Int. J. Human–Computer Studies 184 (2024) 103205

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

A field experiment on gamification of physical activity – Effects on motivation and steps

Elaine Marie Grech [a,*], Marie Briguglio [b], Emanuel Said [a]

[a] Department of Marketing, Faculty of Economics, Management and Accountancy, University of Malta, Msida, Malta
[b] Department of Economics, Faculty of Economics, Management and Accountancy, University of Malta, Msida, Malta

ARTICLE INFO

Keywords:
Gamification
Randomised controlled field experiment
Physical activity
Step counts
Behaviour change
Motivation
Perceived usefulness

ABSTRACT

Gamification is finding growing application in the field of physical activity, promising engaging and motivating experiences that foster behavioural change. Yet, rigorous empirical work substantiating favourable claims is limited. Our study sought to find evidence for the effects resulting from gamification of physical activity on the users' motivation, perceived usefulness, and the behavioural change in terms of step counts. We employed a four-week randomised controlled field experiment to investigate how the choice of different gamification designs affects outcomes. Participants were provided with a wearable physical activity tracker and randomly assigned to either a: 1) competitive gamified group; 2) cooperative gamified group; 3) hybrid gamified group; or 4) control group. Data gathered includes a panel dataset of step counts and self-reported data on the users' intrinsic motivation and perceived usefulness. We found that at the end of the intervention, gamification made no difference to self-assessed intrinsic motivation or perceived usefulness compared to a non-gamified self-tracking experience. Yet, despite the lack of psychological effects, the use of gamification did result in stronger behavioural outcomes relative to the control group - in the form of increased step counts. Indeed, all groups treated with gamification recorded an increase in step counts during the intervention period. Furthermore, amongst the gamified treatments, it was the hybrid design that generated the largest difference in step counts (relative to the control group). The finding that gamification can stimulate a stronger behavioural outcome, but does not evoke a stronger psychological outcome at the end of the intervention merits further investigation as to the mechanisms at play.

1. Introduction

In the domain of physical activity and exercise, there is widespread use of gamification in conjunction with wearables and fitness mobile applications to encourage users to be more physically active and achieve self-improvement goals (Hamari et al., 2018; Neupane et al., 2020; Oinas-Kukkonen, 2013). Gamification attempts to *'translate the engaging aspects of games into other domains of life to create positive experiences and drive desired behaviours'* (Deterding, 2019, p. 1). Existing empirical evidence suggests that the effect of gamification on physical activity is predominantly positive (Johnson et al., 2016; Koivisto and Hamari, 2019a; Mazéas et al., 2022). However, literature has insufficiently scrutinised whether the reported positive outcomes are being realised due to gamification and what type of gamification design is optimal to facilitate the desired behavioural change (Mazéas et al., 2022). Extant empirical evidence is limited, and there is significant heterogeneity between existing studies in terms of the studies' quality and study designs (Mazéas et al., 2022).

With the pursuit of health and well-being outcomes in mind, empirical evidence that confirms the promising outcomes on users' motivation and actual behaviour change through gamification remains crucial (Koivisto and Hamari, 2019a). To extend our understanding of this phenomenon, empirical studies involving randomised, controlled conditions that shed light on both psychological and behavioural outcomes of gamification are encouraged (Koivisto and Hamari, 2019a; Mazéas et al., 2022). Understanding which gamification design or individual game element produces positive effects, and under which circumstances or contexts are these favourable effects realised is thus important. Equally the grasp of how, and to what extent, is gamification able to generate positive psychological and behavioural outcomes that support the users' overall value creation is a valid quest for academics and practitioners alike (Hamari et al., 2014; Koivisto and Hamari,

* Corresponding author.
E-mail address: elaine.m.grech@um.edu.mt (E.M. Grech).

https://doi.org/10.1016/j.ijhcs.2023.103205
Received 6 October 2022; Received in revised form 13 October 2023; Accepted 17 December 2023
Available online 22 December 2023

2019b; Liu et al., 2017; Nacke and Deterding, 2017; Rapp et al., 2019; Schmidt-Kraepelin et al., 2020).

To address these limitations in existing literature, our study investigates:

i) whether the use of gamification stimulates the desired behavioural change, in terms of an increase in step counts;
ii) how the choice of gamification design affects the behavioural change in physical activity; and
iii) how the use of gamification in the context of physical activity influences the users' intrinsic motivation and perceived usefulness.

A randomised controlled field experiment was purposely designed to investigate the effect of gamification of physical activity. Consistent with established classification frameworks on gamification design (Liu et al., 2013; Morschheuser et al., 2017), our study examines three different designs of gamification commonly used in fitness applications, namely: 1) competitive; 2) cooperative; and 3) hybrid (competitive-cooperative) designs. The change in physical activity behaviour is tracked and measured objectively using physical activity trackers. The data gathered includes a panel dataset of step counts to investigate the causal effect of gamification on step counts, as well as self-reported data to examine the effect of gamification on the users' intrinsic motivation and perceived usefulness.

An overview of the theoretical foundations and related work is presented in Section 2. The details pertaining to the research method including the study design and timeline, participants and setting, the study procedure and interventions, data collection and data analysis procedures are presented in Section 3. The results and findings are presented in Section 4. Finally, this paper concludes with a discussion on the findings in relation to existing literature, and potential avenues for future research.

2. Theoretical foundations and related works

2.1. The premise of gamification

Harnessing on the motivational and engaging power of game design characteristics, gamification has the potential to promote motivation towards various utilitarian goals and behaviours (Hamari and Koivisto, 2015a; Seaborn and Fels, 2015; Walz and Deterding, 2015). Gamification features implemented into a system or a service act as stimuli (motivational affordances) designed to provide a gameful experience, that stimulates the users' motivation and psychological states, which ultimately invokes the intended behavioural outcome (Hamari et al., 2014; Huotari and Hamari, 2017). Consistent with this conceptualisation, Liu et al. (2017) emphasise that gamified systems should be designed to address two central goals, namely the experiential (psychological) and instrumental (behavioural) outcomes. The instrumental outcome refers to the intended behavioural response for which gamification is intentionally designed, in this case an increase in step counts. Experiential outcomes refer to various emotional and cognitive responses, such as enjoyment, motivation, usefulness, meaningfulness, and satisfaction (Liu et al., 2017).

Research into the design and effectiveness of gamification is commonly guided by theoretical frameworks related to motivation and behaviour change since gamification is intended to affect the users' self-motivation towards the desired activity (Seaborn and Fels, 2015). A widely researched theoretical framework on human motivation is the self-determination theory (Deci and Ryan, 1985; Ryan and Deci, 2000b). The motivation to engage in an activity or a task can be categorised into intrinsic or extrinsic motivation (Ryan and Deci, 2000a). People are intrinsically motivated when they do an activity for its inherent satisfaction, interest, and enjoyment. By contrast, extrinsic motivation relates to behaviour that is driven by external outcomes, rewards, pressures, or fears (Deci and Ryan, 1985; Ryan and Deci, 2000b, 2000a). The self-determination theory suggests that self-determined behaviour is associated not only with intrinsic sources of motivation, but also with autonomous types of extrinsic motivation where the behaviour is endorsed due to the perceived value or usefulness of the activity and congruence with the individual's personal values and needs (Deci and Ryan, 2008; Ryan and Deci, 2000b). Studies examining whether an activity is intrinsically motivating commonly assess the individuals' subjective experience in terms of enjoyment and interest (Ryan, 1982; Wu and Lu, 2013). Perceived enjoyment and interest reflect the individual's emotional psychological response to the intervention and are considered to be self-report measures of the individuals' *intrinsic motivation* (Ryan, 1982). By contrast, *perceived usefulness* is a widely used measure in studies examining internalisation of extrinsically motivated behaviours (Wu and Lu, 2013). Perceived usefulness facilitates internalisation and integration of extrinsically motivated behaviours (Deci et al., 1994; Ryan and Deci, 2000b).

Self-determination theory (Ryan and Deci, 2000b) identifies three innate psychological needs that support self-motivation, namely the need for competence, autonomy, and relatedness. First, competence refers to the need for challenge, feelings of ability, mastery, and achievement of the task at hand (Rigby, 2014; Ryan and Deci, 2000b). Providing optimal challenges and positive feedback enhances perceived competence (Ryan et al., 2006; Ryan and Deci, 2000b). Second, autonomy refers to the users' freedom of choice in deciding which actions to undertake (Ryan and Deci, 2000b; Vansteenkiste et al., 2010). Providing opportunities for choice and informational feedback enhances autonomy. Conversely, environments that diminish the sense of control on one's actions undermines autonomy and intrinsic motivation (Ryan et al., 2006). Lastly, relatedness refers to the sense of belonging and feelings of connectedness with others, which is promoted through environments characterised by a sense of mutual respect, support, and security (Rigby, 2014; Ryan and Deci, 2000b; Vansteenkiste et al., 2004).

2.2. Gamification design

The rich variety of game elements and mechanics that could be implemented in gamified systems offer endless possibilities for gamification design (Morschheuser et al., 2017). Drawing on the social interdependence theory (Johnson, 2003; Johnson and Johnson, 2005), research frameworks have been established to distinguish between different types of gamification designs (Liu et al., 2013; Morschheuser et al., 2017). Gamification can be classified as i) individualistic, ii) cooperative, iii) competitive, or iv) cooperative-competitive, also referred to as hybrid, inter-team competition, or coopetition (Liu et al., 2013; Morschheuser et al., 2017). The choice of game elements and features utilised determines the type of gamification design of the system (Morschheuser et al., 2017).

An *individualistic gamification design* includes game elements such as private badges or levels to motivate users to achieve personal goals, without causing interdependence amongst individuals (Morschheuser et al., 2017). The other three types of gamification designs include a social-oriented goal setting. A *cooperative gamification design* is based on a positive goal interdependence where users collaborate to achieve a shared goal through cooperative game elements, such as shared puzzles or team challenges (Morschheuser et al., 2017). By contrast, a *competitive gamification design* invokes a negative goal interdependence where users compete against others to achieve a goal through gamification elements, such as competitions, leaderboard, and public rankings (Morschheuser et al., 2017). The combination of competition and cooperation gamification features results in a *cooperative-competitive gamification design*, such as the case of an inter-team competition, where individuals cooperate with their team players to achieve a shared goal, whilst also competing with other teams (Liu et al., 2013; Morschheuser et al., 2017; Tauer and Harackiewicz, 2004).

