short-term effects have been demonstrated.

Giles and colleagues[24] reviewed the effectiveness of financial incentives for behaviour change (n=17 reports of n=16 studies) in a range of health behaviours, including smoking cessation, vaccinations/screening and PA and found some evidence that effect sizes decreased as post-intervention follow-up period and incentive value increased.

Relatively little research has been conducted to date with regards to the provision of financial incentives in the promotion of PA. A recent systematic review which only included three studies was suggestive of short-term effectiveness, with the authors unsurprisingly calling for more research in this area.[25] A recent trial conducted by Shin and colleagues[26] sought to assess the effectiveness of combining an activity tracker and financial incentives in a group of male students, contingent on participants achieving their daily PA goals (process incentive) and achieving their target weight (outcome incentive). The trial concluded that the addition of a financial incentive was effective in increasing PA levels. Christian and colleagues[27] also sought to explore the feasibility of using financial incentives to increase PA levels in a teenage

population, providing £25 of activity vouchers every month for six months. The authors found that not only did the vouchers encourage an increase in PA and significant improvements in (boys') fitness, but they also reduced levels of sedentary behaviour in both sexes and encouraged friends to socialise.

Incentives have been shown to improve participant engagement, and as the aforementioned studies demonstrated, this can lead to increased levels of PA. However, as Mantzari et al.[22] reported, financial incentives can increase attainment of the target levels of behaviour change from the start until the end of the incentives' offer, but subsequently there is a monotonic trend and weakening effect over time. With such mixed findings and few which provide longer-term follow-up data, we cannot yet be certain that increased levels of PA are maintained in the longer-term, whether the behaviour change is maintained once the incentive is removed and indeed if new habits are formed.[22,28,29]

1.2.4 Cost-Effectiveness of Incentive-Based Interventions

To date, studies have explored the cost-effectiveness of incentives across a range of health behaviours. The cost-effectiveness of financial incentives for smoking cessation in pregnancy is encouraging. Boyd and colleagues[30] conducted a smoking cessation trial where women (n=612) were randomized to usual cessation support with or without financial incentive vouchers up to the value of £400. The findings suggested that the financial incentive for smoking cessation in this population was highly cost-effective, with an incremental cost per quality-adjusted life-year (QALY) of £482.

Paul-Ebhohimhen and Avenell[31] reviewed the literature on the use of financial incentives in the treatment of overweight and obesity, and found that few studies had conducted cost-effectiveness analysis. The lack of evidence on cost-effectiveness poses a distinct limitation on its utility for policy making.

1.2.5 Cost-Effectiveness of Incentive-Based Interventions for PA

Whilst the evidence for the cost-effectiveness of incentive-based interventions is limited, studies on interventions that focus on the promotion of PA, particularly within the workplace, are even more sparse.[32] However, encouragingly, a cost-effectiveness analysis performed in our pilot trial showed that the PAL Scheme was potentially cost-effective from both a healthcare and employer's perspective.[1]

In summary, financial incentives alone may not be sufficient to promote maintained PA behaviour change, but may do so when embedded within an evidence-based, theoretically-driven intervention.[33] Further, financial incentive interventions designed using behavioural economic concepts have been shown to be effective for changing behaviour.[34] There is also limited evidence around the cost-effectiveness of interventions to promote PA in the workplace,[33,35] and of the cost-effectiveness of financial incentive-based interventions.[28] Therefore, to address such gaps in the evidence base and following the successful completion of a pilot study,[1] we aimed to conduct a cluster RCT of a complex intervention (the PAL Scheme) incorporating financial incentives to encourage PA and maintained behaviour change.

1.3 THE PAL SCHEME

The PAL Scheme is a multi-component intervention based on concepts similar to those that underpin a high-street loyalty card aimed at encouraging repeated behaviour (i.e. loyalty).[36] Components include the provision of points and financial incentives contingent on the targeted behaviour being achieved (in this case PA), and the provision of feedback on the targeted behaviour, as well as prompting and messaging to encourage the targeted behaviour through a tailored website. Participants can log onto their account on the study website and receive real-time feedback on aspects of their PA, including minutes of activity, recorded remotely by Radio Frequency Identification (RFID) enabled sensors strategically placed in the neighbouring environment. Minutes are converted to points (10 points for 1 minute of activity recorded), and collected points are redeemed for rewards (retail vouchers) sponsored by, and redeemable at, local businesses. In line with the recommendations of recent NICE guidance,[37] the study aimed to gather new evidence on effective and cost-effective workplace PA interventions. Although recent studies show some evidence of effect,[14–16] there are three problems that the current work addresses: 1) very few studies have used objective measures of PA; 2) relatively little is known about the use of financial incentives for workplace PA and free-living activity, and 3) even less is known about their cost-effectiveness.[36] Further, the PAL study addresses key knowledge gaps outlined by NICE,[37] including how individual interventions interact with environmental factors in encouraging people to walk, how to make walking habitual and which factors influence longer term behaviour change. Previous studies have used significant cash payments (up to $750)[21] which are not sustainable for the long-term. Further, to elicit a maintained behaviour change, the intervention should incorporate a phasing strategy to "shift" the focus from extrinsic to intrinsic motivation.[36]

The PAL Scheme has the potential for considerable reach at little cost and the trial has the potential to shed light on sustainable business models using a 'points' based loyalty platform, local businesses 'sponsor' the incentive (retail vouchers) in return for increased footfall to their business. This model is aligned to precepts of the Department of Health Public Health Responsibility Deal.[38] Our previous formative work showed that we can recruit and retain office-based employees to this intervention and that the intervention has the potential to positively influence their PA levels.

1.4 AIMS AND OBJECTIVES

The cluster RCT had the following objectives:

1. To investigate the effectiveness of the intervention to increase employee's PA levels;
2. To investigate if any change in PA behaviour is maintained over time;
3. To conduct cost-effectiveness analyses, Cost Utility Analysis (CUA) and Cost Benefit Analysis (CBA) of the intervention;
4. To investigate how the intervention impacts on other health behaviours and outcomes;
5. To investigate wider work-related effects including sickness absenteeism and work presenteeism;
6. To investigate the mediators of (a) uptake and use of the loyalty card, (b) initiation, and (c) maintenance of behaviour change;
7. To conduct a parallel qualitative study to further identify those who benefitted from the intervention, how and why it worked for them, and explore mediators of behaviour change;
8. To conduct a Discrete Choice Experiment (DCE) to investigate the possible effective levels of incentives for such interventions;
9. To conduct a behavioural economic field experiment on inter-temporal and risk preferences to investigate the relationship between PA behaviour change, discounting and financial incentives.

2 METHODOLOGY

2.1 INTRODUCTION

The multi-centre cluster RCT aimed to evaluate the effectiveness and cost-effectiveness of the PAL Scheme to maintain PA behaviour change. The trial incorporated nested behavioural

economic experiments and a process evaluation and the intervention was tested in an earlier feasibility study.[1]

The trial was designed to target public sector employees in predominantly office-based jobs and had two groups: an Intervention Group and a Waiting-List Control Group. Clusters of participants were randomly allocated to either the Intervention Group or Control Group. Clusters were defined as the smallest organisational unit, for example specific buildings participating within the trial.

This chapter begins by setting out the methodology for the trial in relation to sampling, outcomes and measures, data collection and analyses plans. The methodological approach undertaken for both the process evaluation and health economic evaluation is also described. In parallel to the main health economic evaluation, Section 8 describes behavioural economic experiments (involving Contingent Valuation and DCE modelling) conducted to shed light on participants' Willing-to-accept (WTA) financial incentives to improve PA and its relation to time preference. The chapter concludes by identifying the changes to the original protocol published.

2.2 CLUSTER RANDOMISED TRIAL

2.2.1 Study Population

2.2.1.1 Recruitment of Workplaces

The study targeted public sector employees involved in predominantly office-based occupations whose workplace was within Belfast or Lisburn city centres, Northern Ireland. People in predominantly office-based jobs spend a large proportion of their day physically inactive while public sector organisations have been reported to have higher sickness absence rates than private sector workplaces.[39,40] Public sector organisations were purposively sampled from those within a 2 km radius of the city centre or which could offer PA opportunities within a 2 km radius of their location, and had a minimum of 100 employees in predominantly office-based occupations. Meetings were held with senior management of these organisations to explain the purpose of the study and the practicalities involved if the study were to be implemented within the organisation. Workplaces were recruited between September 2014 and August 2015 and participant recruitment between January 2015 and October 2015.

2.2.1.2 Recruitment of Participants

Recruitment methods included email invitations to employees and posters placed around each workplace advertising the study. Emails and posters included the website address and a web-link was added to the organisations' intranet sites (previously tested in our pilot study).[1] Potential participants were able to access further information (including the Participant Information Sheet) and register their interest to participate on the study website. Potential participants were asked to complete a screening questionnaire via the study website or by telephone, to confirm their eligibility, based on the following inclusion criteria: based at recruited worksite at least four hours/day (within core hours of 8 am-6 pm) on at least three days/week; current contract anticipated to last for the duration of the study (i.e to exclude temporary workers); access to internet at work; able to give informed consent; able to communicate in English; no self-reported recent history of myocardial infarction or stroke or physical limitations that would limit ability to participate in PA (assessed using the Physical Activity Readiness Questionnaire). All participants who met the eligibility criteria and consented to participate were contacted by a member of the research team by telephone or email to complete the baseline assessment. Following this, clusters of participants were randomised to the Intervention or Control Group using computer generated random numbers. Clusters were defined as the smallest organisational unit (e.g. a department or office/floor) within each participating workplace.

2.2.1.3 Sample

During recruitment, a revised power calculation was performed (with the approval of NIHR) assuming a less demanding effect size than in our original protocol. This was proposed in light of more recent literature published and with consideration of our actual baseline data on the mean and variance of cluster size and an intra-correlation co-efficient of 0.029.[41]

In our original protocol, the sample size calculation for the trial was determined using an anticipated effect size of d=0.21 which was based upon a previous meta-analysis of workplace based PA interventions. However, none of the studies which were included in this meta-analysis were incentive-based interventions for PA behaviour change. More recent literature has been published[24,41] including a meta-analysis showing a mean effect size of approximately 1600 steps (d=0.40). Additionally, the TRial of Economic Incentives to Promote Physical Activity (TRIPPA) study,[41] which examined the influence of financial incentives on the

effectiveness of a wireless-upload pedometer to encourage weekly PA goals, was powered to detect a difference of a minimum of 30 minutes of MVPA/week between groups and reflective of a considerably higher effect size than assumed in our original calculation.

Therefore, assuming that our original estimate was too conservative, and after consultation with the Project Team, Trial Steering Committee and NIHR Public Health Research board, the power calculation was updated as follows: For an effect size of 0.40, a study of 330 per group (or 660 in total) would have 90% power at the 5% significance level. Assuming a 15% drop-out, the study would therefore need to randomise 776 participants. Please see section 2.7 for further details regarding deviation of the evaluation from the original protocol.

2.2.2 Randomisation, Concealment & Blinding

Clusters were the smallest work groups or units (e.g. a large open plan office) within each participating organisation. A random allocation sequence was drawn up by the trial statistician and group allocation was stratified to ensure similar number of clusters in both Intervention and Control Groups. Research staff were blinded to group allocation until after data collection was completed. The outcome of the randomisation was communicated to participants by email after baseline measurements were complete.

2.2.2.1 Intervention Group

The PAL Scheme is a complex multi-component intervention based on concepts similar to those that underpin a high-street loyalty card aimed at encouraging repeated behaviour (i.e. loyalty) and is designed to incorporate a range of behaviour change techniques. Components include the provision of 'points' and rewards (financial incentives) contingent on meeting targeted PA behaviour goals (extrinsic motivation, goal-setting). Participants were encouraged to undertake 150 mins/week of PA which is in line with current guidelines. The PAL Scheme integrated a novel PA remote tracking system with web-based monitoring and evidence-based behaviour change tools (i.e. self-monitoring, goal-setting). We carefully considered the design of this complex intervention in line with the MRC guidelines.[42] We have followed advice from the MRC framework in our work, including theoretical work, development of logic model, mapping of intervention onto determinants and measures onto hypothesised mechanisms, pilot/feasibility study including qualitative work.

The six month intervention involved placing sensors (wifi beacons) in the vicinity of participating workplaces at specific locations to encourage PA within a 2 km radius of participants' worksites (i.e. including the provision of prompts/cues to facilitate habit formation). The wifi beacons were placed at locations along footpaths, in local parks, leisure centres, shopping malls, bus stops and train stations. Maps of various walking routes and details about PA opportunities tailored to the workplace were provided on the study website (and instructions on how to perform behaviour). Participant's activity was logged when they passed within an approximate 25 m radius of the wifi sensors with their PAL keyfob when undertaking PA (e.g. walking). This logged the place, date and time of the bout of PA. Participants could log onto their account on the study website and receive real-time feedback on the number of minutes of PA logged by the tracking system. Minutes were converted to 'points' (ten 'points' for one minute of activity recorded), and collected 'points' were redeemable for rewards (downloadable retail vouchers) sponsored by, and redeemable at, local businesses. To reduce the risk of 'gaming', a daily 'points' cap was implemented and the transit times between sensors checked for anomalous values. Bonus rewards and 'Double Points Days' were offered when participants met specific weekly PA targets.

To tailor the intervention, a purposive sample of employees (both men and women across a range of ages) participated in three focus groups (maximum 10 participants/ group) prior to the intervention to explore aspects such as the availability of, and preferences for opportunities for PA in proximity to the workplace. Sensor locations were determined from the feedback received in the focus groups regarding popular walking routes. Employees' opinions on the website interface and the rewards redemption function, for example, were also considered and this permitted final tailoring of its functionality. To determine incentive levels, stated preferences of the participants from the DCEs to assess their mean WTA, Willingness to Pay (WTP) and the trade-offs they would make for the attributes of the incentive programme, prior to the intervention were used. This information helped determine the level of the rewards available for earned 'points'. In addition to the financial incentive element, the intervention had several other components designed to enhance the effectiveness of the incentives. These components were delivered via the study website and designed to have multiple effects: (a) to increase usage of the study website, (b) as effective Behaviour Change Techniques (BCTs) in their own right, and (c) as techniques designed to aid the transition from more extrinsically motivated behaviour to more intrinsically motivated habitual behaviour. The techniques included the provision of regular tailored motivational emails, tailored feedback, information

on walking routes in the vicinity of the participating workplaces and links to other resources such as PA advice and healthy eating guidelines. It also included self-regulation techniques of goal setting, self-monitoring, and prompts to behaviour.

Underpinning theoretical framework: The financial reward component of the intervention was based on principles of Learning Theory by providing an immediate reward (extrinsic motivation) for behaviours that offer health gains in the future. It also contained elements based on other approaches, such as goal setting, prompts, self-monitoring, and habit formation which fit within a self-regulation control theory framework, motivational messages (persuasion), and social support (vicarious experience) which should increase self-efficacy according to Social Cognitive Theory. Social Cognitive Theory also holds that satisfaction with the consequences of behaviour change can act as a reinforcing mechanism, in addition to the reinforcement of financial rewards. Thus, the financial incentive component was embedded in a complex intervention containing evidence-informed BCTs. Figure 7 shown in the appendix presents the logic model underpinning the intervention development.

2.2.2.2 Control Group

Those assigned to the waiting-list control group (n = 388) were offered the opportunity to participate in the intervention after the 12 month follow-up period. Participants in this group completed outcome measures at the same time points as the Intervention Group. The waiting-list control group did receive the intervention following the 12-month data collection period.

2.2.3 Outcomes and Measures

Outcome measures are grouped into primary and secondary outcomes as displayed in Table 1. In summary, the primary outcome was PA behaviour change (mean steps/day) measured using a sealed pedometer (Yamax Digiwalker CW-701, Japan). Participants were asked to wear the pedometer for seven consecutive days on the waistband (dominant hip). They were asked to complete a wear time diary providing details on the dates/times that they removed the monitor and wore the monitor. Participants were advised to remove the monitor when showering, bathing or undertaking any water-based activities. Pedometer data was considered valid if the participant provided ≥250 steps/day for three or more days at each data collection period.[43] Pedometers were sealed at all times to blind participants to the output and prevent reactivity which is standard practice. This device was used as a measurement tool only (i.e. was not part of the intervention) and a standardised measurement protocol used for the intervention and

control group. We followed a standard protocol that has been well-validated[44–48] and successfully employed as a measurement tool in numerous intervention studies.

Secondary outcomes included measures of health and well-being; work-related impacts; proposed mediators of behaviour change; compensatory behaviours; a range of variables for use in the economic evaluation; and a range of variables for use in the process evaluation.

All outcome measures were analysed and, where required, collected by a Postdoctoral Research Fellow (PDRF) blinded to group allocation. Self-reported outcome measures were collected at baseline and six months (unless otherwise stated), via online questionnaires distributed by email and automatically collated via Qualtrics (www.Qualtrics.com).

Table 1. Description of the primary and secondary outcomes and measures

Outcomes	**Measures**
Primary outcomes	
Physical Activity	Participants were asked to wear the pedometer for seven consecutive days and completed a wear time diary. Pedometer data

Mean steps/day objectively measured by a sealed pedometer (to blind participants to the output) worn on the waistband (Yamax Digiwalker CW-701, Japan)	was considered valid if the participant provided ≥250 steps/day for three or more days at each data collection period. Participants were also asked to complete the GPAQ to elucidate the context of PA undertaken.[49] These measures were collected at baseline, six months (immediately post-intervention) and 12 months (six months post-intervention). This schedule allowed us to account for seasonality of PA behaviours.
Secondary outcomes	
Health and Wellbeing General health, mental health, quality of life and mental wellbeing	The following validated self-report measures were completed at baseline and six months via Qualtrics. SF-8:[50] items from the SF-8 questionnaire can be derived to give an indication of both physical and mental health. EQ-5D-5L:[51] is a measure of quality of life and used to derive the health state utility measure based on five dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression (0-100), and the weighted health index. The EQ-5D-5L questionnaire is based on 5 dimensions of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression, and a visual analogue scale (0-100) that assesses the participants' health state. WEMWBS: derived from 14 statements (with higher scores indicating better mental health).[52,53] The WEMWBS comprises 14 positively worded statements, where scores are summed with higher scores indicating greater mental well-being.
Work-related Impacts	WHO HPQ:[54]

Absenteeism and presenteeism	Work absenteeism was measured by asking participants to state the number of day's sick leave in the past six months (collected at baseline and six months). Specific questions from the WHO HPQ[54] were used to measure work presenteeism. This validated method comprises three questions with answers on an 11-point Likert scale asking participants to rate their job performance levels.
Mediators and Moderators Proposed mediators and moderators of PA behaviour change	Common core theoretical constructs of PA behaviour change included outcome expectancy,[55] social norms,[13] self-efficacy,[56] financial motivation,[57] planning,[58] identified regulation,[59,60] integrated regulation,[59,60] intrinsic motivation[59,60] and intention,[61] were collected at baseline and four weeks to assess initiation of behaviour change. Perceptions of workplace environment (measured at baseline only):[62] These included perceptions of workplace environment (attractiveness, safety, accessibility, availability). Perceptions of workplace attractiveness were assessed as the sum of four items (e.g. "In my workplace environment it is pleasant to walk", score range 1-5). Perceptions of workplace safety were assessed as the sum of four items (e.g. "In my workplace environment the roads are dangerous for cyclists", score range 1-5). Perceptions of workplace accessibility were assessed as the sum of three items (e.g. "In my workplace environment there is convenient public transport", score range 1-5). Perceptions of workplace availability were assessed as the sum of three items (e.g. "In my workplace environment there is a park within walking distance", score range 1-5). Objective measures of the workplace environment using GIS data including walkability,[63] street connectivity, access to PA

	opportunities (shops, parks, leisure facilities, train/bus stations) were measured at baseline.[64] The walkability index was computed using four constructs, namely, 1) Net residential density=the ratio of residential units to the land area devoted to residential land use per block group; 2) Intersection density=ratio between the number of true intersections (three or more legs) to the land in km^2; 3) Land use mix=the degree to which a diversity of land use types are present in a land area; 4) Retail floor area ratio. Individual level, cognitive constructs measured to assess maintenance of PA behaviour change were collected at baseline and six months and included planning,[58] self-determined motivation (i.e. identified regulation, integrated regulation and intrinsic motivation),[59,60] habit,[17] recovery and maintenance self-efficacy,[20] and outcome satisfaction.[55,65] Measures of social norms[13] and workplace norms[13] were also collected at baseline and six months. Other measures included web engagement, confidence in using the internet and loyalty card usage. Potential moderators were collected at baseline. These included perceptions of workplace environment (attractiveness, safety, accessibility, availability),[62] age, gender, highest educational level, income, marital status, BMI, SF-8 Mental and Physical Component Scores,[50] EQ-5D Health State and Weighted Health Index,[51] WEMWBS[52,53], WTA (assessed by asking two open-ended questions requiring the respondents to give a minimum amount in £s necessary for compliance with an additional 30 minutes [WTA30] or 60 minutes [WTA60] of outdoor PA), and time discount rates (Q1 Imagine that you are offered either £1,000 today, or £1,005 tomorrow. Which one would you choose?; Q2 Imagine that you are offered either £1,000 today, or £1,050

	tomorrow. Which one would you choose? Coded 1 [low discount rate] if answer to Q1 is "£1,005 tomorrow" and answer to Q2 is "£1,050 tomorrow"; Coded 2 [moderate discount rate] if answer to Q1 is "£1,000 today" and the answer to Q2 is "£1,050 tomorrow"; Coded 3 [high discount rate] if answer to Q1 and Q2 is "£1,000 today").
Compensatory Behaviours Smoking, alcohol and diet	Short FFQ was collected at baseline and six months.[66,67] Self-reported smoking and alcohol behaviours were collected at baseline and six months.
Health Economic Evaluation	Changes in HrQoL (as expressed using QALYs using EQ-5D data) were measured from the participant's perspective. The EQ-5D is a validated measure and has been used extensively for cost-effectiveness analyses. Utilisation of healthcare resources was captured using a specially devised online health and social care resource use data collection form. These measures were completed at baseline and six months. Intervention costs were obtained using a modified template,[68] explicitly discriminating between intervention and research costs. Briefly these include, website development, software (e.g. license fee), hardware (e.g. sensors, loyalty cards), and intervention running costs (e.g. maintenance of sensors), the costs of negotiating incentives from local businesses, and the delivery of vouchers.
Process Evaluation	Informed by the logic model developed from the feasibility study and guided by the evaluation planning framework for public health

	interventions and research.[69] The process evaluation employed a triangulated design using both quantitative and qualitative data.

BMI: Body Mass Index; EQ-5D: Euroqol 5 dimensions; FFQ: Food Frequency Questionnaire; GIS: Geographical Information Systems; GPAQ: Global Physical Activity Questionnaire; HPQ: Health Work Performance Questionnaire; HrQoL; Health-related Quality of Life; PA: Physical activity; QALY: Quality Adjusted Life Year; SF-8: Short Form-8; WEMWBS: Warwick-Edinburgh Mental Wellbeing Scale; WHO: World Health Organisation; WTA: Willingness-To-Accept.

2.2.4 Data Collection

2.2.4.1 Statistical Analysis (Primary and Secondary Outcomes)

Primary and secondary outcomes at six and 12 months (where applicable) were compared between Intervention and Control Groups using ANCOVAs adjusting for baseline values, randomisation stratum (Large 50+ employees, Medium 20-50 employees, Small <20 employees or Schools/Colleges) and season (1=December 2015-April 2016, 2=July-August 2016) with standard errors (SEs) corrected for clustering. Due to large seasonal effects, results are presented before and after adjusting for season.

As specified in our published protocol, the relationship between group assignment and six month mean steps/day (measured using pedometers) was examined for moderating effects with ANCOVAs, adjusting for baseline pedometer steps/day, randomisation stratum and season, with SEs and p-values corrected for clustering. These analyses were then repeated to examine the relationship between group assignment and 12 month mean steps/day for moderating effects.

The moderators and moderator-by-group interactions were included as predictor variables with mean-centering for continuous moderators and creation of dummy variables for dichotomous moderators. Moderators examined included baseline socio-demographic variables (i.e. age, gender, Body Mass Index (BMI), income, education level, marital status), health measures (i.e. Short Form-8 (SF-8) Mental and Physical Component Scores, Quality of Life (EQ-5D) Health State and Weighted Health Index, and the Warwick-Edinburgh Mental Wellbeing Scale (WEMWBS)), perceptions of the workplace environment (i.e. attractiveness, safety, accessibility, availability), WTA and time discount rates.

The level of significance was $p<0.05$ for all analyses. Analyses were carried out using Stata release 13.[70]

2.2.5 Sensitivity Analysis

For a preliminary assessment of possible non-response bias, analysis of the primary outcome (i.e. six month pedometer steps/day, 12 month pedometer steps/day) was repeated after imputation of missing data using multiple imputation by chained equations (MICE) with 20 imputations. The MICE procedure operates on the assumption that data are missing at random.[71] Baseline variables were inspected for potential predictors of missingness of the six

month outcome and predictors of the outcome itself. Baseline "PA self-efficacy",[56] "intrinsic motivation",[59] "habit",[17] "recovery self-efficacy",[20] and "maintenance self-efficacy"[20] scores in addition to marital status (0=married/co-habiting, 1=other) significantly predicted missingness of the six month outcome. Baseline "PA self-efficacy",[56] "intentions",[61] "habit",[17] "outcome expectations",[55] "identified regulation",[59,60] "integrated regulation",[59,60] "intrinsic motivation",[59,60] "financial motivation",[57]"outcome satisfaction",[55,65] "planning",[58] "social norms",[13] group assignment (0=control, 1=intervention), body mass index (BMI), income (0=≤£20k, 1=>£20k) and season (1=December 2015-April 2016, 2=July-August 2016) significantly predicted the six month and 12 month outcome (i.e. mean pedometer steps/day). These variables were used in the imputation model to make predictions about the values of missing data. We then repeated the primary outcome analyses using imputed values and pedometer steps/day at six and 12 months were compared between Intervention and Control Groups using ANCOVAs adjusting for baseline values, randomisation stratum and season with SEs corrected for clustering. Due to large seasonal effects, results are presented before and after adjusting for season. These analyses were carried out using the 'mi estimate' command in Stata. Rubin's rules were used to obtain the final estimates from the estimates derived in the 20 imputed datasets.[72]

2.2.6 Ethics and Consent

Ethical approval was sought and granted from the Office of Research Ethics Committees Northern Ireland (ORECNI) prior to the start of the study (Reference: 14/NI/0090). Fully informed consent was obtained from all participants prior to their inclusion in the study. Participants were asked to confirm that they had read and understood the information sheet prior to agreeing to participate. Participants were also given an opportunity to ask any questions and ensure that these were answered satisfactorily prior to completing the consent form. Research Governance approval was granted from the South Eastern Health and Social Care Trust and Belfast Health and Social Care Trust.

2.3 QUALITATIVE PROCESS EVALUATION

2.3.1 Methods

Participants' experience of being involved in the PAL Scheme, including barriers and facilitators to engagement with the intervention, were explored using focus groups (lasting approximately one hour) at six months (n=9 focus groups) and 12 months (n=6 focus groups). Purposive sampling was undertaken to ensure a representative sample within each focus group (maximum n=10 participants per group). A schedule of open-ended questions was employed to elicit information about reactions to the intervention; barriers to PA; and suggestions for future roll-out of the intervention if proven effective. Additional focus groups were also undertaken with seventeen of the same individuals who participated in the six month focus groups (n=6 focus groups) at 12 months. On this occasion, the open-ended questions focused on PA behaviour change maintenance. Semi-structured interviews were also undertaken with senior managers of six participating employers (n=7 managers) and participating retailers (n=4) to explore their perceptions of being involved in the study. Discussions were facilitated by AG, PhD (female) and analysis of the data was undertaken by LP, PhD (male) both of whom have prior focus group and qualitative methodology experience. The researchers did not have a relationship established with participants prior to study commencement.

2.3.2 Qualitative Analysis

Focus groups and semi-structured interviews were audio-recorded and transcribed verbatim. Transcripts from focus groups and semi-structured interviews with retailers and managers were analysed using thematic analysis.[73] The thematic analysis followed six key steps, including i) familiarisation with the data, ii) systematic coding, iii) grouping of codes to form potential themes, iv) reviewing themes and v) formally naming and vi) defining themes.[73]All transcripts were independently coded by two members of the research team (AG and LP). Meaningful quotes from participants were extracted to highlight typical responses.

2.4 HEALTH ECONOMICS

2.4.1 Methods

2.4.1.1 Resource Use

Resource use data were collected from trial participants via an online questionnaire at baseline and six months. Service use included visits to: General Practice (GP), nurse, physiotherapist, Accident and Emergency (A&E) attendance, outpatient appointment, inpatient stays, medications use, and other services. Unit costs in the financial year 2015/2016 published by

the Personal Social Service Research Unit (PSSRU),[74] and NHS reference costs (where information not available from PSSRU)[75] were attached to each item of resource use.

Table 2 shows the unit costs applied to each resource use item. Costs of service use between the groups were compared with four methods: a. use complete case only, not adjusting for covariates; b. use complete case only, adjusting for covariates; c. using imputed datasets, not adjusting for covariates; and d. using imputed datasets, and adjusting for covariates. Standard errors (SEs) and 95% confidence intervals (CIs) for estimates for all models were adjusted for cluster effects.

Table 2. Unit costs for health-care resource use

Resource use item (unit used in the source)	Unit Cost	Source
NHS resource use		
GP (per contact)	£36.00	PSSRU 2015/16 pg. 145. Per patient contact lasting 9.22 minutes, with qualifications
Practice Nurse (per hour)	£43.00 (£11.11 per contact)	PSSRU 2015/16 pg. 143. Nursing average cost per hour, with qualifications. Duration of contact per patient is 15.5 minutes (PSSRU 2014/15, pg.174, based on the 2006/07 UK general practice survey)[57]
NHS Physiotherapist (per contact)	£49.00	NHS reference cost 2015/16. Allied Health professionals (AHP), physiotherapist, adult, one to one (A08A1). National average unit cost
A&E attendance (per attendance)	£185	NHS Reference costs 2015/16 Average of 'see and treat and convey' (by ambulance) (£236), 'see and treat or refer' (emergency care only) (£181), and A&E attendance by own transport (£138). pg. 18. Table 8: Costs by currency for ambulance services between 2013-14 and 2015-16. pg. 10. Table 2: Table 2: Unit costs by point of delivery,
Outpatient appointment (per attendance)	£117	NHS Reference costs 2015/16 pg. 10. Table 2: Unit costs by point of delivery, 2013/14 to 2015/16. Unit cost per outpatient attendance.
Inpatient stays (per night)	£373.00	NHS reference costs 2015/16 main schedule. Average of cost per elective and non-elective inpatient excess bed days across all currency codes. Elective inpatient excess bed days, average across all currency codes: £395. Non-elective inpatient excess bed days, average across all currency codes: £351.

a. HCHS inflation factor 1.013 (2014/15 PPI 293.1 / 2015/16 PPI 297.0). GP: General Practice; HCHS: Hospital and Community Health Service; NHS: National Health Service; pg: Page; PPI: Pay and Prices Index; PSSRU: Personal Social Service Research Unit; UK: United Kingdom.

2.4.1.2 Intervention Cost

The cost of implementing the PAL Scheme was collected from the study management team and equipment development team. Only the participants allocated to the Intervention Group were assumed to incur the cost of intervention and any 'research–related' cost was stripped from the intervention cost. This ensures the relevance of the cost estimates if the programme were to be 'rolled out' to larger numbers. Equipment costs were annuitised to the six months trial period by spreading the costs over their anticipated life span.

2.4.1.3 Absenteeism and Presenteeism

The effect of PA on absolute absenteeism was used to generate a proxy financial impact on employers. The WHO (World Health Organisation) Health Work Performance Questionnaire (HPQ)[54] was used to capture number of hours absent from work, which was completed by participants at baseline and six months. Impacts on employee absenteeism were measured using question 4 in the questionnaire: 'About how many hours altogether did you work in the past four weeks (28 days)?' and question 2 'How many hours does your employer expect you to work in a typical 7-day week?', using the formula "4*Q2 – Q4".

Individual hourly salary values in the financial year 2016 were attached to the number of hours absent from work. The hourly salary was obtained from NHS pay scales 2016, with the lowest grade at Band 1 (£7.8, equivalent to £15,251 annually), mid-grade at Band 8A (£22.86, equivalent to £44,703 annually), and highest grade at Band 9 (£50.85, equivalent to £99,437) taken separately to reflect the range of potential cost savings for employees at various salary grades.

2.4.1.4 Utility and QALYs

Effectiveness in the CUA was expressed as QALYs. QALYs were estimated using the utility index values generated from the EQ-5D-5L questionnaire collected from participants at baseline and six months follow-up. EQ-5D is a commonly used standardized generic preference-based quality of life measure addressing five domains: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression.[76] The 5L version allows the participants to select the level of problems from five levels: no problem, slight problem, moderate problems, severe problems and extreme problems / unable to do. It was chosen rather than the 3L version for its potentially improved sensitivity and decreased ceiling effect (i.e. there is a high proportion (typically >15%) of respondents reporting the best possible health who are therefore

unable to record any improvement in health status.[77,78]This makes 5L the preferred measure, especially in this population. The employees are drawn from the general population without any specified conditions, and it is thus common to see 'no problem' being answered for all dimensions creating a ceiling effect. EQ-5D-5L responses from the participants were converted to utility scores by mapping the 5L descriptive system data onto the 3L valuation set, using the mapping algorithm developed by van Hout et al.[79] as recommended by the NICE guideline and its latest position statement published in August 2017.[80] The area under the curve method was then used to estimate the QALY score over a 12 month period, following the trapezium rule assuming a linear change in utility between each assessment time point.[81]

2.4.1.5 Missing Data and Multiple Imputation

Data were missing if participants did not complete the online follow-up assessment, or returned an incomplete response. Multiple imputation was conducted for utility values and total costs at aggregate level (as compared to individual dimension and resource use item) to circumvent convergence issues with the imputation model when containing many variables. To maximize the use of the completed resource usage data, missingness of any single item on the resource use questionnaire was assumed to indicate that no resource was used for that item during the assessment period. This assumption has been used in a previous NIHR Health Technology Assessment (HTA) report.[82] The aggregate level missing data were imputed using MICE. Where the cost data have a heavily zero-inflated right skewed distribution, on account of this being a generally healthy population, the predictive mean matching method on log-transformed costs is employed as recommended.[83] A constant (£1 in this case) is added to the raw cost data to avoid problems when log-transforming zero values. After imputation, the cost data were transformed back to the original scale for estimation.[83] Intervention and Control Groups were imputed separately. In addition to cost (on log scale) and utility at baseline and six months, the imputation model also contains several baseline characteristics that are either shown to be statistically significantly related to cost and utility or were primary variables in the analysis, in line with the primary analysis of effectiveness. They were: mean steps/day, age, sex, SF-8 physical and mental scores, cluster, randomisation stratum (categorical variable coded Large, Schools/Colleges, Medium or Small), season (categorical variable indicating six months follow-up data collection time, December 2015-April 2016 or July-August 2016 time for data collection). Fifty imputed datasets were generated and then used to estimate the difference in QALYs and costs, and the ICER.

2.4.2 Statistical Analysis

2.4.2.1 Cost Utility Analysis (CUA)

In the CUA, differences in cost and QALYs between the two groups were estimated using generalised linear models (GLM) which take into account the typically skewed nature of cost and QALY data. Where histogram evidence identifies cost data as zero-inflated right-skewed, a gamma error distribution with log link is recommended. When QALY data are in a distribution with a left-skewed tail bounded by 0.5 (maximum QALY accrued for six months period), decrements of QALYs are predicted in the GLM regression with a gamma error and log link. The method of predicting decrements of QALYs was applied in a previous NIHR HTA report[84] with decrements calculated as the difference between the maximum QALYs that could possibly be accrued within the time horizon of the analysis and the actual QALY gained. This ensures that the QALY variable is right-skewed and left bounded by 0, fitting into a gamma distribution. Covariates in the GLM model were selected based on statistical significance using regressions to predict QALYs and costs with complete cases. As recommended in the literature[85] the chosen covariates included in the model were the same as in the imputation model as mentioned in the previous section (2.4.1.5). Baseline utility scores and baseline costs were also included to adjust for any imbalance between groups.[86]

Mean costs and QALYs for each group were estimated using the method of recycled predictions to adjust for covariates as well as due to the use of log link in the GLM. The ICER was estimated from the difference in cost and QALYs from the GLM regression. A 1000-iteration bootstrapping procedure was conducted to investigate the uncertainty surrounding the ICER estimate and the probability that the intervention was cost-effective under a wide range of hypothetical threshold (£0 - £100,000). Standard errors for differences in cost and QALYs were estimated through the bootstrap, adjusting for cluster effects. These results were visualised in the cost-effectiveness plane and cost-effectiveness acceptability curves (CEAC).

2.4.2.2 Sensitivity Analysis

A number of sensitivity analyses were undertaken to assess the robustness of the base-case results of the CUA to alternative assumptions. These sensitivity analyses are summarized in Table 3.

Table 3. Details of assumptions varied in each sensitivity analysis

Scenario	Element	Position in base-case analysis	Variation for the sensitivity analysis
1	Intervention – voucher cost	Voucher cost £21,000	Three business models are tested: (a) Club Marketing Ltd (b) £10 'shops for all' voucher (c) a 'sustainable business model' via a marketing consultant
2	Least expensive scenario	(a) Voucher cost £21,000 (b) Life span of equipment: • Two year life-span of the standard sensors (plastic bosses with limited number of times for battery replacement) • Key fobs only last for nine month due to unreplaceable battery	(a) Club marketing Ltd (voucher cost - £2,300) (b) Reduced cost of sensors • Life-span extended to five years with upgraded design / 'smarter' sensors that powered on lampost • Key fobs life span extended to five years with battery replaceable design With the technology upgrade, the cost of the completed sensor is estimated to decrease maximum by 50%
3	Missing data	Missing data assumed to be missing at random and a mixed strategy of imputation were conducted	Complete case analysis – missing data assumed to be missing completely at random

The cost of the intervention could easily vary with the development of appropriate technology hence differing deals with local manufacturers and retailers are explored in sensitivity analysis. Costs for retail vouchers can be reduced when identifying different deals with the retailers and thus assumptions were made varying the cost of the vouchers. Three business models have been tested by the trial group. The details are shown as below.

Deal 1: An external company (Club Marketing Co.) was used to negotiate deals with retailers. The negotiation fee was estimated to be £100 per retailer. There were 23 retailers in total so the voucher cost was estimated to be reduced down to £2,300. This leads to a much smaller intervention cost, which is £14.77 per participant over the six months period.

Deal 2: A 'shops for all' voucher was paid at £10 rate for every 400 points. The total number of points earned by the participants was 425,201, leading to a total voucher cost being £10,630.025. The resulting intervention cost was £32.99 per participant.

Deal 3: The third option is a ‘sustainable business model’ which was employed in Dallat et al’s study.[87] A marketing consultant was hired to act as an intermediary with the local business sector and was able to negotiate the provision of these vouchers 'in kind’ from local retailers. The cost is calculated using the hourly salary and the number of hours the consultant worked to negotiate the deals. Dallat et al.’s[87] study reported a cost of £4,077 for hiring the consultant, and this estimate was doubled (£8,154) as the PAL Scheme doubled the number of business partners. Again, this option leads to a smaller intervention cost than the base-case analysis, which is £27.58 per participant.

Costs of sensors and key fobs can also be reduced by assuming an extended life span with ungraded design facilitating or enabling battery replacement; thereby costs with upgraded devices were used in sensitivity analysis. In addition, complete case analysis was conducted assuming data were missing completely at random to assess the impact that imputation strategies had on the incremental cost and QALYs.

2.4.2.3 Cost Benefit Analysis (CBA): A Net Cost Model

In addition to the CUA, a CBA was conducted from the employer’s perspective. This enables the broader impact of the intervention beyond health-related quality of life to be considered, by applying values for differences in absenteeism and presenteeism rates. The CBA employed

a ‘net cost model’ to obtain a comprehensive estimate of intervention costs from an employer’s perspective by subtracting the avoided cost of absenteeism from the direct intervention cost incurred from employers.[88] This provides an estimate of the potential economic benefit of PAL Scheme from the employer’s perspective. Uncertainty of the potential economic benefit of the PAL Scheme was estimated using bootstrapping methods (1000 iterations).

2.5 BEHAVIOURAL ECONOMICS

2.5.1 Methods

Systematic reviews suggest that financial incentive interventions have become increasingly popular in promoting health behaviour change.[89] Although the idea of using financial incentives to encourage healthier behaviours seems acceptable to the public in the UK context,[90] the levels of incentives utilised in previous studies, i.e. the monetary amount thought necessary to trigger behaviour change were not based on prior research, with consequent uncertainty about the effectiveness and cost-effectiveness of incentives as an intervention tool. To investigate the optimal levels of financial incentives required to encourage changes in PA, all participants in our study were sent an online questionnaire at baseline to assess the average level of incentives that participants would be WTA to increase their PA overall as well as for increasing different types of PA. The values obtained in the preliminary Contingent Valuation analysis were later used to fix the level of the rewards available for earned “points”.

While in its infancy, behavioural economics is helping our understanding of the relationship between behaviour change, time discounting and incentives and is providing new insights that have sparked the interests of public health researchers and policy makers.[91] The principle behind financial incentives is to capitalise upon the notion of time preference, i.e. the tendency for an individual to place greater value on present relative to future benefits. Time preference is typically elicited through a binary question that asks an individual to choose between an immediate, smaller reward or a delayed, larger reward. Individuals who choose the smaller, immediate reward are said to have a high time preference and thus, it is hypothesised they, are likely to act upon the present immediate benefits of a (unhealthy) behaviour at the expense of the perceived future benefits (of healthy behaviours). Systematic reviews have shown extensively that time preference plays an important role in shaping some unhealthy behaviours, e.g. smoking and obesity.[92,93]

Two other key concepts possibly related to behaviour change are risk preference and loss aversion. The former refers to participants' willingness to take risks while the latter refers to the tendency of participants to prefer avoidance of losses to winning equivalent gains. A recent advance in the methodology of eliciting time preferences is the joint elicitation of risk and time preferences. Andersen et al.[94] have shown that when discount rates and risk preferences are elicited and analysed together, the discount rates are significantly reduced and are more accurately defined. Major behavioural economic theories have suggested that people are more sensitive to losses than to gains. For instance, Prospect Theory has demonstrated that the utility foregone from a unit monetary loss is twice the utility gain from a unit monetary gain.[95] How loss aversion is related to unhealthy behaviours is rarely investigated. As a result, all participants remaining in our study were invited at the six month follow-up to participate in an incentive-compatible behavioural economic field experiment, scheduled at their workplaces, to elicit their behavioural economic personal traits (i.e. time discounting, risk preference and loss aversion).

Apart from the baseline Contingent Valuation survey (which was necessary to inform plausible levels of incentives), the subsequent economic experiments were conducted on smaller subsamples of the population during follow-up (to keep the participant burden and costs to a minimum). This was planned as ancillary to the main trial analysis, with a primary purpose of shedding light on possible behavioural economic moderators, supplementary to the main mediation-moderation analysis conducted for the trial as a whole.

2.5.1.1 Contingent Valuation

Double-Bounded Dichotomous Choices (DBDC),[96] which is a Contingent Valuation method frequently used in the valuation of non-market goods, was used at baseline to measure participants' WTA financial incentives for increasing their PA. Through an online questionnaire, participants were asked to answer two dichotomous hypothetical questions. In the first question, participants were asked, e.g. "Would you be willing to increase your total amount of physical activity by 30 minutes per week if you were paid £0.50 per week?" If a participant answered "No" or "Don't know" to the first question, indicating that the participant considered £0.50 too low. Then, the follow-up question applied a similar format but with a doubled bid level (i.e. £1.00). Likewise, a "Yes" response to the first question indicated that the £0.50 sufficed so that the proposed amount in the second question was halved to £0.25. We consider three treatments using a between-subject design. First, to investigate whether the

minimum monetary incentives required by participants varied across the range of proposed PA increments, we considered two separate scenarios, i.e. an increase of 30 or 60 minutes of general PA/week. Second, to avoid order effects, participants were randomised into two groups, with one group first presented with the 30-minute PA scenario and then the 60-minute PA scenario; and the other group presented first with the 60-minute PA scenario and then the 30-minute PA scenario. Third, each participant was presented with a starting bid level randomly-chosen from four levels, i.e. £0.10, £0.50, £1.50, and £5.00, for the 30-minute PA scenario and in a second study a starting bid level randomly-chosen from, £0.15, £1.00, £3.00, and £10.00, for the 60-minute PA scenario.. Since what was a reasonable bid set was unknown, we provisionally suggested some initial bid sets which were pre-tested among a pilot group of 98 participants. An updated questionnaire was then sent to the rest of the study participants. Participants who failed to complete the questionnaire were excluded, leading to a sample size of 663 in the final analysis.

2.5.1.2 Discrete Choice Modelling

DBDC provides no information about participants' WTA financial incentives for increasing different types of PA, i.e. walking or cycling to and from places, or moderate-intensity and vigorous-intensity recreational PA. Alternatively, a DCE approach can be applied to provide a more direct route to the valuation of different characteristics or attributes of PA, and of marginal changes in these characteristics, rather than general PA as a whole. The levels for the four attributes are: walking or cycling to and from places (0, 30, 60, and 90 minutes/week), moderate-intensity recreational PA (0, 30, 60, and 90 minutes/week), vigorous-intensity recreational PA (0, 30, and 60 minutes/week), and monetary rewards (£0.50, £1.00, £3.00, £5.00, and £10.00/week). Each choice set had two alternative PA scenarios with different combinations of attributes and levels, and one opt-out option. Combining all attributes and levels results in (4×4×3×5=240) combinations, making it infeasible to present participants with all possible choice sets. Instead, 24 choice sets were generated following a Bayesian efficient design based on the minimisation of the D-error criterion, to reduce participant fatigue while being sufficient to estimate the main effects of attributes.[97] This is accomplished using the N-gene software. We then divided the 24 choice sets into four blocks of six choice sets and each participant received one randomized block of questions. A web-based DCE questionnaire was disseminated to participants at baseline. All participants who completed baseline PA measurement by wearing a pedometer for 7 consecutive days, were categorised as active (≥10,000 steps/day), moderate-active (≥5,000 and <10,000 steps/day), and inactive (<5,000

steps/day). Again, the participants who did not complete the DCE questionnaire were excluded from the analysis (final n=673).

2.5.1.3 Behavioural Economics Experiments

Behavioural economics concepts are useful for understanding the psychological roots of healthy and unhealthy behaviours. We focus on three behavioural economics concepts, namely time preference (i.e. the extent to which future benefits are discounted), risk preference (i.e. the willingness to take risks with greater or lesser positive outcomes), and loss aversion (i.e. the inclination to avoid losses). These behavioural economic experiments were proposed as an adjunct to the main study analysis to yield further insights into possible moderators and mediators of intervention effects. As such they point to avenues that might be fruitful for further public health research.

2.5.1.4 Measurement of Time Preference

To elicit the time preference of participants, an economic experiment methodology first employed by Coller and Williams[98] and further developed by Andersen et al.[94] was utilised. This methodology presents participants with multiple price lists which offer a choice between two real money amounts, option A and option B. Option A pays a smaller amount after a delay (e.g. £250 in a week) whilst option B offers a larger amount after a further delay (e.g. £300 in 25 weeks). The larger payment represents the sooner amount plus the interest from 'saving' the sooner amount within the experiment for the duration of the delay. Each multiple price list consisted of 10 choices between A and B, the sooner amount A and the delay to receiving B remaining constant, with the interest rate increasing as one progresses down through the sequential choices. Two "treatments" were included in this design in order to assess the structural form of the time discounting behaviour of participants, giving a 5x3 design. A front-end delay treatment where option A is either offered immediately or after a short delay (time horizon treatment) allows a test for Quasi-hyperbolic discounting.[99] The basic advantage of Quasi-hyperbolic over the typically-modelled single parameter functional forms (i.e. Mazur-hyperbolic or Exponential discounting), is that it is more flexible and incorporates two parameters each of which represents the two elements of time preference: 1) a conventional discount rate which refers to the extent to which future rewards are discounted; and (2) present-biasedness which means the relative discount rate between two proximal delays is higher than the relative discount rate between two distal dates. There were five front-end delays (one day, three days, one week, two weeks and four weeks) and three time horizons (four weeks, 12

weeks and 24 weeks). Only one front-end delay and two time horizons were randomly drawn, leading to four choice tasks for each participant.

2.5.1.5 Measurement of Risk Preference

Participants were presented with a single multiple price list that consisted of ten decisions between two lotteries: option A or option B. Each option consists of a chance to receive a large amount of money or a smaller amount. Option A is a "safe" option which pays a certain and smaller amount (e.g. £60) whereas option B (the "risky" option) involves the risk to win a substantially higher amount (e.g. £95) but also has the risk of getting a much lower amount (e.g. £50). Each of the participants was presented with two such choice tasks.

2.5.1.6 Measurement of Loss Aversion

The multiple price list format, consisting of ten decisions between two lotteries (option A or option B), was also used to measure loss aversion. Option A pays nothing (e.g. £0) while option B gives the chance to win more (e.g. £22) but also has the chance of losing money (e.g. £11). Each of the participants was presented with two such choice tasks. The maximum possible loss from the task is £20, which is no more than the £20 show-up fee so that participants will always get non-negative payments.

2.5.1.7 Payment Mechanism

Participants received two forms of payment. Firstly, all participants were credited with £20 as a thank you for participating in the sub-study, where this £20 then functions as an endowment to use in the loss aversion tasks. Secondly, real rather than hypothetical money was at stake in the behavioural economic experiments to ensure that participants made considerations that were incentivised and focussed on the choice tasks at hand. Participants had a 10% chance (implemented by rolling a ten-sided die) of receiving payment based on the random choice of one of their decisions in both the discounting task and the risk preference task. In the loss aversion tasks, participants had a chance to gain more money on top of, or lose money, from the £20 that earlier had been credited to them. The balance after the loss aversion task was then paid as cash at the end of the session.

2.5.1.8 Participants

All participants remaining in the study at six month follow-up received an e-mail invitation (including a copy of the Participant Information Sheet) detailing the expected time needed to complete the tasks. A total of 21 sessions were organised during lunch time in the meeting

rooms at the participating organisations, in which 153 participants from the Intervention and 54 from the Control Group took part in this sub-study.

2.5.2 Statistical Analysis

The statistical package STATA[70] was used to carry out all analyses. For the Contingent Valuation, a conditional log-likelihood function for the set of DBDC responses was constructed and ML estimation used to predict the mean WTA. In the DCE analysis, the responses to a choice set were taken as a specific observation. A generalised multinomial logit model[100] was used, which assumes a specification that nests both scale heterogeneity and parameter heterogeneity. The participants who chose the opt-out options for all choice sets, indicating that they did not believe the proposed hypothetical system of rewards, were excluded from the analysis. Interaction terms were used to explore whether preferences for attributes varied by participants' physical inactivity status at baseline. Next, we estimated a double-hurdle model[101] which assumes that the decision to use the PA monitoring system and how long to use it for were independent decisions so that the determinants of the two decisions are allowed to differ. The estimation was accomplished by using the Stata module "*craggit*". Missing data was excluded from the analysis. Also, another double-hurdle model was estimated to investigate the determinants respectively for the decision to increase PA at six month follow-up compared to baseline (i.e. the mean steps/day at six month follow-up was larger than the baseline mean steps/day) and the amount increased. For the analysis of the behavioural economic experiments, we simultaneously elicited present-biasedness, discount rate, risk preference, and loss aversion in an integrated framework using a hierarchical Bayesian methodology.[102] More specifically, a higher discount rate implies a higher level of impatience. And a present-biasedness parameter significantly smaller than one implies a decreasing impatience over time, i.e. a participant has long-term goals that could benefit the participant in the future but no present actions will be taken to achieve the these goals because of self-control problems. The risk aversion and loss aversion parameters take a value larger than 0 and a higher value implies a higher degree of risk aversion and loss aversion, respectively. The estimated behavioural economic parameters were later entered into a double-hurdle model to investigate their influences on the decision to increase PA at 12 month follow-up compared to six month follow-up and the amount increased.

2.6 MEDIATION ANALYSES

2.6.1 Methods

Constructs hypothesised to mediate initiation of PA behaviour were collected at baseline and at four weeks and included: outcome expectations,[55] PA self-efficacy,[56] intention,[61] planning,[58] financial motivation,[57] self-determined motivation (i.e. identified regulation, integrated regulation and intrinsic motivation)[59,60] and social norms.[13]

Constructs hypothesised to mediate maintenance of PA behaviour were collected at baseline and six months and included planning,[58] self-determined motivation (i.e. identified regulation, integrated regulation and intrinsic motivation),[59,60] habit,[17] recovery and maintenance self-efficacy,[20] outcome satisfaction,[55,65] social norms[13] and workplace norms.[13]

2.6.1.1 Engagement and Non-Usage Attrition Analyses

In addition to mediation analysis, factors associated with participant engagement and non-usage attrition were explored for participants in the Intervention Group. 'Engagement' refers to the level of exposure to an intervention, and the amount of skills practice involved (i.e. completing activities or exercises to acquire knowledge or learn behaviour relevant to the target outcome).[104] A participant's level of engagement determines the extent to which they receive the intended intervention, and research on engagement is useful for identifying which intervention components are associated with health outcomes.[105,106] Non-usage attrition refers to the phenomenon of participants ceasing intervention-use, which is particularly observed in the literature to occur in web-based intervention.[107] It is important to understand participants' non-usage patterns and the contributing factors, and to make recommendations for retaining all participants in future intervention studies, since it is difficult to measure intervention effect if participants have not been exposed to the intervention.

Three markers of overall intervention engagement (i.e. daily PA captured via the PAL Scheme PA monitoring system, use of the PAL website, reward redemption) were tracked throughout the six month intervention period and the following variables were derived:

(1) Percentage of intervention days during which participants walked for at least ten minutes[108] captured via the PAL Scheme PA monitoring system over the six month intervention period.
(2) Percentage of intervention weeks during which participants logged onto the PAL Scheme website at least once over the six month intervention period.[109]

(3) Percentage of earned points redeemed over the six month intervention period. Aside from earning points by recording activity via the PA monitoring system, this indicator captured whether participants were interested in redeeming their earned points for financial rewards to incentivise their PA behaviour.

Engagement with the different aspects of the PAL Scheme website was assessed as the frequency of hits on each intervention component for every ten days the participant accessed the website and total number of intervention components accessed on the website at least once (0-6). The six intervention components participants could access on the website were as follows:

(1) Monitoring and feedback: Data and visual representation (i.e. graphics) of the participant's PA over the intervention period for self-monitoring purposes (i.e. self-monitoring and feedback, goal setting);

(2) Rewards: Platform for participants to view their earned and bonus points, information on available rewards, and for points redemption (i.e. immediate reward contingent on behaviour change);

(3) Maps: Maps of sensor locations and example walking routes for planning of PA (i.e. information on when/where to perform PA, action planning);

(4) Health information (PA behaviour only): PA facts and information, health benefits, safety tips and tips for a physically active lifestyle (i.e. provision of information about health benefits of PA);

(5) Health information (Other behaviours): Information related to healthy eating, smoking, alcohol consumption and stress reduction (i.e. provision of information about health benefits of other health behaviours);

(6) Discussion forums: Platform for participants to contact researchers and other participants to ask questions, make enquiries, raise concerns and respond to comments, and provide social support.

Non-usage attrition (assessed separately in relation to use of the PA monitoring system to record activity and website use) was considered to occur if a participant had at least a two week lapse from use.[110,111] Non-usage attrition for recording activity via the PA monitoring system was measured as the number of days until the first two week lapse from recording activity. Website non-usage attrition was measured as the number of days until the first two week lapse from logging onto the website.

2.6.2 Statistical Analysis

All questionnaire items were scaled so that lower values indicated lower levels of the mediator/outcome. The mediation analyses are exploratory and should be interpreted with caution because of issues associated with multiple testing and the fact that the study was powered for the primary PA outcome, and not to detect change in mediators. The level of significance was $p<0.05$ for all analyses. Analyses were carried out using Stata release 13.[70]

2.6.2.1 Initiation and Maintenance of PA Behaviour

Mediators of initiation of PA behaviour (assessed at four weeks: outcome expectancies, PA self-efficacy, intention, planning, financial motivation, identified regulation, integrated regulation, intrinsic motivation, social norms) and mediators of maintenance of PA behaviour (assessed at six months: planning, identified regulation, integrated regulation, intrinsic motivation, habit, recovery and maintenance self-efficacy, outcome satisfaction, social norms and work norms) were compared between Intervention and Control Groups using random-effects regressions adjusting for baseline values of the mediator, randomisation stratum, season and baseline pedometer steps/day with SEs and p-values adjusted for cluster effects. Visual inspection of outcome distributions showed that some outcomes did not conform well to the normal distribution. Therefore, as a sensitivity analysis, these analyses were repeated using ordered-logistic regressions with mediators being categorised arbitrarily into three or four categories.

2.6.2.2 Single Mediator Models

Single mediator models were built to explore the possible mechanisms of behaviour change in the PAL Scheme. Single mediator models were run for all mediators of initiation and maintenance individually based on the structural equation modelling (SEM) based product-of-coefficients approach.[112] In each model, the independent variable (IV) was group assignment, the mediating variable (MV) was the follow-up (i.e. four weeks or six months) score of the mediator, and the dependent variable (DV) was the six month PA outcome. All analyses were adjusted for randomisation stratum, season, baseline values of the mediator and baseline PA, with SEs and p-values corrected for clustering. The significance of indirect effects was determined by 95% CIs estimated using the bias-corrected bootstrap (with 10,000 iterations) procedure recommended by MacKinnon, Lockwood, & Williams[113] and MacKinnon.[114] The maximum-likelihood (ML) method of estimation was used. Model fit was assessed with

reference to the coefficient of determination (CD), and the standardised root mean square residual (SRMR), adjusting SEs and p-values for clustering in Stata. A cut-off value of close to 0.08 for SRMR was required to consider the model a relatively good fit to the data.[115]

2.6.2.3 Single Moderated-Mediation Models

For mediators showing significant indirect effects in single mediator models, and moderators showing significant interaction effects, single moderated-mediation models were run using the moderated product-of-coefficients approach. Research suggests that mediators and moderators should be investigated as they operate together, both to provide a fuller understanding of mechanisms of behaviour change in intervention studies and to promote the advancement of theory.[116,117] Moderation of the a-path was tested by adding the moderator and moderator-by-group interaction as independent variables in the relevant single mediator models. A significant coefficient for the interaction term indicated that moderation of the a-path was present (representing an 'intervention challenge').[117] Moderation of the b-path was tested by adding the moderator and moderator-by-mediator interaction as independent variables in relevant single mediator models. A significant coefficient for the interaction term indicated that moderation of the b-path was present (representing a 'theoretical challenge'). [117]

2.6.2.4 Engagement and Non-Usage Attrition

For the Intervention Group only, random-effects generalised least squares (GLS) regressions were run with six month PA (i.e. mean steps/day) or six month mediators as the DV, and engagement variables (i.e. % intervention days in which participants undertook at least ten minutes of PA captured using the PAL Scheme PA monitoring system, % intervention weeks participants logged onto the PAL Scheme website, % earned points redeemed, frequency of hits on each of the six intervention components which could be accessed on the website for every ten days the participant accessed the website, total number of website sections accessed at least once, total minutes spent on the website) as the IVs. For models including six month PA as the DV, strata, season and baseline mean steps/day were included as covariates with SEs and p-values adjusted for clustering. For models including six month mediators as the DV, strata, season, baseline mean steps/day and baseline values of the mediator were included as covariates with SEs and p-values adjusted for clustering. This determined whether levels of six month PA, or any of the six month mediators were significantly explained by levels of engagement in various aspects of the intervention, when controlling for baseline levels of PA and mediators, group assignment, randomisation stratum, season and cluster. Since changes in

the mediators of initiation (measured at baseline and four weeks) were not related to changes in mean steps/day at six months (see later), they were not subjected to testing as part of the engagement analysis. Engagement variables showing a significant relationship with six month PA in univariable analyses ($p<0.05$) were included in a multivariable model to determine their combined effects on six month PA.

Estimated median lifetime usage (i.e. time after which 50% of participants stopped use) was calculated for use of the PA monitoring system and website use. Baseline measures of socio-demographic variables, psychosocial variables, environmental variables, and PA were investigated as predictors of non-usage attrition of the PA monitoring system to record daily activity and the website using Cox proportional hazards regression analyses. In the first analysis, the time variable was the number of days until the first two week lapse from using the PA monitoring system to record daily activity and the event variable was coded 1 if non-usage attrition occurred or 0 if non-usage attrition did not occur. In the second analysis, the time variable was the number of days until the first two week lapse from logging onto the website and the event variable was coded 1 if non-usage attrition occurred or 0 if non-usage attrition did not occur. Univariate analyses were conducted on all predictor variables and those with $p<0.25$[118] were included in a multivariable model with backwards elimination of the predictor with the highest p-value until all included predictors had $p<0.05$. All analyses included SEs and p-values corrected for clustering, and the Efron procedure was used for handling ties as it is advocated over the Breslow method and can be implemented with models adjusting SEs and p-values for clustering.[119]

2.7 DEVIATIONS OF THE EVALUATION FROM THE ORIGINAL PROTOCOL

2.7.1 Deviations and Rationale

On 31/08/2015 the research team corresponded with the NIHR Public Health Research secretariat to indicate that the initial recruitment target was not met, explaining a variety of reasons why the original target proved problematic. Significant re-structuring of a number of the public sector organisations that were to host the intervention severely impacted on participant recruitment, and the study team undertook various actions to mitigate the shortfall (e.g. by extending the number of sites; and taking advice from Patient and Public Involvement (PPI) members on our Trial Steering Committee regarding alternative recruitment methods within the host organisations).

As recruitment numbers were lower than anticipated from our initial baseline data collection phase in Lisburn (from the re-structured local authority and healthcare Trust), the intervention phase started later than scheduled, in May 2014. It was decided that recruitment would continue across other agreed sites in Belfast {Queen's University Belfast (offering a pool of approx. 3000 staff) and the Stormont Civil Service Estate (offering a pool of approx. 3000 staff)} in order to attempt to meet the recruitment target and the intervention would be implemented on a rolling basis thereafter. Therefore, randomisation and the intervention phases were implemented in July and August 2015 in the Queen's University campus and the Stormont Estate, respectively.

Recruitment continued to be slower than anticipated, even with the addition of our new sites at Queen's University Belfast and the Stormont Estate. This was due to unforeseen circumstances (austerity related re-structuring of local authorities and within the civil service) and therefore we sought approval to recruit from two final sites within the Belfast Health and Social Care Trust (Royal Victoria Hospital and Belfast City Hospital). We subsequently undertook an intense recruitment period in early September 2015. Participant recruitment was completed in January 2016 following the addition of a new worksite (total n=853).

Professor Chris Patterson (study statistician) made a revised power calculation, which took account of more recent literature on effect size estimates and the baseline data on cluster size, cluster variation and intra-class correlation co-efficient (ICC). As reported at the previous Trial Steering Committees, the revised sample size requirement demonstrated minimal impact on power (see Section 2.2.1.3 for further details).

Further, in order to maximise trial retention, all participants received a £10 gift card at the six and 12 month stage for full completion of study outcomes (funded from internal sources). The trial was originally designed to include follow-up of participants to 18 months post-intervention. However, with unavoidable delays in initial recruitment attendant upon the re-structuring of the local authorities and in light of the findings at the six month follow up, it was agreed with NIHR that follow-up should stop at 12 months rather than 18 months as originally planned.

2.8 STAKEHOLDER ENGAGEMENT

Stakeholder engagement in this evaluation has been a key element throughout this study. The purpose of such engagement has been:

1. To inform aspects of the study design
2. To raise awareness of the research in the workplace and its potential benefits
3. To be actively involved in the interpretation of the trial findings and process evaluation
4. To help identify the practical significance of the findings from the trial and implications for further delivery of the scheme or similar schemes
5. To help plan a dissemination strategy.

Such engagement has taken three forms:

2.8.1 Project Team Meetings

From the outset of the trial, the research team has attended and fully engaged in the Project Team Meetings with staff from the Health Trusts, and organisations participating in the trial. This forum has maintained a schedule of twice-yearly meetings throughout the lifespan of the trial. The aim has been to help influence the research at an early stage of development, to raise awareness of the research and support the organisation's involvement in the process.

2.8.2 Stakeholder Members of the Trial Steering Committee

Alongside the above Project Team Meetings, key stakeholders comprising staff from the Health Trusts and participating organisations also contributed directly to the evaluation as members of the Trial Steering Committee. Critically, this has included contributing to the emerging interpretation of the findings and the development of the dissemination strategy.

2.8.3 Dissemination Events

Alongside the dissemination of the key findings of this study at academic conferences, a number of non-academic dissemination activities were planned. Such activities have included the distribution of two annual study ezines to all host organisations, study participants and to the UKCRC Centre of Excellence (Northern Ireland) partners. Dr Gough (key coordinating post-doctoral research fellow for the study) attended a Northern Ireland Public Health Research Network (NI PHRN) workshop to update the network on the PAL Scheme (February 2017). The offer of a health and wellbeing workshop and PAL study debriefing was extended to the 'Champions' or key contacts at each of the participating worksites. The research team also

received interest from a number of organisations that expressed interest in adopting the study's bespoke web platform for promoting PA in the workplace and the study team are currently in talks with the Health and Safety Executive, South Eastern Health and Social Care Trust and Sport Northern Ireland regarding this. Talks have also been held with the University Research and Enterprise team regarding licensing of the web platform for such purposes. Further, plans have been made to archive the results of dissemination activities on the British Science Association "Collective Memory" database (http://collectivememory.britishscienceassociation.org/).

2.8.4 Process Evaluation

The aim of the process evaluation was to examine i) what the participants' exposure was to the intervention; ii) the extent to which the intervention was implemented across the participating organisations; iii) how, for whom and under what circumstances the intervention could bring about behaviour change; iv) how, for whom and under what circumstances the intervention might maintain behaviour change; (v) whether there were any unintended consequences of the intervention.

As described above, this has involved in-depth engagement with all of the key stakeholders to ascertain their experiences of and perspectives on the programme. The qualitative insights gained from the participants and also from the employers are demonstrated in section 4.4.4 of this report.

3 RESULTS FROM THE TRIAL

3.1 RECRUITMENT

A total of 1,209 employees expressed an interest in participating in the study and were assessed for eligibility. Reasons for exclusion included: not based at recruited worksite at least four hours/day on at least three days/week (n=93); current contract did not last the duration of the study (n=1); no longer wished to participate pre-randomisation (n=107); did not provide baseline data (n=150); unable to contact (n=5).

Eight hundred and fifty three individuals were randomized to either the Intervention (n=457) or the Control Group (n=396). The study flow diagram is presented in Figure 1.

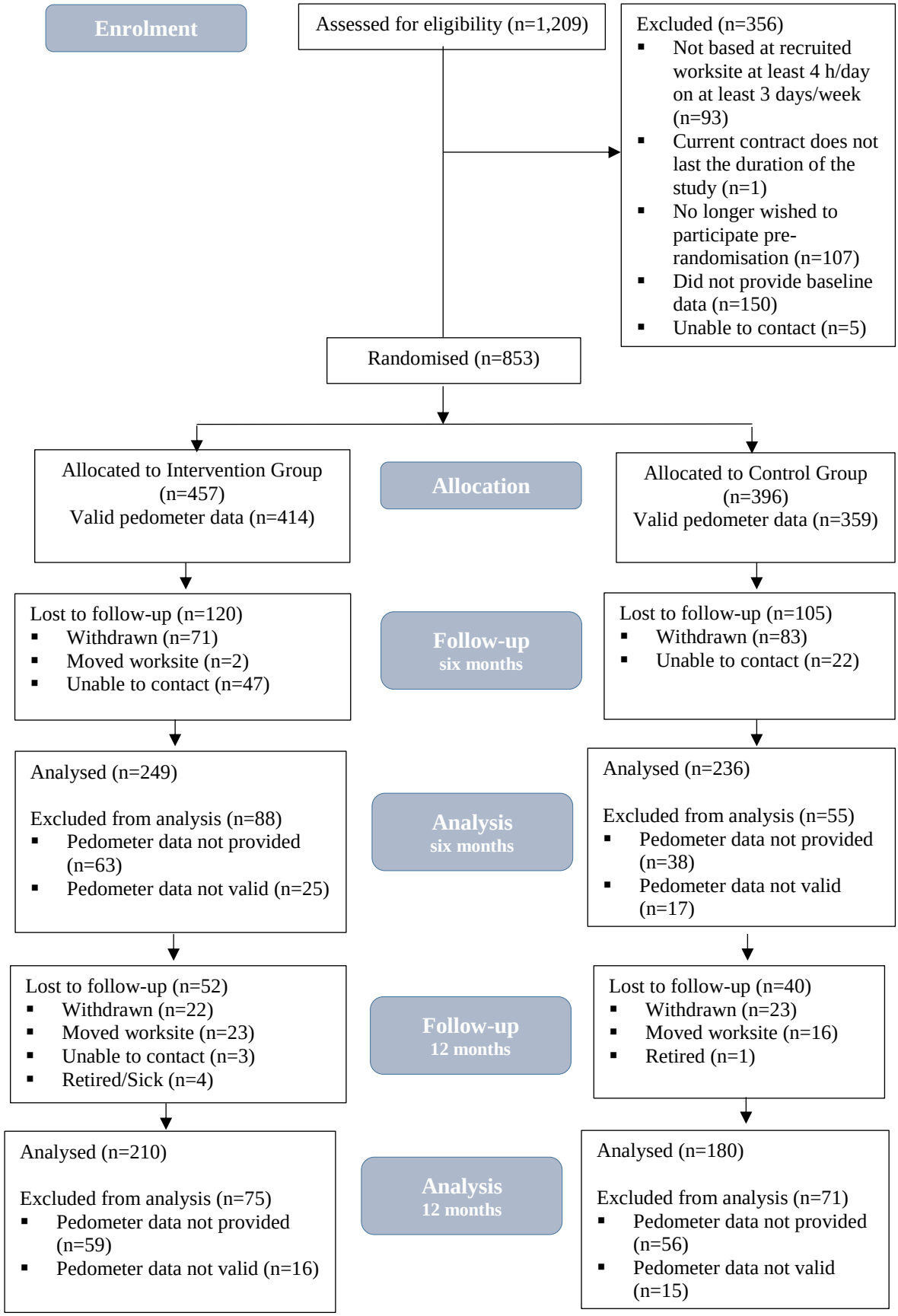

Figure 1. CONSORT Flow diagram

3.1.1 Withdrawals

Participants noted as withdrawals were those who formally, via email, telephone, or in person, expressed their wish to no longer take part in the trial. At the six month post-baseline, 154 participants (n=71 Intervention Group) had withdrawn from the study. At 12 months post-baseline, a further 45 participants had withdrawn, of which 22 participants were in the Intervention Group. Stated reasons for withdrawals included lack of time and/or interest.

3.1.2 Lost to Follow Up

Participants noted as 'lost to follow up' were those who were unable to be contacted due to moving worksite location, or who did not reply to email/phone contact by the research team. At the six months post-baseline, 71 participants were lost to follow up, 49 of whom were in the Intervention Group. Across the two groups, 69 participants were unable to be contacted and two participants had moved location. At 12 months post-baseline, a further 47 participants (n=30 Intervention Group) were lost to follow up, of which, five had retired or were off sick, 39 had moved location and a further three were unable to be contacted by the research team.

3.1.3 Harms and Adverse Effects

There were no harms or adverse effects reported by any participant throughout the conduct of the trial.

3.2 SAMPLE CHARACTERISTICS

A total of 853 participants from 37 clusters were recruited and randomised into two groups (n=457 Intervention Group, n=396 Control Group). Table 4 shows baseline characteristics of the clusters and participants stratified by group. The mean age of participants was 43.6 (standard deviation (SD) 9.6) years and 71% were female. At baseline, mean steps/day were 7,826 (SD 3,425), measured over seven consecutive days. Twenty nine percent of participants were categorised as having 'low' PA levels according to self-report GPAQ.