Gamified fitness applications, such as Nike+ Run Club, MapMyRun, Fitbit, Strava and Pacer, 'connect' individuals to a community of users who are also performing similar activities. It is common for the design of gamified applications to include a social setting and a strong presence of social features (Hamari and Koivisto, 2013; Koivisto and Hamari, 2019b; Neupane et al., 2020). Earlier studies suggest that social factors are determinants of the attitude and use of gamified fitness applications (Hamari and Koivisto, 2013, 2015b). The use of gamification and being part of a group or team (a social setting) are determining factors that positively influence the users' adherence to physical activity applications (Yang et al., 2020). Prior research has identified that the most common game elements used in health and fitness applications include social influences, challenges, goals, collaboration, and competition (Cotton and Patel, 2019). Similarly, a systematic review conducted by Neupane et al. (2021) identified that many gamified fitness tracker applications use a combination of game elements that include social influence, competition, and challenges.

Research shows that interpersonal social contexts lead to higher levels of performance compared to individualistic contexts (Johnson and Johnson, 1989; Stanne et al., 1999). An interpersonal social setting presents different gamification design opportunities where several behavioural processes come into play. In a social setting, users are exposed to various forms of social influences, social comparison, and social support opportunities. Social influence refers to the process where individuals' attitudes, beliefs and behaviours are altered due to the presence or actions of other people in their social environment (Kelman, 1958). The behaviour of an individual could be influenced and guided by the norms of the reference group, including both descriptive and injunctive subjective norms (Ajzen, 1991; Cialdini and Trost, 1998). Social comparison refers to the process by which people compare their performance with others (Buunk and Gibbons, 2007). Through upwards and downward social comparison, individuals self-evaluate their standing relative to others in the group with the intention to improve themselves and / or to enhance their self-esteem (Festinger, 1954). Social comparison is commonly implemented in gamified applications through leaderboard rankings which display the users' performance relative to other users of the gamified application stimulating competition amongst the users (Schmidt-Kraepelin et al., 2018; Wu et al., 2015).

Different designs can influence individuals' motivation and behaviour in several ways (Johnson and Johnson, 1989; Tauer and Harackiewicz, 2004). Competitive environments can increase the individuals' desire to do well and provide a sense of challenge and excitement (Epstein and Harackiewicz, 1992; Tauer and Harackiewicz, 2004). As a result, individuals become more engaged and involved in the task, thereby fostering intrinsic motivation (Tauer and Harackiewicz, 2004). Competitions have the potential to satisfy the innate psychological need for competence providing a sense of achievement and satisfaction (Tauer and Harackiewicz, 2004). These effects are consistent with the self-determination constructs that promote motivation (Deci et al., 1999). Studies show that positive feedback during the competition enhances intrinsic motivation (Reeve and Deci, 1996; Tauer and Harackiewicz, 2004). Empirical evidence also shows that adding a competition element increases task performance, even when intrinsic motivation is not found to be a mediating factor (Landers et al., 2019). However, as argued by Santhanam et al. (2016) not all competitions are equally motivating. Competitive environments could also be demotivating for low achievers or when the individuals' level of skill is unbalanced (Epstein and Harackiewicz, 1992; Liu et al., 2013; Santhanam et al., 2016). Focusing on winning rather than the task itself undermines intrinsic motivation (Deci et al., 1981).

Cooperative designs also provide opportunities that enhance motivation and task performance (Tauer and Harackiewicz, 2004). The concept of a group provides a nurturing environment for social support and relatedness through which other needs can be better met (Martin and Dowson, 2009). Being part of a group working towards a shared goal begets a sense of social relatedness, which has been identified as one of the constructs fostering intrinsic motivation (Ryan and Deci, 2000b; Tauer and Harackiewicz, 2004). Cooperative environments could provide opportunities for social support that enhance motivation and promote the desired behaviour (Ryan and Deci, 2000b). Positive feedback on achieving the shared goal evokes feelings of competence and mastery, and in turn, intrinsic motivation (Ryan et al., 2006; Tauer and Harackiewicz, 2004). However, cooperation could also undermine motivation when individuals feel loss of autonomy, if joint commitment from the group members is lacking, if the group members fail to achieve the shared goal or if individuals perceive the shared goal as externally controlling (Johnson and Johnson, 1989; Tauer and Harackiewicz, 2004).

The cooperative-competitive design (hybrid) provides opportunities that enable individuals to foster positive relations and support amongst their team members, while getting involved into the competitive spirit to perform better than other teams (Liu et al., 2013; Tauer and Harackiewicz, 2004). Taken together, the feelings of relatedness, social support, as well as competence give rise to motivation (Ryan and Deci, 2000b). The hybrid design of cooperation-competition creates an environment which supports the individualism promoted through competition, as well as the collectivism and interdependence that exists in cooperative designs. Studies show that the simultaneous occurrence of competition and collaboration resulting in an inter-team competition led to even greater benefits than pure competition or pure collaboration in sports (Tauer and Harackiewicz, 2004), and in other gamification domains, such as crowdsourcing (Morschheuser et al., 2019). Thus, we hypothesise (see Section 2.5) that the use of a hybrid gamified design will generate the strongest positive effect on behaviour change.

2.3. Empirical evidence on the effect of gamification on physical activity

Despite the ever-growing body of literature on gamification, empirical evidence on the effect of gamification of physical activity remains limited (Koivisto and Hamari, 2019a; Mazéas et al., 2022). Existing literature primarily focused on whether gamified interventions of physical activity result in positive outcomes (Koivisto and Hamari, 2019a). While the success of gamification differs, the majority of empirical studies in this domain report positive results (Johnson et al., 2016; Koivisto and Hamari, 2019a; Mazéas et al., 2022). Overall, the effect of gamification on physical activity behaviour is described as ranging from a small to a medium positive effect in the short term (Mazéas et al., 2022). The long-term effect is even more volatile with very small to small effects being reported in literature (Mazéas et al., 2022). Some studies suggest that the positive effect on physical activity declines over time during the intervention (Gremaud et al., 2018; Patel et al., 2019; Thorsteinsen et al., 2014) or is not maintained in the long-term when the gamification stimulus is removed (Maher et al., 2015). The positive effects reported for gamified interventions are considerably higher when the gamified intervention is compared to inactive control groups (for instance individuals on waiting lists), rather than active control groups, such as individuals using a non-gamified version of the application (Mazéas et al., 2022).[1] Nonetheless, there is also evidence through randomised controlled studies reporting null effects for gamified interventions of physical activity (Direito et al., 2015; Edney et al., 2020; Kurtzman et al., 2018; Zuckerman and Gal-Oz, 2014).

The lack of randomised controlled studies isolating the effect of specific game elements or gamification designs limits our understanding on which gamification design is optimal (Koivisto and Hamari, 2019a; Mazéas et al., 2022). The type of gamified interventions investigated in previous studies varies greatly and interventions are hardly comparable across the different studies (Koivisto and Hamari, 2019a; Mazéas et al.,

[1] Effect size of 0.58 in the case of gamification versus inactive control group; and an effect size of 0.23 in the case of gamification versus active control group.

2022). Nevertheless, the literature provides some indications in this regard. Gamified interventions incorporating competitive game design elements were reported to have a significant positive effect on physical activity behaviour in several studies (Gremaud et al., 2018; Patel et al., 2019; Thorsteinsen et al., 2014; Tu et al., 2019). A competitive gamified intervention amongst overweight and obese adults facilitated the highest physical activity levels, when compared to an individualistic gamified design that included social support and a collaborative gamified design (Patel et al., 2019). Tu et al. (2019) reported that making the fitness application more social by incorporating a competitive design including a leaderboard ranking, adding friends amongst the team members, and providing opportunities for social support is more effective than an individualistic design that includes badges, points, achievements, level of progression and virtual goods. Gremaud et al. (2018) reported that a gamified intervention incorporating competitive game elements significantly increased physical activity behaviour amongst sedentary office workers, compared to a non-gamified self--tracking experience. Self-reported measures of physical activity indicate that other gamified interventions involving competitive game elements also resulted in positive effects on physical activity (Thorsteinsen et al., 2014).

While a competitive gamified design seems to be the most common gamified design investigated in gamification literature, empirical evidence also indicates that there is potential to increase physical activity through cooperative (Chen and Pu, 2014; Patel et al., 2017) and hybrid gamified designs (Chen and Pu, 2014; Lin et al., 2006; Mo et al., 2019). A cooperative gamified intervention led to a significant increase in physical activity amongst family members when compared to a control group (Patel et al., 2017). A gamified intervention involving teams and social support improved self-reported physical activity measures, as well as related cognitive constructs compared to the control group (Mo et al., 2019). Another study conducted by Lin et al. (2006) shows that both an inter-team competition (hybrid design) and an individualistic gamified design were effective at increasing physical activity. To our knowledge, the study by Chen and Pu (2014) was the only study investigating competitive, cooperative, and hybrid designs of gamification on physical activity. The authors developed *HealthyTogether*, a gamified mobile application encouraging dyads of friends to engage in physical activity together. In the absence of a control group, the authors compared physical activity during a one-week intervention period with a one-week baseline period. The cooperative gamified design facilitated the highest increase in physical activity, followed by the hybrid gamified design. Albeit positive, the effect of the competitive gamified design was not significant (Chen and Pu, 2014).