Table 4. Mean (SD) baseline characteristics of participants according to group

	Intervention Group	**Control Group**
Characteristics of clusters	N = 19	N = 18
Number of participants; mean (range)	24 (4 to 147)	22 (4 to 81)
Randomisation stratum, clusters (n, % participants)		
Small (<20 employees)	11 (114, 25%)	11 (105, 27%)
Medium (20-50 employees)	5 (167, 37%)	4 (123, 31%)
Large (>50 employees)	1 (147, 32%)	2 (144, 36%)
Schools	2 (29, 6%)	1 (24, 6%)
Characteristics of participants	n=457	n=396
Age (years)	44.0 (9.3)	43.0 (10.0)
Female gender; n (%)	329 (72%)	278 (70%)
BMI (kg/m^2)	27.2 (5.6)	26.6 (5.3)
Income >£20k; n (%)	341 (75%)	291 (73%)
Education some higher level; n (%)	295 (65%)	270 (68%)
Marital status married/co-habiting; n (%)	313 (68%)	274 (69%)
Objective PA: pedometer steps (steps/day)	7,977 (3,602)	7,650 (3,204)
Objective: physical activity category, n (%)		
High (>7,500 steps/day)	204 (45%)	167 (42%)
Moderate (>2,500-≤7,500 steps/day)	199 (44%)	184 (46%)
Low (<2,500 steps/day)	11 (2%)	8 (2%)
GPAQ: minutes of work PA (minutes/week)	42 (138)	58 (151)
GPAQ: minutes of MVPA (minutes/week)	296 (342)	344 (333)
GPAQ: physical activity category, n (%)		
High	70 (15%)	76 (19%)
Moderate	140 (31%)	130 (33%)
Low	141 (31%)	104 (26%)
SF-8: Mental Component Score	48.0 (8.9)	47.7 (9.3)
SF-8: Physical Component Score	52.5 (6.6)	52.7 (7.0)
EQ-5D: Health State	82.4 (13.8)	83.8 (14.3)
EQ-5D: Weighted Health Index	0.89 (0.11)	0.89 (0.12)
WEMWBS: Mental wellbeing scale	50.2 (8.2)	50.3 (8.9)
Physical Activity Self-Efficacy scale	2.91 (0.97)	2.92 (0.94)
HPQ: Four week absolute absenteeism	5.04 (41.3)	3.48 (50.0)
HPQ: Absolute presenteeism	80.3 (13.6)	81.0 (13.4)
HPQ: Combined relative absenteeism and absolute presenteeism	8.84 (12.57)	8.56 (7.32)

EQ-5D: EuroQol five dimensions; GPAQ: Global Physical Activity Questionnaire; HPQ: Health and Work Performance Questionnaire; MVPA: moderate- to vigorous-intensity physical activity; NHS: National Health Service; PA: physical activity; SD: standard deviation; SF: Short Form; WEMWBS: Warwick-Edinburgh Mental Wellbeing Scale.

3.3 FINDINGS AT INTERVENTION END (SIX MONTHS POST-BASELINE)

3.3.1 Primary Outcomes

At six month post-baseline, mean steps/day were 6,990 (SD 3,078) for the Intervention Group and 7,576 (SD 3,345) for the control group. There was a significant difference between the Intervention and Control Groups in mean steps/day (*b*=-336, 95% CI: -612 to -60, p=0.02) after adjustment for baseline values, randomisation stratum and season, and correction for cluster effects (Table 5). Possible reasons for this result are explored in the Sensitivity Analysis section 3.4.1.1 below.

At six months post-baseline, there was a significant difference between the Intervention Group compared to the Control Group for minutes/week of self-reported work-related PA (*b*=-33.3, 95% CI: -65.44 to -1.24, p=0.04). There was no significant difference between groups for minutes/week self-report MVPA (*b*=4.12, 95% CI: -47.07 to 55.31, p=0.88).

Table 5. Mean (SD) month six outcomes according to group and ANCOVA results before and after adjusting for season

Outcome	Intervention Group[a]		Control Group[a]		Analysis of covariance[b]		Analysis of covariance[c]	
	n	Mean (SD)	n	Mean (SD)	b (95% CI)	P-value	b (95% CI)	P-value
Primary outcome								
Objective PA: pedometer steps (steps/day)	249	6,990 (3,078)	236	7,576 (3,345)	-519 (-931, -107)	0.01	-336 (-612, -60)	0.02
Secondary outcomes								
GPAQ: minutes of work PA (minutes/week)	253	16.32 (57.55)	235	43.72 (129.33)	-33.67 (-65.32, -2.02)	0.04	-33.34 (-65.44, -1.24)	0.04
GPAQ: minutes of MVPA (minutes/week)	231	291.84 (255.76)	221	350.14 (329.41)	-2.75 (-63.36, 57.85)	0.93	4.12 (-47.07, 55.31)	0.88
SF-8: Mental Component Score	269	48.43 (8.90)	244	47.12 (9.44)	1.07 (-0.27, 2.42)	0.12	1.17 (-0.23, 2.56)	0.10
SF-8: Physical Component Score	269	50.87 (8.37)	244	50.99 (7.64)	0.43 (-1.26, 2.12)	0.62	0.64 (-0.79, 2.08)	0.38
EQ-5D: Health State	262	77.95 (16.16)	239	78.26 (16.08)	0.67 (-2.35, 3.70)	0.66	0.70 (-2.45, 3.86)	0.66
EQ-5D: Weighted Health Index	262	0.83 (0.14)	239	0.84 (0.15)	0.005 (-0.02, 0.03)	0.71	0.01 (-0.01, 0.04)	0.37
WEMWBS: Mental wellbeing scale	266	50.34 (8.38)	243	49.42 (8.24)	1.15 (0.18, 2.12)	0.02	1.34 (0.48, 2.20)	<0.01
HPQ: Four week absolute absenteeism	247	4.04 (66.11)	227	7.01 (51.40)	-2.25 (-12.49, 8.00)	0.67	-2.59 (-12.79, 7.62)	0.62
HPQ: Absolute presenteeism	261	78.62 (14.58)	236	78.47 (14.21)	0.82 (-1.79, 3.44)	0.54	1.48 (-0.43, 3.38)	0.13
HPQ: Combined relative absenteeism and absolute presenteeism	246	9.52 (24.63)	226	7.91 (4.94)	-0.03 (-0.91, 0.85)	0.95	-0.02 (-0.92, 0.87)	0.96

[a]Month six outcomes (unadjusted)

[b]ANCOVA comparison of six month means in Intervention vs Control Group adjusted for baseline values of the outcome, and randomisation stratum and corrected for clustering

[c]ANCOVA comparison of six month means in Intervention vs Control Group adjusted for baseline values of the outcome, randomisation stratum and season and corrected for clustering

ANCOVA: Analysis of Covariance; CI: confidence interval; EQ-5D: EuroQol five dimensions; GPAQ: Global Physical Activity Questionnaire; MVPA: moderate- to vigorous-intensity physical activity; NHS: National Health Service, PA: physical activity; SD: standard deviation; SE: standard error; SF: Short Form; WEMWBS: Warwick-Edinburgh Mental Wellbeing Scale.

3.3.1.1 Sensitivity Analyses

At six months post-baseline, mean steps/day in the imputed dataset were 6,914 (SD 3,142, n=457) for the Intervention Group and 7,408 (SD 3,253, n=369) for the Control Group. There was a significant difference in the Intervention Group compared to the Control Group (*b*=-526, 95% CI: -948 to -104, p=0.02) after adjustment for baseline values, randomisation stratum and season, and correction for cluster effects (Table 6).

Since the intervention has marginally but significantly lowered "steps per day" (in the Intervention Group) we abductively explored possible reasons for this. In one recruitment site, there were some periodic technical glitches affecting the sensors in the environment (which recorded participants' PA). These glitches were routinely detected by the network and the sensor was replaced usually within 24-48 hours. However some Intervention Group participants may have been temporarily demotivated when they knew that their activity was not being monitored for these short periods. Thus the main analysis was repeated excluding the affected clusters (n=11). The magnitude of the estimated difference in mean steps/day between Intervention and Control Groups was reduced and the difference was non-significant (*b*=-172, 95% CI: -472 to 128, p=0.26) after adjustment for baseline values, randomisation stratum and season, and correction for cluster effects.

Table 6. Mean (SD) outcomes at months six and 12 according to group and ANCOVA results before and after adjusting for season, with imputation of missing values on the six or 12 month outcomes

Outcome	Intervention Group[a]		Control Group[a]		Analysis of covariance[b]		Analysis of covariance[c]	
	N	Mean (SD)	N	Mean (SD)	b (95% CI)	P-value	b (95% CI)	P-value
Six month objective PA: pedometer steps (steps/day)	457	6,914 (3,142)	369	7,408 (3,253)	-574 (-1,109, -41)	0.04	-526 (-948, -104)	0.02
12 month objective PA: pedometer steps (steps/day)	457	7,522 (3,489)	369	7,739 (3,340)	-543 (-1,207, 120)	0.11	-558 (-1,203, 87)	0.09

[a]Month six or month 12 outcomes (unadjusted)

[b]ANCOVA comparison of six month/12 month means in Intervention vs Control Group adjusted for baseline values of the outcome, and randomisation stratum and corrected for clustering

[c]ANCOVA comparison of six month/12 month means in Intervention vs Control Group adjusted for baseline values of the outcome, randomisation stratum and season and corrected for clustering.

ANCOVA, Analysis of covariance; CI, confidence interval; EQ5D, EuroQol, five dimensions; GPAQ, Global Physical Activity Questionnaire; MVPA, moderate- to vigorous-intensity physical activity; NHS, National Health Service, PA, physical activity; SD, standard deviation; SE, standard error; SF, short form; WEMWBS, Warwick-Edinburgh Mental Wellbeing Scale.

3.3.2 Exploratory Subgroup Analyses

Moderation analyses in Table 7 showed that perceptions of the availability of PA opportunities in the Workplace Environment (WE: Availability) moderated the relationship between group assignment and six month mean steps/day, as indicated by a significant coefficient for the interaction term (*b*=-280; 95% CI: -538, -23; p=0.03). This indicated that the negative effect of assignment to the Intervention Group on six month mean steps/day was greater at higher levels of workplace environment availability (WE: Availability). Thus, there was a greater decrease in mean steps/day (baseline-six months) for Intervention Group participants compared to Control Group participants when perceptions of the availability of PA opportunities in the workplace environment were higher. It should be noted that owing to issues with multiple testing, this result should be interpreted with caution. The results of all other moderation analyses were non-significant.

Table 7. Results of moderation analyses (random-effects regressions of mean steps/day at six months or 12 months on baseline mean steps/day, group assignment moderator and moderator-by-group interaction)

	Six month mean steps/day			12 month mean steps/day		
	Moderator X Group interaction			Moderator X Group interaction		
MODERATOR	**n**	***b* (SE)**	**P-value**	**n**	***b* (SE)**	**P-value**
Age	454	10.7 (21.0)	0.61	366	-10.5 (31.8)	0.74
Gender (male)[a]	456	-310.0 (330.6)	0.35	368	709.1 (684.7)	0.30
BMI	442	56.4 (39.1)	0.15	358	61.3 (46.5)	0.19
Income (>£20k)[a]	449	-92.7 (591.8)	0.88	363	-25.7 (748.5)	0.97
Education (some higher level)[a]	449	374.1 (384.7)	0.33	363	272.7 (534.2)	0.61
Marital status (married/co-habiting)[a]	448	-181.6 (454.4)	0.69	362	-183.0 (758.2)	0.81
SF-8 Mental Component Score	449	6.48 (29.0)	0.82	363	-1.61 (44.3)	0.97
SF-8 Physical Component Score	449	9.12 (26.5)	0.73	363	-49.9 (38.5)	0.20
EQ5D: Health State	426	-5.17 (14.7)	0.72	347	9.95 (22.0)	0.65
EQ5D: Weighted Health Index	427	1,444 (2,045)	0.48	347	-1,352 (2,829)	0.63
WEMWBS: Mental wellbeing scale	449	-7.88 (32.0)	0.81	363	30.3 (42.6)	0.48
WE: Attractiveness	448	-73.8 (51.7)	0.15	363	-50.5 (115.9)	0.66
WE: Safety	448	-76.3 (85.6)	0.37	363	294.2 (122.2)	0.02
WE: Accessibility	448	-173.1 (102.0)	0.09	363	57.8 (176.5)	0.74
WE: Availability	448	-280.3 (131.3)	0.03	363	-49.3 (194.1)	0.80
WTA30	405	-2.91 (39.7)	0.94	329	25.0 (52.5)	0.63
WTA60	401	-5.48 (47.8)	0.91	325	-7.44 (57.4)	0.90
Time discount2 (2 vs 1)[a]	419	162.7 (384.0)	0.67	340	889.9 (670.2)	0.18
Time discount3 (3 vs 1)[a]	419	-630.0 (385.1)	0.10	340	-1,216 (515.5)	0.02

[a]Dichotomous moderators

NB. Results are adjusted for strata, season and baseline mean steps/day with cluster-adjusted standard errors and p-values.

3.4 FINDINGS FROM SIX MONTHS FOLLOW-UP (12 MONTHS POST-BASELINE)

3.4.1 Primary Outcomes at Six Months Follow-Up (12 Months Post-Baseline)

At 12 months, mean steps/day were 7,790 (SD 3,462) for the Intervention Group and 8,203 (SD 3,401) for the Control Group. There was a non-significant difference between the Intervention and Control groups in steps/day (*b*=-570, 95% CI: -1,267 to 127, p=0.11) after adjustment for baseline values, randomisation stratum and season, and correction for cluster effects (Table 8).

At 12 months, there were no significant differences between the Intervention and Control Groups for minutes/week of self-reported work-related PA (*b*=7.00, 95% CI: -12.55 to 26.56, p=0.48) and minutes/week self-report MVPA (*b*=77.04, 95% CI: -7.94 to 162.0, p=0.08) (Table 8).

Table 8. Mean (SD) 12 month outcomes according to group and ANCOVA results before and after adjusting for season

Outcome	Intervention Group[a]		Control Group[a]		Analysis of covariance[b]		Analysis of covariance[c]	
	N	Mean (SD)	N	Mean (SD)	b (95% CI)	P-value	b (95% CI)	P-value
Primary outcome								
Objective PA: pedometer steps (steps/day)	210	7,790 (3,462)	180	8,203 (3,401)	-561 (-1,243, 120)	0.11	-570 (-1,267, 127)	0.11
Secondary outcomes								
GPAQ: minutes of work PA (minutes/week)	207	43.2 (146.6)	182	44.7 (121.7)	1.77 (-22.3, 25.8)	0.89	7.00 (-12.6, 26.6)	0.48
GPAQ: minutes of MVPA (minutes/week)	188	387.4 (385.9)	169	354.4 (314.7)	75.2 (-8.29, 158.7)	0.08	77.0 (-7.94, 162.0)	0.08

[a]Month 12 outcomes (unadjusted)

[b]ANCOVA comparison of 12 month means in Intervention vs Control Group adjusted for baseline values of the outcome, and randomisation stratum and corrected for clustering.

[c]ANCOVA comparison of 12 month means in Intervention vs Control Group adjusted for baseline values of the outcome, , randomisation stratum and season and corrected for clustering.

ANCOVA, Analysis of covariance; CI, confidence interval; GPAQ, Global Physical Activity Questionnaire; MVPA, moderate- to vigorous-intensity physical activity; PA, physical activity; SD, standard deviation; SE, standard error.

3.4.1.1 Sensitivity Analyses

At 12 months, mean steps/day in the imputed dataset were 7,522 (SD 3,489, n=457 for the Intervention Group and 7,739 (SD 3,340, n=369) for the Control Group. There was no significant difference in the Intervention Group compared to the Control Group (*b*=-558, 95% CI: -1,203 to 87, p=0.09) after adjustment for baseline values, randomisation stratum and season, and correction for cluster effects (Table 6).

3.4.2 Exploratory Subgroup Analyses

Moderation analyses (see Table 7) showed that perceptions of the safety of the Workplace Environment (WE: Safety) moderated the relationship between group assignment and 12 month mean steps/day, as indicated by a significant coefficient for the interaction term (*b*=294.2; 95% CI: 55, 534; p=0.02). This indicated that the negative effect of assignment to the Intervention Group on 12 month mean steps per fay was greater at lower levels of workplace environment safety (WE: Safety). Thus, there was a greater decrease in mean steps/day (baseline-12 months) for Intervention Group participants compared to Control Group participants when perceptions of the safety of the workplace environment were lower. Moderation analyses also showed that the "Time discount3" variable (high discount rate versus low discount rate) moderated the relationship between group assignment and 12 month mean steps/day, as indicated by a significant coefficient for the interaction term (*b*=-1,216; 95% CI: -2,226, -205; p=0.02). This indicated that the negative effect of assignment to the Intervention Group on 12 month mean steps/day was greater when discount rates were higher. Thus, there was a greater decrease in mean steps/day (baseline-12 months) for Intervention Group participants compared to Control Group participants when participants' discounts rates were higher (i.e. they prefer an immediate reward with a lower monetary value rather than having to wait to get a reward with a slightly higher monetary value tomorrow). Owing to issues with multiple testing, these results should be interpreted with caution. The results of all other moderation analyses were non-significant.

3.5 SECONDARY OUTCOMES

There was a significant difference between groups for the WEMWBS (*b*=1.34, 95% CI: 0.48 to 2.20, p<0.01) with wellbeing scores higher in the Intervention Group participants. There were no significant differences between groups for SF-8 mental (*b*=1.17, 95% CI: -0.23 to 2.56,

p=0.10) and physical (*b*=0.64, 95% CI: -0.79 to 2.08, p=0.38) component scores. There were no significant differences between groups for the EQ-5D-5 L health state (*b*=0.70, 95% CI: -2.45 to 3.86, p=0.66) and weighted health index scores (*b*=0.01, 95% CI: -0.01, 0.04, p=0.37). There was no significant difference between groups in four week absolute work absenteeism (*b*=-2.59, 95% CI: -12.8 to 7.62, p=0.62), absolute presenteeism (*b*=1.48, 95% CI: -0.43 to 3.38, p=0.13) and combined relative absenteeism and absolute presenteeism (*b*=-0.02, 95% CI: -0.92 to 0.87, p=0.96) (Table 5).

4 RESULTS FROM THE PROCESS EVALUATION

4.1 INTRODUCTION

The intervention produced a small but significant decline from baseline to six month follow-up in mean steps/day (objectively measured) in the Intervention Group compared with the Control Group. The results highlight a need to explore and understand the processes underlying these findings.

4.2 IMPLEMENTATION

4.2.1 Participation and Reach

9,031 participants were invited to participate across four main sites (Lisburn, Stormont Estate, Queen's University Belfast, Belfast City Hospital), as shown in Figure 1. In total, 853 participants were recruited and randomised. The demographic characteristics of the study population are shown in Table 4.

4.3 INTERVENTION FIDELITY AND 'DOSE'

Standardised training manuals were provided to research staff to guide the intervention delivery and study conduct (including collection of outcome data and all email communication with participants). Participants had access to an online discussion forum, where queries or issues in relation to the day-to-day running of the intervention were voiced. Commonly reported issues, consistent across study sites, included issues with the sensors not accurately recording time spent doing PA across the work day, discrepancies between times recorded for other participants undertaking the same activity, the location of the sensors and the type of rewards available to participants.

On average, 1,000 minutes (SD 987.08) of PA was recorded per participant via the PAL Scheme PA monitoring system across the intervention period (n=422 participants). Physical activity was captured by the PAL monitoring system on 43.41 (SD 33.21) days over the six month intervention, equating to PA being recorded on approximately one third (32.71%, SD 25.05) of intervention days. With regards to intervention 'dose', participants walked for at least 10 minutes on 32.80 (SD 28.87) days across the intervention period (Table 9).

Participants in the Intervention Group self-reported spending 2.31 (SD 2.12) hours/day on the internet. There was no significant difference in reported internet use (hours/day) between the Intervention and Control Groups. Participants reported high confidence in using the internet at baseline (Intervention Group 8.67, SD 1.74), ranked from 1-10 on Likert scale.

Across the intervention period, on average participants earned 932.37 (SD 825.33) points by tracking their PA using the PA monitoring system. In addition, participants' earned, on average, an additional 75.21 (SD 106.30) points from the scheme's double point days. Participants redeemed 39.33% (SD 42.51) of points accumulated throughout the intervention for vouchers.

Table 9. Engagement with PA monitoring system and study website

Engagement	N	Mean (SD)
Physical activity monitoring system		
Number of days participants recorded PA for at least 10 minutes (week 4)[a]	422	5.99 (5.09)
Number of days participants recorded PA for at least 10 minutes (month 6) [a]	422	32.80 (28.87)
% of days participants recorded PA for at least 10 minutes (week 4) [a]	422	29.95 (28.87)
% of days participants recorded PA for at least 10 minutes (month 6) [a]	422	24.71 (21.76)
Study website		
Total minutes	418	1170.68 (2048.27)
% of days participants logged onto website (month 6)	418	17.39 (20.06)
% of weeks participants logged onto website (month 6)	418	37.75 (32.45)
Frequency of hits: Feedback and monitoring [b]	418	13.7 (3.54)
Frequency of hits: Rewards [b]	418	5.7 (4.47)
Frequency of hits: Maps [b]	418	3.4 (3.97)
Frequency of hits: Health information (PA) [b]	418	0.5 (1.72)
Frequency of hits: Health information (other) [b]	418	1.2 (3.17)
Frequency of hits: Discussion Forum [b]	418	1.9 (4.23)
Frequency of hits: Study information [b]	418	4.0 (5.23)
% of total website time spent on Feedback and Monitoring	418	44.66 (31.16)
% of total website time spent on Rewards	418	17.45 (23.78)
% of total website time spent on Maps	418	15.32 (24.07)
% of total website time spent on Health Information (PA)	418	1.36 (7.34)
% of total website time spent on Health information (other)	418	3.04 (10.18)
% of total website time spent on Discussion Forum	418	3.14 (10.21)
% of total website time spent on Study Information	418	10.25 (18.58)

[a] Data captured using the PA monitoring system

[b] Total number of hits for every ten days the participant accessed the website

PA: Physical activity.

4.4 RESPONSIVENESS

The aim of the thematic analyses was to tease out the main themes and sub-themes which were considered to be important factors for engagement/disengagement with the intervention and maintenance of PA. The themes pertaining to their overall perceptions of the intervention were very much positive, including with respect to mental health and wellbeing, social aspects and physical fitness and health. In terms of the key behaviour change and intervention components, themes included the issue of rewards by means of a financial incentive, but also disincentives linked to 'voucher type'; technology as a means of tracking progress, support and feedback, but also the negative aspect relating to technological issues with sensors and key fobs (see Section 4.4.2 for further details).

4.4.1 Participants

In total, nine focus groups were conducted at the end of the intervention period across the study sites: Lisburn (n=3 FGs; n=21 participants), Stormont Estate (n=2 FGs; n=12 participants), Queen's University Belfast (n=2 FGs; n=16 participants) and Belfast City Hospital (n=2 FGs; n=16 participants). Sixty six participants took part in the focus group discussions (n=52/78% females; mean age 43.8 (SD 9.0) years). At 12 months, 17 participants who had taken part in the six month focus groups also took part in the 12 month discussions (n=9/53% female; mean age 46.1 (SD 9.3)) years.

The following coding was applied to quotes to distinguish between the different focus groups and semi-structured interviews undertaken: FG denotes the focus group session (1-9) at six months, FG12 denotes the focus group session (12.1-12.6) at 12 months, MR: Male respondent, FR: Female respondent, R: Retailer, SM: Senior manager.

4.4.2 General Perceptions of the Intervention

Overall feedback from participants about the PAL Scheme was positive. Many participants cited how the scheme had helped them engage in more PA:

"FG3, FR: It's been a very positive thing to get people out walking and to encourage them to walk more, with all the wee perks and to do a bit extra every day" and to develop habits that could be maintained once the intervention had ended:

"FG8, FR: So it's been a brilliant success to get me walking and it's now part of my daily routine."

4.4.2.1 Type of PA

The type of PA upon which the intervention was based was also viewed as positive by participants, with many citing how convenient walking and/or jogging were compared with other activities:

"FG3, FR: …it doesn't suit everybody to go to a class, to be there at a certain time… that's the good thing about walking and running, that you can do it at any time it suits…";

"FG9, FR: It wasn't hard, that's what I quite liked about it. You could do as much or as little as you wanted, and that was quite nice."

4.4.2.2 Format of the PAL Scheme

Many participants felt the format of the scheme, targeted at changing PA across the working day, was restrictive. Participants highlighted that the timing of the intervention did not enable them to account for how active they had been across the whole day:

"FG6, FR: What I found with it was that I was really doing the most of the stuff. I was a bit aggrieved because I was doing a lot in the morning which wasn't recorded on the fob, you didn't record those things."

Some participants also found the intervention limiting in that it didn't take account of other types of PA:

"FG3, FR: I think it's very restrictive in just having walking or running, because we do have very good staff membership for the gym, which a lot of the staff do avail of but you didn't get any recognition for it…".

4.4.2.3 Increasing PA and Health Benefits

A number of key themes emerged when participants discussed their reasons for signing up the PAL Scheme. Increasing PA across the work day was cited by many participants as their main reason for signing up:

"FG1. FR: I felt I needed more exercise, I knew that it was beneficial, I thought it was a good idea and would help me do more";

"FG7, FR: Mine was to get up from my desk and get out to walk more, that was my reason, more self-motivation…".

In addition to being more active, the associated health benefits of this behaviour change were also noted by participants, with many commenting that they signed up to lose weight or improve their health:
"FG6, FR: For me, it was health and fitness and good habit forming, plus you could see it was just better for you to get away from the desk, mentally."

Some participants felt that they were already active during the working day, and the scheme was worthwhile to gain something for their efforts:
"FG9, FR: I did walk a lot anyway, at lunchtime and before work and stuff and I sort of thought I might as well get something for it. I was doing it anyway…", while others noted helping out with research was a factor that encouraged them to sign up:
"FG1, MR: I do a lot of cycling and running and record it … I just figured that there's a bit of data that maybe someone else could use, they may as well have my data."

The convenience of an intervention being delivered within the workplace was also an influencing factor in encouraging people to take part:
"FG8, FR: I think generally if you can incentivise anything that's within work hours, I know personally for me, with working with a young family, if I can get that activity within those hours from 8 to 6 then it will really work for me…".

4.4.2.4 Key Barriers and Facilitators to the PAL Scheme

Key barriers and facilitators to the PAL Scheme and being active throughout the working day were identified through the focus group discussions. Work demands, time and the weather were cited by participants as barriers to being more active at work and participating in the PAL Scheme. The structure of the working day meant it was not always possible for participants to be active during lunchtime:
"FG6, MR: I find, especially in this place, you don't always get the chance to get out at lunchtime anyway. Once I started, if I didn't get out at lunchtime I would be more inclined to make sure that I did get out in the evening for a walk."

In addition, increased workload had an impact on the time participants had available to be active:
"FG9, FR: But as term got busier and we got into September it just got more and more difficult to find the time to fit it all in."

Participants highlighted how the weather was also a barrier, deterring them taking part in the scheme on occasions:
"FG6, MR: We walked in a group, so I think if it rained the girls got their hair wet, so that sort of stopped them, really. If it rained we didn't walk basically."

With regards to facilitating increasing PA through the PAL Scheme across the working day, a number of key themes emerged. Access to nice areas for PA was noted as a facilitator:
"FG6, FR: We're very fortunate in this area, we've this lovely green park and you can walk out there and walk back."

Support within the workplace from colleagues was another facilitator of PA, with participants commenting on the importance of sharing experiences of being involved in the scheme:
"FG8, FR: Even if you didn't want to walk in a group, just chatting about it "How much did you do...?" I think that works for a lot of people."

In addition to support from colleagues, the attitude of employers was noted as another important influence:
"FG3, FR: I think that this council is very good as an employer, as they do put on the likes of the classes at lunchtime, you have access to the gym and everything."

4.4.2.5 General Benefits of the PAL Scheme

A number of themes emerged from the focus group discussions on the benefits participants received from their participation in the PAL Scheme. Benefits of the intervention included increased levels of PA, health benefits, social benefits and increased productivity. Participants highlighted how the scheme had instigated changes in their usual routine, leading to increased PA across the day:
"FG4, MR: If I got the train in I would actually make my route longer to make sure that I passed a few points in the way in and on the way home, rather than just straight here."
A number of health benefits were listed by participants, including the impact being more active had on their physical wellbeing:
"FG9, FR: I wanted to lose weight. I've lost 2 stone since I started, but I've done lots of other activity. But it spurred me on and gave me the motivation, and actually I built up my fitness so that I could do some of the other activities."

In addition to the positive impact the intervention had on physical health, participants also noted the benefits in relation to mental health and wellbeing:
"FG8, FR: It's quite nice too, because even the mindfulness idea, as you're walking along on your own you're more mindful of what you're doing, and you're into health issues there too."
With regards to how PA behaviour change was maintained post-intervention, there were mixed responses from focus group participants. A number of participants noted how the PAL Scheme had changed their routine from doing no PA across the working day to now being more active:
"FG4, FR: I would walk more now; I wouldn't have thought of doing it at all."

4.4.2.5.1 Incorporating PA into Habits and Daily Life

In addition, being involved in the intervention encouraged participants to make small changes to their daily habits to facilitate PA within the working day: *"FG9, FR: ... I do at least as much now, even after it's finished, as I did when I was doing it, and in fact I probably do more now and I'm probably much more conscious of what I do because of even trying to think up how I was going to make those points in a day. I keep walking, I still walk up and down from the school every day, and I do it at least five times a week, if I can, to build up 150 minutes."*

Other participants commented on how being involved in the intervention had made them think about their PA, and that they were still making efforts to increase their activity as a result of being involved in the PAL Scheme:
"FG6, FR: I haven't continued as much but I am really determined to, because I am aware of the fact when I'm not walking and that your health is not quite as good, so I have made more of a conscious effort recently to try and get back into it, especially now that the days are a little bit longer and it's a bit brighter outside. I kept up other exercise, other than walking as well, quite consistently, even more so than I would have done previously."

For some, the extra PA that they were doing during the PAL Scheme stopped once the concept of the points and rewards was removed:
"FG7, FR: I did it when the vouchers were there but now I wouldn't...I would go outside but I wouldn't actually make an effort now. So I would never dream of going for a walk or anything during work."

4.4.2.5.2 Social Aspects of the PAL Scheme

Participants commented on the social aspects of being involved in the PAL Scheme, with many citing the intervention as a good way of socialising with colleagues:

"FG6, FR: There was a social element to it too, like someone would say 'oh, do you want to go outside and get some points?'"

Participating in the Intervention Group identified a number of benefits specific to the workplace. Within the focus group discussions, participants noted that the intervention provided them with the opportunity to have a proper break during the work day:

"FG3, FR: Yeah, usually you have your lunch here, sometimes at your desk, so you aren't getting away from it, so it's good to get outside and get fresh air and you're more awake in the afternoons then because you have been out."

Furthermore, participants felt the opportunity to be active during the day had a positive effect on their productivity:

"FG4, MR: I probably worked harder because I'd normally be sleeping after lunch."

As well as increasing levels of PA across the work day, participants highlighted the additional effect being involved in the PAL Scheme had on other types of PA they were involved in:

"FG7, FR: It triggered off for me to go and join a club, a rowing club, because you weren't really aware that you weren't doing as much and you were sitting quite a bit at the desk, and you thought 'right okay' …".

In addition, the PAL Scheme increased awareness of other improvements individuals could be making to their overall health and wellbeing:

"FG8, MR: Like others, it stimulated me to look at other parts of my lifestyle and try to amend them, to make small changes as well."

4.4.3 Perceptions of Intervention Components and Behaviour Change Techniques

The key themes in relation to participants' thoughts on the different intervention components of the PAL Scheme for behaviour change are summarised in Table 10. These include the use of incentives, sensors, maps, feedback, fobs, email and social support.

With regard to the sensors, participants reported some issues which affected the day-to-day running of the intervention, including issues with the accuracy, placement and restrictiveness of the sensors (Table 10). Participants had mixed opinions on a number of website components, including the maps and email prompts. The feedback element of the website was a positive feature for many, with participants highlighting that the visual feedback often prompted them to be more active.

A number of themes emerged in relation to rewards (retail vouchers) offered to participants in exchange for points they had accumulated during their participation in the PAL Scheme. Some participants felt the vouchers did not provide a sufficient reward for the effort and activity undertaken:
"FG5, FR: You'll get incentives joining gyms anyway, so there wasn't really any incentive in the incentives, really, unless it was a free cup of tea. Walking ten miles for a cup of tea. It needed to be more; the incentives just weren't enough for me" whilst other participants felt the vouchers were limited: *"FG3, FR: The vouchers were an issue, when you go to get them there wasn't the ones you wanted, such as Argento. Just there was only so many available."*

Participants noted that the issues around the format and availability of the vouchers had an impact on their interest in other elements of the intervention:
"FG5, FR: And I found then that I lost interest because the points weren't really going anywhere. And then whatever points I had, the incentive wasn't good enough for the points I had."

Table 10. Results from the process evaluation regarding participant perceptions of the intervention components

Intervention components	Quote
Sensors	"FG2, MR: I do quite a bit of walking and I was just outside of the area and I kind of felt I was changing my walk just so I could walk past the sensors, whereas it would be handy if it wasn't quite so dependent on the sensors." "FG5, MR: Yeah, they were definitely occasions, I think everybody experienced it at some stage. It hadn't recorded parts of your journey and the sensors hadn't been working. There was other things as well. If you went so far and you didn't come across a second sensor, it was only calculating your journey from… between two sensors rather than the total time of the journey." "FG7, FR: You sort of had to plan, which sometimes I did plan, I thought then...I know you were doing it for your research but I thought this is just me planning my route round this way to do it, as opposed to this would be my normal walk."
Maps	"FG7, FR: I thought the wee maps were good. I looked at them, because you could think 'oh, I could go out now but I have only half an hour.' So you could have a look and see where you could walk…" "FG8, FR: I found they weren't that easy to read…"
Feedback	"FG5, FR: It was good to go in and see how far you had walked, how long you had been out walking." "FG6, FR: but, for me, whenever I was in the habit of it, it was seeing the circle be green, if I got to Friday and there was a chunk left, I would think 'oh, I'd better get out and try to make it to the full amount'."

Fob	"FG8, FR: But there was a number of times I kept forgetting the key fob. I had my keys on it and then they keys were on the desk and you would walk out without it, so if the sensors were the clip on belt variety." "FG2, FR: I quite like the idea of a fob that you maybe wore, because I don't carry my phone around... something connected to the phone, that could in some ways be more cumbersome, or something where the two would be synchronised…"
Email prompts	"FG9, FR: The weekly emails helped as well, just to spur you on a Monday morning, just to remind you that you were involved and you had to participate." "FG6, MR: I got so many emails, I got annoyed with them, the ones that were 'You should be out walking'."
Support	"FG2, MR: … I found they were always really quick at getting back to me, even if I rang or I emailed, it was really good." "FG9, FR: … answered the emails really quickly and were really responsive and that helped keep the motivation going as well, so we do appreciate what you did."

FG: Focus group; FR: Female respondent; MR: Male respondent.

4.4.3.1 Focus Groups at 12 Month Post-Baseline

The main aim of the focus groups conducted at 12 months was to explore behaviour change maintenance. On the whole, participants referred to their maintained or at least attempt to maintain PA levels and desire to keep up with that which they had achieved during the six month intervention;

"FG12.4, MR: I know I started well but I think I wasn't consistent the whole way through because of other factors that were happening at work at the time, but I tried to maintain a level, I was trying to sort of hit the target a week.. after the scheme ended I did try and keep it going".

The detailed walking routes provided during the scheme were also still utilised after the intervention period ended;

"FG12.4, MR:.. so, I would do that route more or less exactly everyday whenever I do go out", "FG12.3, FM: ..., and in the mornings I'm still doing a lot of the same routes".

The notion of building PA into their day was a key aspect to maintaining levels of PA;

"FG12.1 FR: Yeah, and I'll tell you another thing, while even waiting on the train on the platform I'll be walking up and down the platform, rather than just stand there, I'll walk.";

As was making conscious efforts to move more; *"FG12.1, FR: Well now, if I'm cooking, I would tend to do that step, from one side to the other or back and forth, so as I'm not just standing still".*

This was also noted with regard to the idea of looking for opportunities to be active beyond the six month intervention period;

"FG12.4, MR It's something I can work into my lifestyle, and even when you don't have it you're sort of thinking well is there an opportunity to perhaps go for a walk at lunch time to just keep that activity going"; and other commented on looking for additional opportunities in their day to be active;

"FG12.4, MR: .. I can adapt my pattern and I spotted two other opportunities, the beginning of the working day and the end of the working day, where I can do something that fit's in".