So far, only few studies examined psychological outcomes resulting from gamified interventions of physical activity in conjunction to behavioural outcomes as the following observations suggest (Mazéas et al., 2022). Following a gamification intervention, positive psychological outcomes in terms of self-efficacy, physical activity intention, and knowledge were reported together with positive behavioural outcomes in terms of walking behaviour (Dadaczynski et al., 2017). Conversely, in response to a gamification intervention Direito et al. (2015) reported no net effects on perceived enjoyment, psychological need satisfaction and self-efficacy, as well as no effect on physical activity behavioural outcomes. Gamification literature posits that psychological outcomes resulting from gamification affordances mediates behavioural outcomes (Huotari and Hamari, 2017). The findings from Dadaczynski et al. (2017) and Direito et al. (2015) support this conceptualisation, however further research that empirically investigates the psychological and behavioural outcomes of gamification in the context of physical activity is necessary (Mazéas et al., 2022).

2.4. Shortcomings

There are mainly five limitations and methodological shortcomings in existing empirical evidence on the effect of gamification of physical activity (Johnson et al., 2016; Koivisto and Hamari, 2019a; Mazéas et al., 2022).

First, extant literature is unable to conclude that the reported positive effect on physical activity emerges from the implementation of gamification itself (Mazéas et al., 2022). Positive effects attributed to gamification in existing literature (Koivisto and Hamari, 2019a; Mazéas et al., 2022) involve location-based games (Broom and Flint, 2018; Kaczmarek et al., 2017), and exergames (Farrow et al., 2019; Garde et al., 2016; Geelan et al., 2016; Höchsmann et al., 2019), some of which also involve augmented and mixed realities. Although these interventions include game elements, these developments are conceptually different from gamification (Deterding et al., 2011; Fogg, 2002; Huotari and Hamari, 2012; Koivisto and Hamari, 2019b).

Second, several studies (Dadaczynski et al., 2017; Harris, 2018; Maher et al., 2015; Thorsteinsen et al., 2014) investigating the effect of gamification on physical activity rely on self-reported data using diaries or questionnaires, rather than objective data gathered through pedometers, accelerometers, or other sensor-based technologies. Subjective self-reported measures are based on the individuals' recollection of events which may not be as precise as those recorded through objective measures (Fiedler et al., 2021; Prince et al., 2008). For instance, Edney et al. (2020) report that while self-reported measures of physical activity indicated a significant positive effect, objective data gathered for the same study confirm that the intervention did not actually change physical activity levels.

Third, apart from the fact that the number of rigorous empirical studies investigating gamification of physical activity are rather limited, existing studies vary greatly in terms of the motivational affordances included, the type of gamification design, as well as the outcome measures being investigated (Koivisto and Hamari, 2019a; Mazéas et al., 2022). The effect of gamification on physical activity tends to be difficult to compare as it is measured on diverse outcomes. Whilst the most common objective behavioural outcome measure is the daily step count, other measures used in literature include minutes of moderate-to-vigorous physical activity, active minutes, and walking time (Mazéas et al., 2022). Harmonising and standardising the gamified interventions, and the outcome measures on which the effect of gamification is investigated would be beneficial to compare like with like (Nacke and Deterding, 2017).

Fourth, literature indicates that the empirical studies that did not include a control group report greater positive outcomes than studies that adopted a randomised controlled design (Koivisto and Hamari, 2019a). In the absence of a control group, studies rely on comparing levels of physical activity measured during the intervention period with baseline levels of physical activity taken prior to the intervention. Furthermore, even though some studies (Dadaczynski et al., 2017; Maher et al., 2015) included a control group, one cannot ascertain whether the gamified intervention or the physical activity tracker accounted for the behavioural change since the control group did not have access to a pedometer. Notwithstanding the operational issues, it is recommended that future studies employ full randomisation and control conditions (Koivisto and Hamari, 2019a) with multiple groups to isolate the effect of gamification elements (Mazéas et al., 2022).

Fifth, there is scarce evidence on how gamified behavioural interventions of physical activity affect psychological outcomes together with behaviour change outcomes (Mazéas et al., 2022). Apart from the outcome measures investigating the behavioural change of physical activity, it would be beneficial to also investigate the psychological outcomes to better understand the mechanisms related to behaviour change (Mazéas et al., 2022).

In summary, there is wide variation in the study design and quality of existing empirical studies, lack of controlled designs, diversity in the study populations, varied targeted outcome measures, considerable statistical heterogeneity, and a high risk of bias in some of the reviewed studies which limits the conclusions that can be made (Johnson et al., 2016; Koivisto and Hamari, 2019a; Mazéas et al., 2022).

2.5. Hypotheses

Through this randomised controlled field experiment, we investigate the effect of gamification of physical activity. Reminiscent to the popularity of social elements used in mobile fitness applications and consistent with established classification frameworks on gamification design (Liu et al., 2013; Morschheuser et al., 2017), this study investigates three types of socially oriented gamification designs, namely competitive, cooperative, and hybrid designs. Based on the theoretical foundations and literature discussed in this section, we set out to test the following hypotheses:

Hypothesis 1. Gamification improves physical activity - Gamified groups will report higher step counts than the control group during the intervention period.

Hypothesis 2. Hybrid design will facilitate the strongest effect on step counts.

Hypothesis 3a. Gamified groups will report higher intrinsic motivation than the control group.

Hypothesis 3b. Gamified groups will report higher perceived usefulness than the control group.

3. Method

3.1. Study design

Our study involved a four-arm randomised controlled field experiment, examining the effect of three gamified interventions versus a control group. The treatment groups included three different gamification designs: 1) competition; 2) cooperation; and 3) a hybrid design involving competitive and cooperative elements. This study involved a parallel group design, where each participant was randomised to one group throughout the experimental period. The four-week experimental period consisted of one-week baseline period, followed by a three-week intervention period. The randomised controlled experiment was conducted between January and March 2020 following the timeline[2] set out in Fig. 1.

3.2. Participants and setting

The study population included academic researchers and postgraduate research students. Previous research suggests that people involved in academia typically lead a sedentary lifestyle that does not meet the recommended levels of physical activity, leading to higher risks of non-communicable diseases related to lack of physical activity (Cooper and Barton, 2016). The study was conducted in Malta, a country with a prevailing rate of insufficient physical activity and high incidence of overweight and obese people (WHO, 2022).

Participants were recruited between December 2019 and January 2020 using a non-probabilistic convenience sampling method. Following an email invitation and a post on social media, interested participants were informed about the study (including its objectives, duration, and requirements) and submitted informed consent through an online form.

Eligible participants:

- had over 18 years of age,
- did not use a smartwatch or a wearable to monitor their physical activity during the previous 12-month period,[3] and
- had no health issues (such as heart condition, chest pain, bone or joint pain, or dizziness) that they are aware of, which could prevent them from engaging in physical activity.

Conditions like pregnancy and doctors' advice to not engage in physical exercise precluded individuals' participation.

3.3. Sample size

Sample size was calculated a priori following the recommended guidelines on sample size estimation for randomised controlled trials suggested by Chow et al. (2017). On the basis of previous studies (Gremaud et al., 2018), the expected difference in daily steps between the gamified intervention groups and those in an active control group using a wearable device was around 2000 steps per day. The standard deviation was assumed to be about 2500 steps per day. The sample size required to establish superiority of the gamified interventions compared to the control group was based on a targeted power of 80 % ($1 - \beta = 0.8$) at 5 % significance level ($\alpha = 0.05$) with equal allocation between the groups ($k = 1$). These considerations led to a sample size calculation of 20 participants for each treatment group and control group respectively. Thus, for a four-arm randomised controlled experiment a total sample size of 80 participants was required to detect between-group differences on the daily step count.

3.4. Randomisation

Following the eligibility screening criteria, a Unique Reference Number (URN) was assigned to all participants to ensure anonymity throughout the study. Using an online random sequence generator (random.org), eligible participants ($n = 80$) who provided informed consent were randomly allocated to the control or one of the treatment groups using a 1:1:1:1 ratio. Participants were blinded to group allocation and groups were colour-coded to hide the identity of each group from participants.

3.5. Procedure and interventions

All participants attended a group information meeting (see Fig. 1). Separate information meetings were held for each group of participants to avoid cross-contamination between groups. During the information meeting, all participants were given a smartwatch (Xiaomi Mi Band) to monitor their step counts. Earlier studies (Tam and Cheung, 2019; Xie et al., 2018) show that these wearable devices are adequately reliable in measuring step counts, and hence these were preferred against other brands of pedometers due to their cost and battery lifespan.[4] The use of wearable devices permits the collection of objective step count data that is more reliable than physical activity reported through self-reporting approaches (Fiedler et al., 2021).

All participants were instructed on how to pair and sync the smartwatch with the corresponding mobile application, and to wear the device at all times. During the set-up of the wearables and the corresponding application installed on their smartphones, all participants were allowed to choose a personalised daily step target. Goal setting is a commonly used feature in self-tracking motivational technologies (Aldenaini et al., 2020a, Aldenaini et al., 2020b) that supports

[2] The experiment was completed prior to COVID-19 outbreak in Malta.

[3] Even though eligibility criteria excluded participants who used a smartwatch during the previous 12-month period, 22.5% of participants claimed that they had prior usage of wearables at some point before the study.

[4] The battery lasts approximately two weeks. Participants had to charge their wearable device only once during the experiment and were advised to do so during the night.

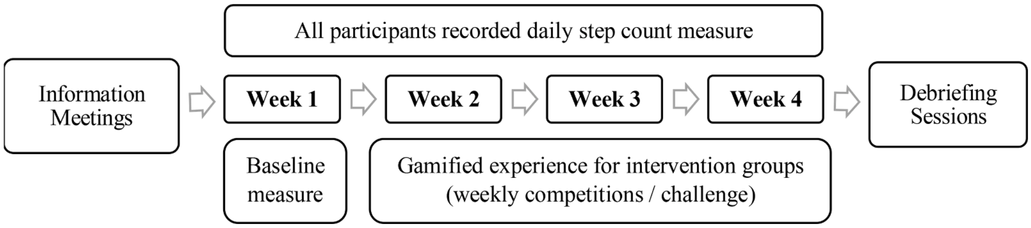

Fig. 1. Study timeline.

users' intrinsic motivation and self-regulation (Latham and Locke, 1991).