Habit, be that continuing with the established walking routes or continuing with their lunchtime walk or step goals, alongside seeking out opportunities to be active, were common threads throughout the focus groups for maintained PA behaviour change after the scheme had ended.

4.4.4 Senior Managers of Participating Employees

In total, semi-structured interviews were also undertaken with senior managers of six participating employers (n=7 managers) and participating retailers (n=4). Feedback about the intervention was positive from employers, with themes emerging from the interviews with regard to employee benefits; both by means of getting active, (a focus on physical and mental health) and also relating to workplace productivity. Senior manager perceptions of the intervention were aligned with participants regarding time as a key barrier to PA in the workplace.

4.4.4.1 Senior Managers Overall Perceptions of the PAL Scheme

Employers' overall thoughts on the PAL Scheme were positive, with managers highlighting the benefits of getting their employees more active:

"SM2: We know that walking is better than being sedentary and it reduces lots of different diseases and things. So in principle it's very good."

"SM2: It's the health benefits… that people are more active. You know yourself, it reduces risks of lots of different types of diseases…"

"SM4: …it's the mental health and people are enjoying the brighter evenings and the sunshine" for their employees.

Wider benefits of the intervention were also noted for the employer in relation to productivity:

"SM4: A lot of them commented recently on feeling more productive in the afternoon when they've had a chance to get some fresh air and a chance to stretch their legs. We've seen a great stretch in productivity."

In addition, senior managers commented on how the intervention aligned with other initiatives within the workplace:

"SM3: I think it's very valuable to add some evidence to the research that's already been done, and obviously it fits very nicely with us in terms of workplace health…" .

In addition, employers commented on the benefits of the intervention in terms of making contacts with other businesses and organisations: *"SM6: … suppose in terms of the business side of it too, in terms of the partnerships with the business side, in terms of incentives, that's of value too."*

Senior managers highlighted a number of barriers to promoting PA within the workplace. Time and resources were commonly cited as barriers to being active within the working day: "*SM1: Time. I think that's the reality, is we are all increasingly… work pressure is probably the most significant barrier, certainly in my opinion and from my own personal experience…".*

Employers noted the need for a change of attitudes at management level to facilitate such schemes: "*SM1: … question comes from their line managers: 'Where are the staff going to get the time for this?' And our job is really to convince them of the bigger picture…".*

4.4.5 Retailers

Semi-structured interviews were undertaken with four participating retailers. A number of key themes emerged with regards to why retailers became involved with the PAL Scheme, including benefits for their business and the broader benefits of the scheme for participants. Of the participating retailers interviewed, all cited the potential benefits being involved in the PAL Scheme could have for their business.

4.4.5.1 Potential of the PAL Scheme to Improve Business

The PAL Scheme had the potential to increase use and sales within participating businesses:
"R1: …somebody that hadn't been into the shop before, they've got a voucher so they might be encouraged to come in, they might like what they see when they come in…".

It also provided businesses with an opportunity to promote their services, and entice participants into their businesses over other similar businesses surrounding the workplace:
"R1: The opportunity to go down Botanic Avenue and go into those cafes and restaurants is always there. So this was an opportunity for me to promote the cafes and restaurants on the site."

4.4.5.2 Potential Benefits of the PAL Scheme for Participants

Participating retailers also acknowledge the benefits such a scheme could have on the health and wellbeing of participants, and cited this as another factor behind their willingness to participate in the scheme:
"R2: …doing more exercise is kind of what everybody needs to be doing but people don't do it, and encouraging them in some way, I think is definitely a good idea.".

4.4.5.3 Incentives and Rewards

A number of themes emerged in relation to the incentive-based element of the scheme, including the (i) concept of incentive-based PA; (ii) the type of rewards offered; and, (iii) value of the rewards. All participating retailers felt incentivising participants to increase their PA was a good idea, both from a behaviour change perspective:

"R2: …if I knew there was something at the end of it that you're going to or you're working towards getting money off or working towards a different voucher then it makes it seem more worthwhile." and also for the added benefit it could have for businesses:

"R3: We like it from the idea of the retail point of view because they don't have to pay anything and they're getting free promotion, which everybody should appreciate…".

The type of rewards offered varied across the participating businesses, with different reward structures being perceived to have a different influence from the retailers' points of view. Some retailers offered a cash value voucher as their reward, with a view that participants would spend more when cashing in their voucher:

"R4: So a £10 voucher, there's a fair chance if they come in they will upscale in terms of what they're going to buy. So you'll hopefully get it in the long-term." There was a general consensus among retailers that the rewards offered to participants were not 'overly generous'.

Retailers highlighted a number of issues that may influence upscaling the present intervention or taking part in similar schemes in the future. All retailers spoke about the sustainability of the incentive-based approach, and highlighted that they would take part in future schemes if they were getting a good return from a business perspective:

"R4: … future for me going, is down the voucher route. So yeah, it's sustainable providing people come and use it."

Retailers also cited requests from other organisations and schemes as a potential barrier, with most retailers constantly asked to provide rewards and sponsorship:

"R4: But the down side is we as a shop again are always asked for vouchers and sponsorship and you have to pick what you can go for."

Questions regarding the impact and financial gains from participating in the scheme were cited as another key theme in relation to future participation. All retailers highlighted the need for

feedback from participants and tracking of how the rewards were used, which was not monitored within the present study by the researchers or the businesses themselves. Retailers noted that they would make efforts to track voucher use and spending in future. In addition, retailers highlighted that tracking use could also provide them with an opportunity to gain feedback from new customers: *"R1: … I would just like some feedback from the participants of the scheme in terms of their experience… on a purely customer focused level, did they have a positive experience when they used the cafes…".*

5 RESULTS FROM THE HEALTH ECONOMICS ANALYSIS

5.1 OVERVIEW

The CEA of the PAL Scheme is presented in this chapter. Overall, the intervention induced a small (but significant) decline in mean steps/day by the Intervention Group compared to the Control Group. Thus, the economic evaluation takes the form of a within trial CUA *and* CBA, comparing the Intervention Group with the Waiting-List Control Group. In line with NICE methods for the development of NICE public health guidance the CUA adopted a public sector perspective.[120] Costs included intervention costs (apportioned per participant) and health-care resource use. Health outcomes were expressed in terms of QALYs accrued over the six month trial period in the CUA. Results are presented using an ICER estimated by dividing the adjusted difference in mean costs between groups by the adjusted difference in mean QALYs between groups. ICER estimates were compared with a £20,000 - £30,000 per QALY threshold applied by NICE.[120] The CBA was undertaken from an employer's perspective by employing a 'net-cost model'[88] by incorporating not only the intervention cost but also the avoided costs of absenteeism and productivity loss due to sick days. It is a technique recognised in the NICE guideline for economic evaluations for public health interventions to allow broader benefit of the intervention to be considered in decision-making.[78] All analyses were undertaken according to the principle of intention-to-treat and in STATA/SE 12.0 (StataCorp, College Station, TX, USA).[121]

5.1.1 Intervention Cost

The intervention costs are detailed in Table 11. Intervention costs included the retail vouchers, sensors and keyfobs, electrical engineering staff and travel for maintaining the sensors, and a computer scientist for monitoring and managing the data from the sensors and maintaining the

website. Cost for placement of sensors on the lampposts was not included as it is refundable by local government. Vouchers cost £21,000 in total. These voucher costs could be greatly reduced if there was more meaningful engagement with local authority / city management groups to facilitate better deals with the retailers. Thus the costs of vouchers were varied in sensitivity analysis. The life span of the sensors was estimated to be two years as the screw bosses were made of plastic, which limits the number of battery replacements. With upgraded screw bosses, the life span can be prolonged to over five years. The battery used in the key fobs cannot be replaced so the life span for the current key fobs in the trial is the life of the battery, i.e., nine months. This short life span will be extended to five years with a battery replaceable design. Therefore, in sensitivity analyses, the costs for sensors with upgraded screw bosses and battery replaceable key fobs is considered, to explore the influence of upgraded technology on the cost-effectiveness of the intervention. The cost per participant for each disaggregable item is shown in Figure 2, from which the retail voucher and sensor together accounted for over 90% of the total cost.

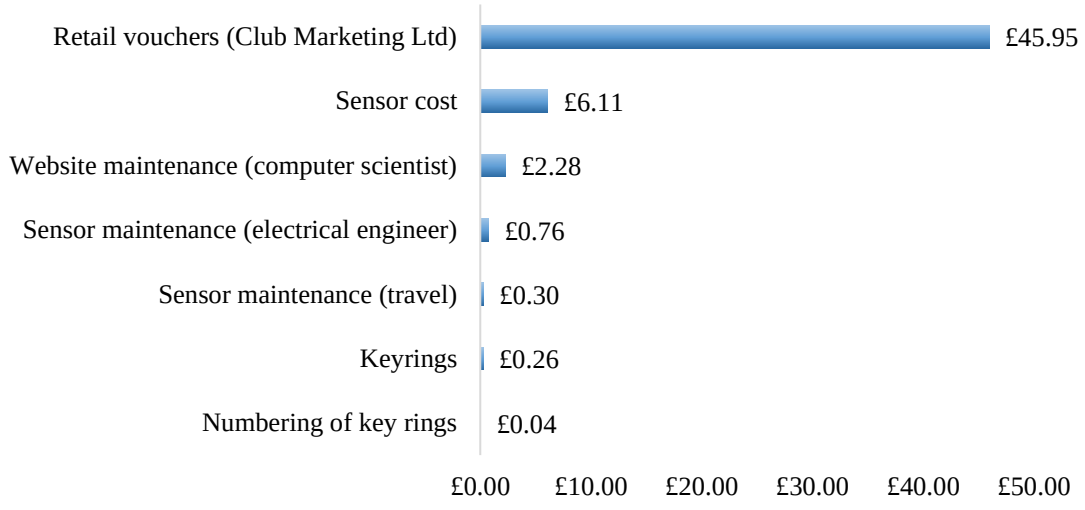

Figure 2. Cost of PAL scheme over six months (per participant)

Table 11. Intervention costs

Cost Items	Unit cost	Quantity	Total cost for six months
Retail vouchers (Club Marketing Ltd)	/	/	£21,000
Equipment			
Keyrings	£0.35	503 (457 participants *(1+10%))[a]	£117.37[b]
Numbering of key rings	£0.04	503 (457 participants *(1+10%))[a]	£16.91
Sensor cost	£151.55	70	£2,792.16[c]
Staff time and travel			
Sensor maintenance (electrical engineer)	£13.33 /hour	26 hours (1 hr/week * 26 weeks)	£346.58
Sensor maintenance (travel)	£0.40/ mile	338 miles (13 miles/week *26 weeks)	£135.20
Website maintenance (computer scientist)	£13.33 /hour	78 hours (3 hr/week * 26 weeks)	£1,039.74
Total (six month)			**£25,447.96**
Cost per participant in the Intervention Group (n=457)			**£55.68**

Note: a. Only for the participants in the Intervention Group and account for a 10% lost/broken/ replacements.

b. Nine months life span was estimated for key rings, and therefor cost for six months was calculated as a proportion of the nine months total cost.

c. Two years life span for sensor was estimated for the sensors. The cost was annuitized at 3.5% discount rate to estimate the cost of one-year use of the sensors. Formular: E=K/[(1-(1+r)-n)/r]. K=£151.55*70=10,608.50, r=3.5%, n=2. E=5584.32. To estimate the six months cost, half of the first year cost was taken due to no discount within each year.

5.1.2 Missing Data

The number and percentage of missing data for each collected NHS resource use item and EQ-5D-5L questions at baseline and six months' assessment point are shown in Table 12 for Intervention and Control Groups respectively. Baseline for the resource use refers to the six month period before randomisation completed by participants retrospectively at baseline. All of the missing data for the EQ-5D-5L responses were due to whole questionnaires missing, which can be seen from the uniform proportion of missing data for the five dimensions. A similar proportion of missing was observed across the randomised groups and for both cost and EQ-5D-5L data. Approximately 15% of data were missing at baseline and 40% of data were missing at the six month follow-up.

Table 12. Missing data for the resource use and EQ-5D-5L dimensions in the Intervention and Control Group

Data items collected form the trial	Intervention (n=457)				Control (n=396)			
	Baseline		Six month		Baseline		Six month	
	No. of missing	%	No. of missing	%	No. of missing	%	No. of missing	%
NHS resource use								
GP	73	15.97	196	42.89	63	15.91	161	40.66
Nurse	73	15.97	206	45.08	62	15.66	165	41.67
Physiotherapist	73	15.97	206	45.08	63	15.91	164	41.41
AE	73	15.97	207	45.30	63	15.91	166	41.92
Outpatient	73	15.97	205	44.86	63	15.91	166	41.92
Nights in Hospital	73	15.97	207	45.30	62	15.66	170	42.93
EQ-5D-5L dimensions								
Mobility	67	14.66	195	42.67	57	14.39	157	39.65
Self-care	67	14.66	195	42.67	57	14.39	157	39.65
Usual activities	67	14.66	195	42.67	57	14.39	157	39.65
Pain and discomfort	67	14.66	195	42.67	57	14.39	157	39.65
Anxiety and depression	67	14.66	195	42.67	57	14.39	157	39.65

A&E: Accident and Emergency; EQ-5D-5L: Euroqol 5 dimension, 5 level; GP: General practitioner; NHS: National Health Service.

5.1.3 Resource Use and Costs

Table 13 reports the use of each resource item (mean number of visits, SD, median, min and max) accrued over the six month primary follow-up period in the Intervention and the Control Group. Overall, there was no statistically significant differences in the use of each resource item between the groups.

Table 13. NHS and social care resource use per participant over six months (complete case)

Resource use (No. of visits)	Intervention (N=457)						Control (N=396)						p value
	N	Mean	SD	Median	Min	Max	N	Mean	SD	Median	Min	Max	
GP	261	1.21	1.36	1	0	7	235	1.24	1.53	1	0	10	0.78
Nurse	251	0.71	2.08	0	0	20	231	0.54	1.25	0	0	10	0.28
Physiotherapist	251	0.51	1.49	0	0	8	232	0.76	2.77	0	0	23	0.20
A&E	250	0.10	0.38	0	0	3	230	0.07	0.29	0	0	2	0.48
Outpatient	252	0.43	0.99	0	0	8	230	0.57	1.28	0	0	8	0.17
Inpatient	251	0.04	0.28	0	0	3	229	0.04	0.22	0	0	2	0.98
Nights in Hospital	250	0.16	1.86	0	0	28	226	0.13	0.86	0	0	7	0.86

A&E: Accident and Emergency; GP: General Practitioner; NHS: National Health Service; SD: standard deviation.

Table 14 shows other services used by the participants in each group. Other services and medications are not included in the total resource use due to the similar amount of service use between groups, their randomness (i.e., they are most likely unrelated to the intervention), and the difficulty matching unit cost with the free-text inconsistent reporting.

Table 14. Other services used by the participants in the trial

Intervention (n= 457)		Control (n=396)	
Services	**Freq. (times)**	**Services**	**Freq. (times)**
Services used in both groups			
Dentist	77	Dentist	86
Counselling	22	Counselling	11
CBT	1	Counselling/CBT	6
Chiropractor	1	Chiropractor	2
Diabetic clinic	1	Diabetic clinic	2
Breast screening	1	Breast screening	1
Day inpatient	3	Day inpatient	1
Optician	7	Optician	1
Orthopaedic referral	1	Orthopaedic surgery	1
Private physiotherapist	1	Private physiotherapist	1
Other services			
MRI scanning	4	Pain clinic	5
Haematology	3	Podiatry	4
Kinesiology	3	Asthma clinic	1
Blood donation	2	Cardiac rehab	1
Mammogram	2	Hearing test	1
Sports massage	2	Knee consultant	1
Bone scan	1	Operation on broken wrist	1
Dermatology	1	Private orthopaedic surgeon	1
Dietician at hospital	1	Surgeon	1
Flu jab	1	Treatment	1
Medical day case unit	1	Work's doctor	1
Therapist	1		
X-ray	1		
Total number of services	138		130

CBT: Cognitive Behavioural Therapy; MRI: Medical Resonance Imaging.

Table 15 compares the cost of service use between the groups with four regression approaches as specified in method section 2.4.1.1. There is no statistically significant difference in the resource use between groups. Despite this, the decrease of resource use in the Intervention Group compared to the Control Group was larger (approximately £60) with the gamma distribution adjusting for covariates than the OLS (approximately £15) unadjusted model, in both complete case, and multiply imputed data.

Table 15. Total cost of health care resources per participant over the six month follow-up

	Mean	**SE[d]**	**95%CI lower[d]**	**95%CI upper[d]**	**P-value**
Complete case[a], unadjusted					
Intervention	200.00	49.65	99.31	300.69	0.829
Control	212.41	28.21	155.21	269.62	
Complete case[a], adjusted[b]					
Intervention	193.25	32.42	129.71	256.80	0.326
Control	249.89	44.92	161.84	337.94	
MI[c], unadjusted					
Intervention	200.50	41.39	114.99	286.01	0.755
Control	216.47	26.72	161.27	271.67	
MI[c], adjusted[b]					
Intervention	190.14	31.21	128.67	251.60	0.235
Control	247.70	39.49	170.16	325.24	

CI: Confidence interval; EQ-5D-5L: Euroqol, 5 dimension, 5 levels; MI: Multiple imputation; SE: Standard error; SF-8: Short Form 8;

a. Sample size for complete case analysis: unadjusted: 499; adjusted: 429.
b. Adjusted for covariates: Strata, season, age, sex, baseline cost, baseline EQ-5D-5L utility value, mean steps at baseline, SF-8Physical, SF-8Mental
c. MI – multiple imputation with chained equations
d. All standard errors and 95% CI estimates were adjusted for clusters.

5.1.4 Health-related Quality of Life (HrQOL)

The EQ-5D-5L health utility values and visual analogue scale (VAS) for each group at baseline and six months are shown in Table 16 (complete case) and Table 17 (after imputation). The completeness at six month follow-up was 57% (262/457) for Intervention Group and 52.3% (239/457) for Control Group. This was lower than the rate of completeness at baseline, which was 85.3% (390/457) and 74.2% (339/457) for Intervention and Control Group, respectively. 45.0% (328/729) of participants recorded a state of 'full health' at baseline and 29.1% (146/501) at six months. The utility values for both groups declined over the six months from approximately 0.89 to 0.83. There was no statistical significant difference between groups at both baseline and six months.

Table 16. EQ-5D-5L VAS and index value (complete case)

	n	mean	SD	min	max	P value (unadjusted)*
Baseline index score						
Intervention	390	0.887	0.113	0.316	1	0.736
Control	339	0.890	0.120	0.320	1	
Six month index score						
Intervention	262	0.828	0.143	0.310	1	0.650
Control	239	0.835	0.150	0.238	1	
Difference between baseline and six month						
Intervention	242	-0.053	0.125	-0.586	0.259	0.401
Control	219	-0.064	0.117	-0.453	0.346	
Baseline VAS						
Intervention	389	82.42	13.83	38	100	0.317
Control	339	83.76	14.27	0	100	
Six month VAS						
Intervention	262	77.95	16.16	19	100	0.827
Control	239	78.26	16.08	19	100	

EQ-5D-5L: Euroqol 5 dimension, 5 level; SD: standard deviation; VAS: Visual Analogue Scale.

*adjusted for cluster.

Table 17. EQ-5D-5L index value (after imputation)

	Mean	SE*	95% lower	95% upper	P-value (unadjusted)*
Baseline index score					
Intervention	0.887	0.006	0.875	0.899	0.661
Control	0.891	0.008	0.876	0.906	
Six month index score					
Intervention	0.829	0.010	0.809	0.850	0.886
Control	0.827	0.011	0.805	0.849	

EQ-5D-5L: Euroqol 5 dimension, 5 level; SE: standard error

*cluster adjusted.

5.1.5 Cost-Utility Base-Case Analysis

The cost-effectiveness results for the intervention are presented in Table 18. Overall, the intervention was approximately £25 more costly but had no effect on QALYs. The average cost per participant was £253.5 (95% CI £188.4, 318.6) in the Intervention Group and £227.64 (95% CI £170.9, 284.4) in the Control Group. Mean QALYs accrued over the six months trial period were 0.4157 (95% CI 0.4077, 0.4238) for the Intervention Group and 0.4158 (95% CI 0.4057, 0.42600) for the Control Group, leading to a 0.0000891 (95% CI -0.008, 0.008) lower QALY gain in the Intervention Group compared to the Control Group.

Table 18. Cost-effectiveness results (after imputation) with six month follow-up

Treatment group	Cost (£)[a]			QALY[b]		
	Mean	**SE**	**95% CI**	**Mean**	**SE**	**95% CI**
Intervention (n=457)	253.49	33.06	188.41, 318.57	0.4157	0.0041	0.4077 , 0.4238
Control (n= 396)	227.64	28.90	170.86, 284.43	0.4158	0.0052	0.4057, 0.4260
Difference (95% CI)	25.85	28.44	-29.89, 81.60	-0.0000891	0.004	-0.008, 0.008
ICER	-£ 290,178 per QALY					
95% CI for ICER (from bootstrap)	-£480,011.8 to -£100,336.1 per QALY					

CI: Confidence interval; ICER: Incremental cost-effectiveness ratio; QALY: Quality Adjusted Life Year; SE: Standard error

a. adjusted cost, per-participant intervention cost included in the cost of intervention group.
b. Adjusted QALY.

The cost-effectiveness plane for the base-case analysis is shown in Figure 3. The dyads come from the 1000 bootstrap iterations. The X-axis represents the bootstrapped incremental QALYs between the randomisation groups and the y-axis represents the incremental costs. The bootstrapped pairs of cost difference and QALY difference were across the four quadrants, indicating greater uncertainty around the estimate of ICER. The bootstrapped pairs were almost evenly (47% vs. 53%) distributed across the Y axis in the first and second quadrant, caused by the close to zero mean incremental QALY. The existence of dyads in the third and fourth quadrants indicates that there was a small possibility (=13.4%) that the intervention is cost saving. The proportion of the simulated pairs distributed in each quadrant is presented in Table 19.

The cost-effectiveness acceptability curve for the base-case analysis is shown in Figure 4. The probability that the intervention was cost-effective at the £30,000 threshold was 34.6%. This corresponds to the proportion of dyads on the right side of the sand colour line in the first, third and fourth quadrant of the cost-effectiveness plane in Figure 3.

Table 19. Distribution of the bootstrapped pairs of cost difference and QALY difference in the four quadrants

Quadrant	Description	Probability
1st, Northeast	Intervention is more costly and more effective	38.6%
2nd, Northwest	Intervention is more costly and less effective	48.0%
3rd, Southwest	Intervention is less costly and less effective	5.0%
4th, Southeast	Intervention is less costly and more effective	8.4%

QALY: Quality Adjusted Life Year.

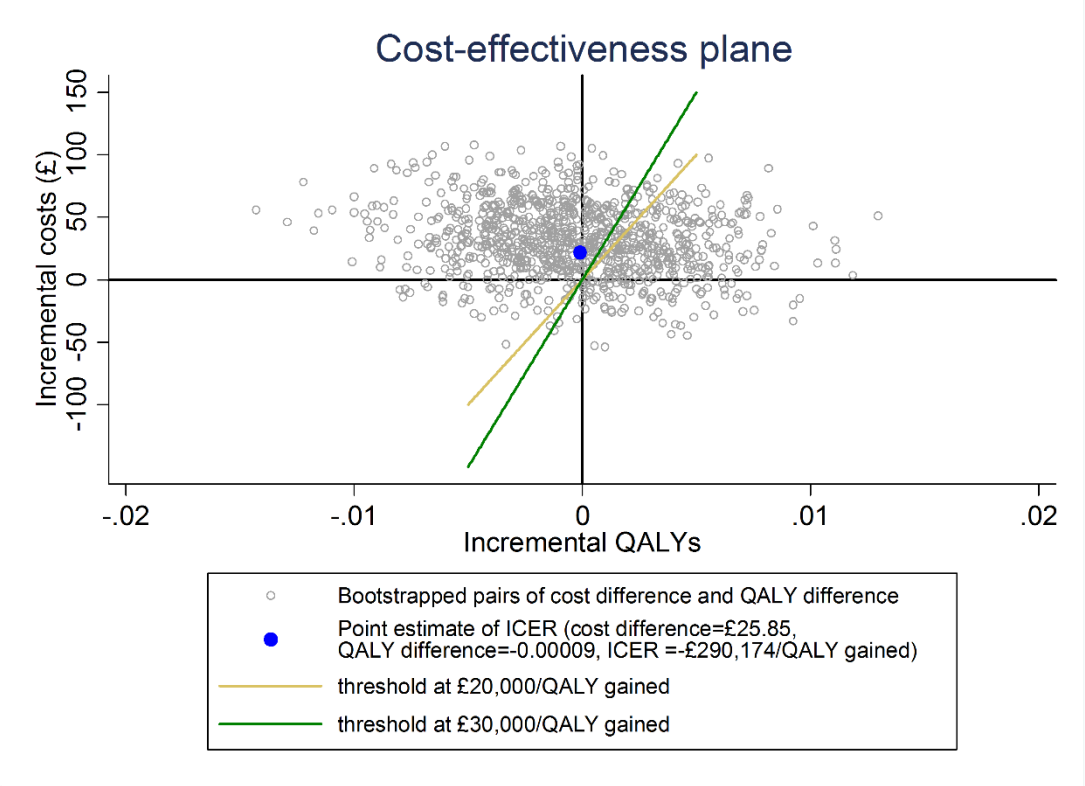

Figure 3. Cost-effectiveness plane representing 1000 bootstrapped cost difference and QALY difference pairs

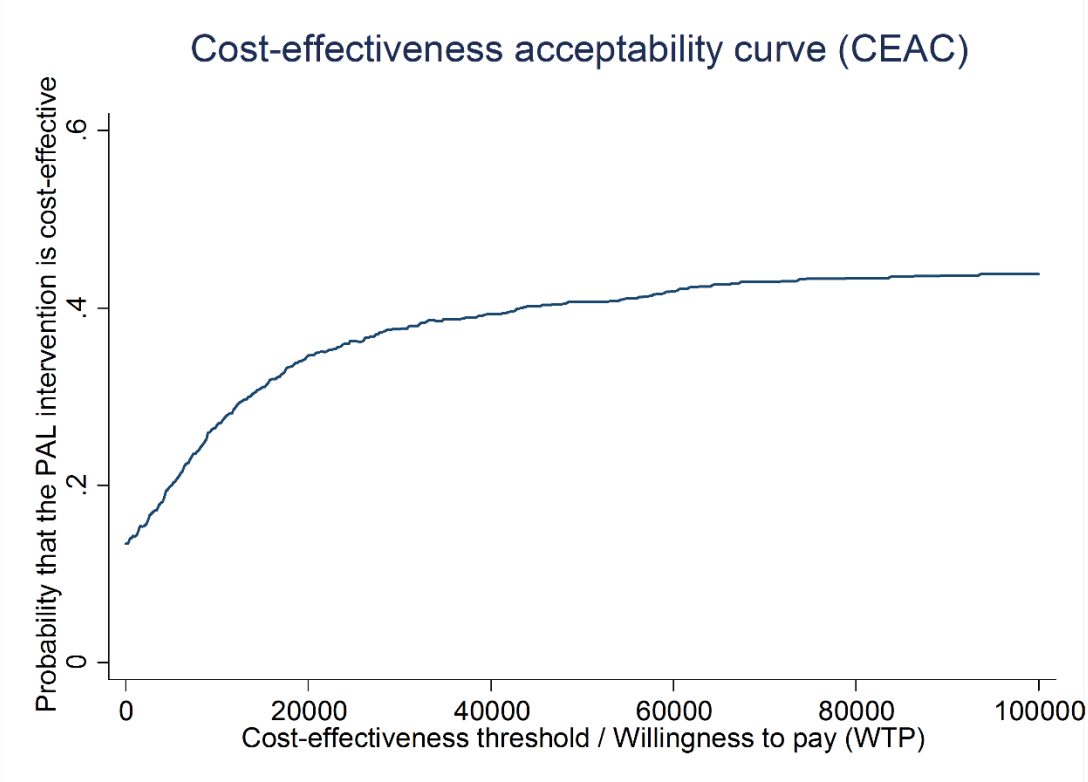

Figure 4. Cost-effectiveness acceptability curve (CEAC)

5.1.6 Sensitivity Analysis

Results for the sensitivity analysis are presented in Table 20.

Table 20. Summary results of the sensitivity analysis of health economic evaluation

Analysis	Incremental cost (intervention – control) (95%CI) (£)	Incremental QALYs (intervention – control) (95%CI)	ICER (£/QALY gained)	Probability of being C/E at threshold of £20,000/QALY gained (%)
Base-case analysis				
Base-case (six month)	25.85 (-29.89, 81.60)	-0.0000891 (-0.008, 0.008)	-290,173.9	34.6
Voucher business mode				
1. Club Marketing Ltd	- 32.71 (-89.84 24.43)	-0.0000891 (-0.008, 0.008)	367,102	60.8
2. £10 shops for all	-5.34 (-61.21, 50.53)	-0.0000891 (-0.008, 0.008)	599,13.87	48
3. Marketing consultant	-13.19 (-69.27, 42.89)	-0.0000891 (-0.008, 0.008)	148,040.2	52.1
Least expensive scenario				
Upgraded devices, club marketing	-42.61 (-100.41, 15.19)	-0.0000891 (-0.008, 0.008)	47,828.1	69.3
Missing data				
Complete case analysis (n=427), gamma family, log link	36.24 (-60.52, 133.00)	-0.00047 (-0.013, 0.012)	-76,978.1 (-1,923,213, 1,769,257)	35.7

ICER: Incremental cost-effectiveness ratio; QALY: Quality Adjusted Life Year.

5.1.7 CBA Employing Net-Cost Model

It was reported earlier in Table 5 that the absolute absenteeism hours over a four week period for Intervention and Control Group was 4.04 (66.11) and 7.01 (51.40) respectively (p=0.62). The difference was estimated to be 2.59 hours over a four week period after adjusting for stratum, season and cluster (p=0.62). This equates to 15.54 hours pro-rata for a six month time period. After attaching the hourly salary values, the avoided cost of absenteeism and the net cost of intervention for each representative salary grade is shown in Table 21. This estimated cost saving for the employers ranged from £66 to £735 depending on the wage rate employed.

Table 21. Estimation of avoided cost of absenteeism and net cost of intervention using the 'net cost model'

Grade	Avoided cost of absenteeism (£)	Net cost of intervention (Intervention cost minus the avoided cost of absenteeism)
Lowest (NHS Band 1, £7.8/hr)	£121.21	-£65.53
Mid (NHS Band 8A, £22.86/hr	£355.24	-£299.56
Highest (NHS Band 9, £50.85/hr)	£790.21	-£734.53

Hr: Hour; NHS: National Health Service.

Given the difference in absolute absenteeism was not statistically significant, Figure 5 presents the uncertainty of the potential economic benefit from the employer's perspective for each representative salary grade. At current intervention cost (=£55.68), the probability that the PAL intervention is cost-saving from employer's perspective is 64% for the high salary group, 62% for the middle salary group, and 57% for the low salary group.

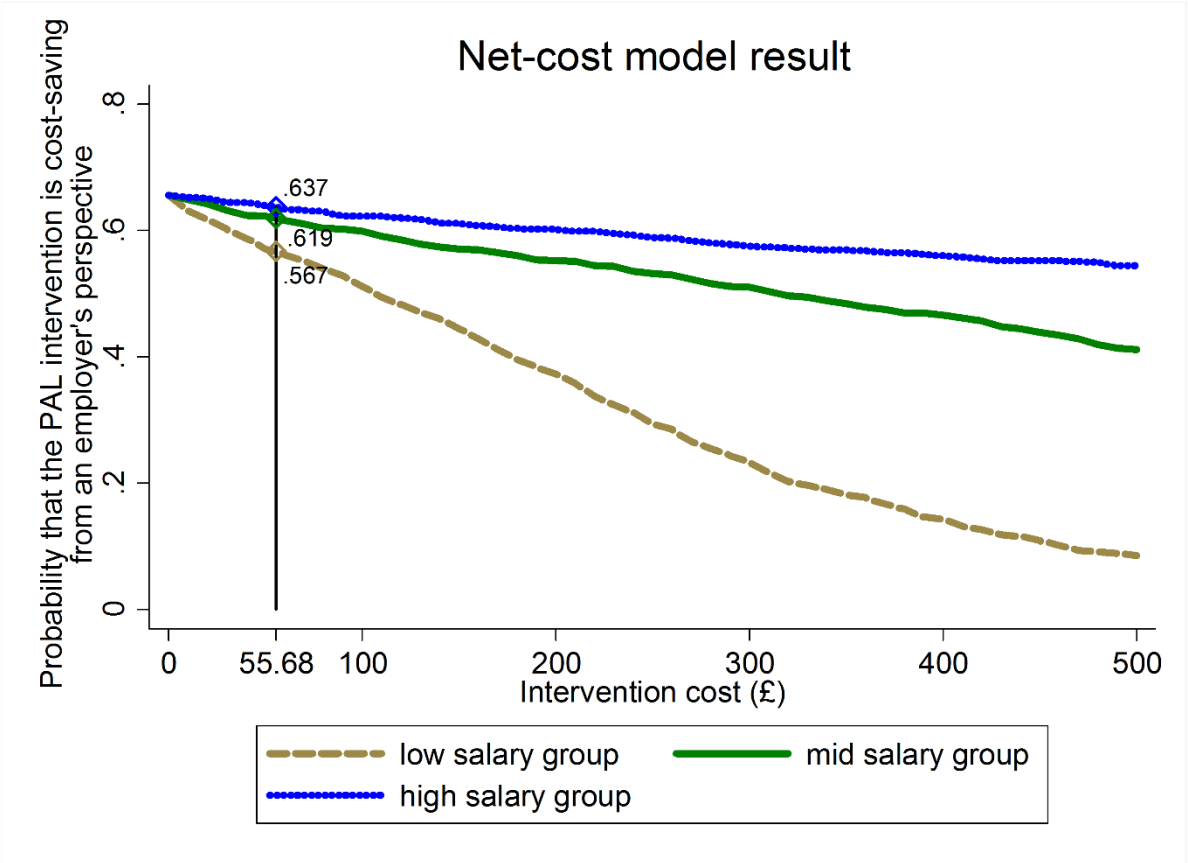

Figure 5. Uncertainty surrounding the potential economic benefit from the employer's perspective

6 RESULTS FROM THE BEHAVIOURAL ECONOMICS ANALYSIS

6.1 RESULTS FROM THE BEHAVIOURAL ECONOMIC ANALYSES

Contingent Valuation was used to elicit the minimum financial incentives required by participants to increase general PA. Next, discrete choice modelling was utilised to examine the monetary value required for increasing different types of PA, i.e. walking or cycling to and from places, moderate-intensity, and vigorous-intensity recreational PA.

Two double-hurdle models were then estimated, with the first examining which individual socio-demographic characteristics were associated with the decision to increase PA at six month follow-up compared to baseline, and analysis of the extent to which PA increased; the second investigated the influence of behavioural economic traits and other variables on the decision to increase PA at 12 month follow-up (compared to six month) and the amount by which PA increased.

6.1.1 Contingent Valuation

On average, participants' WTA financial incentives was £1.38 (95% CI: £1.16, £1.61) for increasing general PA for 30 minutes and was £2.80/week (95% CI: £2.32, £3.27) for 60 minutes/week, respectively.

6.1.2 Discrete choice modelling

In general, the average amounts required by participants (WTA) for increasing walking or cycling to and from places, moderate-intensity and vigorous-intensity recreational PA were £2.88/hour [95% CI: £2.33; £3.43], £1.02/hour [95% CI: £0.68; £1.37], and £3.29/hour [95% CI: £2.72; £3.86], respectively. In particular, the minimum monetary incentives necessary for increasing PA differed significantly according to whether a participant was inactive or active at baseline. Participants from the inactive group at baseline required a mean of £5.45/hour (95% CI: £4.39, £6.51) to walk or cycle (e.g. active travel behaviour), in contrast to their counterparts from the moderate-active group (£3.31/hour; [95% CI: £2.32, £4.29]) and active group (£3.19/hour; [95% CI: £1.69, £4.68]). Similarly, inactive participants required £3.24 [95% CI: £2.30, £4.17] and £4.79 [95% CI: £3.25, £6.33] to undertake an additional 60 minutes of

moderate and vigorous PA, respectively, values which were significantly higher than those required for the moderate-active group (£0.92 [95% CI: £0.15, £1.68] and £3.51 [95% CI: £2.24, £4.78], respectively) and active group (£1.82 [95% CI: £0.73, £2.90] and £1.84 [95% CI: £-0.04, £3.72], respectively).

6.1.3 Double-hurdle models

The estimated double-hurdle model (n=372 from the Intervention Group) shown in Table 22 shows that none of the socio-demographic variables (e.g. age, gender, education, and marital status etc.) were significantly associated with any use of the PA monitoring system. However, compared to active participants (baseline pedometer steps more than 10,000 steps/week), moderately-active [-50.453 (25.144), $p<0.05$] participants had significantly earlier non-usage attrition. Being in the inactive group [-55.126 (32.055), $p<0.10$] was only marginally related to non-usage attrition. Additionally, older participants [2.783 (1.261), $p<0.05$] had a longer duration of usage before attrition than younger participants and those who were married/co-habituating [-49.524 (24.748), $p<0.05$] had significantly shorter duration to non-usage attrition than single participants.

Table 22. Cragg's double-hurdle model showing the determinants of the decision to record daily activity via the PA monitoring system and the duration to non-usage attrition

Determinants	The decision to use PA monitoring system		The duration to non-usage attrition	
	Coefficients (S.E.)	**P-value**	**Coefficients**	**P-value**
Age (years)	-0.005 (0.008)	0.496	**2.783 (1.261)***	**0.027**
Gender				
(reference group: Male)				
Female	0.035 (0.161)	0.828	-32.794 (25.434)	0.197
Income				
(reference group: ≤£20,000 per year)				
£20,001 to £39,961	-0.040 (0.209)	0.849	-3.064 (29.773)	0.918
>£39,961	-0.243 (0.231)	0.294	-46.686 (36.591)	0.202
Education attainment				
(reference group: no higher education)				
Higher education	0.134 (0.150)	0.368	3.473 (22.568)	0.878
Marital status				
(reference group: unmarried/co-habiting)				
Married/co-habiting	0.054 (0.167)	0.745	-49.524 (24.748)*	0.045
Physical activity				
(reference group: physically-active group (baseline physical activity steps≥10,000/day))				
Moderately-active (≥5,000 and <10,000 steps/day)	-0.267 (0.168)	0.113	-50.453 (25.144)*	0.045
Inactive (<5,000 steps/day)	-0.186 (0.214)	0.385	-55.126 (32.055)	0.085
Strata: (reference group: small (<20)				
Strata: school	-0.480 (0.322)	0.137	-39.580 (67.343)	0.557
Strata: medium (20-50)	-0.039 (0.172)	0.818	52.423 (27.672)	0.058
Strata: large (50+)	0.028 (0.192)	0.883	30.591 (31.134)	0.326
N	372			
Log-likelihood	*-1575.10*			

Notes: * indicates p<0.05.

The second Cragg's double hurdle model in Table 23 shows the influence of socio-demographic characteristics on the decision to increase PA at six month follow-up compared to baseline and the amount by which PA increased. Those participants who were moderately-active [0.519 (0.159), $p<0.001$] and inactive [1.067 (0.197), $p<0.001$] at baseline were more likely to increase their PA at the six month follow-up, compared to the participants who had already been active at baseline. Moreover, participants from the organizations with more than 50 participants [0.400 (0.164), $p<0.05$] were more likely to increase PA than participants in those organisations with a smaller number (<20) of participants. Next, participants from the medium [0.386 (0.179), $p<0.05$] and high income [0.417 (0.196), $p<0.05$] groups had a significantly higher likelihood of increasing PA at six months, in comparison to the participants from the low-income group. Finally, the variables that were significantly related to the amount by which PA increased (among those that for whom increases were observed at six months) were: higher education [0.590 (0.204), $p<0.01$] and medium income [-0.611 (0.300), $p<0.05$], suggesting that participants with university education improved more than those without higher education and that participants with medium income improved less than did low-income participants.

Table 23. Cragg's double hurdle model showing the determinants of the decision to increase PA at six month follow-up and the amount of PA increased

Determinants	The decision to increase PA at six month follow-up compared to baseline		The amount of PA increased	
	Coefficients (S.E.)	**P-value**	**Coefficients**	**P-value**
Age (years)	-0.004 (0.007)	0.527	-0.014 (0.010)	0.188
Gender (reference group: Male)				
Female	0.193 (0.141)	0.172	-0.136 (0.228)	0.553
Income (reference group: ≤£20,000 per year)				
£20,001 to £39,961	0.386 (0.179)*	0.031	-0.611 (0.300)*	0.041
>£39,961	0.417 (0.196)*	0.034	-0.629 (0.327)	0.055
Education attainment (reference group: no higher education)				
Higher education	0.164 (0.133)	0.215	0.590 (0.204)**	0.004
Marital status (reference group: unmarried/co-habiting)				
Married/co-habiting	0.195 (0.156)	0.212	-0.072 (0.251)	0.774
Physical activity (reference group: physically-active group (baseline physical activity steps≥10,000/day))				
Moderately-active (≥5,000 and <10,000 steps/day)	0.519 (0.159)***	0.001	0.131 (0.283)	0.644
Inactive (<5,000 steps/day)	1.067 (0.197)***	0.000	0.296 (0.312)	0.343
Strata: (reference group: small (<20)				
Strata: school	0.287 (0.287)	0.317	-0.334 (0.423)	0.429
Strata: medium (20-50)	0.031 (0.156)	0.843	-0.166 (0.244)	0.497
Strata: large (50+)	0.400 (0.164)*	0.015	-0.269 (0.244)	0.270
Intervention Group	-0.163 (0.126)	0.196	0.079 (0.189)	0.674
N	467			
Log-likelihood	-548.74			

Note: * indicates p<0.05, ** indicates p<0.01, and *** indicates p<0.001. 35.8% participants increased their PA at the six month follow-up compared to their baseline PA. The dependent variable is the logarithm transformation of amount of PA increased.

6.1.4 Behavioural economic experiments

Present-biasedness, discount rate, risk aversion, and loss aversion among the 207 (153 from the Intervention Group and 54 from the Control Group) individuals who participated in the behavioural economic experiments were estimated to be 0.995 [95% CI: 0.989, 1.000], 0.18 [95% CI: 0.14, 0.23], 0.50 [95% CI: 0.37, 0.63], and 1.32 [95% CI: 0.26, 4.27], respectively. The estimated discount rate was comparable to the result (0.18) reported in a similar study conducted in Northern Ireland[122] but was much higher than the discount rate estimates elicited in similar experimental settings among a general Danish population (0.10).[123]Meanwhile, the present-biasedness parameter was estimated to be 0.995 and is significantly smaller than 1, indicating that our sample was relatively present-biased, in contrast to the general Danish population which did not manifest present-biasedness. Moreover, risk preference was estimated to be 0.50, indicating that our sample was more willing to take risks compared to the general Danish population (0.65). Finally, the loss aversion estimates were similar to the figure (1.35) reported in West Africa[124] and were lower than in Vietnam (2.63).[125]

Table 24 shows the results of Cragg's double hurdle model that accounts for the effects of behavioural economic characteristics on the decision to increase PA at the 12 month follow-up compared to six month follow-up and the amount by which PA increased. None of the four behavioural economic characteristics were associated with the actual increment of PA at the 12 month follow-up compared to six month follow-up. Neither did present-biasedness nor discount rate play a role in either of the hurdle models. When it comes to the amount of PA increased, however, participants who were more loss-averse [1.189 (0.523), $p<0.001$] improved their PA significantly more than the less loss-averse participants. Obviously we caveat these interpretations with due regard to the self-selection of participants into this subsample and cannot make broader inferences concerning the entire set of trial participants at baseline.

Table 24. Cragg's double hurdle model showing the influences of behavioural economics concepts on the decision to increase PA at 12 month follow-up compared to six month follow-up and the amount of PA increased

Determinants	The decision to increase PA at 12 month follow-up compared to six month follow-up		The amount of PA increased	
	Coefficients (SE)	**P-value**	**Coefficients**	**P-value**
Age (years)	0.026(0.013)	0.055	0.011(0.014)	0.436
Gender (reference group: Male)				
Female	-0.209(0.253)	0.409	0.268(0.243)	0.269
Income (reference group: ≤£20,000 per year)				
£20,001 to £39,961	-0.258 (0.398)	0.518	0.280 (0.362)	0.439
>£39,961	-0.407 (0.434)	0.348	0.182 (0.409)	0.656
Education attainment (reference group: no higher education)				
Higher education	-0.073 (0.244)	0.764	-0.090 (0.234)	0.701
Marital status (reference group: unmarried/co-habiting)				
Married/co-habiting	0.051 (0.278)	0.854	0.750 (0.289)**	0.009
Physical activity (reference group: physically-active group (baseline physical activity steps≥10,000/day))				
Moderately-active (≥5,000 and <10,000 steps/day)	0.091 (0.259)	0.724	-0.404 (0.243)	0.097
Inactive (<5,000 steps/day)	-0.256 (0.333)	0.441	-0.570 (0.331)	0.085
Strata: (reference group: small (<20)				
Strata: medium (20-50)	-0.026 (0.261)	0.920	0.496 (0.266)	0.062
Strata: large (50+)	0.097 (0.293)	0.742	-0.083 (0.283)	0.770
Intervention Group	-0.522 (0.260)	0.045	-0.016 (0.245)	0.948
Behavioural economics concepts				
Present-biasedness	-31.078 (32.336)	0.337	-26.021 (29.126)	0.372
Discounting rate	0.023 (1.059)	0.983	0.510 (1.067)	0.633
Risk preference	-0.827 (0.692)	0.232	1.131 (0.593)	0.056
Loss aversion	0.506 (0.573)	0.378	1.189 (0.523)***	0.000
N	149			
Log-likelihood	-205.66			

Notes: * indicates p<0.05, ** indicates p<0.01, and *** indicates p<0.001. 42.8% participants increased their PA at the 12 month follow-up compared to six month follow-up. The dependent variable is the logarithm transformation of amount of PA increased.

Table 25 illustrates the associations between PA at six and 12 month follow-up and their possible behavioural economic determinants, respectively. A higher discount rate [-0.662 (0.329), $p<0.05$] was associated with a lower PA at the six month follow-up. This is in line with previous time preference studies suggesting that people with higher discount rates were less likely to resist the temptation for an immediate gratification which might harm their future health.[93,94] However, discount rate was not associated with the PA level observed at the 12 month follow-up [-0.073 (0.315), $p>0.10$]. The remaining behavioural economic parameters did not play a role in determining PA measured either at six or the 12 month time points.

Table 25. OLS regression showing the influences of behavioural economics concepts on PA at six and 12 month follow-up

Determinants	PA at six month follow-up		PA at 12 month follow-up	
	Coefficients (SE)	**P-value**	**Coefficients (SE)**	**P-value**
Age (years)	0.004(0.004)	0.303	0.010(0.004)*	0.014
Gender (reference group: Male)				
Female	-0.038(0.074)	0.605	-0.067(0.073)	0.360
Income (reference group: ≤£20,000 per year)				
£20,001 to £39,961	0.267 (0.111)*	0.017	0.287 (0.116)*	0.015
>£39,961	0.137 (0.124)	0.270	0.169(0.128)	0.191
Education attainment (reference group: no higher education)				
Higher education	0.123 (0.074)	0.095	0.105 (0.072)	0.147
Marital status (reference group: unmarried/co-habiting)				
Married/co-habiting	0.069 (0.083)	0.404	0.058 (0.082)	0.481
Physical activity (reference group: physically-active group (baseline physical activity steps≥10,000/day))				
Moderately-active (≥5,000 and <10,000 steps/day)	-0.557 (0.080)***	0.000	-0.465 (0.076)***	0.000
Inactive (<5,000 steps/day)	-0.792 (0.098)***	0.000	-0.754 (0.099)***	0.000
Strata: (reference group: small (<20)				
Strata: medium (20-50)	-0.054 (0.078)	0.492	-0.001 (0.077)	0.987
Strata: large (50+)	0.093 (0.090)	0.304	0.007 (0.088)	0.939
Intervention Group	-0.160 (0.077)*	0.040	-0.217 (0.075)**	0.005
Behavioural economics concepts				
Present-biasedness	-7.287 (9.178)	0.428	-12.063 (9.074)	0.186
Discounting rate	-0.662 (0.329)*	0.046	-0.073 (0.315)	0.817
Risk preference	-0.108 (0.210)	0.608	0.134 (0.197)	0.500
Loss aversion	-0.293 (0.166)	0.080	-0.012 (0.167)	0.943
N	174		157	
R^2	0.42		0.42	

Notes: * indicates p<0.05, ** indicates p<0.01, and *** indicates p<0.001. The dependent variable is the logarithm value of the number of daily mean steps.

7 RESULTS FROM THE MEDIATION ANALYSES

The results of the main outcome evaluation show that there was a small but significant decline in mean steps/day at six months relative to baseline, for the Intervention Group compared to Controls and no significant differences at 12 months. Several authors have claimed that investigation of causal mechanisms is more valuable when interventions are not successful as information about which components (if any) were successful, and which were not, can inform the design of future studies.[126,127]Therefore, mediation analyses were conducted to explore potential reasons for this result and to determine whether there were any intervention components with beneficial effects for PA.

Means and SDs for all putative mediator variable scores are presented in Table 26. Descriptive statistics are presented according to time-point (i.e. baseline, four weeks or six months) and group.

Table 26. Baseline, four week and six month scores on mediator variables and perceptions of workplace environment

	Baseline				Four weeks				Six months			
	Intervention		Control		Intervention		Control		Intervention		Control	
VARIABLES	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**n**	**Mean (SD)**	**n**	**Mean (SD)**
PA self-efficacy (1-5)	439	2.91 (0.97)	376	2.92 (0.94)	344	2.83 (0.89)	319	2.80 (0.92)				
Intentions (1-7)	435	5.38 (1.68)	375	5.37 (1.75)	343	5.42 (1.58)	321	5.09 (1.77)				
Outcome expectations (1-5)	418	3.37 (0.62)	354	3.36 (0.64)	318	3.25 (0.67)	292	3.27 (0.60)				
Financial motivation (1-7)	439	1.71 (1.16)	376	1.79 (1.28)	345	2.12 (1.36)	320	2.06 (1.42)				
Planning (1-4)	414	2.37 (0.69)	363	2.45 (0.69)	344	2.29 (0.70)	319	2.32 (0.75)	255	2.35 (0.74)	235	2.32 (0.71)
Social norms (1-7)	414	3.87 (1.20)	357	4.04 (1.14)	346	3.93 (1.16)	317	3.78 (1.29)	253	3.90 (1.13)	235	3.90 (1.12)
Identified regulation (1-5)	438	3.81 (0.87)	375	3.92 (0.83)	346	3.99 (0.78)	319	3.89 (0.83)	262	3.93 (0.82)	239	3.91 (0.87)
Integrated regulation (1-5)	439	3.12 (1.13)	373	3.27 (1.11)	344	3.37 (1.08)	319	3.24 (1.10)	258	3.41 (1.10)	238	3.31 (1.12)
Intrinsic motivation (1-5)	438	3.52 (0.99)	376	3.63 (0.97)	346	3.70 (0.88)	320	3.58 (0.94)	259	3.70 (0.91)	239	3.63 (0.97)
Habit (1-5)	437	2.89 (1.32)	375	3.08 (1.24)					256	3.18 (1.40)	235	2.87 (1.45)
Workplace norms (1-5)	439	3.20 (0.82)	377	3.20 (0.85)					260	3.19 (0.76)	237	3.14 (0.83)
Recovery self-efficacy (1-4)	438	2.36 (0.82)	375	2.34 (0.79)					261	2.41 (0.73)	238	2.41 (0.70)
Maintenance self-efficacy (1-4)	438	2.79 (0.86)	376	2.77 (0.89)					262	2.69 (0.83)	237	2.69 (0.75)
Outcome satisfaction (1-5)	404	3.85 (0.68)	352	3.87 (0.65)					257	3.87 (0.62)	233	3.80 (0.69)
WE: Attractiveness (4-20)	438	10.83 (2.71)	377	10.37 (3.14)								
WE: Safety (4-20)	438	10.63 (2.47)	377	10.39 (2.68)								
WE: Accessibility (3-15)	439	9.27 (2.14)	377	9.14 (2.07)								
WE: Availability (3-15)	439	10.36 (2.00)	376	10.27 (2.11)								
WE (14-70)	437	38.10 (6.22)	376	37.19 (7.07)								

WTA30	352	3.26 (4.70)	315	2.72 (4.03)
WTA60	345	5.37 (5.99)	312	4.22 (4.67)

PA: physical activity; SD: standard deviation; WE: Workplace environment; WTA: Willingness-to-accept.

7.1 MEDIATOR OUTCOMES

Table 27 shows the results of random-effects regressions for individual mediators, at four weeks and six months and these are briefly described below. These results should be interpreted with caution as no adjustment has been made for multiple testing, meaning that type 1 errors may occur.

7.1.1 Initiation of PA

At four weeks post-baseline, there were significant differences between the Intervention and Control Groups for intentions (*b*=0.29, SE=0.13, p=0.02), social norms (*b*=0.23, SE=0.08, p<0.01), identified regulation (*b*=0.14, SE=0.06, p=0.01), integrated regulation (*b*=0.23, SE=0.07, p<0.01), and intrinsic motivation (*b*=0.18, SE=0.06, p<0.01). There were non-significant between-group differences for PA self-efficacy (p=0.31), outcome expectations (p=0.58), financial motivation (p=0.46) and planning (p=0.75) (Table 27). Thus the Intervention Group participants showed increases above Control Group participants in intentions, social norms, identified regulation, integrated regulation and intrinsic motivation between baseline and four weeks, approximately equivalent to 0.29, 0.23, 0.14, 0.23 and 0.18 points on their original scale, respectively. Sensitivity analyses conducted using logistic regressions showed that the significance level changed for some mediators. Specifically, logistic regressions indicated that there was a significant between-group difference for financial motivation (exp(*b*)=0.37, SE=0.15, p=0.02) and no significant between-group difference for social norms (p=0.11).

7.1.2 Maintenance of PA

At six months, there were significant differences between the Intervention and Control Groups for identified regulation (*b*=0.11, SE=0.05, p=0.02), integrated regulation (*b*=0.26, SE=0.08, p<0.01), intrinsic motivation (*b*=0.17, SE=0.06, p<0.01), and habit (*b*=0.48, SE=0.12, p<0.01). There were non-significant between-group differences for planning (p=0.18), social norms (p=0.42), workplace norms (p=0.06), recovery self-efficacy (p=0.80), maintenance self-efficacy (p=0.80), and outcome satisfaction (p=0.56) (Table 27). Thus the Intervention Group participants showed increases above Control Group participants in identified regulation, integrated regulation, intrinsic motivation and habit between baseline and six months, approximately equivalent to 0.11, 0.26, 0.17 and 0.48 points on their original scale, respectively. Sensitivity analyses conducted using logistic regressions showed that the significance level changed for some mediators. Specifically, logistic regressions indicated that

there was a significant between-group difference for workplace norms (exp(*b*)=0.37, SE=0.18, p<0.05).

Table 27. Results of random-effects regressions of individual mediators, at four weeks or six months, on its baseline value and group assignment

	Four week mediators			Six month mediators		
MEDIATOR	**n**	**b (SE)**	**P-value**	**n**	**b (SE)**	**P-value**
PA self-efficacy	597	0.09 (0.08)	0.31			
Intentions	595	0.29 (0.13)	0.02			
Outcome expectations	528	-0.03 (0.06)	0.58			
Financial motivation	600	0.11 (0.15)	0.46			
Planning	575	0.02 (0.05)	0.75	436	0.08 (0.06)	0.18
Social norms	576	0.23 (0.08)	<0.01	434	0.10 (0.12)	0.42
Identified regulation	598	0.14 (0.06)	0.01	459	0.11 (0.05)	0.02
Integrated regulation	595	0.23 (0.07)	<0.01	454	0.26 (0.08)	<0.01
Intrinsic motivation	599	0.18 (0.06)	<0.01	456	0.17 (0.06)	<0.01
Habit				448	0.48 (0.12)	<0.01
Workplace norms				456	0.13 (0.07)	0.06
Recovery self-efficacy				457	0.02 (0.06)	0.80
Maintenance self-efficacy				459	-0.02 (0.09)	0.80
Outcome satisfaction				427	0.03 (0.05)	0.56

PA: Physical activity; SE: Standard error

NB. Results are adjusted for strata, season, baseline pedometer steps/day and baseline mediator values with cluster-adjusted standard errors and p-values (*b*=coefficient for group assignment variable, i.e. Intervention versus Control).

7.1.3 Single Mediator Models

Table 28 shows the results of single mediator models based on the product-of-coefficients approach with CIs for the indirect effect formed using the bias-corrected bootstrap, and adjustment for randomisation strata, season, baseline values of the mediator and baseline pedometer steps/day, with SEs and p-values corrected for clustering. An explanatory diagram for single mediator models is presented in Figure 6.

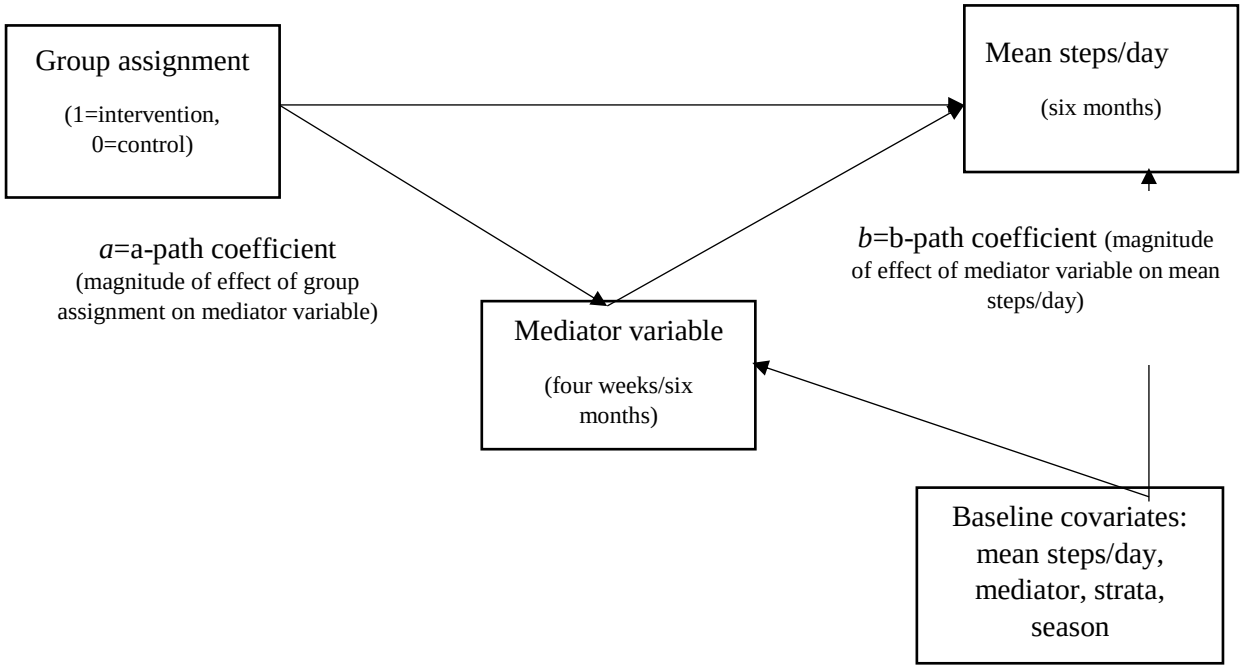

Figure 6. Explanatory diagram for single mediator models showing an indirect effect of group assignment on six month mean steps/day through a four week/six month mediator

7.1.3.1 Initiation of PA Behaviour

There were no significant indirect effects in any of the four week single mediator models (left side of table 28). SRMR values were close to zero for all models, indicating a good fit[115], and CD values ranged from 0.56-0.76.

7.1.3.2 Maintenance of PA Behaviour

The right side of Table 28 shows that there were significant, positive indirect effects in single mediator models including the following mediators: integrated regulation (*b*=94.70, 95% CI: 18.69, 204.4), intrinsic motivation (*b*=59.00, 95% CI: 3.09, 154.5), and habit (*b*=198.7, 95% CI: 84.32, 369.9). SRMR values were close to zero for all models, indicating a good fit,[115] and CD values ranged from 0.62-0.76. Thus, whilst the total intervention effect on pedometer steps/day at six months was negative, this was partially mitigated by increases in integrated regulation, intrinsic motivation and habit (baseline-six months), which were positively related to six month steps/day, as a result of being assigned to the Intervention Group. Assignment to the Intervention Group resulted in increased intrinsic motivation, integrated regulation and habit; as such, participants who experienced an increase of these constructs (baseline-six months), also experienced less of a decline in PA. Therefore, an increase in such constructs lessened the decline in PA.

Table 28. Indirect effects in single mediator models with bias-corrected confidence intervals

	Six month mean steps/day					
	Four week mediators			Six month mediators		
MEDIATOR	**n**	**Indirect effect (SE)**	**95% CI**	**n**	**Indirect effect (SE)**	**95% CI**
PA self-efficacy	417	32.74 (28.20)	-2.87, 116.0			
Intentions	415	-11.19 (30.18)	-89.87, 38.52			
Outcome expectations	363	1.84 (14.41)	-16.63, 49.68			
Financial motivation	420	2.55 (17.92)	-27.61, 51.77			
Planning	403	-3.06 (17.50)	-60.43, 19.87	382	50.64 (42.34)	-17.31, 153.3
Social norms	405	-33.26 (32.69)	-131.2, 8.09	382	22.02 (32.71)	-23.41, 115.8
Identified regulation	417	13.09 (35.48)	-50.45, 98.09	403	35.11 (36.90)	-32.39, 116.9
Integrated regulation	416	22.91 (38.42)	-46.88, 111.4	399	94.70 (46.27)	18.69, 204.4
Intrinsic motivation	418	5.99 (34.79)	-62.39, 81.47	400	59.00 (36.81)	3.09, 154.5
Habit				394	198.7 (70.73)	84.32, 369.9
Workplace norms				400	-36.10 (32.58)	-129.6, 5.61
Recovery self-efficacy				402	-2.47 (17.11)	-55.60, 20.98
Maintenance self-efficacy				403	-3.41 (19.55)	-57.08, 28.95
Outcome satisfaction				376	28.04 (29.82)	-11.61, 112.7

CI: Confidence interval; PA: Physical activity; SE: Standard error

NB: results are coefficients and cluster-adjusted standard errors and p-values from single mediator models. Bias-corrected bootstrap confidence intervals for the indirect effect are reported (10,000 reps). IV=Group assignment, MV=follow-up scores of mediators, DV=follow-up scores of outcome (mean steps/day). All paths are adjusted for strata, season, baseline values of the mediator and baseline pedometer steps/day. Unstandardised coefficients are extracted.

7.1.4 Single Moderated-Mediation Models

Single mediator models including six month scores of integrated regulation, intrinsic motivation, and habit (with six month mean steps/day as the DV) were tested for moderation of the a-path and b-path respectively by perceptions of workplace environment availability (WE: Availability). Significant a-path moderation was found for the single mediator model with integrated regulation as the mediator (*b*=-0.10, SE=0.05, p<0.05), and the single mediator model with habit as the mediator (*b*=-0.17, SE=0.05, p<0.01). There was a non-significant moderator-by-group interaction coefficient (i.e. a-path moderation) for the single mediator model with intrinsic motivation as the mediator (p=0.19). There was a non-significant moderator-by-mediator interaction coefficient (b-path moderation) for the single mediator models with integrated regulation (p=0.36), intrinsic motivation (p=0.06) and habit (p=0.44) as the mediator. SRMR values ranged from 0.00-0.05, indicating a good fit,[115] and CD values ranged from 0.66-0.77. Thus the apparent mitigation of the intervention effect, by virtue of higher levels of intrinsic motivation and habit was dependent on participant's perception of workplace environment availability for PA. It is unclear if those with higher levels of such constructs benefited more with regard to PA due to having more PA opportunities near their workplace or as a result of their perception of such opportunities. This is representative of an 'intervention challenge', meaning future intervention studies should give consideration to how to increase levels of integrated regulation and habit for participants with higher perceptions of the availability of PA opportunities in the workplace environment.

7.1.5 Engagement and Non-Usage Attrition

Table 29 shows descriptive statistics for six month engagement and non-usage attrition. Engagement variables which were significant predictors of six month mean steps/day in univariable analyses were included in a multivariable model which showed that frequency of hits on the monitoring and feedback component of the website across the six month intervention period (*b*=50.2, SE=24.5, p=0.04) and percentage of earned points redeemed across the six month intervention period (*b*=9.1, SE=3.3, p<0.01) were positively related to six month mean steps/day, whilst frequency of hits on the discussion forum component of the website across the six month intervention period (*b*=-69.3, SE=26.6, p<0.01) was negatively related to six month mean steps/day (Table 30). None of the other variables were significant predictors of six month mean steps/day in univariable analyses.

Table 29. Descriptive statistics for six month engagement and non-usage attrition

VARIABLES	Six months N	Six months Mean (SD)
Engagement		
% intervention days participants walked for at least 10 mins captured via the PA monitoring system [a]	422	24.7 (21.8)
% intervention weeks participants logged onto the website[b]	418	37.8 (32.5)
% earned points redeemed[c]	422	39.3 (42.5)
Frequency: Monitoring and feedback[d]	418	13.7 (3.5)
Frequency: Rewards[d]	418	5.7 (4.5)
Frequency: Maps[d]	418	3.4 (4.0)
Frequency: Health info. (PA)[d]	418	0.5 (1.7)
Frequency: Health info. (Other)[d]	418	1.2 (3.2)
Frequency: Discussion forums[d]	418	1.9 (4.2)
Total no. sections (website)[e]	418	3.9 (1.5)
Total minutes (recording daily activity via PA monitoring system)	422	1,000 (987)
Total minutes (PAL website)	418	1,171 (2,048)
Non-usage attrition		
Days to non-usage attrition (recording daily activity via PA monitoring system)[f]	422	53.7 (61.2)
Days to non-usage attrition (PAL website)[g]	418	31.7 (43.4)
No. participants with non-usage attrition for recording daily activity via PA monitoring system (%)		375 (89%)
No. participants with PAL website non-usage attrition (%)		403 (96%)

[a]Percentage of days participants were recorded walking for at least 10 mins captured via the PA monitoring system.

[b]Percentage of weeks participants logged onto the website at least once.

[c]Percentage of total accumulated points which the participant had redeemed by six months.

[d]Frequency of hits (i.e. total number of hits for every ten days the participant accessed the website).

[e]Number of sections accessed on website at least once (0-6).

[f]Number of days until first two week lapse from recording daily activity via PA monitoring system.

[g]Number of days until first two week lapse from logging onto the website.

NB. Numbers are Mean (SD) unless otherwise stated.

Table 30. Results of random-effects regressions with six-month pedometer steps/day as the dependent variable and engagement indicators as independent variables

ENGAGEMENT VARIABLES	Univariable models			Multivariable model		
	n	*b* (SE)	P-value	n	*b* (SE)	P-value
Engagement indicators						
% intervention days participants walked for at least 10 mins captured via the PA monitoring system[a]	231	4.21 (8.54)	0.62			
% intervention weeks participants logged onto the website[b]	234	4.39 (6.01)	0.47			
% earned points redeemed[c]	231	8.26 (4.07)	0.04	230	9.1 (3.3)	<0.01
Website sections						
Monitoring and feedback[d]	234	66 (18)	<0.01	230	50.2 (24.5)	0.04
Rewards[d]	234	14 (36)	0.70			
Maps[d]	234	-47 (44)	0.28			
Health information: PA[d]	234	35 (160)	0.83			
Health information: Other[d]	234	26 (66)	0.70			
Discussion forums[d]	234	-77 (27)	<0.01	230	-69.3 (26.6)	<0.01
Number of sections[e]	234	-32 (117)	0.78			

[a]Percentage of days participants were recorded walking for at least 10 mins captured via the PA monitoring system.
[b]Percentage of weeks participants logged onto the website at least once.
[c]Percentage of total accumulated points which the participant had redeemed by six months.
[d]Frequency of hits (i.e. total number of hits for every ten days the participant accessed the website).
[e]Number of sections accessed on website at least once (0-6).

NB. Results are adjusted for strata, season, baseline pedometer steps/day with cluster-adjusted standard errors and p-values.

The percentage of days (Table 31) that participants were recorded walking for at least 10 mins captured via the PA monitoring system was a significant predictor of integrated regulation (*b*=0.007, SE=0.002, p<0.01), intrinsic motivation (*b*=0.003, SE=0.001, p=0.04), habit (*b*=0.007, SE=0.003, p=0.01) and outcome satisfaction (*b*=0.005, SE=0.02, p=0.01). The percentage of weeks that participants logged onto the website at least once was a significant predictor of integrated regulation (*b*=0.004, SE=0.002, p=0.02), intrinsic motivation (*b*=0.002, SE=0.0010, p=0.04), habit (*b*=0.004, SE=0.002, p=0.03) and outcome satisfaction (*b*=0.004, SE=0.001, p=0.01). The percentage of earned points redeemed was a significant predictor of six month social norms (*b*=0.003, SE=0.001, p=0.01).

Table 31. Results of random-effects regressions with individual mediators as dependent variables and engagement in difference intervention components as independent variables

	% intervention days participants walked for at least 10 mins captured via the PA monitoring system [a]			% intervention weeks participants logged onto the website[b]			% earned points redeemed[c]		
OUTCOME	n	*b* (SE)	P-value	n	*b* (SE)	P-value	n	*b* (SE)	P-value
Pedometer steps/day	231	4.21 (8.54)	0.62	234	4.39 (6.01)	0.47	231	8.26 (4.07)	0.04
Planning	228	0.003 (0.002)	0.18	229	0.003 (0.002)	>0.05	228	0.0008 (0.001)	0.56
Social norms	230	0.004 (0.003)	0.19	230	0.002 (0.003)	0.36	230	0.003 (0.001)	0.01
Identified regulation	241	0.002 (0.002)	0.33	243	0.002 (0.0008)	0.06	241	0.0001 (0.0007)	0.87
Integrated regulation	238	0.007 (0.002)	<0.01	240	0.004 (0.002)	0.02	238	-0.0002 (0.0009)	0.82
Intrinsic motivation	238	0.003 (0.001)	0.04	240	0.002 (0.0010)	0.04	238	0.0004 (0.001)	0.72
Habit	233	0.007 (0.003)	0.01	235	0.004 (0.002)	0.03	233	-0.00004 (0.002)	0.98
Workplace norms	239	-0.0004 (0.002)	0.83	241	0.001 (0.002)	0.55	239	0.0007 (0.0009)	0.45
Recovery self-efficacy	241	0.004 (0.003)	0.15	243	0.002 (0.001)	0.18	241	0.0002 (0.0009)	0.81
Maintenance self-efficacy	242	0.002 (0.003)	0.55	244	0.002 (0.001)	0.27	242	-0.0005 (0.0005)	0.36
Outcome satisfaction	225	0.005 (0.002)	0.01	225	0.004 (0.001)	0.01	225	0.001 (0.0007)	0.11

[a]Percentage of days participants were recorded walking for at least 10 mins captured via the PA monitoring system.
[b]Percentage of weeks participants logged onto the website at least once.
[c]Percentage of total accumulated points which the participant had redeemed by six months.

NB. Results are adjusted for strata, season, baseline pedometer steps/day and baseline mediators (for models including six month mediator variables as the DV) with cluster-adjusted standard errors and p-values.

Frequency of hits (Table 32) on the monitoring and feedback component of the website was a significant predictor of six month planning (*b*=-0.04, SE=0.01, p<0.01), identified regulation (*b*=0.02, SE=0.01, p=0.04) and integrated regulation (*b*=0.03, SE=0.01, p=0.02). Frequency of hits on PA health information was a significant predictor of integrated regulation (*b*=0.11, SE=0.05, p=0.03). Frequency of hits on discussion forums was a significant predictor of six month social norms (*b*=-0.03, SE=0.01, p=0.03) and integrated regulation (*b*=-0.02, SE=0.01, p=0.03). Number of sections accessed on the website at least once was a significant predictor of identified regulation (*b*=0.07, SE=0.02, p<0.01) and integrated regulation (*b*=0.09, SE=0.03, p<0.01).