3.5.1. Control group

Participants in the control group were equipped with a physical activity tracker that enabled them to monitor their daily step count. An active control group ensured that the observed effect on physical activity was not the result of having a wearable device to monitor physical activity.

3.5.2. Interventions

A gamified platform (pointagram.com) was used to design a separate gamified experience for each treatment group.[5] All participants could access the gamified platform through an application that was installed on their smartphone or through a web browser.

The design of the gamified interventions was guided by the taxonomy of gamification concepts utilised in health applications (Schmidt-Kraepelin et al., 2018) and gamification design frameworks (Buckley et al., 2018; Liu et al., 2013; Morschheuser et al., 2017) identified in literature. The game elements and principles implemented in the interventions, and the desirable outcomes are detailed in Table 1.

A set of commonly occurring game elements associated with the constructs of the self-determination theory identified in literature (Buckley et al., 2018) were incorporated in all the gamified interventions to afford an appealing and motivating experience that supports the users' intrinsic motivation. These included points, badges, progress feedback and opportunities for social support, comparison, and interaction on the newsfeed section of the gamified application. Participants were asked to enter their daily step count to achieve points (one point for each step count recorded). Based on their daily step count, individual badges were awarded at increments of 2 K step counts, starting from a 2 K badge going up to a 20 K badge. Participants could see others' performance (social comparison) and interact with other participants in their respective group through posts, comments, and likes (social interaction and support). Screenshots of the gamified application are presented in Figs. 2–4.

Furthermore, based on the classification of gamification features proposed by Morschheuser et al. (2017) that is grounded on the social interdependence theory (Johnson, 2003), each gamified intervention incorporated specific game elements to create 1) a competitive gamified design; 2) a cooperative gamified design; and 3) a hybrid (competitive-cooperative) gamified design. Participants in the **competitive treatment group** had a weekly individual competition, where the accumulated points were visible on a leaderboard (Fig. 4) and the top three players were awarded a virtual trophy. By contrast, participants in the **hybrid treatment group** had a weekly team competition (participants were randomly assigned in teams of four participants each) where the accumulated points of each team were visible on a leaderboard (Fig. 4) and the top three teams were also awarded virtual trophies.

Finally, the **cooperation treatment group** had a weekly group

Table 1
Gamification design of the interventions.

Gamification design	
Gamification design principles, elements and mechanics	*Applicable to ALL gamified interventions:* • *Points:* users earn individual points for step count recorded (one step = one point) • *Badges:* users earn individual badges for achieving higher daily step counts • *Progression status*: progression bar indicating the progress and remaining effort required to achieve the next badge. • *Opportunities for social interaction and support:* users can post comments, send likes to each other comments and notifications on the newsfeed section of the gamified application • *Opportunities for social comparison:* users can see others' performance, progress and achievements • *User identity*: users are anonymised and represented by a URN code • *Feedback:* users are notified when they earn points and badges through a notification on the gamified application • *Episodical:* competitions and challenges/quests run from Monday to Sunday, users' progress in the competition / challenge resets every week *Applicable to the Competitive Gamified Design (Player* vs. *Player competition):* • *Leaderboard:* showing the ranking of all the players • *Virtual trophies:* awarded to the top three players with the highest step counts *Applicable to the Hybrid Gamified Design (Team* vs. *Team competition):* • *Teams:* players were randomised to teams of 4 players each • *Leaderboard:* showing the ranking of all the players • *Virtual trophies:* awarded to the top three teams with the highest step counts *Applicable to the Cooperative Gamified Design (Shared group challenge/quest):* • *Visualisation /Plot:* a map showing a pirate making his way to reach the treasure chest with a countdown timer indicating the time left for the participants to complete the challenge. Users' step counts are reflected in the progress that the pirate made towards the treasure chest. • *Quest/Challenge:* Quest to reach a target step count (shared goal) which automatically opens the pirate treasure chest before the end of the week.
Desirable outcomes	A positive behavioural change in physical activity, in terms of an increase in step counts A positive outcome on the users': • Intrinsic motivation, in terms of interest and enjoyment • Perceived usefulness of the experience

[5] Each treatment group had a separate gamified interface on the platform, so the participants in the gamified groups would not become aware of the other groups.

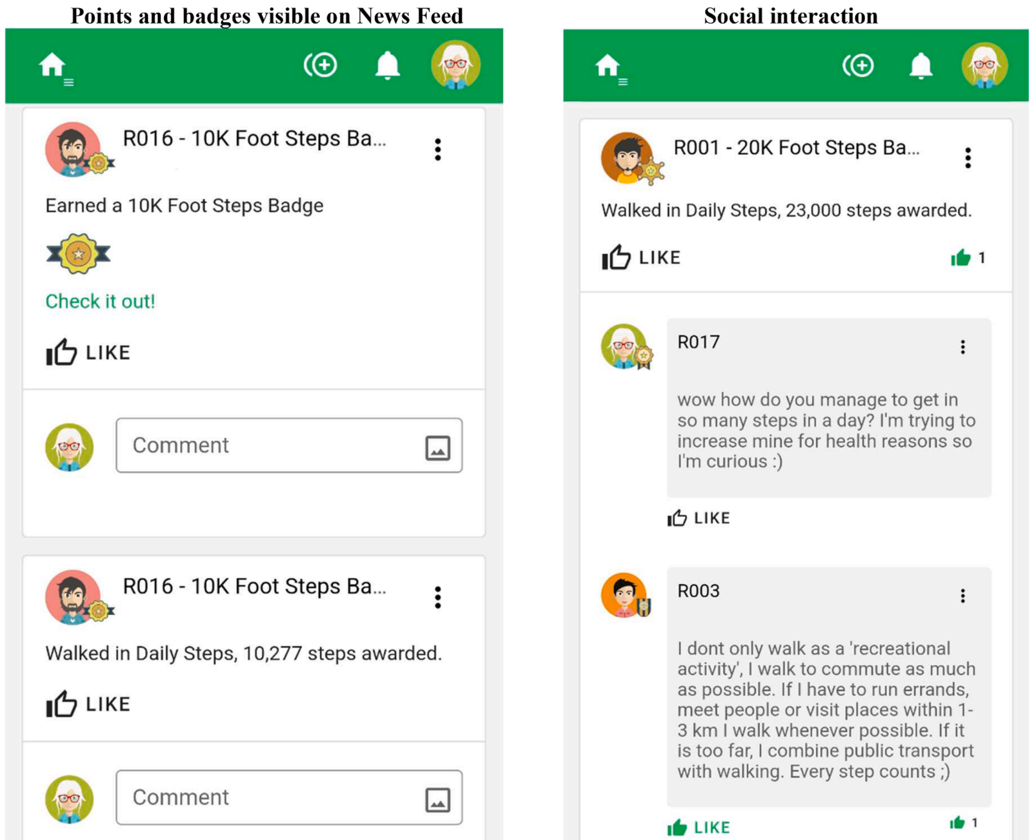

Fig. 2. Screenshots of the gamified application.

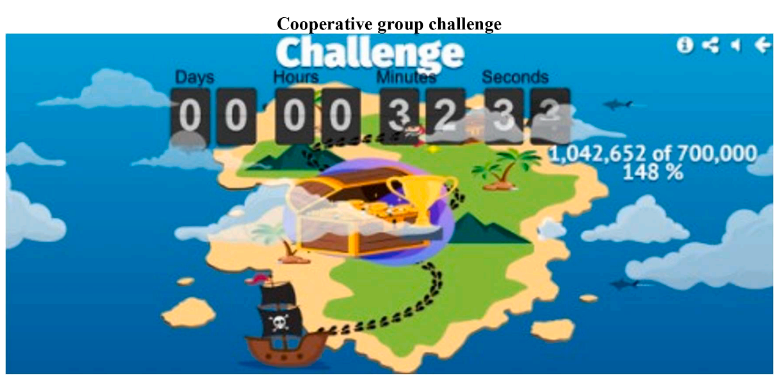

Fig. 3. Screenshot of the cooperative group challenge.

challenge (quest) to reach a target step count (shared goal) by the end of the week. Their steps were accumulated and depicted on a visualisation of a pirate making his way to reach the treasure chest on an island, with a countdown timer indicating the time left for the participants to complete the challenge (Fig. 3). The group target step count was 700 K steps for the first week (based on approximately 5 K daily step count per participant) and then increased every week based on the equivalent of 7.5 K and 10 K daily step counts per participant as a group target. All the challenges and competitions were scheduled to run on a weekly basis from Monday to Sunday.

Recent literature (Neupane et al., 2020) published on the taxonomy of game elements utilised in gamified fitness applications verifies that the game elements and design principles implemented for this study are amongst the most commonly adopted gamification design strategies in industry practice.

At the end of the study, all the participants were asked to return their wearable devices. At this stage, the step count data recorded through the participants' smartwatches was collected through the corresponding application of the Xiaomi Mi Bands (called MiFit app[6]) that was installed on the participants' smartphones. Two researchers retrieved the step count data from the MiFit app installed on each smartphone and independently validated the step counts of each participant. This process was done to confirm the panel dataset of step counts that was analysed for this study, ensure that no data was lost and minimise human error.

During the debriefing meeting, all participants were also asked to complete a pen-and-paper questionnaire to measure the participants' interest and enjoyment during this experience as a measure of the users' intrinsic motivation, the perceived usefulness of the experience, as well

[6] After the experiment period, Xiaomi's MiFit app has been renamed to Zepp life app.

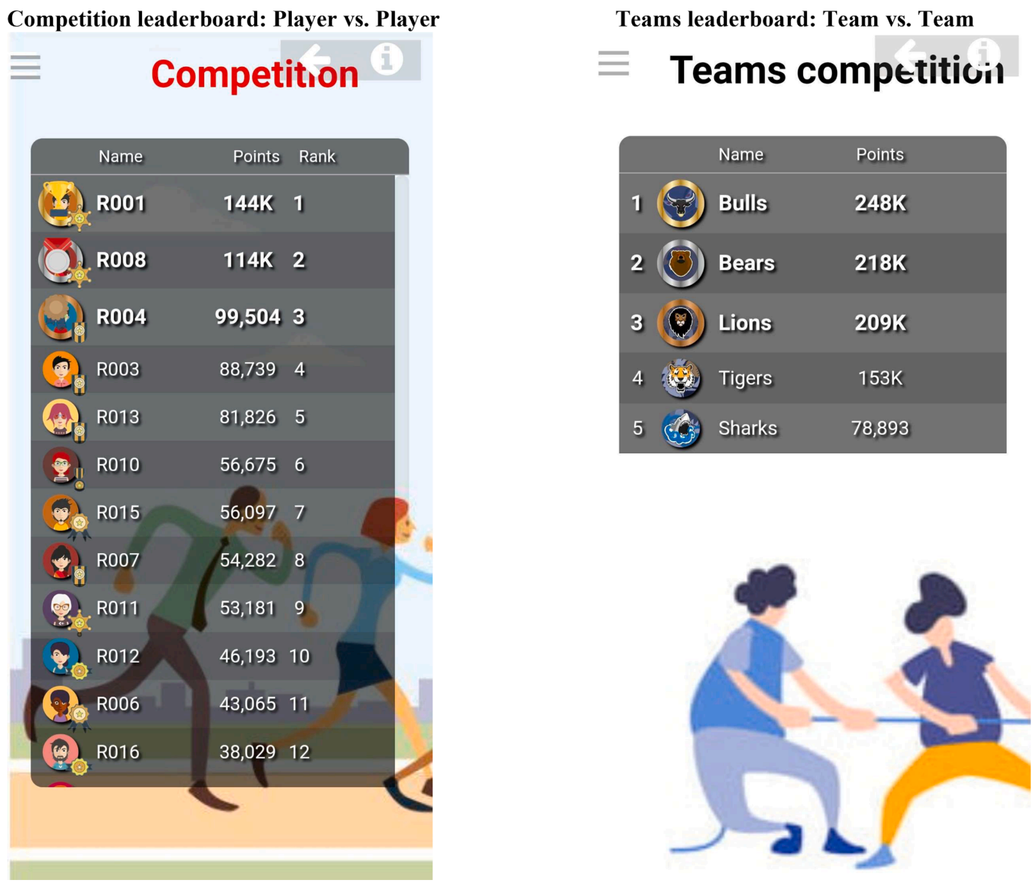

Fig. 4. Screenshots showing the leaderboards used in the competition and hybrid gamified groups.

as demographic information.

3.6. Outcome measures

The behavioural outcome was measured in terms of the change in step counts from baseline through the intervention period. This study relies solely on the step count data gathered from the wearable physical activity trackers. Step counts were recorded daily as a continuous variable. The panel dataset considered for this study included four weeks of step count data, each week starting on Monday. Earlier studies suggested that the most reliable measures are achieved when monitoring of step count data starts on Monday (Sigmundová et al., 2013). The first week of step count data was considered as the baseline measure, during which no treatment was administered, whilst the following three weeks of step count data were during the intervention period.

Psychological outcomes were measured in terms of the participants' intrinsic motivation (based on the users' interest and enjoyment) and the users' perceived usefulness of the experience. Self-reported data on these measures was gathered at the end of the intervention period using adaptations of the intrinsic motivation inventory (IMI), a validated instrument based on the self-determination theory (McAuley et al., 1989; Ryan, 1982). The Interest / Enjoyment sub-scale is considered as a validated self-reported measure of intrinsic motivation (McAuley et al., 1989; Ryan, 1982). The Perceived Usefulness subscale of the IMI refers to the perceived value of an activity which facilitates internalisation and self-regulation of activities that are found to be useful (Deci et al., 1994; McAuley et al., 1989; Ryan, 1982). Table 2 sets out the details on the scale items for each construct.

3.7. Statistical analysis

3.7.1. Data cleaning and handling of missing data

The dataset of step counts analysed for this study was retrieved from the wearable physical activity trackers (objective measure) to avoid any data loss and ensure accuracy. Nonetheless, the authors looked at whether the participants interacted with the gamified application on a daily basis to submit their step counts. All participants from the three different gamification groups interacted and utilised the gamified

Table 2
Measure, items and scale.

Measure	Source	Item wording	Notes
Intrinsic Motivation *(Emotional response)*	Adapted from the Interest / Enjoyment sub-scale of the IMI	I enjoyed doing this experience very much This experience was fun to do I thought this was a boring experience (R)	7-point Likert scale anchored 'Not at all true' to Very true'
Perceived Usefulness *(Cognitive response)*	Adapted from the Value / Usefulness sub-scale of the IMI	I believe this experience was of some value to me I think that doing this experience was useful to increase my physical activity I think doing this experience helped me to increase my physical activity	7-point Likert scale anchored 'Not at all true' to Very true'

application during the three-week intervention period. Participants' adherence to submit their daily step count on the gamified application was high, ranging from 92 % in the first week, 89 % during the second week, and 85% during the third week. The adherence rates split by group are presented in the Supplementary Material. There were no significant differences between the different gamification groups in terms of adherence to submit their daily step counts.

The authors also examined the data for any mismatch between the step counts entered on the gamified application by participants versus the step counts measured through wearable device. The mismatch between the data entered by participants and that recorded by the wearable devices was minimal. In total, only 3.4 % (43 out of 1260 observations) were mismatched. Mismatched observations account for 1.0 % in the hybrid group, 4.5 % in the cooperation group, and 4.8 % in the competition group. To avoid any reliance on human error and ensure accuracy, the data analysis considered the data gathered from the wearable devices.

In line with the intention-to-treat principle, all randomised participants were included in the analysis. Step count data on the days when smartwatches were given to participants during the initial information meetings was discarded since this did not capture full-day data of the physical activity of participants. The following two days of initial wearable use (which were on weekend days) were also excluded from the baseline estimate due to potential higher activity during initial wearable use. This approach is similar to that adopted in previous studies (Patel et al., 2017, 2019).

Missing step count data throughout the experiment period accounted for only 1.6 % of the total observations (35 out of 2240 participant-days). This proportion is much lower than reported in other studies (Chokshi et al., 2018; Patel et al., 2017, 2019) where missing step data ranged from 19 % to 29 % in physical activity interventions with longer timeframes. Days with unrecorded steps could result if a participant did not wear the wearable device, or the device did not synchronise with the smartphone application. Research on pedometer monitoring indicates that three days of step count data within a week can provide a sufficient reliable estimate of physical activity (Tudor-Locke et al., 2005). Missing step counts were imputed with the weekly mean step count. The mean daily step count for each week was derived by summing up the daily step count for each respondent and dividing it by the number of days on which step counts were recorded.

Previous literature suggests that daily step count values that are less than 1000 steps do not reflect full day data activity and should thus be excluded and imputed (Kurtzman et al., 2018; Patel et al., 2017, 2019). In this study, step counts values less than 1000 accounted for only 1 % of the total observations, and these were imputed at the weekly mean step count.[7]

3.7.2. Data analysis of behavioural outcome

Data was restructured into the long data format and analysed using Generalized Linear Mixed-Effect Models (GLMM) in STATA™ (version 16.1, StataCorp). To test the first hypothesis (*H1* - Gamified groups will report higher step counts than the control group during the intervention period) the effect of gamification (*Treatment*) was estimated using the longitudinal analysis of covariance method (Twisk et al., 2018). The analysis included an adjustment for the baseline values of the outcome variable as recommended in literature (Twisk et al., 2018), even though the differences at baseline are attributed to chance and random fluctuations. The effect of gamification at different timepoints during the intervention period was examined by extending the longitudinal analysis of covariance model to include *Time* as a main effect (fixed variable) and the interaction between *Time* and *Treatment*. The generalized mixed-effects model analysis also included a random intercept to adjust for the repeated observations over time at individual level and was estimated using a robust estimator of variance. For the second hypothesis *(H2: Hybrid design expected to facilitate the strongest effect on step counts*), the effect of each respective treatment group (competition, cooperation, and hybrid) during the intervention was estimated by including *Group* as the treatment variable. Differences were considered statistically significant at a level of 10 %. Effect sizes were computed using Hedge's g (also known as the corrected effect size) since this is preferable to Cohen's d in the case of small sample sizes (Hedges and Olkin, 1985). To test the robustness of our findings, the analysis was repeated including gender as a covariate in the model.

3.7.3. Data analysis of experiential outcomes

The constructs' reliability for *Intrinsic Motivation* and *Perceived Usefulness* were measured using Cronbach's alpha (α), composite reliability (CR), and average variance extracted (AVE). All the convergent validity metrics obtained were checked against the thresholds (Cronbach's $\alpha > 0.7$, CR > 0.7, and AVE > 0.5) as suggested in literature (Hair et al., 2010). High construct reliability indicated that internal consistency exists amongst the scale items used to measure a specific construct (Hair et al., 2010).

Descriptive statistics including means and standard deviation were computed for each experiential outcome for the control and gamified groups. Following that, a Shapiro-Wilk test was conducted to determine whether the score distribution of each construct follows a normal distribution. To test the third hypothesis set out for this study (*H3: Gamified groups expected to report higher intrinsic motivation and higher perceived usefulness than the control group*), an analysis of the differences in the means between the groups was carried out for each construct. A Mann-Whitney test was carried out for each construct to test whether there were significant differences in the means reported between the control and the gamified groups. Furthermore, a Kruskal-Wallis test was conducted for each construct to check for any significant differences between the means of each treatment group.

4. Data and results

4.1. Sample characteristics

Eighty participants completed the pre-screening and provided voluntary informed consent for participation. All randomised participants completed the study till the end without any dropouts during the study duration. We analysed the data of all the participants according to the intention-to-treat principle. The baseline participants' demographic characteristics are presented in Table 3. These descriptive results summarise the frequencies and respective percentages for categorical variables, and the means and standard deviation for continuous variables for each respective group. The p-values presented are the results of the Chi-Square tests for categorical variables and independent t-tests for continuous variables carried out to check for significant differences between the sub-groups.