Table 32. Results of random-effects regressions with individual mediators as dependent variables and frequency of hits on each section of the website as independent variables

		Monitoring and feedback[a]		Rewards[a]		Maps[a]		Health information: PA[a]		Health information: Other[a]		Discussion forums[a]		Number of sections[b]	
OUTCOME	n	*b* (SE)	P-value	*b* (SE)	P-value	*b* (SE)	P-value	*b* (SE)	P-value	*b* (SE)	P-value	*b* (SE)	P-value	*b* (SE)	P-value
Planning	229	-0.04 (0.01)	<0.01	-0.005 (0.01)	0.59	-0.01 (0.01)	0.55	-0.05 (0.05)	0.36	-0.05 (0.03)	0.06	0.002 (0.01)	0.80	0.03 (0.04)	0.43
Social norms	230	0.01 (0.02)	0.61	-0.02 (0.02)	0.31	-0.03 (0.02)	0.09	-0.05 (0.06)	0.46	0.01 (0.03)	0.68	-0.03 (0.01)	0.03	0.04 (0.04)	0.29
Identified regulation	243	0.02 (0.01)	0.04	-0.01 (0.01)	0.61	0.005 (0.01)	0.52	0.06 (0.04)	0.08	0.04 (0.04)	0.33	0.01 (0.005)	0.15	0.07 (0.02)	<0.01
Integrated regulation	240	0.03 (0.01)	0.02	0.0005 (0.01)	0.97	-0.02 (0.01)	0.10	0.11 (0.05)	0.03	0.06 (0.06)	0.27	-0.02 (0.01)	0.03	0.09 (0.03)	<0.01
Intrinsic motivation	240	0.01 (0.01)	0.29	-0.01 (0.01)	0.67	0.002 (0.01)	0.88	0.06 (0.05)	0.23	0.05 (0.05)	0.34	-0.004 (0.01)	0.61	0.04 (0.03)	0.12
Habit	235	-0.0004 (0.02)	0.99	-0.01 (0.02)	0.79	0.01 (0.02)	0.50	0.06 (0.06)	0.36	-0.06 (0.07)	0.40	-0.02 (0.01)	0.07	0.01 (0.04)	0.73
Workplace norms	241	0.005 (0.01)	0.70	-0.001 (0.01)	0.96	-0.001 (0.01)	0.95	0.06 (0.04)	0.18	0.01 (0.04)	0.71	0.003 (0.01)	0.72	0.03 (0.03)	0.28
Recovery self-efficacy	243	-0.01 (0.01)	0.46	-0.01 (0.01)	0.68	-0.01 (0.01)	0.11	0.05 (0.04)	0.24	0.03 (0.03)	0.30	0.002 (0.01)	0.80	0.03 (0.02)	0.10
Maintenance self-efficacy	244	-0.01 (0.01)	0.53	-0.01 (0.01)	0.40	-0.02 (0.02)	0.31	-0.03 (0.07)	0.65	-0.02 (0.04)	0.59	-0.01 (0.01)	0.06	0.003 (0.04)	0.94
Outcome satisfaction	225	-0.01 (0.01)	0.38	0.01 (0.01)	0.57	-0.01 (0.01)	0.44	-0.04 (0.04)	0.31	-0.02 (0.02)	0.35	0.01 (0.01)	0.11	0.05 (0.04)	0.16

[a]Frequency of hits (i.e. total number of hits for every ten days the participant accessed the website).
[b]Number of sections accessed on website at least once (0-6).

NB. Results are adjusted for strata, season, baseline pedometer steps/day and baseline mediators (for models including six month mediator variables as the DV) with cluster-adjusted standard errors and p-values.

The estimated median lifetime usage (i.e. after which 50% of participants stopped use) was 25.5 days for use of the PA monitoring system to record daily activity and 12.5 days for use of the website. Non-usage attrition of the PA monitoring system to record daily activity occurred for 89% of participants (n=375), and website non-usage attrition occurred for 96% of participants (n=403). The results of univariable and multivariable Cox regression analyses examining baseline measures of sociodemographic variables, psychosocial variables, environmental variables and PA as predictors of non-usage attrition are presented separately for use of the PA monitoring system to record daily activity and use of the website. The multivariable analysis for use of the PA monitoring system to record daily activity showed that having higher levels of identified regulation at baseline (hazard ratio=0.88, 95% CI: 0.81, 0.97), having higher levels of recovery self-efficacy at baseline (hazard ratio=0.88, 95% CI: 0.80, 0.98), or having a lower perception of the safety of the workplace environment for PA at baseline (hazard ratio=1.06, 95% CI: 1.02, 1.11) reduced the risk of attrition. The multivariable analysis for website use showed that being older (hazard ratio=0.99, 95% CI: 0.98, 1.00), unmarried/not co-habiting (hazard ratio=0.79, 95% CI: 0.63, 1.00), having higher values on the EuroQol (EQ-5D) weighted health index (hazard ratio=0.32, 95% CI: 0.12, 0.86), having higher levels of financial motivation at baseline (hazard ratio=0.93, 95% CI: 0.87, 0.99), having a lower perception of the attractiveness of the workplace environment for PA at baseline (hazard ratio=1.06, 95% CI: 1.01, 1.11), or having a higher perception of the availability of PA opportunities in the workplace environment at baseline (hazard ratio=0.94, 95% CI: 0.90, 0.98) reduced the risk of attrition.

8 DISCUSSION AND CONCLUSIONS

8.1 INTRODUCTION

Chapter 7 first summarises key findings and gives an overview of how the PAL Scheme delivery and fidelity may have impacted the findings and their generalisability, and then offers a more detailed examination of how the pre-planned moderator and mediator analyses might help us interpret the study.

8.2 KEY FINDINGS

Results from the study demonstrate that assignment to the Intervention Group resulted in a small but significant decline in the primary outcome (i.e. mean steps/day at six months) relative to baseline, compared to the Waiting-List Control Group. This result was observed despite the fact that pedometer steps/day were lower for participants in the Waiting-List Control Group compared to the Intervention Group at baseline. Self-reported minutes/week of workplace PA (measured using the GPAQ) also declined among intervention compared to Control participants, but there was **no** significant change for total MVPA.

The overall results of this study (i.e. that a workplace scheme incorporating an offer of small financial incentives had a negative impact on PA at six months) are in contrast to the findings of earlier systematic reviews. Previous workplace PA interventions have shown promise[128–130] but though the literature suggests that financial incentives can promote *short-term* (i.e. at least to six months) increases in PA[25,89,131,132] and other healthy behaviours ,[22,24,132,133] it seems that increases are rarely maintained, or dissipate, after this point.[22,24,25,89,131,133]

This was the case, for example, in the TRIPPA study.[41] This study compared the use of activity trackers (i.e. self-monitoring) with or without financial incentives to increase PA in office employees. It found that provision of personal or charitable cash incentives, in addition to an activity tracker, significantly increased PA compared to a control group but not compared to an activity tracker (only) group at six months. At 12-months, whilst activity trackers and charitable donations fostered further PA increases, PA in the personal cash incentive group was not significantly different from that in the control group, and had actually decreased compared to the activity tracker (only) group. Albeit not in line with our own pilot work,[1] some of the literature highlights a concern that programmes that offer rewards are attractive only to individuals who are already active, which would place bounds on the potential effectiveness of our own intervention targeted at inactive employees. Since rewards-based programmes are unlikely to be a good use of resources if dominated by active employees, a recent national study of full-time employees in Singapore used stated-preference techniques to understand who is more likely to participate in rewards-based

PA programmes, quantifying uptake by both sufficiently active and inactive employees. However, the authors found that there was high demand for PA rewards programmes for both active and inactive employees.[134]

At 12 months, those in the Intervention Group still had a lower mean step/day total than those in the Control Group, though the difference was not significant. The focus groups conducted at 12 months further highlighted that at least some of those who maintained the PA levels that they achieved during the intervention had been previously active before the study commenced. The idea that the PAL Scheme would prompt participants to look for opportunities to embed PA as a habit within a typical working day, was to a certain extent, aligned with the findings of our mediation analyses, whereby the Intervention Group participants showed increases (above Control Group participants) in identified regulation, integrated regulation, intrinsic motivation and habit between baseline and six months. These changes were propelled through to the 12 month follow-up. While the overall findings around the primary outcome (mean steps/day) were negative, and null for self-reported MVPA, the results are more complex and from a public health perspective harder to interpret given the encouraging signals with regard to improved wellbeing and reduced absenteeism. While many public health programme evaluations struggle with issues around implementation and "dose" (both of which may have had a bearing on our findings), Tannahill and Kelly[135] have discussed some of the complexities that evaluators face when interpreting workplace-based interventions. By their account, workplaces are part of a complex system and contextual factors within such a complex system will have a significant bearing on generalisability.

Moreover, recent work by Nooijen and colleagues[136] has suggested that those who move into a sedentary occupation, may compensate by exercising more outside of the workplace. It might therefore be reasonable to think that some of those participants who became active during the early phase of the PAL Scheme, subsequently took up other forms of PA (gym membership, sports and evening walks [suggested by some in our qualitative findings]) outside of working hours, and then later, in compensation, placed less emphasis on workplace-based PA. Whilst this was not explicitly gauged in our study, as Nooijen and colleagues concluded, "mutual influences on PA of different

contexts should be considered."[136] Moreover, a novelty or "observation" effect may have affected the Control Group more than the Intervention Group, in that although the Intervention Group were more continuously aware of their activity, wearing a pedometer may have had a greater impact on the Control Group who only wore it when steps were being recorded at six and 12 months.[137,138]

8.3 PROGRAMME DELIVERY

It is notable that the PAL Scheme received generally positive comments from participants, employers and retailers. However, a number of key issues emerged from the qualitative process evaluation.

There was a consensus amongst participants and senior managers that the employer has a role to play in ensuring their employees have access to opportunities and support for being more active within the work environment. During the study, a range of programmes were concurrently delivered within the trial workplaces, over which we had no discretion, including health promotion talks and exercise classes, as well as other ongoing initiatives such as cycle to work schemes. Both the Control and Intervention Groups would have had variable exposure to these parallel programmes. However, only six employers outlined having policies in place in relation to staff health and wellbeing. Horodyska and colleagues[139] have compiled a checklist for developing practice and reporting research on workplace interventions and policies and a key aspect for success relates to integration and institutionalisation of wellbeing programs. Whilst many of the participating organisations in the study had ongoing workplace health promotion schemes, there were some which did not and the PAL Scheme was seen as a significant opportunity for PA. Our qualitative work also highlighted how participants would have liked the PAL Scheme to continue, stating that it would have had organisational support.

Given the many barriers to being active during the working day reported by our participants, it is not clear how effective existing schemes and policies within the workplace are, or how they are being implemented. This evaluative work highlighted the need for a 'top down' approach, for employees to feel they were supported by management in being more active within the workplace.

A number of issues were identified in relation to the day-to-day running of the scheme. There were occasional technical issues with the sensors and fobs which had a knock on effect on points accumulated. In addition, participants felt the scheme was restrictive in terms of the location of the sensors and the daily time limits surrounding the scheme.

A strong theme to emerge from the focus group discussions related to the feedback on PA that the intervention provided to participants, with many noting that their main motivation was to complete their personal daily PA goals. However, many participants felt the rewards element of the scheme needed to be improved, with a greater choice of vouchers, (of higher value), needed if participants were to feel the effort was worthwhile. Further, less than 40% of participants redeemed their points for vouchers which may suggest that the vouchers were not a suitable incentive for a proportion of participants. However, the types of vouchers available were based on those which were most popular in our pilot trial (same target population) and verified in pre-intervention focus groups with the target population

8.4 LIMITATIONS

With sparse literature on PA workplace-based behaviour change interventions, the PAL Scheme is one of the few studies which has used a randomised controlled trial design and has assessed the longer term effects of the intervention on objectively measured PA up to 12 months post-baseline.

Nevertheless we faced some important limitations. As recruitment rates were lower than anticipated, there was a delay in commencing the intervention as well as a need for additional recruitment sites. The overall levels of retention were lower than anticipated but not too different from some other recent work-place based studies. For example van Dongen et al[140] managed to record outcome data in 56% of trial participants at 12 months (n=232) while Puig-Ribera[141] managed to obtain outcome data on 67% of their subjects at 19 weeks (n=178). The loss to follow-up and attrition in our study were, in part, a side effect of current austerity related re-structuring of the local authorities and within the civil service which was ongoing throughout the trial.

This resulted in upheaval or uncertainty about job location and security for some participants.

There were clearly some technical issues affecting intervention fidelity. Some of the sensors developed battery problems which affected the recording and accuracy of the PA behaviour captured, and required replacement. Recruitment of our intended sample size is also a limitation of this work. Similarly, previously our pilot work did not have a true-control group, nor did it use the same technological design for the PA monitoring system or the same sensors.[1] Thus, while pilot work was informative and provided promise, such methodological differences and the possible chance finding due to the underpowered nature of a pilot study may partially account for discrepancies in expectation and reality.

8.5 GENERALISABILITY

The study design – a cluster RCT - was robust and following the amendment of the target sample size, should have been sufficiently powered (90% power at the 5% significance level) for an effect size of d=.40. It is difficult to directly compare the profile of our study participants with those in other studies but the mean steps/day in our Control Group at baseline is not dissimilar to values reported for other office workers.[142] Clearly the variation in step counts among such staff, even in UK settings, will be driven by a complex mixture of social, environmental and organisational factors.

8.6 INTERPRETATION AND IMPLICATIONS OF THE HEALTH ECONOMICS FINDINGS

The results showed that the intervention was not likely to be cost-effective for employees working in office settings, from the NHS and PSSRU's perspectives over the six month time horizon. The Intervention Group consumed less health care resources compared to the Control Group, however the difference (£57.56 per person for the six months trial period, please see the Table 15, 'MI, adjusted' model) was not statistically significant after adjusting for covariates. This indicates the large variations between individuals' health care resources use and uncertainty surrounding this variable. There were no statistically significant differences in the utility values between the Intervention Group and Control Group at both baseline and six months

follow-up, leading to no difference in QALY gained between the groups. Despite the close to zero mean QALY gained, the probability that the intervention might be cost-effective was 34%, indicating uncertainty in the estimation. This uncertainty was visualised in the cost-effectiveness plane which showed the bootstrapped pairs covered all of the four quadrants.

Uncertainty in the ICER estimates results from both the uncertainty around the incremental cost and, primarily, the incremental QALY. There is a small chance (13.4%) that the intervention is a cost-saving strategy from the public sector perspective. This is due to the combination effect of small intervention costs and the possible reduction in resource use by the intervention. The CEAC shows that the probability that the intervention might be cost effective is almost constant with the increase of the willingness to pay threshold, suggesting that the uncertainty is mainly surrounding the QALY gained. The ICER had a very large magnitude with a negative sign, which can be attributed to the close to zero negative mean QALY gained in the Intervention Group compared to the Control Group. This also leads to a wide confidence interval for the estimate of ICER as a result of the relatively large variation of the QALY gained compared to its mean estimate.

8.6.1 Sensitivity Analysis Results

The sensitivity analysis showed that when an alternative business model was used or the cost of sensors was reduced, the incremental cost of the intervention became negative. The probability of the intervention being cost effective increased to around 60%, however this may not be considered as a dominant scenario given the "benefit" of the intervention on utility was still negative and unclear.

A potential ceiling effect of EQ-5D-5L was observed in this population, with over 40% and 30% of participants responding *"full health"* at baseline and six months respectively. This may limit the sensitivity of this instrument for assessing the potential benefit of interventions such as this in public health settings as there is little room for utility values to get higher in a generally healthy population.

8.6.2 Cost Benefit Analysis/Net Cost Model Results

Adopting an 'employee' perspective, there was a net cost saving for the employers due to reduced absenteeism in the Intervention Group. This net cost saving ranged from £66 to £735 depending on the wage rate employed. By exploring the uncertainty surrounding this net cost saving estimate, we determined that the probability that the intervention is cost-saving at current intervention cost (£55.68) is highest (=64%) in the high salary group and lowest (=57%) in the low salary group. If the intervention cost reduces in the future, with the aforementioned alternative business model or upgraded technology for the sensors, the probability to be cost-saving will increase and be capped at 65.4% (the intercept of the three lines with y-axis in Figure 5). This capped value is due to the great uncertainty (p=0.62) surrounding the benefit in terms of the difference in absolute absenteeism rates between the groups.

A number of issues arising from the study merit further research from a health economic perspective. Although we used utilities directly elicited from participants using the EQ-5D, we observed small positive benefits in terms of well-being scores in the Intervention Group which were at odds with the negative impact on the primary outcome. Studies with longer term follow-up could usefully explore further how to aggregate or reconcile the various health and productivity effects of PA programmes in workplace settings, especially since, in the longer term, they may impact on the individual and the employer in different ways. To date, there have been few attempts to model the impact of workplace wellness programmes that are based on financial incentives. Basu and Kiernan[143] constructed a novel modelling framework that identified how to optimize the marginal return on investment from workplace programmes that were based on incentivizing behaviour change (including those targeting PA). By integrating "demand curves" that captured individual differences in response to any given incentive level with employee demographic and risk factor data (from the US) they demonstrated that an employer (with ~3000 employees) would see optimal return on investment by offering incentives of $367 per year (i.e. somewhat larger than in the PAL Scheme) but that this return on investment (ROI) was very sensitive to self-selection (if already active participants enrolled). Even with the optimally designed programmes, the actual degree of behaviour change varied widely, with most participants still undertaking low to moderate levels of PA despite the

incentive; and >70% of participants not engaging in PA on a weekly basis (in other words, they would be intermittent participants). Thus, it is right that future research in this area take a broad perspective and attempts to integrate individual and employer perspectives.

8.7 Interpretation and implications of the Behavioural Economics Analyses

It was shown that participants were generally prepared, prior to the intervention, to receive financial incentives to increase their PA, which accords with the findings of Giles et al.[24] who also observed that the use of financial incentives was acceptable to the public. Further, we found that high-value incentives were preferred over low-value incentives, which is consistent with findings reported in Farooqui et al.[144] based on a sample of 802 Singaporean adults. However, the opposite trend was reported when UK respondents were asked to decide on the provision of financial incentives from public funds.[24] Next, the stated incentive levels in our study (£2.80 per hour) are small but are comparable to the levels used, (although not empirically justified), in some other PA promotion schemes {($4.27 per hour,[145]) and in a weight loss programme ($3.75/week}.[146] Although our trial reported negative effects on PA, having utilised DCE results to determine the level of the rewards, we suggest that it is still useful to incorporate the stated preference method into the design of similar schemes, as argued by Tinelli et al.[147] to better understand respondents' preferences for incentives and to increase their chances of demonstrable effectiveness and cost-effectiveness.

Physically inactive participants, especially those who were highly inactive (daily steps less than 5,000) at baseline, were more susceptible to early non-usage attrition. Moreover, physically-inactive participants increased PA less at the 12 month follow-up, compared to physically-active participants. Because of these trends, physically-inactive participants still had lower PA at both six and 12 month follow-up than those who had already been active at baseline. This phenomenon can be (partly) explained by our analysis indicating that inactive participants would require substantially higher financial incentives to increase their PA. It may also be the case that they had relatively higher discount rates which were significantly and negatively associated with PA.

Future PA interventions should target especially the highly inactive population who were revealed to be the most difficult to be incentivised for behavioural change.

Our sample manifested relatively high discount rates, impling that incentivisation of our sample was rather difficult and our sample was present-biased implying that participants with high discount rates are more likely to pursue immediate gratification rather than execute their salutary intentions for their future selves. Since these behavioural economic experiments were undertaken among participants who remained in the study at the six month follow-up, incentivisation might have been even more difficult among the drop-outs because of *inter alia* even higher levels of discount rates. Next, although discount rates were not associated with behavioural change, i.e. the actual increment of PA at the 12 month follow-up compared to six month follow-up, higher discount rates did result in significantly lower PA at the six month follow-up. Consequently, reducing a participant's discount rate might potentially promote the participant's PA. Recent studies suggest that discount rates can be altered or modulated by therapeutic cognitive, behavioural, or structural environmental manipulation.[148]

Higher loss aversion was associated with a larger PA increase at the 12 month follow-up compared to the six month follow-up. This reconciles with the idea of 'deposit contracts' asking participants to deposit forfeitable money which can only be redeemed if a certain behavioural change target is met. It was already reported that smoking cessation programs were more effective in deposit-based programs than reward-based programmes.[149] As a result, utilising the concept of loss aversion could be a useful part of interventions to promote PA.

8.8 Interpretation and implications of the Mediation and Moderation Analyses

The main results of the trial were negative at six months and null for the primary outcome at 12 months and in order to explore potential reasons for these results, we carried out a range of mediation and moderation analyses as specified in our protocol.

Out of 19 tests of intervention effects on mediators (i.e. tests of the significance of the relationship between group assignment and the mediator), nine were significant

(n=9/19; 47%). Of 19 formal tests for indirect effects (i.e. the multiplication of the coefficient of the path relating group assignment to the mediator and the coefficient of the path relating the mediator to PA), three were significant (n=3/19; 16%). However, it should be noted that p-values have not been adjusted for multiple testing. This finding suggests that whilst the intervention was effective in changing some mediating constructs targeted by the intervention, mediator changes were generally not translated to changes in PA behaviour. For constructs showing significant mediating effects on PA behaviour, indirect effects were in the opposite direction to the overall intervention effect on PA. Specifically, whilst the total effect of assignment to the Intervention Group (versus Control Group) on PA behaviour at six months was **negative**, indirect effects through significant mediators were in the **positive** direction. Therefore, the intervention's negative effect on PA at six months may have been mitigated by modest PA increases achieved through increases in some of our measured mediators. However, it is apparent that the results of these tests of the behaviour change theories on which the intervention was designed were not sufficient to explain the overall change in PA behaviour at six months.

In summary, none of the putative mediators of initiation (measured at baseline and four weeks) significantly mediated the relationship between group assignment and PA behaviour change at six months. However, there were significant increases for Intervention Group participants compared to Control Group participants in intentions, social norms, identified regulation, integrated regulation and intrinsic motivation between baseline and four weeks. Therefore, it is apparent that whilst assignment to the Intervention Group led to increases in some putative mediators of initiation, these increases were not related to PA behaviour change at six months. Two potential reasons are: (1) changes in these mediators of initiation do not induce change in positive PA behaviour; (2) changes in mediators of initiation are not carried through to PA behaviour change at six months (i.e. perhaps PA behaviour change caused by changes in these mediators occurred at an earlier time-point).

Several of the mediators of maintenance (measured at baseline and six months) significantly mediated the relationship between group assignment and PA behaviour change at six months (i.e. integrated regulation, intrinsic motivation, habit) in the

positive direction (i.e. there were increases in these constructs between baseline and six months for intervention participants compared to controls and changes in these mediators were positively related to changes in PA behaviour). There were also significant increases for Intervention Group participants compared to Control Group participants in identified regulation, integrated regulation, intrinsic motivation and habit between baseline and six months. This was notwithstanding the overall negative intervention effect, suggesting that participants who were able to increase their levels of these constructs (e.g. by focusing on the scheme's self-regulation, social or environmental aspects rather than the financial incentives or scheme limitations) experienced less of a decline in PA. There is support in the theoretical and empirical literature to suggest that integrated regulation, intrinsic motivation and habit are important mediators of PA change and for maintenance of behaviour change.[150–157] However, results of the mediation analyses suggest that the behavioural changes observed in the present study were not achieved by means of the mechanisms posited by the behaviour change theories on which the intervention was designed. Therefore, these behaviour change theories were not sufficient to explain the overall change in PA behaviour. Finally, the fact that changes in mediators of initiation were not translated to changes in PA behaviour whilst changes in some mediators of maintenance were related to changes in PA behaviour implies that participants were in the behavioural maintenance phase at six months. This is in line with the definition of behaviour change maintenance adopted in a recent systematic review investigating maintenance in PA interventions targeting young and middle-aged adults.[158]

These considerations highlight what has been emphasized by Threfall et al.,[159] that interpretation of intervention effects (whether positive, negative or null) is rarely a-theoretic. In the sections that follow, we offer some theoretical contextualisation of these moderation and mediation findings.

8.8.1 Perception of the Availability of PA Opportunities in the Workplace Environment

Perception of availability of PA opportunities in the workplace environment was a significant moderator of the relationship between group assignment and PA change between baseline and six months. Specifically, higher levels of perceived availability

of PA in the workplace environment were associated with a greater decrease in PA between baseline and six months for an Intervention Group participant compared to Control. Perceptions of workplace environment safety was a significant moderator of the relationship between group assignment and PA change between baseline and 12 months, with lower levels of perceived safety associated with a greater decrease in PA between baseline and 12 months for an Intervention Group participant compared to Control Group participant. Previous studies have also supported the idea that perceptions of the environment are important moderators of PA behaviour change.[160,161]However, the direction of the moderating effect of perceptions of availability of PA in the workplace environment on PA at six months may appear counterintuitive (i.e. the intervention's negative effect on PA was stronger when perceptions of availability were high). Perhaps participants who perceived that there were plentiful opportunities for PA around their workplace felt less reliant on the intervention to identify opportunities for PA or were less willing to confine their PA within the limited geographical range of the study's sensors. This result merits exploration in future research.

8.8.2 Participant Discount Rate

Participant discount rate was a significant moderator of the relationship between group assignment and PA change between baseline and 12 months, with higher (versus lower) discount rates associated with a greater decrease in PA between baseline and 12 months for an Intervention Group participant compared to Control. Thus, when participants were more impatient (i.e. preferred an immediate reward with a slightly lower monetary value to a reward with a short delay and a slightly higher monetary value) there was a greater decrease in PA for Interventions versus Controls. This pattern of results may make intuitive sense for the PA outcomes at 12 months (if, for a subgroup at least, increases in intrinsic motivation were not sufficient to compensate over the longer follow-up period when the extrinsic rewards had been removed after the six month intervention) but cannot explain the decline in PA at six months (the primary outcome).

8.8.3 Financial Incentive and Motivation

Self-Determination Theory (SDT) posits that increasing intrinsic motivation (i.e. based on an individual's own values) leads to more sustainable behaviour change than can be achieved by inducing extrinsic motivation.[162] The finding of increases in intrinsic forms of motivation (i.e. identified regulation, integrated regulation and intrinsic motivation) for Intervention Group participants compared to Control Group participants at four weeks and six months, and no intervention effect for financial (extrinsic) motivation at four weeks, has important implications for some of the main contentious issues highlighted in the literature on the use of financial incentives for achieving behaviour change. Firstly, extrinsic (e.g. financial) rewards are not thought to produce long-term behavioural changes since withdrawal of the reward causes the behaviour to be extinguished.[163] Secondly, the use of financial incentives has been thought to have a 'crowding out' effect on intrinsic motivation for behaviours which are already internalised.[162] For example, the qualitative data from our study showed that some participants, who were already physically active, reported that they joined the scheme to receive rewards for behaviour that they already engaged in. Based on this argument we may have expected to observe decreases in intrinsic motivation for Intervention Group participants versus Control Group participants. Conversely, the review conducted by Promberger and Marteau[162] suggested that there was no current evidence to support the notion that extrinsic rewards or incentives undermine any intrinsic motivation in relation to health-related behaviours. In line with the findings of that review, our results suggest that using financial incentives in a behaviour change intervention does not necessarily diminish intrinsic motivation and can actually increase it if rewards are delivered as part of a complex behaviour change intervention with multiple components.[29]

The general consensus is that financial incentives may be useful in acting as a 'kick-start' to PA but are less useful for promoting long-term behaviour change. The qualitative work in our study would attest to this, with some participants commenting on the positive aspects of getting people out walking, and to do a little extra every day. Under Learning Theory, previous behaviour is predictive of subsequent behaviour.[164,165] This may be explained by the persistence of factors (e.g. thoughts or rewards) which cause the behaviour to be repeated, or by the formation and

strengthening of habits (i.e. learned behavioural responses to certain stimuli and repetition in similar contexts).[155] Thorndike's Law of Exercise states that the strength of the connection between a behavioural response and a given situation is directly related to the number of times the connection has occurred previously and to the potency and duration of those connections.[166] By the use of contingent reinforcement it was hypothesised that repetition of the behaviour-reward cycle would lead to workplace PA becoming a 'learned' behaviour. According to SDT, external regulation occurs when the individual engages in the behaviour to gain reward (e.g. financial incentives) or to avoid punishment by others.[167] However, the reward offered must be deemed worthwhile by them personally. For example, previous studies have found that monetary value[131] or type (e.g. individual versus group-based)[168] of the reward can impact its effectiveness for behaviour change. There was no overall effect of group assignment on changes in financial motivation for participants in this study and focus group interviews indicated that, for some participants at least, vouchers failed to live up to expectations (e.g. insufficient monetary value, lacking in variety, undesirable goods or services, inconvenient expiration dates, or that desirable vouchers 'ran out'). Therefore, the incentive may not have been attractive enough to entice participants to start (and by repetition, to 'learn') PA behaviour in the first place. This is despite the fact that these aspects of the study were designed using 'best practice' of pre-intervention development focus groups with participants to discuss all important features of the intervention (e.g. types of rewards, appropriate level of rewards, and placement of sensors).

8.8.4 Intrinsic Motivation (Integrated Regulation and Intrinsic Motivation) and PA

The results of mediation analyses suggested that intrinsic forms of motivation (i.e. integrated regulation and intrinsic motivation) were important mediators of the relationship between group assignment and PA behaviour change at six months. SDT, for example, suggests that individuals experience greater vitality, self-motivation and wellbeing when the basic psychological needs for autonomy, competence and relatedness to others are fulfilled within the social context.[169] Focus groups highlighted that participants wanted more options with respect to rewards, and how and when they engaged in PA. Perhaps participants did not perceive a high degree of autonomy

support (i.e. opportunities for choice with minimal use of pressure or demands)[170] in the scheme to fully allow internalisation of PA behaviour.

8.8.5 Habit Formation and PA

Mediation analyses indicated that the strongest indirect effects of group assignment on six month PA occurred through habit formation. The emphasis placed on the intervention's self-regulation techniques supports the observation that people develop habits after a period of successful self-regulation. The process of habit formation incorporates an exchange between automatic and self-regulatory processes within the context of a range of external and internal stimuli.[157] Habits are formed when the behaviour becomes regular, automatic and ingrained in the self-concept.[17,152] Prompts in the social (e.g. email reminders sent by the study team and having work colleagues participating in the intervention) and physical (e.g. seeing sensors surrounding the workplace) environments could have aided habit formation. Moreover, focus groups at 12 months post-baseline sought to establish if participants were still engaging with PA and, if so, for what reason. The underlying rationale was to establish if behaviour change maintenance had occurred. Many of those who participated in the focus group discussion at 12 months commented about how they had continued to incorporate PA into their working day; from utilising the walking routes provided by the PAL Scheme, to walking to meetings in short proximity rather than driving and generally looking for new ways to fit PA into their day. This is also in line with several habit theories[171–173] which hypothesise that environmental cues associated with prior and new behaviours determine whether the new behaviour is maintained. Moderation analyses also indicated that participants' perceptions of the environment impacted on the effectiveness of the intervention in line with theory which suggests that supportive social and physical environments facilitate behaviour change maintenance by lowering the opportunity cost of the behaviour.[157] When participants utilised the scheme's self-regulatory, social and environmental aspects to increase their PA habits, this seemed to mitigate the overall negative impact on PA.

8.8.6 Discount Rates and Intervention Effectiveness

Moderation analyses also indicated that participants' derived discount rates impacted on the "effectiveness" of the intervention, with higher discount rates (i.e. higher levels of impatience) associated with greater decreases in the primary outcome for Intervention Group participants. This is line with the findings of the study's additional behavioural economics analyses (Section 8.2) and also with the qualitative work from which we noted how vexed some participants were upon not achieving enough points to redeem a certain voucher or, that a priority voucher had been time expired (and withdrawn from offer) before they had collected enough points, and so they subsequently 'gave up'. Although we cannot infer causality, these associations suggest that when designing financial incentive based interventions, future studies should consider using information on participants' discount functions to tailor incentive structures (e.g. timing, contingency and value) in order to achieve behavioural initiation and maintenance.[174] Other intervention strategies (e.g. goal-setting,[175] provision of health information[176]) could also be usefully targeted depending on information available on individual participants' time preferences.

The concept of adaptive versus static goals has also been explored in recent PA literature, whereby Zhou and colleagues[177] employed automated machine learning-based personalised daily step goals delivered via a mobile phone app. Their study sought to differentiate PA behaviours in those who received personalised daily step goals compared to a control group who had a constant step goal of 10,000 steps. Whilst the PAL Scheme incorporated personalisation via the website where users could track and monitor progress, a more advanced method for deriving adaptive goals might be beneficial for PA. Adams and colleagues,[178] for example, also found that dynamic adaptive goal setting may outperform static approaches for goal setting for PA (steps), with slower decreases in PA steps from baseline to intervention end. The use of adaptive goals may be particularly useful in a workplace setting. Where some employees may be able to incorporate PA into their commute to work, others - as we identified from the qualitative work -may find the workplace to be their only opportunity to be physically active. Recent research on goals does not generally account for the potential time discounting confound. "Nearer" outcomes are more motivating, and are proposed to explain differences in motivation over time in human

behaviour as well.[179] One interpretation is that more temporally distant rewards are less motivating specifically because delayed rewards are discounted (and potentially also riskier). In this view, goal-gradient behaviour (exerting more effort for sooner rewards, depending on how far away you are from your goal) is another source of evidence for time discounting. Urminsky, Goswami and Lewis (2014)[180] reported discount rates which were lower when separately accounting for goal gradient effects. With this in mind, the effects of adaptive goals, attendant on differing feedback schedules, could potentially be more or less motivating in people with different discount rates and may cater for patterns of behaviour resulting from changes in day-to-day circumstances. This could usefully be the subject of future research.

8.8.7 PA Self-Efficacy

Whilst constructs such as PA self-efficacy are frequently shown to be important mediators or predictors of PA behaviour[181–186] PA self-efficacy and phase-specific self-efficacy (i.e. maintenance and recovery self-efficacy) were not shown to be significant mediators in the present analysis. There is emerging literature disputing the causal role of self-efficacy in behaviour change[187] and several systematic reviews show that the supporting evidence is lukewarm at best. Rhodes & Pfaeffli[188] found limited evidence for self-efficacy and outcome expectations, but stronger support for self-regulatory processes and social support as mediators of behaviour change (focusing on initiation) in PA intervention studies targeting adult non-clinical populations. Notably, whilst self-regulation constructs have not been examined as putative mediators in this study, the qualitative analysis revealed that participants appreciated the self-regulation (e.g. self-monitoring, planning, goal-setting) and social aspects of the PAL Scheme. Another review conducted by Teixeira and colleagues[151] revealed that PA self-efficacy was related to short-term PA outcomes (i.e. at less than 12 months) in ten of 15 tests (67%), whilst outcome expectations were not related to PA outcomes in one test. Importantly, these results did not rely on formal mediation tests, highlighting the need for robust mediation analysis in PA intervention studies.

8.8.8 Engagement and Non-Usage Attrition

The results of GLS regressions showed that engagement in specific components of the intervention (i.e. the frequency of using the website self-monitoring and feedback component, the percentage of earned points redeemed, the frequency of using website discussion forums) was significantly related to PA at six months. These findings suggest that when participants focused more on the intervention's positive aspects (e.g. self-monitoring), this mitigated the intervention's overall negative impact on PA behaviour. Previous research also shows that self-regulation techniques and self-monitoring are useful strategies for changing PA behaviour.[14,35,188–192] Redeeming a higher proportion of earned points for rewards was also associated with slightly higher PA levels at six months. Therefore, it appears that when participants found the financial incentives desirable and redeemed their accumulated points, this also mitigated the overall negative impact of the scheme on PA behaviour. Previous studies have also shown that in order for financial incentives to be successful in inducing behaviour change, the reward on offer must be deemed to be valuable to the individual participant. For example, monetary value[131] or type (e.g. individual versus group-based)[168] of reward can have an impact on effectiveness for behaviour change, and this was alluded to in feedback from focus groups. Examination of the relationship between engagement in different aspects of the intervention and PA at six months showed that a higher frequency of accessing the discussion forum component of the website was associated with a significant decline in PA at six months. Discussion forums were included on the PAL Scheme website as a means of providing social support for behaviour change. For example, it was expected that participants would use these forums to contact researchers and interact with other participants to support behaviour change. However, this component of the scheme was primarily used by participants as a vehicle to report technical issues and queries, rather than as a means of engaging with other users in a support network. Therefore this finding may indicate that participant frustration with some of the intervention's perceived limitations contributed to the overall negative impact on PA. One previous study that also found negative intervention effects on PA behaviour concluded that reduced support for the intervention over time was a contributing factor.[193]

8.8.9 Intervention Fidelity

The literature highlights the importance of intervention fidelity and of ensuring that variations from the original design can be assessed since non-significant findings may be attributable to intervention delivery rather than intervention design.[193] Carroll et al.'s theoretical framework provides the means of measuring intervention fidelity and understanding its role in the process of intervention implementation.[194] For example, the authors highlight that several potential moderators of the fidelity with which an intervention is implemented include intervention complexity, use of facilitation strategies, quality of intervention delivery and participant responsiveness. Future intervention studies could potentially improve intervention outcomes by improving intervention fidelity with reference to such theoretical frameworks. Finally, systematic reviews also show that the effectiveness of web-based interventions may be enhanced by including various modes of intervention delivery (e.g. text messages, email),[195,196] which the PAL Scheme did (e.g. through the use of email prompts). However, we have been unable to include data on whether participants received and read these emails in our engagement analysis. Future work may seek to delve into further analyses around the delivery, open and click rate of emails relating to an intervention.