Out of 80 participants, 56 % ($n = 45$) were female, 52 % ($n = 42$) were young adults (20 - 34 years), while 45 % ($n = 36$) were middle-aged (35 – 54 years), and 65 % ($n = 52$) were in full-time employment. The participants' average BMI was 25.3. The distribution between the control and the gamified groups was relatively well-balanced in terms of all the demographic variables, with the exception of having less participants in the control group who had children under the age of sixteen, even though randomisation was employed. The distribution of participants between the groups was relatively well-balanced in terms of their level of physical activity at baseline, suggesting a balanced mix of participants ranging from a sedentary or low activity lifestyle to a highly active lifestyle. The classification of the baseline step counts into physical activity levels ranging from a sedentary lifestyle to a highly active lifestyle is based on established pedometer-determined physical activity

[7] Details about the missing step data for each group and the number of participant-days with step count data less than 1K step count per day are presented in the Supplementary Material.

Table 3
Sample characteristics for control versus gamification group.

Sample characteristics	Control group	Gamification group	p value
Gender, n (%)			
Male	11 (55.0 %)	24 (40.0 %)	0.24
Female	9 (45.0 %)	36 (60.0 %)	
Age Groups, n (%)			
Young adulthood (20 - 34 years)	13 (65.0 %)	29 (48.3 %)	0.36
Middle aged (35 - 54 years)	7 (35.0 %)	29 (48.3 %)	
Older adulthood (55+ years)	0 (0 %)	2 (3.3 %)	
Employment Status, n (%)			
Full-time employed	14 (70.0 %)	38 (63.3 %)	0.59
Part-time employed	4 (20.0 %)	13 (21.7 %)	0.88
Full-time student	7 (35.0 %)	20 (33.3 %)	0.89
Part-time student	8 (40.0 %)	18 (30.0 %)	0.41
Have children under 16 years, n (%)	1 (5.0 %)	17 (28.3 %)	0.03
Have a steady relationship, n (%)	13 (65.0 %)	45 (75.0 %)	0.39
Have sufficient income, n (%) *(see note)*	18 (90.0 %)	52 (88.1 %)	0.82
BMI Pre Intervention, mean (SD)	26.59 (4.23)	24.80 (4.69)	0.13
Familiarity with technology, mean (SD)	5.30 (1.92)	5.65 (1.40)	0.46
Baseline level of physical activity, n (%)			
Sedentary lifestyle (<5000 steps/day)	2 (10.0 %)	11 (18.3 %)	0.48
Low active (=> 5000 and <7500 steps/day)	7 (35.0 %)	17 (28.3 %)	
Somewhat active (=>7500 and <10,000 steps/day)	7 (35.0 %)	15 (25.0 %)	
Active lifestyle (=>10,000 and <12,500 steps/day)	1 (5.0 %)	11 (18.3 %)	
Highly active lifestyle (=>12,500 steps/day)	3 (15.0 %)	6 (10.0 %)	

Note: $N = 80$; 1 respondent provided no data to the question related to income.

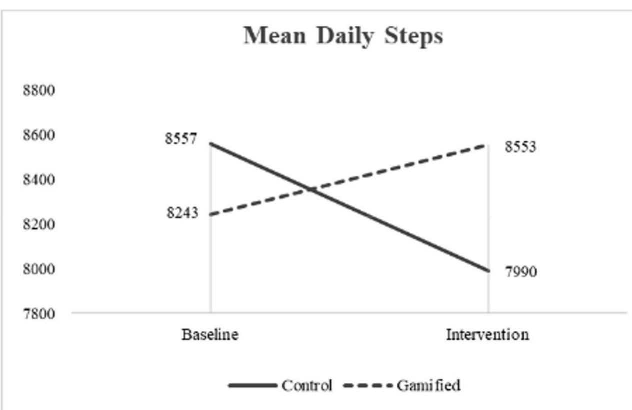

Fig. 5. Mean daily step count at baseline versus intervention period.

levels identified in literature (Tudor-Locke and Bassett, 2004).

4.1.1. *Behavioural outcome*

The mean (SD) baseline daily step count for the control group was 8557 (SD = 2916), and 8243 (SD = 3837) for the gamified groups. Differences in baseline levels between the groups following randomisation are due to chance (Roberts and Torgerson, 1999). Nonetheless, an independent *t*-test on these baseline levels confirms that there were no statistical differences in the mean daily step counts of the control and the gamified groups (t(78) = 0.334; $p = 0.739$).

During the intervention period, the mean daily step count of the gamified groups increased, whilst that of the control group declined (Fig. 5). The result from the longitudinal data analysis of covariance (see Table 4) shows that gamification resulted in a positive effect on the daily step count (adjusted difference from control = 811; 95 % CI = 57 to 1565; $p = 0.035$) supporting Hypothesis 1: Gamified groups will report higher step counts than the control group during the intervention period (see Table 4). The effect size of 0.25 is interpreted as a small effect (Ellis, 2010; Hedges and Olkin, 1985; Lenhard and Lenhard, 2016). To increase the robustness of our findings, the analysis was repeated including gender as a covariate in the model. The results remain unchanged; gamification increased the mean daily step count (adjusted difference from control = 772 steps; 95 % CI= 17 to 1528; $p = 0.045$) and gender was not a significant covariate ($p = 0.482$).

The analysis of step counts at different timepoints during the intervention period indicated that both the control and the gamified groups exhibit a similar progressive decay in the mean daily steps during the intervention period (see Fig. 6). In the absence of gamification, there is a statistically significant decline of 1055 steps in the control group by the end of the intervention (t(19) = −2.674, $p = 0.015$, Hedge's $g = 0.37$). The mean daily step count of the control group decreased week after week ending at 87.7 % of their baseline level by the end of the study.

Table 4
Gamification effect on mean daily step count.

Timepoint	Control mean (SD)	Gamified mean (SD)	Overall treatment effect: adjusted between-groups difference (95 % CI)	p-value	Hedge's g
Baseline	8557 (2916)	8243 (3837)	811 (57 to 1565)	0.035	0.25
Intervention	7990 (2146)	8553 (3690)			

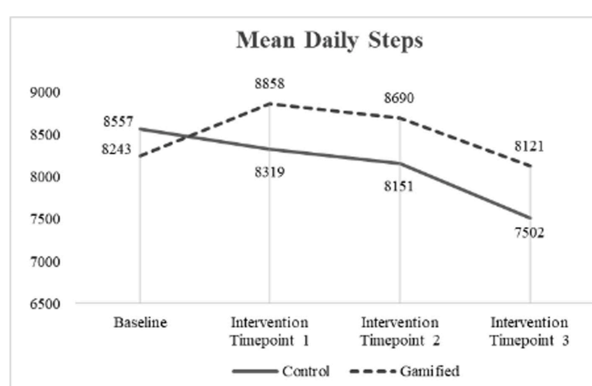

Fig. 6. Mean daily steps at different timepoints.

Table 5
Sample characteristics – Demographics characteristics for each treatment group.

Demographic characteristics	Control group (n = 20)	Competition group (n = 20)	Cooperation group (n = 20)	Hybrid group (n = 20)	p value
Gender, n (%)					
Male	11 (55.0 %)	5 (25.0 %)	13 (65.0 %)	6 (30.0 %)	0.03
Female	9 (45.0 %)	15 (75.0 %)	7 (35.0 %)	14 (70.0 %)	
Age Groups, n (%)					
Young adulthood (20 - 34 years)	13 (65.0 %)	6 (30.0 %)	11 (55.0 %)	12 (60.0 %)	0.30
Middle aged (35 - 54 years)	7 (35.0 %)	13 (65.0 %)	8 (40.0 %)	8 (40.0 %)	
Older adulthood (55+ years)	0 (0 %)	1 (5.0 %)	1 (5.0 %)	0 (0 %)	
Employment Status, n (%)					
Full-time employed	14 (70.0 %)	10 (50.0 %)	12 (60.0 %)	16 (80.0 %)	0.22
Part-time employed	4 (20.0 %)	4 (20.0 %)	5 (25.0 %)	4 (20.0 %)	0.97
Full-time student	7 (35.0 %)	6 (30.0 %)	8 (40.0 %)	6 (30.0 %)	0.89
Part-time student	8 (40.0 %)	4 (20.0 %)	7 (35.0 %)	7 (35.0 %)	0.56
Have children under 16 years, n (%)	1 (5.0 %)	6 (30 %)	7 (35.0 %)	4 (20 %)	0.11
Have a steady relationship, n (%)	13 (65.0 %)	15 (75.0 %)	15 (75.0 %)	15 (75.0 %)	0.86
Have sufficient income, n (%)	18 (90.0 %)	17 (89.5 %)	17 (89.5 %)	18 (90.0 %)	0.95
BMI Pre Study, mean (SD)	26.59 (4.23)	24.44 (5.42)	25.20 (3.48)	24.76 (5.15)	0.48
Familiarity with technology, mean (SD)	5.30 (1.92)	5.35 (1.66)	5.85 (1.46)	5.75 (1.02)	0.59
Baseline level of physical activity, n (%)					
Sedentary lifestyle (<5000 steps/day)	2 (10.0 %)	3 (15.0 %)	4 (20.0 %)	4 (20.0 %)	0.83
Low active (=> 5000 and <7500 steps/day)	7 (35.0 %)	6 (30.0 %)	5 (25.0 %)	6 (30.0 %)	
Somewhat active (=>7500 and <10,000 steps/day)	7 (35.0 %)	4 (20.0 %)	4 (20.0 %)	7 (35.0 %)	
Active lifestyle (=>10,000 and <12,500 steps/day)	1 (5.0 %)	4 (20.0 %)	5 (25.0 %)	2 (10.0 %)	
Highly active lifestyle (=>12,500 steps/day)	3 (15.0 %)	3 (15.0 %)	2 (10.0 %)	1 (5.0 %)	

Note: $N = 80$; 1 respondent provided no data to the question related to income.

Over the course of the intervention, both groups experienced a decline in the physical activity recorded (decline in step count over time = 408 steps; 95 % CI= −756 to −60; p = 0.022). However, notwithstanding the progressive decay in the physical activity levels during the intervention period, the relative advantage of the gamified groups in comparison to the control group is sustained throughout the intervention period (Fig. 6). Participants in the gamified group did 839 more steps per day compared to those in the control group by the end of the intervention (adjusted difference from control = 839, 95 % CI = 4 to 1673, p = 0.049, Hedge's g = 0.28). This provides further evidence that there is potential to leverage gamification to increase physical activity and that the relative advantage is sustained in the short-term.

In order to estimate which gamification design was the most effective at increasing physical activity, the effect of each gamified design (competitive, cooperative and hybrid designs) was analysed. Baseline participants' characteristics in terms of the demographic characteristics and pre-intervention physical activity levels for each treatment group versus the control group are set out in Table 5. The p-values presented are the results of the Chi-Square tests for categorical variables and one-way ANOVA tests for continuous variables carried out to check for significant differences between the sub-groups. When comparing the distribution of participants across all sub-groups, there was a relative gender imbalance (more females in the competition and hybrid groups). Nonetheless, gender did not prove to be significantly different when comparing the control versus the gamified groups (see Table 3).

At baseline all groups had a mean daily step count in the range of 7600 to 8950 daily steps. The mean baseline daily steps (SD) were 8557 (SD = 2916) for the control group, 8202 (SD = 3512) for the cooperation group, 7608 (SD = 3123) for the hybrid group and 8920 (SD = 4766) for the competition group. To assess for significant differences between the groups in the baseline step counts, a one-way ANOVA test was carried out. The result shows that there were no statistically significant differences in the baseline step counts (F(3,76) = 0.469, p = 0.705).

As set out in Fig. 7, we observe that during the intervention period, the mean daily step count of all the gamified groups increased, whilst that of the control group declined. The results from the longitudinal analysis of covariance shows that the hybrid gamified group was the most effective at increasing step counts (adjusted difference from control = 981 steps; 95 % CI= −45 to 2008; p = 0.061), supporting Hypothesis 2: Hybrid design will facilitate the strongest effect on step

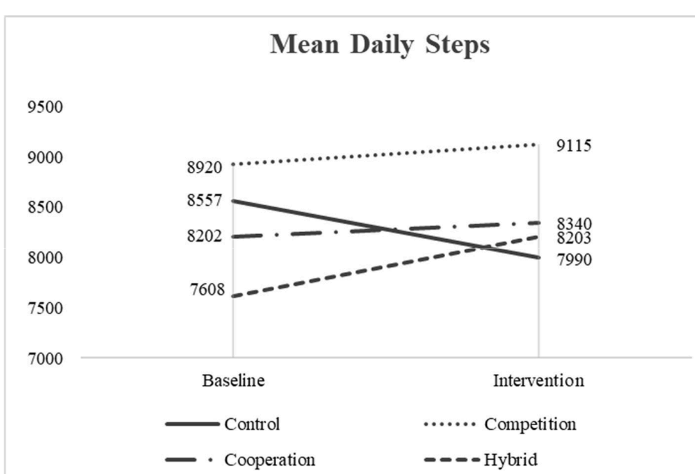

Fig. 7. Mean daily step count for each treatment group.

Table 6
Effect of different gamified interventions on the mean daily step count.

Timepoint	Control mean (SD)	Competition mean (SD)	Cooperation mean (SD)	Hybrid mean (SD)
Baseline	8557 (2916)	8920 (4766)	8202 (3512)	7608 (3123)
Intervention	7990 (2146)	9115 (4449)	8340 (3391)	8203 (3238)
Adjusted difference from control		817	637	981
95 % CI		−138 to 1772	−328 to 1602	−45 to 2008
p value		0.094	0.196	0.061
Hedge's g		0.23	0.23	0.39

counts (see Table 6). The increase in step counts is significant for the hybrid and competitive gamified groups at 0.1 level of significance. The hybrid gamified design, in which participants were involved in an inter-team competition had the strongest effect (Hedge's $g = 0.39$). Smaller effects were reported for the pure competitive (Hedge's $g = 0.23$) and pure cooperative (Hedge's $g = 0.23$) gamified designs.

To increase the robustness of our findings, the analysis was repeated including gender as a covariate in the model. The hybrid design was still the most effective gamification design (adjusted difference from control = 928 steps; 95 % CI= −86 to 1943; $p = 0.073$), and gender was not a significant covariate ($p = 0.540$).

4.1.2. Intrinsic motivation and perceived usefulness

This study measured the participants' *intrinsic motivation* and *perceived usefulness* of the experience at the end of the intervention period. Each construct subscale was found to be reliable, indicating internal consistency amongst the scale items used to measure each specific construct. Intrinsic Motivation sub-scale (α = 0.735; CR = 0.859; AVE = 0.677) resulted in a scale with $M = 6.22$ and SD = 0.867, and the Perceived Usefulness sub-scale (α = 0.808; CR = 0.891; AVE = 0.734) resulted in a scale with $M = 5.46$ and SD = 1.288. The descriptive statistics including the means and the standard deviation for both constructs are presented in Table 7. As observed in the results below (see Fig. 8 and Fig. 9), both the control group and the gamified group reported similar positive experiential outcomes.

Table 7
Means and standard deviation for users' intrinsic motivation and perceived usefulness.

	Intrinsic Motivation		Perceived Usefulness	
Group	mean	SD	mean	SD
Gamified	6.16	0.901	5.52	1.324
Control	6.40	0.746	5.28	1.186

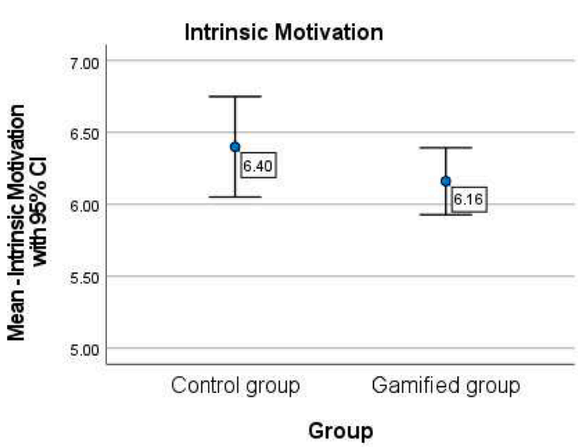

Fig. 8. Error bar chart - Intrinsic motivation (control vs. gamified group).

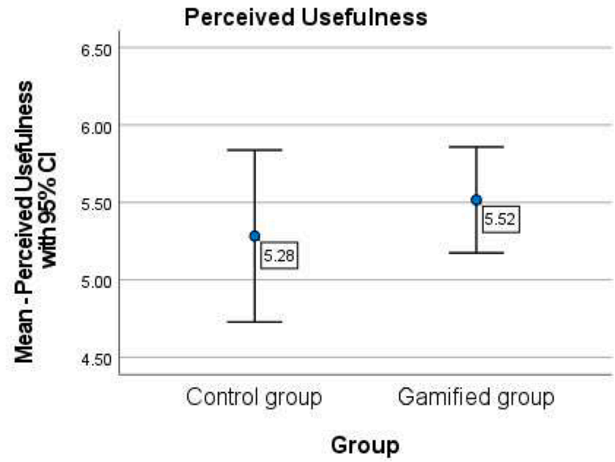

Fig. 9. Error bar chart - Perceived usefulness (control vs. gamified group).

A Shapiro-Wilk test was conducted to determine whether the score distribution of each construct follows a normal distribution. As presented in Table 8, the score distributions do not follow the normal distribution, thus non-parametric inferential statistical tests were appropriate.

Mann-Whitney test results revealed no statistically significant differences between the control group and the gamified group in intrinsic motivation scores ($U = 513.50$, $z = -0.988$, $p = 0.323$) and perceived usefulness scores ($U = 509.00$, $z = -1.017$, $p = 0.309$). Therefore, both Hypothesis 3a and Hypothesis 3b listed below are rejected.

H3a. Gamified groups will report higher intrinsic motivation than the control group

H3b. Gamified groups will report higher perceived usefulness than the control group

Furthermore, descriptive statistics for both experiential outcome constructs for each treatment group were computed (see Table 9, Figs. 10 and 11). A Kruskal-Wallis H test results revealed no statistically significant differences between the groups for both Intrinsic Motivation ($\chi^2(3) = 1.160$, $p = 0.657$), and Perceived Usefulness ($\chi^2(3) = 1.969$, $p = 0.579$).

Table 8
Shapiro-Wilk test result for intrinsic motivation and perceived usefulness constructs.

	Shapiro-Wilk		
Variable	Statistic	Df	Sig.
Intrinsic Motivation	0.838	80	<0.001
Perceived Usefulness	0.920	80	<0.001

Table 9
Means and standard deviation for users' intrinsic motivation and perceived usefulness.

	Intrinsic Motivation		Perceived Usefulness	
Group	Mean	SD	Mean	SD
Control	6.40	0.746	5.28	1.186
Cooperation	6.08	0.910	5.47	1.126
Hybrid	6.25	0.830	5.68	1.348
Competition	6.15	0.994	5.40	1.520

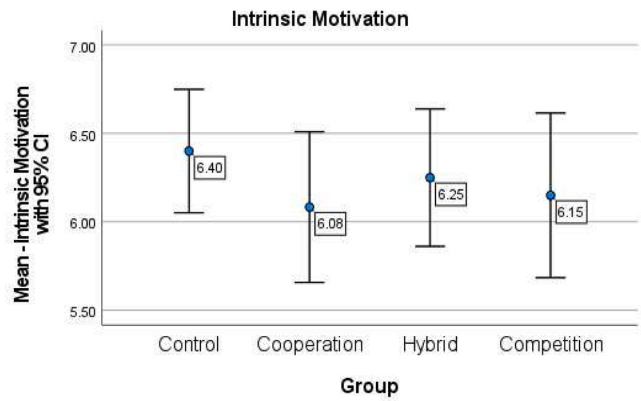

Fig. 10. Error bar chart - Intrinsic motivation (control vs. different treatment groups).