GLS regressions also showed that use of particular intervention components was associated with increases in some mediators between baseline and six months for Intervention Group participants. A higher frequency of using the PAL Scheme PA monitoring system (i.e. behavioural practice) and a higher frequency of accessing the study website was associated with higher levels of integrated regulation, intrinsic motivation and habit between baseline and six months (i.e. the constructs for which mediation analyses showed significant positive indirect effects). This suggests that a higher level of overall engagement with the intervention (as opposed to engagement with specific intervention components) was related to higher levels of these constructs for Intervention Group participants. Accessing a higher number of intervention components on the website at least once was associated with increases in identified regulation and integrated regulation, although there was no association with PA levels. Previous work has also identified that overall intervention engagement, and also greater use of intervention components was associated with increases in the primary outcome.[196,197]

Further evidence for the importance of the intervention's self-regulation component was provided by the results of the study's engagement analysis (i.e. a higher frequency of accessing the website's feedback and monitoring component was also associated with increases in integrated regulation and intrinsic motivation) which broadly support the findings of the mediation and qualitative analyses. Previous systematic reviews of web-based interventions also highlight web-based self-monitoring as a potentially effective technique.[197] In comparison, engagement with the financial incentive component of the intervention (i.e. redeeming a higher proportion of earned incentives for rewards or a higher frequency of accessing the website's rewards component) was not found to be related to levels of identified regulation, integrated regulation and intrinsic motivation. Thus, we have further supporting evidence to conclude that the use of financial rewards does not necessarily 'crowd out' or diminish intrinsic motivation when delivered as part of a complex multi-component behaviour change intervention.

8.8.10 Scheme Non-Usage Attrition

High levels of non-usage attrition in the present study may, to some, reflect dissatisfaction with the intervention's perceived limitations contributing to the decline in PA outcomes at follow-up. It should also be noted that the definition of non-usage attrition that we employed (i.e. occurring at the time of the first two week lapse from intervention use) in the present study may be viewed as somewhat arbitrary. Though several studies of web-based PA interventions have adopted this definition, it may be less applicable to the analysis of non-usage attrition in workplace interventions for which a two week lapse from intervention use may occur for many different reasons: e.g. if a participant is on annual leave, out of the office on work-related travel, or is absent from work, for two weeks or more through illness. Time to non-usage attrition for use of the PA monitoring system and use of the website was shown to be significantly predicted by some of the baseline socio-demographic, psychosocial and environmental variables (e.g. age, marital status, health status, financial motivation, identified regulation, recovery self-efficacy and perceptions of the workplace environment) in multivariable Cox regression models. These results are consistent with findings in previous intervention studies, demonstrating that older participants

(versus younger),[110,198–202] unmarried participants (versus married/co-habiting)[203] and participants with higher reported health status at baseline (versus lower health status)[204] have decreased risk for non-usage. Individuals with higher levels of financial motivation may have been encouraged to continue participation in the intervention over time in order to continue benefitting from financial rewards. Existing theory suggests that identified regulation and recovery self-efficacy are constructs which are important for long-term behavioural maintenance.[169,170,204–208] For example, it makes sense that individuals with higher recovery self-efficacy should experience a longer period of consistent intervention use before encountering their first two week break since they are quicker to recover from a lapse. Notably, the results of our behavioural economics analysis (see Section 8) indicated that individuals who were more present-biased (i.e. who had less self-control) were also at higher risk for stopping intervention use. Furthermore, some subgroups of participants who were found to be less present-biased in the behavioural economics analysis were also found to be at reduced risk for non-usage attrition in survival analyses (e.g. older participants, single participants). This may point to possible interactions between age or marital status and levels of present-bias.

8.8.11 Mediation and Moderation Analyses: Implications for Future Research

Future PA intervention studies should make use of self-regulation techniques (e.g. self-monitoring, goal-setting), with social and environmental prompts. This may promote habit formation as the new PA behaviour, originally carried out consciously, becomes more automatic. This should also allow for the internalisation of the behaviour, enabling it to be more sustainable long-term. Monitoring of behavioural outcomes is an additional self-regulation technique which should be explored, as it can potentially encourage participants to focus on their satisfaction with more internal outcomes of behaviour change (and therefore to maintain it long-term). As Bouton aptly summarised, ‘successful learning of a new behaviour does not permanently replace an earlier one’.[209] Moreover, behaviour change context is also key to maintained change and as such, new behaviours (in the case of the current work, increased PA) may be easily disrupted. Whilst the PAL scheme utilised cues with regard to emails and environmental cues with the placement of sensors, more specific work on the types of contextual cues to aid with behaviour change should be explored.

To promote the internalisation and maintenance of PA behaviour change, future intervention studies should support all of the basic psychological needs posited as important in SDT (i.e. competence, autonomy, relatedness to others).[169,170,206] Studies should also examine how the social and physical environment can influence participants' experience of PA interventions, and behaviour change (e.g. enjoyment, autonomy). Finally, behaviour change theories should be developed which accurately capture the process of PA behaviour change for specific individuals. This is reliant upon well-designed intervention studies investigating causal mechanisms by means of mediation and moderation analyses, and exploring how these constructs operate together.

8.8.12 Engagement and Non-Usage Attrition: Implications for Future Research

Participant engagement and non-usage attrition are challenges faced in public health intervention research. Therefore, future intervention studies should measure levels of engagement and non-usage with a view to making recommendations for retaining groups of participants who are at the highest risk for non-usage and lack of intervention engagement. This is particularly important for studies of web-based interventions which are known to be particularly susceptible to non-usage and lack of participant engagement.[107] Better guidelines on how to measure intervention engagement are needed. Future intervention studies should also consider how to ensure that those with high risk for non-usage attrition keep engaging in the intervention. This is important to ensure participants receive the full potential benefits of the intervention.

Our engagement analysis indicated that when participants found the financial incentive attractive and redeemed their points for rewards, there was a positive impact on PA. This suggests that offering financial incentives can potentially encourage participants to increase their PA behaviour. However, future research is warranted investigating the ability of financial incentives to initiate PA and facilitate long-term behaviour change. There is little evidence on optimal monetary values, varieties (e.g. retail vouchers, more generic gift cards, charity donations, group rewards) or timing of incentives, for achieving health behaviour change. Interventionists should consider how to retain participants who are initially less financially motivated in studies whose

main component is the offer of financial rewards. Moreover, where trials of complex interventions such as the PAL Scheme result in a negative primary finding, future work needs to determine which element of the intervention 'does not work' or whether there is a "failure" that signifies sub-optimal fidelity, dose, implementation or of the underlying programme Logic Model. For studies of PA interventions requiring behavioural practice in the outdoor environment (as was the case for the current study), participants' perceptions of the environment with respect to PA are important influencing factors which should be mitigated to optimise engagement and minimise non-usage.

8.9 OVERALL LIMITATIONS AND STRENGTHS

There is a long standing belief that positive results are favoured by scientific journals and that this may contribute to "publication bias". On the other hand, some journals claim now to select articles for publication based on their contribution to the literature and welcome null results that challenge conventional wisdom or prior expectations.[210] The results from our trial certainly challenged prior expectations. However, it is notoriously hard to disprove any hypothesis, and so negative studies must have the precision and strength of design to be reasonably persuasive. The revised sample size and power calculations for our trial were approved by NIHR, and the effects size on which they were premised was within the category of "moderate" effect size observed in past PA interventions.[211] Even though attrition and loss to follow-up in our study was higher than predicted, (though comparable to some other published workplace PA trials, such as those of van Dongen et al.[140] and Puig Ribera et al.[141] we might have concluded that accepting the null hypothesis at 12 months was justified. However, the primary outcome was specified at six months and the non-significant difference in mean steps/day at 12 months (~500 steps) was, arguably, still of a magnitude which could be of public health importance. We would probably have claimed as much had the direction of the intervention effect been positive. If the significant negative effect at six months is indicative of a true negative effect *of this intervention in this setting*, one might still ask whether the result is generalizable. Is this recruited sample of public sector office workers in Northern Ireland representative of office workers elsewhere in the UK? We have no reason to conclude that it isn't,[14–16,34] but are mindful of the modelling undertaken by Basu and Kiernan[143] which demonstrated that two key factors

impacting the success of workplace-based financial incentives for behaviour change are (i) *who* participates and (ii) the *levels* of incentives. Thus we need to examine carefully the strengths and limitations of our mediation and moderation analyses to gain a better understanding of our findings.

The authors acknowledge the well-known self-report biases associated with the use of questionnaires. Where possible, well validated questionnaires were used. There are some limitations regarding the use of a pedometer for measuring our primary outcome. These include the inability to consider wear time in the processing of the data. The use of <250 steps/day as an indicator of non-wear could have removed participants who were inactive, or included participants who wore the monitor for a short period and erroneously classified them as inactive. The use of a pedometer meant we were not able to detect when in the day the PA was accumulated. Using a research grade accelerometer could have addressed some of these limitations. Obviously careful consideration was given to our choice of primary outcome and the appropriate measurement instrument. This inevitably involved a trade-off between reliability, efficiency and the cost of PA measurement in a sample of this size (n=1390 was our initial sample size). Firstly, we wanted a measure of total PA - which is captured by the pedometer - not just MVPA. However, the estimated cost for using accelerometers would be £400,000 (Actigraph GT3X ~£250 each) compared to £75,796 for the pedometers. After careful consideration, the study team agreed that pedometers provided an objective, valid and reliable measure of PA at a reasonable cost for this sample size.

We were able to specifically capture workplace PA using the remote sensing monitoring system (sensors and keyfobs) and the GPAQ which incorporates a workplace PA domain. However, we used a well-validated pedometer, followed a standardized measurement protocol (including sealing the pedometer to prevent reactivity), and we supplemented the data with the GPAQ and daily PA monitoring using the remote sensing system. A broader question remains whether a maximum of seven days of physical activity measurement at baseline, six months and 12 months is sufficient to adequately document behaviour change. However, at present this is the most commonly used approach.[151]

The mediation analyses controlled for baseline values and improves on previous studies which have usually relied on cross-sectional data. We followed recommended practice and used bias-corrected bootstrap CIs in assessing the significance of the indirect effects.[113,114]Studying the relationship between intervention engagement, PA outcomes, and mediators is a useful adjunct to other aspects of the trial's process evaluation that can help us gain additional insights and perspective on the overall results. Previous studies have investigated whether engagement (or adherence) is predictive of behavioural outcomes[198,212–214] without consideration of how they may relate to the mechanisms of behaviour change. A strength of our analysis is the examination of several distinct indicators of engagement with different components of the intervention. The wealth of our qualitative data supporting and complementing the quantitative mediator analysis is a key asset that helps us look at "pictures as a whole" to find the missing pieces of the health improvement intervention 'jigsaw' described by Tannahill and Kelly.[135] Qualitative analysis is to a certain extent, subjective and reliant on the researchers' interpretation of the data.[215] However, all transcripts in our trial were independently coded by two members of the research team (one member was directly involved in facilitating the focus group discussions).

Our study was not powered to detect changes in mediating variables and the power to detect changes in some mediators may have been low. Mediators were based on self-report questionnaire items and may have been subject to measurement error. While all mediator measures were based on previously used validated measurement instruments, the internal consistency of some mediators was low. Inspection of residual plots revealed some analyses may have violated normal-distribution assumptions, and sensitivity analyses using logistic regressions showed there were changes in the significance of results for some mediators. Whilst previous research on engagement in intervention studies has compared engagement between an intervention group and a comparison group,[111] we were unable to include comparable engagement (e.g. recording of daily activity and web usage) data from our Control Group because of the nature of the Waiting-List Control. Thus we are cautious not to over-interpret any comparison between the mediators of engagement and those disclosed by our main effects analysis. Furthermore, whilst measures of the frequency of hits on different

sections of the website may indicate the participant's level of interest in a specific intervention component, they do not capture how well the participant processed the information.

Another potential limitation is that only baseline variables were investigated as predictors of non-usage attrition. Possibly, other factors such as changing levels of satisfaction with the intervention, or of PA behaviour itself, may have influenced attrition rates. However, our goals were broadly similar to previous studies of predictors of non-usage attrition,[201] the aim being to better describe the groups who continue engaging in an intervention at enrolment. Our definition of non-usage attrition (i.e. occurring at the time of the first two week lapse from intervention use) may have contributed to the high levels of non-usage attrition observed due to some participants potentially taking a two week period of annual leave or other absence from work. In general, defining non-usage attrition in workplace interventions is difficult as any period of non-use may not be indicative of an intentional lapse from using the intervention but may indicate temporary absence from the workplace. Finally, it should be remembered that mediation analysis, and analyses pertaining to intervention engagement and non-usage attrition, are exploratory. Because of low power and multiple testing, they need to be interpreted with caution.

There are two key strengths to our behavioural economics analysis. First, two stated preference methods (i.e. Contingent Valuation, and DCE) were used to investigate the optimal levels of financial incentives required to encourage changes in PA, whereas previous studies on financial incentives have eschewed such empirical *a priori* groundwork into the appropriate levels of financial incentives. Our approach also resonates with the argument that a better understanding of participant preferences through DCEs is useful for developing or more effectively delivering interventions.[147] Next, in these adjunct behavioural economic studies, we have used an objective measure of PA (i.e. pedometer measurement) to avoid the biases and inaccuracies of self-reported measures. Also, using real incentives in the form of cash payments in the economic experiments is widely interpreted as helping to ensure that participants respond carefully and truthfully in the experiments, by comparison with hypothetical elicitations where the participant burden is not related to real rewards. In addition, discount rates, present-biasedness, risk preferences and loss aversion were jointly

estimated using a tried and tested procedure to eliminate the bias that afflicts analyses that attempt their separate estimations.[94]

However, the behavioural economics analysis also has several limitations. First, because of the nature of incentivised and experimentally-controlled economic experiments, they take approximately one hour to complete. Consequently, the experiments were not conducted during the baseline assessment to keep the participant burden to a minimum. However, the participants who had withdrawn from the study before the six month follow-up measurement (i.e. when the experiments were conducted), were not invited to participate in the experiments and this was specified in our ethics approval. This may have resulted in a biased sample[147] if the participants lost to follow up had systematically higher or lower discount rates. Next, the principle behind the time preference theory suggests that individuals who have higher discount rates and who are more present-biased are more difficult to be incentivised so that they should be provided with greater financial incentives to encourage them to commence behavioural change. This hypothesis was not tested in our study, but it opens new avenues for future research.

Table 33 provides an overview of the possible competing explanations for the results of the trial (negative primary outcome) that have been mentioned throughout the report. Broadly, these cover contextual factors, intervention design and methodological issues.

Table 33: Overview of competing explanations for results of the trial (negative primary outcome)

Competing Explanation	Source of Evidence	Strength of Evidence	Author Commentary
Contextual Factors			
Significant organisational restructuring	-Senior management interviews -Participant focus groups	+ supportive evidence	During our recruitment phase, a number of participating organisations undertook significant re-structuring due to the then current economic austerity, resulting in uncertainty regarding job security and job location; a time when employee health and wellbeing was at its most vulnerable. The impact of this was evident in our qualitative data which highlighted how motivation can be more usefully seen as a property of systems (incorporating technologies, organisation and action) rather than just of individuals
Seasonal effect	-All analyses adjusted for season	- evidence against	Our analyses included an adjustment for seasonal effects which would suggest that the six month difference in steps/day between the intervention and control groups not due to a seasonal effect
Intervention Design			
Suboptimal choice of incentives	-Participant focus groups -39% of points reimbursed for vouchers	+ supportive evidence	The types of vouchers available were based on those which were most popular in our pilot trial (same target population) and verified in pre-intervention focus groups with the target population
Suboptimal level of incentives	-Participant focus groups -39% of points reimbursed for vouchers	+ supportive evidence	The level of incentive (i.e. points value) was determined by a pre-intervention Contingent Valuation survey with all participants. The determined levels were then verified in pre-intervention focus groups with the target population.
Incorrectly specified logic model	-Mediation analyses	+ supportive evidence	Our results showed increased internal forms of motivation (i.e. identified regulation, integrated regulation and intrinsic motivation) for the intervention group compared to the control group at four weeks and six months.

			Our results demonstrated that there was no intervention effect for financial (extrinsic) motivation at four weeks. Whilst the intervention group showed increases in some hypothesised mediators of initiation, these increases were not related to PA behaviour at six months. However, hypothesised mediators of maintenance were related to PA behaviour at six months. As reported above, our results run counter to predictions of self-determination theory suggesting that there is a role for financial incentives and self-regulation interventions. However, the amount of variance explained by our measured mediators of behaviour change was low. Therefore, future studies would need to examine other potential mediators of behaviour change to shed further light on these associations.
Technical issues causing frustration in intervention group	-Participant focus groups	+ supportive evidence	There were occasional technical issues with the sensors and fobs which had a knock on effect on points accumulated. There were clearly some technical issues affecting intervention fidelity. Some of the sensors developed battery problems which affected the recording and accuracy of the PA behaviour captured, and required replacement.
Intervention too short to support behaviour change maintenance	-Author hypothesis	- evidence against	Our intervention was six months in duration which is in line with other PA interventions attempting to elicit and support PA maintenance (a period consonant with the mean in a recent review).
Boredom with intervention	-Participant focus groups - Usage data from the PA monitoring system showed that the minutes of activity recorded on the system declined over the six month intervention period (participants logging at least	+ supportive evidence	New rewards, walking routes, and double point's days were regularly introduced in an attempt to keep the format fresh and appealing but this did not improve workplace PA.

	10 minutes of activity via the PA monitoring system on 25% of all possible intervention days).		
Study participants in the intervention group could have felt patronised or infantilised by certain components of the intervention which may have been considered 'control mechanisms' by some. For example, steps being counted and recorded online, regular email prompts.	- Reviewer hypothesis	- evidence against	These specific intervention components are not "control mechanisms" but are evidence-based behaviour change techniques that were embedded within the intervention to help support PA behaviour change initiation and maintenance. The rationale for each intervention component is detailed in section 2.2.2.1 and the logic model (Figure 7). Further, there was also no evidence from our focus group discussions to suggest that the participants felt "patronised" or "infantalised". Rather our focus group findings suggested that participants valued the opportunity to self-monitor their PA behaviour and receive regular feedback (both evidence-based behaviour change techniques). They also found the maps very helpful in letting them know where the sensors were placed. However, it is possible that those who attended a focus group were more likely to be positive which is a common criticism of qualitative research.
Methodological Considerations			
Selection bias	- Reviewer hypothesis	- evidence against	Our eligibility criteria was very broad; participants were randomly allocated using suitable methods; cluster randomisation limited contamination between groups; the groups were balanced at baseline for key demographic characteristics; demographic characteristics of the recruited sample are similar to the demographics of public sector employees in Northern Ireland.
Physical activity measurement tool	- Reviewer hypothesis	- evidence against	A well validated, objective measure of physical activity was employed following a validated, standardised protocol.
Inadequate sample size	-Author hypothesis	+ supportive evidence	The revised sample size and power calculations for our trial were approved by the study funder, and the effect size on which they were

			premised was within the category of “moderate” effect sizes observed in past physical activity interventions.
Hawthorne effect in control group	- Reviewer hypothesis	- evidence against	At 6 months, the intervention group decreased mean steps/day by 947 steps, and the control group by 398 steps from baseline. At 12 months, the intervention group maintained a reduction in steps from baseline (decrease of 552 steps), with the control group returning to baseline levels (increase of 98 steps from baseline).
Confounding	-Author hypothesis	- evidence against	All analyses were adjusted for baseline values, randomisation stratum (size of participating organisation) and season with standard errors (SEs) corrected for clustering. Analysis was by intention to treat.

8.10 RESEARCH RECOMMENDATIONS

Several recommendations for intervention design, trial design and future research have been made throughout the report. Table 34 provides a comprehensive summary of these recommendations. Briefly, key methodological recommendations include accounting for negative findings and conflicting results between primary and secondary outcomes; the need to account for context; understanding causal mechanisms; and engagement and non-usage attrition. Intervention design recommendations include the use of financial incentives and self-regulation techniques; the design of financial incentives-based interventions; the role of SDT and the need to integrate individual and employer perspectives.

Table 34: Methodological and Intervention Design Recommendations

Consideration	Recommendation
Methodological	
Negative findings	Determine which element of the intervention 'did not work' or whether there is a "failure" that signifies sub-optimal fidelity, dose, implementation or of the underlying programme logic model.
Conflicting findings from primary and secondary outcomes	Need to develop methods on how best to balance positive and negative results when primary and secondary outcomes are discordant.
Taking account of context	Future interventions must take account of context in their intervention and evaluation framework.
Understanding causal mechanisms	Need for a framework based on consensus about how mediation should be measured and tested in trials of complex interventions. Such a framework should include the use of formal mediation tests, the embedding of evidence-based techniques for changing hypothesised mediators and the need to investigate constructs with particular relevance for initiation *and* maintenance of behaviour change.
	Future research should examine mediators of adverse effects so we can better understand unintended consequences and negative findings.
	Behaviour change theories should be developed which accurately capture the process of PA behaviour change for specific individuals. This is reliant upon well-designed intervention studies investigating causal mechanisms by means of mediation and moderation analyses, and exploring how these constructs operate together.
Engagement and non-usage attrition	Need to measure levels of engagement and non-usage to understand participants' non-usage patterns and the contributing factors with a view to making recommendations for retaining groups of participants who are at the highest risk for non-usage and lack of intervention engagement.
	Need to consider how to ensure that those with high risk for non-usage attrition keep engaging in the intervention.
	Need better guidelines on how to measure intervention engagement.
Intervention Design	

Use of financial incentives	Our findings suggest that using financial incentives within a complex behaviour change intervention with multiple components collectively does not necessarily diminish and may facilitate intrinsic motivation.
	Our findings suggest that the provision of financial incentives does not necessarily increase financial (extrinsic) motivation.
	Our findings run counter to predictions of self-determination theory suggesting that there is a role for financial incentives and self-regulation interventions.
Design of financial incentive component	The use of time preference theory should be considered when designing financial incentive-based interventions.
	More research is needed on optimal monetary values, types (e.g. retail vouchers, generic gift cards, charity donations, group rewards) and timing of incentives, for achieving health behaviour change.
Use of self-regulation techniques	Our findings suggest that future PA interventions should make use of self-regulation techniques (e.g. self-monitoring, goal-setting), with social and environmental prompts. This may promote habit formation and should also allow for the internalisation of the behaviour, enabling it to be more sustainable long-term.
Self Determination Theory (SDT)	To promote the internalisation and maintenance of PA behaviour change, future intervention studies should support all of the basic psychological needs posited as important in SDT (i.e. competence, autonomy, relatedness to others).
	Studies should also examine how the social and physical environment can influence participants' experience of PA interventions, and behaviour change (e.g. enjoyment, autonomy).
Integration of individual and employer perspectives	Given the many barriers to being active during the working day reported by our participants, it is not clear how effective existing schemes and policies within the workplace are, or how they are being implemented. This evaluative work highlighted the need for a 'top down' approach, for employees to feel they were supported by management in being more active within the workplace.

8.11 CONCLUSION

In summary, the PAL Scheme intervention was not more effective than waiting-list control. Reduced health care costs, reduced absenteeism and improved mental wellbeing in the intervention group are somewhat noteworthy, and results suggest that the intervention could be cost beneficial for employers. Finally, we believe our results pose several scientific and real world implementation challenges that are too infrequently exposed in public heath intervention trials.

ACKNOWLEDGEMENTS

The study team would like to acknowledge the involvement of several members from the Institute of Electronics, Communications and Information Technology (ECIT), Queen's University Belfast for their role in developing the PA monitoring system and study website. They include Prof William Scanlon, Dr Adrian McKernan, Dr William Scully, Mr Ryan McConville and Mr Philip McShane. Club Marketing Services Ltd and Lisburn City Centre Management were instrumental in their role liaising with retail partners; and we acknowledge support from project partners, the South Eastern Health and Social Care Trust and Lisburn and Castlereagh City Council. The study team also wish to acknowledge the involvement of Dr Jennifer Badham (Queen's University Belfast) for her assistance with cleaning the main trial data.

The authors thank the study participants, many of whom contributed to multiple parts of the study in multiple years; and for guidance and support from members of the independent study steering committee.

CONTRIBUTION OF AUTHORS

Ruth Hunter (Lecturer in Physical Activity and Public Health, Centre for Public Health, Queen's University Belfast) was joint Principal Investigator on the project, led the design and execution of the study, oversaw all subsequent study analyses and publications, and led the preparation of the final report.

Aisling Gough (Postdoctoral Research Fellow, Centre for Public Health, Queen's University Belfast) was the project manager and lead qualitative researcher for the six and 12 month follow-up time-points, led the qualitative analyses and publications, and contributed to others, co-ordinated the preparation of the final report.

Jennifer M. Murray (Postdoctoral Research Fellow, Centre for Public Health, Queen's University Belfast) was a Ph.D student working on the project, assisted with recruitment, data collection and data analyses, led the mediation analyses and publications, and contributed to others, led the preparation of the mediation analyses of the final report.

Jianjun Tang (Postdoctoral Research Fellow, Centre for Public Health, Queen's University Belfast) led the behavioural economic aspect of the project including recruitment, data collection and data analyses, assisted with recruitment and data collection of the main trial, led the behavioural economic analyses and publications, and contributed to others, led the preparation of the behavioural economic analyses of the final report.

Sarah F. Brennan (previous Postdoctoral Research Fellow, Centre for Public Health, Queen's University Belfast) was the project manager for the first 2 years of the project and led recruitment and data collection up until the six month follow-up time-point, contributed to the final report.

Oliver J. Chrzanowski-Smith (previous Research Assistant, Centre for Public Health, Queen's University Belfast) was a Research Assistant for the first 2 years of the project and assisted with recruitment and data collection up until the six month follow-up time-point, contributed to the final report.

Angela Carlin (previous Research Assistant, Centre for Public Health, Queen's University Belfast) was a Research Assistant for the final year of the project and assisted with data collection of the quantitative and qualitative data at 12 months follow-up time-point, contributed to the final report.

Chris Patterson (Professor of Medical Statistics, Centre for Public Health, Queen's University Belfast), co-investigator, led and overseen the conduct and interpretation of the quantitative data for the primary and secondary analyses, and mediation analyses, contributed to the final report.

Alberto Longo (Senior Lecturer in Environmental Economics, School of Biological Sciences, Queen's University Belfast), co-investigator, led and overseen the conduct and interpretation of the behavioural economic data, contributed to the final report.

George Hutchinson (Professor of Environmental Economics, School of Biological Sciences, Queen's University Belfast), co-investigator, led and overseen the conduct and interpretation of the behavioural economic data, contributed to the final report.

Lindsay Prior (Professor Emeritus of Sociology, Centre for Public Health, Queen's University Belfast), co-investigator, led and overseen the conduct and interpretation of the qualitative data, contributed to the final report.

Mark A. Tully (Senior Lecturer in Physical Activity and Public Health, Centre for Public Health, Queen's University Belfast), co-investigator, advised on intervention design, physical activity measurement and analyses, contributed to the final report.

David P. French (Professor of Health Psychology, School of Psychological Sciences, University of Manchester), co-investigator, overseen the theoretical aspect of the intervention, advised on the measurement of mediators of behaviour change, significantly involved in the interpretation of the mediation analyses, contributed to the final report.

Jean Adams (Senior Research Fellow, CEDAR, University of Cambridge), co-investigator, advised on the development of the incentive element of the intervention, contributed to the final report.

Emma McIntosh (Professor of Health Economics, Health Economics and Health Technology Assessment, University of Glasgow), co-investigator, led and overseen

the conduct and interpretation of the health economic data, contributed to the final report.

Yiqiao Xin (Research Assistant, Health Economics and Health Technology Assessment, University of Glasgow), led the health economic analyses, contributed to the final report.

Frank Kee (Professor of Public Health Medicine, Centre for Public Health, Queen's University Belfast), joint Principal Investigator, led the design and execution of the study, oversaw all subsequent study analyses and publications, and led the preparation of the final report.

DATA SHARING STATEMENT

All data requests should be submitted to the corresponding author for consideration. Access to anonymised data may be granted following review and necessary agreements being in place.

REFERENCES

1. Hunter RF, Tully MA, Davis M, Stevenson M, Kee F. Physical activity loyalty cards for behavior change: a quasi-experimental study. *Am J Prev Med*. 2013;45(1):56-63.

2. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet*. 2012;380(9838):219-229.

3. Ding D, Lawson KD, Kolbe-Alexander TL, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. *Lancet*. 2016;388(10051):947-951.

4. Allender S, Foster C, Scarborough P, Rayner M. The burden of physical activity-related ill health in the UK. *J Epidemiol Community Health*. 2007;61(4):344-348.

5. Department of Health. Start active, stay active. Report on physical activity in the UK. https://www.gov.uk/government/publications/start-active-stay-active-a-report-on-physical-activity-from-the-four-home-countries-chief-medical-officers. Published 2011. Accessed March 28, 2018.

6. Farrell L, Hollingsworth B, Propper C, Shields MA. *The Socioeconomic Gradient in Physical Inactivity in England. CMPO Working Paper Series No. 13/311*.; 2013. http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/wp311.pdf. Accessed September 11, 2018

7. Sport NI. The Northern Ireland Sport and Physical Activity Survey 2010 (SAPAS). http://www.sportni.net/sportni/wp-content/uploads/2013/03/SAPASReport.pdf. Published 2011. Accessed March 28, 2018.

8. Department for Communities. Experience of sport and physical activity in Northern Ireland: Findings from the Continuous Household Survey 2015/2016. https://www.communities-ni.gov.uk/sites/default/files/publications/communities/experience-of-sport-and-physical-activity-in-northern-ireland-201516.pdf. Published 2016. Accessed March 28, 2018.

9. National Institute for Health and Clinical Excellence (NICE). Physical activity: Local government briefing. https://www.nice.org.uk/advice/lgb3/chapter/Introduction. Published 2012. Accessed March 28, 2018.

10. McEachan RR, Lawton RJ, Jackson C, Conner M, Meads DM, West RM. Testing a workplace physical activity intervention: a cluster randomized controlled trial. *Int J Behav Nutr Phys Act*. 2011;8(1):29.

11. Working for a healthier tomorrow. Presented to the Secretary of State for Health and the Secretary of State for Work and Pensions Dame Carol Black's

Review of the health of Britain's working age population 17th March 2008. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/209782/hwwb-working-for-a-healthier-tomorrow.pdf. Published 2008. Accessed March 28, 2018.

12. Price Waterhouse Coopers. Building the case for wellness. https://www.gov.uk/government/publications/work-health-and-wellbeing-building-the-case-for-wellness. Published 2008. Accessed March 28, 2018.

13. Ball K, Jeffery RW, Abbott G, McNaughton SA, Crawford D. Is healthy behavior contagious: associations of social norms with physical activity and healthy eating. *Int J Behav Nutr Phys Act*. 2010;7(1):86.

14. Abraham C, Graham-Rowe E. Are worksite interventions effective in increasing physical activity? A systematic review and meta-analysis. *Health Psychol Rev*. 2009;3(1):108-144.

15. Conn VS, Hafdahl AR, Cooper PS, Brown LM, Lusk SL. Meta-analysis of workplace physical activity interventions. *Am J Prev Med*. 2009;37(4):330-339.

16. Rongen A, Robroek SJW, van Lenthe FJ, Burdorf A. Workplace health promotion: a meta-analysis of effectiveness. *Am J Prev Med*. 2013;44(4):406-415.

17. Verplanken B, Orbell S. Reflections on past behavior: a self-report index of habit strength. *J Appl Soc Psychol*. 2003;33(6):1313-1330.

18. O'Malley GC, Baker PR, Francis DP, Perry I, Foster C. Incentive-based interventions for increasing physical activity and fitness (protocol). *Cochrane Database Syst Rev*. 2012;1.

19. Higgins R, Murphy B, Worcester M, Daffey A. Supporting chronic disease self-management: translating policies and principles into clinical practice. *Aust J Prim Health*. 2012;18(1):80-87.

20. Scholz U, Sniehotta FF, Schwarzer R. Predicting physical exercise in cardiac rehabilitation: the role of phase-specific self-efficacy beliefs. *J Sport Exerc Psychol*. 2005;27(2):135-151.

21. Volpp KG, Troxel AB, Pauly M V., et al. A randomized, controlled trial of financial incentives for smoking cessation. *N Engl J Med*. 2009;360(7):699-709.

22. Mantzari E, Vogt F, Shemilt I, Wei Y, Higgins JPT, Marteau TM. Personal financial incentives for changing habitual health-related behaviors: a systematic review and meta-analysis. *Prev Med (Baltim)*. 2015;75:75-85.

23. Olsho LE, Klerman JA, Wilde PE, Bartlett S. Financial incentives increase fruit and vegetable intake among Supplemental Nutrition Assistance Program participants: a randomized controlled trial of the USDA Healthy Incentives Pilot. *Am J Clin Nutr*. 2016;104(2):423-435.

24. Giles EL, Robalino S, McColl E, Sniehotta FF, Adams J. The effectiveness of financial incentives for health behaviour change: systematic review and meta-

analysis. Baradaran HR, ed. *PLoS One*. 2014;9(3):e90347.

25. Molema CCM, Wendel-Vos GCW, Puijk L, Jensen JD, Schuit AJ, de Wit GA. A systematic review of financial incentives given in the healthcare setting; do they effectively improve physical activity levels? *BMC Sports Sci Med Rehabil*. 2016;8(1):15.

26. Shin DW, Yun JM, Shin J-H, et al. Enhancing physical activity and reducing obesity through smartcare and financial incentives: a pilot randomized trial. *Obesity*. 2017;25(2):302-310.

27. Christian D, Todd C, Hill R, et al. Active children through incentive vouchers – evaluation (ACTIVE): a mixed-method feasibility study. *BMC Public Health*. 2016;16(1):890.

28. Jochelson K. *Paying the Patient: Improving Health Using Financial Incentives*. London; 2007.

29. Marteau TM, Ashcroft RE, Oliver A. Using financial incentives to achieve healthy behaviour. *BMJ*. 2009;338:b1415.

30. Boyd KA, Briggs AH, Bauld L, Sinclair L, Tappin D. Are financial incentives cost-effective to support smoking cessation during pregnancy? *Addiction*. 2016;111(2):360-370.

31. Paul-Ebhohimhen V, Avenell A. Systematic review of the use of financial incentives in treatments for obesity and overweight. *Obes Rev*. 2008;9(4):355-367.

32. National Institute for Health and Clinical Excellence (NICE). *Workplace Health Promotion: How to Encourage Employees to Be Physically Active: A Rapid Review of Economic Literature*.; 2007. https://www.nice.org.uk/guidance/ph13/documents/workplace-physical-activity-economics-review2. Accessed September 11, 2018

33. Kohl HW, Craig CL, Lambert EV, et al. The pandemic of physical inactivity: global action for public health. *Lancet*. 2012;380(9838):294-305.

34. Hutchinson AD, Wilson C. Improving nutrition and physical activity in the workplace: a meta-analysis of intervention studies. *Health Promot Int*. 2012;27(2):238-249.

35. Bird EL, Baker G, Mutrie N, Ogilvie D, Sahlqvist S, Powell J. Behavior change techniques used to promote walking and cycling: a systematic review. *Heal Psychol*. 2013;32(8):829-838.

36. Hunter RF, Brennan SF, Tang J, et al. Effectiveness and cost-effectiveness of a physical activity loyalty scheme for behaviour change maintenance: a cluster randomised controlled trial. *BMC Public Health*. 2016;16(1):618.

37. National Institute for Health and Clinical Excellence (NICE). Walking and cycling: local measures to promote walking and cycling as forms of travel or recreation. https://www.nice.org.uk/guidance/ph41. Published 2012. Accessed March 28, 2018.

38. Department of Health. *Healthy Lives, Healthy People: Our Strategy for Public Health in England*.; 2010. https://www.gov.uk/government/publications/healthy-lives-healthy-people-our-strategy-for-public-health-in-england. Accessed September 11, 2018

39. Black C, Frost B. Health at work-an independent review of sickness absence. https://www.gov.uk/government/publications/review-of-the-sickness-absence-system-in-great-britain. Published 2011. Accessed February 1, 2017.

40. Hamilton MT, Hamilton DG, Zderic TW. Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease. *Diabetes*. 2007;56(11):2655-2667.

41. Finkelstein EA, Haaland BA, Bilger M, et al. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. *Lancet Diabetes Endocrinol*. October 2016:219-229.

42. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. *Developing and Evaluating Complex Interventions: New Guidance*.; 2008.

43. Tudor-Locke C, Burkett L, Reis JP, Ainsworth BE, Macera CA, Wilson DK. How many days of pedometer monitoring predict weekly physical activity in adults? *Prev Med (Baltim)*. 2005;40(3):293-298.

44. Bassett DR, Ainsworth BE, Leggett SR, et al. Accuracy of five electronic pedometers for measuring distance walked. *Med Sci Sports Exerc*. 1996;28(8):1071-1077.

45. Schneider PL, Crouter SE, Lukajic O, Bassett DR. Accuracy and reliability of 10 pedometers for measuring steps over a 400-m walk. *Med Sci Sports Exerc*. 2003;35(10):1779-1784.

46. Tudor-Locke C, Craig CL, Brown WJ, et al. How many steps/day are enough? For adults. *Int J Behav Nutr Phys Act*. 2011;8(1):79.

47. Tudor-Locke C, Williams JE, Reis JP, Pluto D. Utility of pedometers for assessing physical activity: convergent validity. *Sport Med*. 2002;32(12):795-808.

48. Corder K, Brage S, Ekelund U. Accelerometers and pedometers: methodology and clinical application. *Curr Opin Clin Nutr Metab Care*. 2007;10(5):597-603.

49. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *J Phys Act Health*. 2009;6(6):790-804.

50. Ware JE, Kosinski M, Dewey JE, Gandek B. *How to Score and Interpret Single-Item Health Status Measures: A Manual for Users of the of the SF-8 Health Survey*. Lincoln RI: QualityMetric Incorporated; 2001.

51. EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy (New York)*. 1991;16:199-208.

52. Lloyd K, Devine P. Psychometric properties of the Warwick–Edinburgh Mental Well-being Scale (WEMWBS) in Northern Ireland. *J Ment Heal*. 2012;21(3):257-263.

53. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. *Health Qual Life Outcomes*. 2007;5(1):63.

54. Kessler RC, Barber C, Beck A, et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). *J Occup Environ Med*. 2003;45(2):156-174.

55. Finch EA, Linde JA, Jeffery RW, Rothman AJ, King CM, Levy RL. The effects of outcome expectations and satisfaction on weight loss and maintenance: correlational and experimental analyses--a randomized trial. *Heal Psychol*. 2005;24(6):608-616.

56. Marcus BH, Selby VC, Niaura RS, Rossi JS. Self-efficacy and the stages of exercise behavior change. *Res Q Exerc Sport*. 1992;63(1):60-66.

57. Moller AC, Buscemi J, McFadden HG, Hedeker D, Spring B. Financial motivation undermines potential enjoyment in an intensive diet and activity intervention. *J Behav Med*. 2014;37(5):819-827.

58. Sniehotta FF, Schwarzer R, Scholz U, Schüz B. Action planning and coping planning for long-term lifestyle change: theory and assessment. *Eur J Soc Psychol*. 2005;35(4):565-576.

59. Markland D, Tobin V. A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. *J Sport Exerc Psychol*. 2004;26(2):191-196.

60. Wilson PM, Rodgers WM, Loitz CC, Scime G. "It's who I am … Really!' The importance of integrated regulation in exercise contexts. *J Appl Biobehav Res*. 2007;11(2):79-104.

61. Fishbein M, Ajzen I. *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading,Addison-Wesley; 1975.

62. Ogilvie D, Mitchell R, Mutrie N, Petticrew M, Platt S. Perceived characteristics of the environment associated with active travel: development and testing of a new scale. *Int J Behav Nutr Phys Act*. 2008;5(1):32.

63. Frank LD, Sallis JF, Saelens BE, et al. The development of a walkability index: application to the Neighborhood Quality of Life Study. *Br J Sports Med*. 2010;44(13):924-933.

64. Dalton AM, Jones AP, Panter JR, Ogilvie D. Neighbourhood, route and workplace-related environmental characteristics predict adults' mode of travel to work. *PLoS One*. 2013;8(6):e67575.

65. Rothman AJ, Sheeran P, Wood W. Reflective and automatic processes in the initiation and maintenance of dietary change. *Ann Behav Med*. 2009;38 Suppl 1:S4-17.

66. Harrington J, Perry I, Lutomski J, et al. SLÁN 2007: survey of lifestyle, attitudes and nutrition in Ireland. Dietary habits of the Irish population. *Psychol Reports*. 2008;6.

67. Harrington J, Perry IJ, Lutomski J, et al. Living longer and feeling better: healthy lifestyle, self-rated health, obesity and depression in Ireland. *Eur J Public Health*. 2010;20(1):91-95.

68. Thompson S, Wordsworth S. *UK Working Party on Patient Costs. An Annotated Cost Questionnaire for Completion by Patients.*; 2001.

69. Steckler A, Linnan L, Israel B. *Process Evaluation for Public Health Interventions and Research*. San Francisco, CA: Jossey-Bass; 2002.

70. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP. 2013.

71. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res*. 2011;20(1):40-49.

72. Rubin DB, Wiley InterScience (Online service). *Multiple Imputation for Nonresponse in Surveys*. Wiley; 1987.

73. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101.

74. Curtis L, Burns A. Unit Costs of Health and Social Care: PSSRU, Univesity of Kent at Canterbury. https://www.pssru.ac.uk/project-pages/unit-costs/. Published 2016. Accessed March 28, 2018.

75. National Health Service (NHS). National Health Service Reference Costs. https://www.gov.uk/government/organisations/department-of-health-and-social-care. Published 2016. Accessed March 28, 2018.

76. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736.

77. Feng Y, Devlin N, Herdman M. Assessing the health of the general population in England: how do the three- and five-level versions of EQ-5D compare? *Health Qual Life Outcomes*. 2015;13(1):171.

78. Agborsangaya CB, Lahtinen M, Cooke T, Johnson JA. Comparing the EQ-5D 3L and 5L: measurement properties and association with chronic conditions and multimorbidity in the general population. *Health Qual Life Outcomes*. 2014;12(1):74.

79. van Hout B, Janssen MF, Feng Y-S, et al. Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets. *Value Heal*. 2012;15(5):708-715.

80. National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal. https://www.nice.org.uk/process/pmg9/chapter/the-reference-case#measuring-

and-valuing-health-effects. Published 2013. Accessed March 28, 2018.

81. Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford University Press; 2017.

82. McMillan A, Bratton DJ, Faria R, et al. A multicentre randomised controlled trial and economic evaluation of continuous positive airway pressure for the treatment of obstructive sleep apnoea syndrome in older people: PREDICT. *Health Technol Assess (Rockv)*. 2015;19(40):1-188.

83. MacNeil Vroomen J, Eekhout I, Dijkgraaf MG, et al. Multiple imputation strategies for zero-inflated cost data in economic evaluations: which method works best? *Eur J Heal Econ*. 2016;17(8):939-950.

84. Paton NI, Stöhr W, Oddershede L, et al. The Protease Inhibitor Monotherapy Versus Ongoing Triple Therapy (PIVOT) trial: a randomised controlled trial of a protease inhibitor monotherapy strategy for long-term management of human immunodeficiency virus infection. *Health Technol Assess*. 2016;20(21):1-158.

85. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377-399.

86. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ*. 2005;14(5):487-496.

87. Dallat MAT, Hunter RF, Tully MA, Cairns KJ, Kee F. A lesson in business: cost-effectiveness analysis of a novel financial incentive intervention for increasing physical activity in the workplace. *BMC Public Health*. 2013;13(1):953.

88. Lahiri S, Faghri PD. Cost-Effectiveness of a Workplace-Based Incentivized Weight Loss Program. *J Occup Environ Med*. 2012;54(3):371-377.

89. Barte JCM, Wendel-Vos GCW. A systematic review of financial incentives for physical activity: the effects on physical activity and related outcomes. *Behav Med*. 2017;43(2):79-90.

90. Giles EL, Becker F, Ternent L, Sniehotta FF, McColl E, Adams J. Acceptability of financial incentives for health behaviours: a discrete choice experiment. Mitra N, ed. *PLoS One*. 2016;11(6):e0157403.

91. Bickel WK, Moody L, Higgins ST. Some current dimensions of the behavioral economics of health-related behavior change. *Prev Med (Baltim)*. 2016;92:16-23.

92. Barlow P, McKee M, Reeves A, Galea G, Stuckler D. Time-discounting and tobacco smoking: a systematic review and network analysis. *Int J Epidemiol*. 2017;46(3):860-869.

93. Story GW, Vlaev I, Seymour B, Darzi A, Dolan RJ. Does temporal discounting explain unhealthy behavior? A systematic review and reinforcement learning perspective. *Front Behav Neurosci*. 2014;8:76.

94. Andersen S, Harrison GW, Lau MI, Rutström EE. Eliciting risk and time preferences. *Econometrica*. 2008;76(3):583-618.

95. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. In: *Handbook of the Fundamentals of Financial Decision Making: Part I*. ; 2013:99-127.

96. Hanemann M, Loomis J, Kanninen B. Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. *Am J Agric Econ*. 1991;73(4):1255-1263.

97. Ferrini S, Scarpa R. Designs with a priori information for nonmarket valuation with choice experiments: A Monte Carlo study. *J Environ Econ Manage*. 2007;53(3):342-363.

98. Coller M, Williams MB. Eliciting individual discount rates. *Exp Econ*. 1999;2(2):107-127.

99. Laibson D. Golden Eggs and Hyperbolic Discounting. *Q J Econ*. 1997;112(2):443-478.

100. Fiebig DG, Keane MP, Louviere J, Wasi N. The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Mark Sci*. 2010;29(3):393-421.

101. Cragg JG. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*. 1971;39(5):829-844.

102. Ferecatu A, Onculer A. Heterogeneous risk and time preferences. *J Risk Uncertain*. 2016;53(1):1-28.

103. Marcus BH, Rossi JS, Selby VC, Niaura RS, Abrams DB. The stages and processes of exercise adoption and maintenance in a worksite sample. *Heal Psychol*. 1992;11(6):386-395.

104. Saul JE, Amato MS, Cha S, Graham AL. Engagement and attrition in Internet smoking cessation interventions: insights from a cross-sectional survey of "one-hit-wonders." *Internet Interv*. 2016;5:23-29.

105. Saunders RP, Evans MH, Joshi P. Developing a process-evaluation plan for assessing health promotion program implementation: a how-to guide. *Health Promot Pract*. 2005;6(2):134-147.

106. Thomson JL, Tussing-Humphreys LM, Goodman MH, Zoellner JM. Engagement indicators predict health changes in a lifestyle intervention. *Am J Health Behav*. 2015;39(3):409-420.

107. Eysenbach G. The law of attrition. *J Med Internet Res*. 2005;7(1):e11.

108. World Health Organization. *Global Recommendations on Physical Activity for Health*. World Health Organization; 2010.

109. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JEWC. Persuasive system design does matter: a systematic review of adherence to web-based interventions. *J Med Internet Res*. 2012;14(6):17-40.

110. Guertler D, Vandelanotte C, Kirwan M, Duncan MJ. Engagement and nonusage attrition with a free physical activity promotion program: the case of 10,000 Steps Australia. *J Med Internet Res*. 2015;17(7):e176.

111. Kolt GS, Rosenkranz RR, Vandelanotte C, et al. Using Web 2.0 applications to promote health-related physical activity: findings from the WALK 2.0 randomised controlled trial. *Br J Sports Med*. 2017;0:1-8.

112. Preacher KJ, Hayes AF. Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behav Res Methods*. 2008;40(3):879-891.

113. MacKinnon DP, Lockwood CM, Williams J. Confidence limits for the indirect effect: distribution of the product and resampling methods. *Multivariate Behav Res*. 2004;39(1):99-128.

114. MacKinnon D. *Introduction to Statistical Mediation Analysis*. Taylor & Francis; 2007.

115. Hooper D, Coughlan J, Mullen M. Structural equation modelling: guidelines for determining model fit. *Electron J Bus Res Methods*. 2008;6(1):53-60.

116. Muller D, Judd CM, Yzerbyt VY. When moderation is mediated and mediation is moderated. *J Pers Soc Psychol*. 2005;89(6):852-863.

117. Rothman AJ, Baldwin AS. A person x intervention strategy approach to understanding health behavior. In: *The Oxford Handbook of Personality and Social Psychology. Oxford University Press*. ; 2012.

118. Hosmer Jr. DW, Lemeshow S, Sturdivant RX. Model-building strategies and methods for logistic regression. In: *Applied Logistic Regression. New York: Wiley; 2013*. John Wiley & Sons, Inc.; 2013:89-151.

119. Borucka J. Methods for handling tied events in the Cox proportional hazard model. *Stud Oeconomica Posnaniensia*. 2014;2(2):91-106.

120. National Institute for Health and Care Excellence (NICE). *Methods for the Development of NICE Public Health Guidance (Third Edition) [PMG4]. 6 Incorporating Health Economics*.; 2012. https://www.nice.org.uk/process/pmg4/chapter/incorporating-health-economics. Accessed September 11, 2018

121. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP. 2012.

122. Tang J, Hutchinson WG, Chilton SM, Hunter R, Lau MI, Kee F. Exponential or Hyperbolic? Identifying and testing the predictive power of time preference over unhealthy behaviours. *Soc Sci Res Netw*. August 2016.

123. Andersen S, Harrison GW, Lau MI, Rutström EE. Discounting behavior: a reconsideration. *Eur Econ Rev*. 2014;71:15-33.

124. Liebenehm S, Waibel H. Simultaneous estimation of risk and time preferences among small-scale cattle farmers in West Africa. *Am J Agric Econ*. 2014;96(5):1420-1438.

125. Tanaka T, Camerer CF, Nguyen Q. Risk and time preferences: linking experimental and household survey data from Vietnam. *Am Econ Rev*. 2010;100(1):557-571.

126. Cheong J, Mackinnon DP, Khoo ST. Investigation of mediational processes using parallel process latent growth curve modeling. *Struct Equ Model*. 2003;10(2):238-262.

127. Whittle R, Mansell G, Jellema P, van der Windt D. Applying causal mediation methods to clinical trial data: What can we learn about why our interventions (don't) work? *Eur J pain*. 2017;21(4):614-622.

128. Dugdill L, Brettle A, Hulme C, McCluskey S, Long AF. Workplace physical activity interventions: a systematic review. *Int J Work Heal Manag*. 2013;1(1):20-40.

129. Malik SH, Blake H, Suggs LS. A systematic review of workplace health promotion interventions for increasing physical activity. *Br J Health Psychol*. 2014;19(1):149-180.

130. Schröer S, Haupt J, Pieper C. Evidence-based lifestyle interventions in the workplace--an overview. *Occup Med (Lond)*. 2014;64(1):8-12.

131. Finkelstein EA, Linnan LA, Tate DF, Birken BE. A pilot study testing the effect of different levels of financial incentives on weight loss among overweight employees. *J Occup Environ Med*. 2007;49(9):981-989.

132. Tanham J, Murphy MH, Breslin G. Using financial incentives to increase physical activity, weight loss and well being: a randomized control trial. *Work Pap Heal Sci*. 2014;1(9):1-5.

133. Cahill K, Perera R, Kate C, Rafael P. Competitions and incentives for smoking cessation. *Cochrane database Syst Rev*. 2011;(4):CD004307.

134. Ozdemir S, Bilger M, Finkelstein E. Stated uptake of physical activity rewards programmes among active and insufficiently active full-time employees. *Appl Health Econ Health Policy*. 2017:1-10.

135. Tannahill A, Kelly MP. Layers of complexity in interpreting evidence on effectiveness. *Public Health*. 2013;127(2):164-170.

136. Nooijen CFJ, Del Pozo-Cruz B, Nyberg G, Sanders T, Galanti MR, Forsell Y. Are changes in occupational physical activity level compensated by changes in exercise behavior? *Eur J Public Health*. January 2018.

137. Godin G, Belanger-Gravel A, Amireault S, Vohl M-C, Perusse L. The effect of mere-measurement of cognitions on physical activity behavior: a randomized controlled trial among overweight and obese individuals. *Int J Behav Nutr Phys Act*. 2011;8:2.

138. Spence JC, Burgess J, Rodgers W, Murray T. Effect of pretesting on intentions and behaviour: a pedometer and walking intervention. *Psychol Health*. 2009;24(7):777-789.

139. Horodyska K, Luszczynska A, Hayes CB, et al. Implementation conditions for

diet and physical activity interventions and policies: an umbrella review. *BMC Public Health*. 2015;15:1250.

140. van Dongen JM, Proper KI, van Wier MF, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev*. 2011;12(12):1031-1049.

141. Puig-Ribera A, Bort-Roig J, González-Suárez AM, et al. Patterns of impact resulting from a 'Sit Less, Move More' web-based program in sedentary office employees. Zhang H, ed. *PLoS One*. 2015;10(4):e0122474.

142. Department of Finance and Personnel. Personnel statistics for the 11 NI departments. https://www.gov.uk/government/organisations/department-of-finance-and-personnel-for-northern-ireland. Published 2010. Accessed March 28, 2018.

143. Basu S, Kiernan M. A simulation modeling framework to optimize programs using financial incentives to motivate health behavior change. *Med Decis Mak*. 2016;36(1):48.

144. Farooqui MA, Tan YT, Bilger M, Finkelstein EA. Effects of financial incentives on motivating physical activity among older adults: results from a discrete choice experiment. *BMC Public Health*. 2014;14:141.

145. Finkelstein EA, Brown DS, Brown DR, Buchner DM. A randomized study of financial incentives to increase physical activity among sedentary older adults. *Prev Med (Baltim)*. 2008;47(2):182-187.

146. Leahey TM, LaRose JG, Mitchell MS, Gilder CM, Wing RR. Small incentives improve weight loss in women from disadvantaged backgrounds. *Am J Prev Med*. 2018;54(3):E41-E47.

147. Tinelli M, Ryan M, Bond C. What, who and when? Incorporating a discrete choice experiment into an economic evaluation. *Health Econ Rev*. 2016;6:31.

148. Koffarnus MN, Jarmolowicz DP, Mueller ET, Bickel WK. Changing delay discounting in the light of the competing neurobehavioral decision systems theory: a review. *J Exp Anal Behav*. 2013;99(1):32-57.

149. Halpern SD, French B, Small DS, et al. Randomized trial of four financial-incentive programs for smoking cessation. *N Engl J Med*. 2015;372(22):2108-2117.

150. Teixeira PJ, Carraça E V, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. *Int J Behav Nutr Phys Act*. 2012;9(1):78.

151. Teixeira PJ, Carraça E V, Marques MM, et al. Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. *BMC Med*. 2015;13(1):84.

152. Kassavou A, Turner A, Hamborg T, French DP. Predicting maintenance of attendance at walking groups: testing constructs from three leading maintenance theories. *Heal Psychol*. 2014;33(7):752-756.

153. Kaushal N, Rhodes R. Exercise habit formation in new gym members: a longitudinal study. *J Behav Med*. 2015;38(4):652-663.

154. Wood W, Rünger D. Psychology of habit. *Annu Rev Psychol*. 2016;67:289-314.

155. Lally P, Gardner B. Promoting habit formation. *Health Psychol Rev*. 2013;7(sup1):S137-S158.

156. Gardner B, de Bruijn G-J, Lally P. A systematic review and meta-analysis of applications of the Self-Report Habit Index to nutrition and physical activity behaviours. *Ann Behav Med*. 2011;42(2):174-187.

157. Kwasnicka D, Dombrowski SU, White M, Sniehotta F. Theoretical explanations for maintenance of behaviour change: a systematic review of behaviour theories. *Health Psychol Rev*. February 2016:1-39.

158. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. *Soc Sci Med*. 2017;192:125-133.

159. Threlfall AG, Meah S, Fischer AJ, Cookson R, Rutter H, Kelly MP. The appraisal of public health interventions: the use of theory. *J Public Health (Bangkok)*. 2015;37(1):166-171.

160. King AC, Toobert D, Ahn D, et al. Perceived environments as physical activity correlates and moderators of intervention in five studies. *Am J Heal Promot*. 2006;21(1):24-35.

161. Rhodes RE, Dickau L. Moderators of the intention-behaviour relationship in the physical activity domain: a systematic review. *Br J Sports Med*. 2013;47(4):215-225.

162. Promberger M, Marteau TM. When do financial incentives reduce intrinsic motivation? Comparing behaviors studied in psychological and economic literatures. *Heal Psychol*. 2013;32(9):950-957.

163. Johnston M, Sniehotta F. Financial incentives to change patient behaviour. *J Health Serv Res Policy*. 2010;15(3):131-132.

164. Johnston M. What more can we learn from early learning theory? The contemporary relevance for behaviour change interventions. *Br J Health Psychol*. 2016;21(1):1-10.

165. Miller N, Dollard J. *Social Learning and Imitation.* New Haven, CT: Yale University Press.; 1941.

166. Thorndike EL. Animal Intelligence. In: *Reproduced in Classics in the History of Physchology, an Internet Resource (Ed. C.D. Green).* ; 1911. http://psychclassics.yorku.ca/Thorndike/Animal/chap5.htm. Accessed September 11, 2018

167. Deci EL, Eghrari H, Patrick BC, Leone DR. Facilitating internalization: the self-determination theory perspective. *J Pers*. 1994;62(1):119-142.

168. Kullgren JT, Troxel AB, Loewenstein G, et al. Individual- versus group-based financial incentives for weight loss. *Ann Intern Med*. 2013;158(7):505-514.

169. Ryan RM. Self-determination theory and wellbeing. *Wellbeing Dev Ctries*. 2009;1:1-2.

170. Deci EL, Ryan RM. *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer Science & Business Media; 1985.

171. Hofmann W, Friese M, Wiers RW. Impulsive versus reflective influences on health behavior: a theoretical framework and empirical review. *Health Psychol Rev*. 2008;2(2):111-137.

172. Hunt SM, Martin CJ. Health-related behavioural change—a test of a new model. *Psychol Health*. 1988;2(3):209-230.

173. Verplanken B, Aarts H. Habit, attitude, and planned behaviour: is habit an empty construct or an interesting case of goal-directed automaticity? *Eur Rev Soc Psychol*. 1999;10(1):101-134.

174. Lawless L, Drichoutis AC, Nayga RM. Time preferences and health behaviour: a review. *Agric Food Econ*. 2013;1(1):17.

175. Lim S-L, Bruce AS. Can't wait to lose weight? Characterizing temporal discounting parameters for weight-loss. *Appetite*. 2015;85:8-13.

176. Roberto CA, Kawachi I. Use of psychology and behavioral economics to promote healthy eating. *Am J Prev Med*. 2014;47(6):832-837.

177. Zhou M, Fukuoka Y, Mintz Y, et al. Evaluating machine learning–based automated personalized daily step goals delivered through a mobile phone app: randomized controlled trial. *J Med Internet Res mHealth uHealth*. 2018;6(1):e28.

178. Adams MA, Sallis JF, Norman GJ, Hovell MF, Hekler EB, Perata E. An adaptive physical activity intervention for overweight adults: a randomized controlled trial. Bacurau RF, ed. *PLoS One*. 2013;8(12):e82901.

179. Gjesme T. Goal distance in time and its effects on the relations between achievement motives and performance. *J Res Pers*. 1974;8(2):161-171.

180. Urminsky O, Goswami I, Lewis T. *"Impatient to Achieve or Impatient to Receive: How the Goal Gradient Effect Underlies Time Discounting. University of Chicago Working Paper."*; 2014.

181. Kaewthummanukul T, Brown KC. Determinants of employee participation in physical activity: critical review of the literature. *Am Assoc Occup Heal Nurses J*. 2006;54(6):249-261.

182. Rovniak LS, Anderson ES, Winett RA, Stephens RS. Social cognitive determinants of physical activity in young adults: a prospective structural equation analysis. *Ann Behav Med*. 2002;24(2):149-156.

183. Burke V, Beilin LJ, Cutt HE, Mansour J, Mori TA. Moderators and mediators of behaviour change in a lifestyle program for treated hypertensives: a randomized controlled trial (ADAPT). *Health Educ Res*. 2008;23(4):583-591.

184. Darker CD, French DP, Eves FF, Sniehotta FF. An intervention to promote walking amongst the general population based on an "extended" theory of planned behaviour: a waiting list randomised controlled trial. *Psychol Heal*. 2010;25(1):71-88.

185. Dutton GR, Tan F, Provost BC, Sorenson JL, Allen B, Smith D. Relationship between self-efficacy and physical activity among patients with type 2 diabetes. *J Behav Med*. 2009;32(3):270-277.

186. Sharma M, Sargent L, Stacy R. Predictors of leisure-time physical activity among African American women. *Am J Health Behav*. 2005;29(4):352-359.

187. French DP. The role of self-efficacy in changing health-related behaviour: cause, effect or spurious association? *Br J Health Psychol*. 2013;18(2):237-243.

188. Rhodes RE, Pfaeffli LA. Mediators of physical activity behaviour change among adult non-clinical populations: a review update. *Int J Behav Nutr Phys Act*. 2010;7(1):37.

189. Bravata DM, Smith-Spangler C, Sundaram V, et al. Using pedometers to increase physical activity and improve health: a systematic review. *JAMA*. 2007;298(19):2296-2304.

190. Hobbs N, Godfrey A, Lara J, et al. Are behavioral interventions effective in increasing physical activity at 12 to 36 months in adults aged 55 to 70 years? A systematic review and meta-analysis. *BMC Med*. 2013;11:75.

191. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. *Heal Psychol*. 2009;28(6):690-701.

192. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour--and are they the same? *Health Educ Res*. 2011;26(2):308-322.

193. Azevedo LB, Burges Watson D, Haighton C, Adams J. The effect of dance mat exergaming systems on physical activity and health – related outcomes in secondary schools: results from a natural experiment. *BMC Public Health*. 2014;14(1):951.

194. Horner S, Rew L, Torres R. Enhancing intervention fidelity: a means of strengthening study impact. *J Spec Pediatr Nurs*. 2006;11(2):80-89.

195. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci*. 2007;2(1):40.

196. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res*. 2010;12(1):e4.

197. Carolan S, Harris PR, Cavanagh K. Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *J Med Internet Res*.

2017;19(7):e271.

198. Heesch KC, Mâsse LC, Dunn AL, Frankowski RF, Mullen PD. Does adherence to a lifestyle physical activity intervention predict changes in physical activity? *J Behav Med*. 2003;26(4):333-348.

199. Glasgow RE, Nelson CC, Kearney KA, et al. Reach, engagement, and retention in an internet-based weight loss program in a multi-site randomized controlled trial. *J Med Internet Res*. 2007;9(2):e11.

200. Verheijden MW, Jans MP, Hildebrandt VH, Hopman-Rock M. Rates and determinants of repeated participation in a web-based behavior change program for healthy body weight and healthy lifestyle. *J Med Internet Res*. 2007;9(1):e1.

201. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial web-based weight loss program. *J Med Internet Res*. 2010;12(4):81-96.

202. Wanner M, Martin-Diener E, Bauer G, Braun-Fahrländer C, Martin BW. Comparison of trial participants and open access users of a web-based physical activity intervention regarding adherence, attrition, and repeated participation. *J Med Internet Res*. 2010;12(1):e3.

203. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. *J Med Internet Res*. 2013;15(8):e162.

204. Lim S, Norman R, Clifton P, Noakes M. Weight loss and attrition in overweight and obese young women during a 36- week internet-based lifestyle intervention. *J Obes Weight Loss Ther*. 2014;4(4).

205. Jahangiry L, Shojaeizadeh D, Montazeri A, Najafi M, Mohammad K, Yaseri M. Adherence and attrition in a web-based lifestyle intervention for people with metabolic syndrome. *Iran J Public Health*. 2014;43(9):1248-1258.

206. Deci EL, Ryan RM. Facilitating optimal motivation and psychological well-being across life's domains. *Can Psychol*. 2008;49(1):14-23.

207. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. *Appl Psychol*. 2008;57(1):1-29.

208. Schwarzer R. *Self-Efficacy in the Adoption and Maintenance of Health Behaviors: Theoretical Approaches and a New Model.* Washington DC: Hemisphere Publishing Corp; 1992.

209. Bouton ME. Why behavior change is difficult to sustain. *Prev Med (Baltim)*. 2014;68:29-36.

210. Wilcox AJ. A positive approach to negative results. *Epidemiology*. 2014;25(2):165.

211. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. *Am J Public Health*.

2011;101(4):751-758.

212. Zbikowski SM, Hapgood J, Smucker Barnwell S, McAfee T. Phone and web-based tobacco cessation treatment: real-world utilization patterns and outcomes for 11,000 tobacco users. *J Med Internet Res*. 2008;10(5):e41.

213. Glasgow RE, Christiansen SM, Kurz D, et al. Engagement in a diabetes self-management website: usage patterns and generalizability of program use. *J Med Internet Res*. 2011;13(1):e9.

214. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 Steps Australia. *J Med Internet Res*. 2012;14(1):e4.

215. Krueger RA, Casey MA. *Focus Groups: A Practical Guide for Applied Research*. Thousand Oaks, CA: Sage Publications; 2000.

APPENDICES

QUALITATIVE PROCESS EVALUATION STUDY CONTEXT

The majority of participating workplaces had ongoing health improvement programmes or policies, with only one workplace reporting no such programmes or policies in place. Examples of ongoing programmes within the participating workplaces included: free health checks, fitness classes, health talks by local charities or other organisations and cycle to work schemes. Six participating workplaces explicitly mentioned policies within their organisation relevant to health improvement, with examples including 'Health and wellbeing strategy', 'Healthy Eating Policy' and 'Staff Welfare policies'. One participating organisation did not have such policies in place. The majority of participating workplaces provided free car parking for staff within walking distance to the workplace.

The majority of participants reported travelling more than 1 mile to their place of work (92.9%). The average distance travelled by participants was 10.70 (SD 9.92) miles. Individual perceptions of the workplace environment are shown in Table 35. There was a small but significant difference between Intervention and Control Groups for perceived workplace attractiveness at baseline (p=0.028).

Table 35. Perceptions of workplace environment

Perceived workplace measure	Intervention		Control	
	N	Mean (SD)	N	Mean (SD)
Workplace Attractiveness (4-20)	438	10.17 (2.71)	377	10.63 (3.14)
Workplace Safety (4-20)	438	10.37 (2.47)	377	10.61 (2.68)
Workplace Accessibility (3-15)	439	6.73 (2.14)	377	6.86 (2.08)
Workplace Availability (3-15)	439	5.65 (2.00)	376	5.73 (2.11)
Work place Environment (14-70)	437	32.90 (6.22)	376	33.81 (7.07)

Environmental measures, including neighbourhood walkability (for buffers of 200m, 400m, 800m and 1600m) and distance to nearest green space are shown in Table 36. The average distance to the nearest greenspace across all participating sites was 160m. Please see Table 1 for descriptions of the constructs of the perceptions of workplace environment.

Table 36. Objectively measured environmental variables for each workplace

Workplace	Neighbourhood Walkability				Distance to nearest greenspace (m)
	200m	400m	800m	1600m	
Queen's University Belfast	4.713929	5.589759	4.525275	6.746569	83.88
Lisburn City Council	0.91465	1.02877	-1.4841	-1.57334	25.02
Stormont Estate	-3.991	-4.3998	-6.93502	-6.83894	93.24
Belfast City Hospital	5.323861	5.174881	4.492633	7.360332	299.21
Northern Ireland Housing Executive	0.193904	-0.03975	-0.01453	-0.48188	260.72
Lisburn Jobs and Benefits Office	0.21769	-0.28801	0.182381	-0.70459	215.45
Northern Ireland Fire and Rescue Service	-0.36682	0.10832	-0.72576	-0.74107	418.31
Northern Ireland Environment Agency	0.32344	0.540713	-0.19459	-0.67335	278.65
Wallace High School	-2.62456	-2.5966	-1.27271	-0.40642	55.57
Friends School	-4.55482	-4.13064	-0.18929	-0.53366	36.48
South Eastern Regional College	3.599555	1.994418	1.088759	-0.77367	85.57
Lagan Valley Hospital	-3.74985	-2.98204	0.526789	-1.37984	67.75

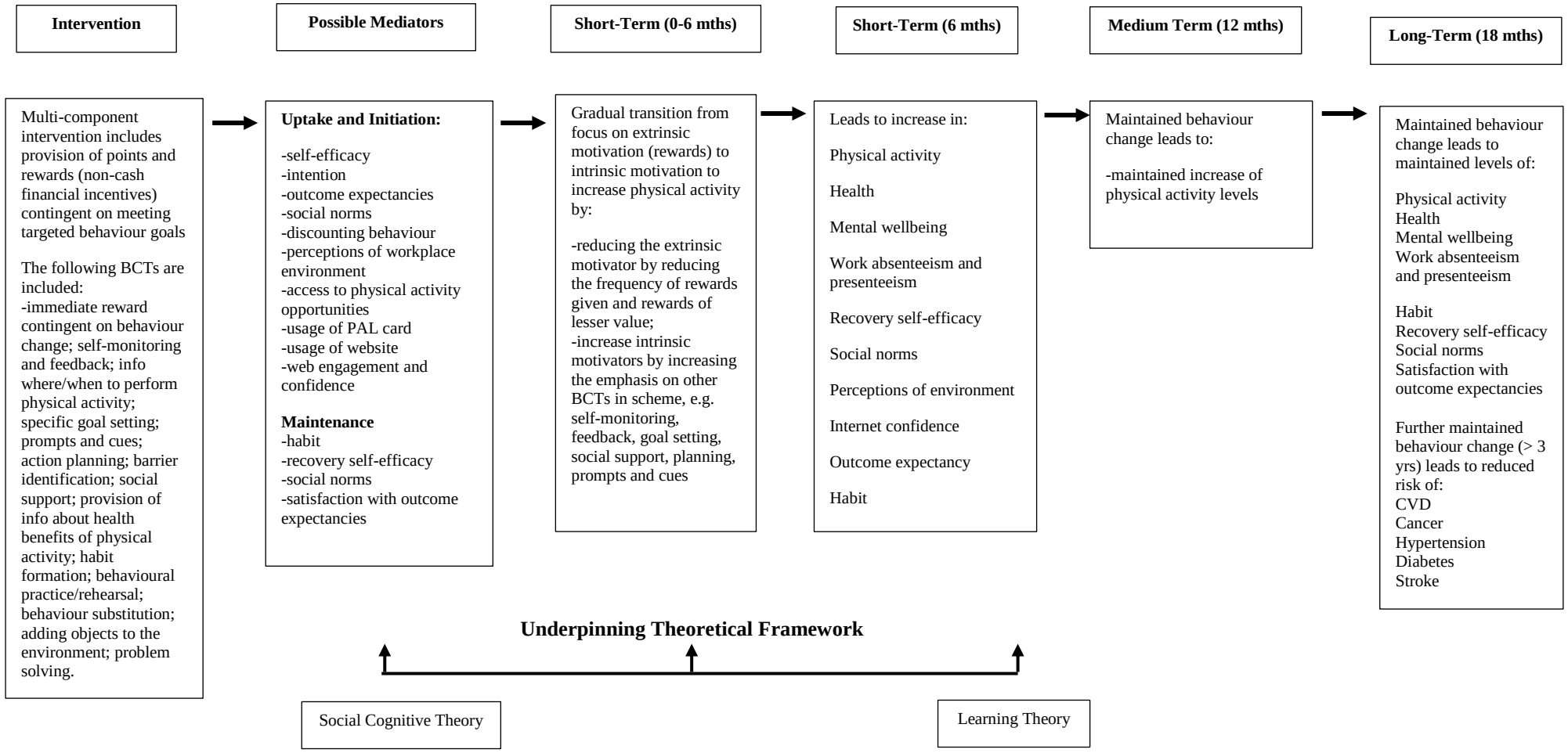

Figure 7. Logic model of the Physical Activity Loyalty scheme