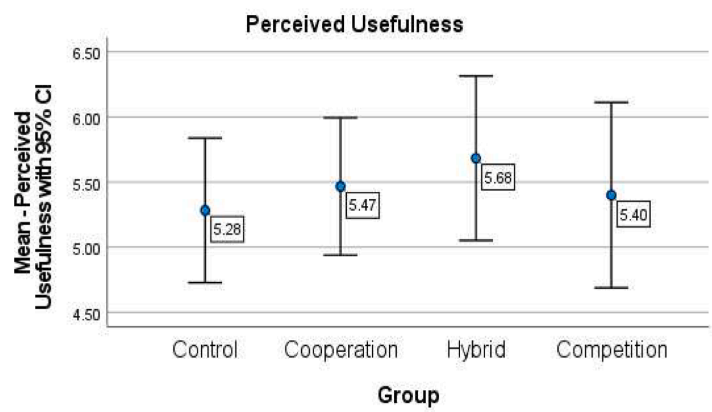

Fig. 11. Error bar chart - Perceived usefulness (control vs. different treatment groups).

5. Discussion and conclusion

In view of the need for rigorous experimental studies to isolate and estimate the effects of gamification (Johnson et al., 2016; Koivisto and Hamari, 2019a), this study involved a randomised controlled field experiment to investigate the effect resulting from gamification of physical activity on the users' motivation, perceived usefulness, and the behavioural change in step counts. The hypotheses, outcomes and conclusions of this study are summarised in Table 10.

Our results show that gamification improves physical activity behavioural outcomes. Relative to a non-gamified self-tracking experience of physical activity, a gamified experience resulted in significant increase in step counts (Hedge's $g = 0.25$). The positive effect is considered to be a small effect (Ellis, 2010; Hedges and Olkin, 1985; Lenhard and Lenhard, 2016) supporting the findings of a recent meta-analysis synthesising existing empirical evidence in this domain (Mazéas et al., 2022). While a progressive decay trend in the mean daily step count is evident in both the gamified groups and the control group, the relative advantage of the gamified group in comparison to the control group is sustained throughout the intervention period. This finding provides further evidence that there is potential to leverage gamification to increase physical activity.

While literature has mainly focused on whether gamification works (Nacke and Deterding, 2017), this study extends the contribution to literature by providing insight into which gamification design facilitated the strongest behavioural change. The results revealed that a hybrid gamification design using an inter-team competition facilitated the highest increase in step counts (Hedge's $g = 0.39$). The pure competitive

Table 10
Summary of hypotheses, outcomes, and conclusions.

Hypothesis prediction	Statistical outcome	Conclusions
H1: Gamified groups will report higher step counts than the control group during the intervention period	Gamification resulted in an increase in the daily step count (adjusted difference from control = 811; 95 % CI = 57 to 1565; $p = 0.035$).	**H1 supported.** Gamification led to a behavioural change in physical activity, namely an increase in step counts.
H2: Hybrid design will facilitate the strongest effect on step counts	Hybrid group was the most effective at increasing step counts (adjusted difference from control = 981 steps; 95 % CI= −45 to 2008; $p = 0.061$).	**H2 supported.** A combination of competitive and collaborative features is more effective than including only cooperative or competitive gamification features.
H3: Gamified group will report higher intrinsic motivation and perceived usefulness than the control group	No statistically significant differences between the control group and the gamified group in terms of intrinsic motivation ($U = 513.50$, $z = -0.988$, $p = 0.323$) and perceived usefulness ($U = 509.00$, $z = -1.017$, $p = 0.309$).	**H3 rejected.** At the end of the intervention, gamification did not evoke a stronger response in terms of intrinsic motivation and perceived usefulness relative to a non-gamified self-tracking experience.

and cooperative gamification designs resulted in smaller positive effects (Hedge's $g = 0.23$). Supporting cooperation along competition as suggested in previous literature (Rapp, 2015) promises better outcomes. This result is consistent with earlier studies like Tauer and Harackiewicz (2004) where the combination of cooperation and competition in sports led to greater benefits, and Morschheuser et al. (2019) who looked at the effects of gamification in the crowdsourcing domain. However, the result of this study contrasts with previous empirical evidence investigating gamification of physical activity (Chen and Pu, 2014), which found that a cooperative design facilitated higher physical activity than a hybrid design.

Comparing the study design, procedures and setting with previous studies may shed light on the differences in the findings, as well as provide valuable insights for future intervention studies. This study was conducted amongst people with no pre-existing social connections. Moreover, participants were assigned a unique code to ensure anonymity. A previous randomised controlled trial (Patel et al., 2019) involving participants who had no pre-existing social connections found that competition was the most effective strategy compared to social support and cooperation (the hybrid design was not included in this trial). On the other hand, the study conducted by Chen and Pu (2014) reporting that a cooperative design was more effective than a hybrid design was implemented amongst dyads of close friends. This observation suggests that cooperative designs may be more effective when individuals have close social connections, while competitive designs seem to be more effective when participants do not have pre-existing social connections or in anonymised settings.

Considering the level of physical activity of participants in this study, all the groups in this study had a baseline step count in the range of 7600 to 9000 steps per day which is classified as a 'somewhat active lifestyle' (Tudor-Locke and Bassett, 2004). In a previous study conducted by Patel et al. (2019), the greatest increase in step counts was recorded amongst participants with low baseline step counts that had less than 7500 steps per day. Similarly, positive effects of gamification reported by Gremaud et al. (2018) involved sedentary office workers. This suggests that the effect on physical activity could be higher amongst participants who have a predominantly sedentary lifestyle.

In addition to the behavioural outcome of gamification, this study examined psychological outcomes in terms of the users' intrinsic motivation and perceived usefulness. The findings from this study reveal that while the use of gamification stimulated the desired behavioural change in step counts, it did not evoke stronger psychological responses relative to a non-gamified self-tracking experience at the end of the intervention. Literature suggests that enjoyment and perceived usefulness of gamification declines with use (Koivisto and Hamari, 2014). Thus, future work should consider more frequent measurements of these outcomes during the intervention period, rather than at the end of the intervention. More granular data would enable researchers to identify any potential variations, trends or patterns over time related to the users' psychological responses. The results suggest that gamification can act as a stimulus or a nudge that begets desired behavioural outcomes (in this case an increase in step counts), without stimulating different psychological responses (relative to non-gamified self-tracking interventions). This contribution is theoretically interesting and merits further research into the mechanisms at play.

Our study extends existing literature on the effect of gamification in the domain of physical activity - an area of research where rigorous empirical evidence is limited (Koivisto and Hamari, 2019a; Mazéas et al., 2022). To our knowledge this is the first randomised controlled field experiment investigating gamification of physical activity with competitive, cooperative, and hybrid gamified designs. The behavioural effect of gamification was examined on step count data gathered through physical activity trackers, rather than relying on self-reported measures. Our study also involved an active control group using a non-gamified self-monitoring experience to isolate the behavioural effect of gamification. We also took into consideration that physical activity is more accurately measured having a physical activity tracker as a wearable device, rather than relying on the sensor-based technologies of smartphones. In view of the lack of knowledge on which gamification elements or designs produce positive effects (Koivisto and Hamari, 2019a; Mazéas et al., 2022), we opted for a four-arm randomised controlled experiment to test the effect of different designs of gamification. Furthermore, while previous experimental research examining the effect of gamification on physical activity behaviour have largely neglected the users' psychological responses to gamified interventions (Mazéas et al., 2022), our study investigated both psychological and behavioural outcome measures resulting from gamification of physical activity.

Three key limitations to this study could be addressed in future studies. First, the small sample size limited the potential of further subgroups to test specific game elements and further mechanisms. Also, having a larger dataset would support more covariates in the analysis. Finding significant effects despite the sample size strengthens the evidence that gamification could motivate behaviour change in physical activity. Some ideas that could be tested in future studies include the possibility of introducing new game elements during the intervention, offering tangible rewards rather than virtual rewards when targets are achieved, having participants' identities disclosed (provided that ethical issues are complied with) and giving more opportunities for social interaction amongst participants.

Second, although this study includes a randomised controlled field experiment over a period of four weeks, it is still considered as a relatively short timeframe and longer interventions are encouraged in future studies. Understandably, there are challenges to conduct randomised controlled trials using wearable devices (to achieve objective data) with large sample sizes and longer timeframes. However, the accumulation of knowledge from rigorous empirical studies on the effect of gamified interventions on health-related behaviours would have practical relevance. To achieve a comprehensive evaluation, it would be beneficial to employ rigorous studies that consider a combination of outcomes resulting from gamification (Nacke and Deterding, 2017; Rapp et al., 2019).

Third, at the time of planning this study, an off-the-shelf fitness application that catered for the scope of the study with three distinct gamified experiences (competition, cooperation, hybrid) could not be identified. Therefore, gamified interventions were purposely designed using a gamification platform (Pointagram). However, an automatic integration of the step counts gathered through the wearable devices to the gamified application was not possible, and thus participants were asked to submit their daily step counts on the gamified application. Participants' adherence to submitting their daily step counts was high. Also, the mismatch between the data inputted on the gamified application and the data generated from the wearable devices was minimal. To avoid any reliance on human error and ensure accuracy, the data analysed for this study was gathered from the wearable devices. Nonetheless, advances in technological developments including data integration and data analytics could facilitate the design of gamified systems and the data insights that could be generated.

To conclude, results from this randomised controlled field experiment show that gamification induced a positive behavioural change in terms of step counts, especially with the implementation of a hybrid (competitive-cooperative) gamified design. While the use of gamification stimulated the desired behavioural change, the psychological responses to gamification (albeit positive) were not stronger when compared to a non-gamified self-tracking experience at the end of the intervention period. The finding that gamification can stimulate a stronger behavioural outcome, but does not evoke a stronger psychological outcome at the end of the intervention merits further investigation. Future work could investigate further the link between psychological outcomes and behavioural outcomes and explore how gamification mechanisms work.

Funding details

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijhcs.2023.103205.

References

Ajzen, I., 1991. The theory of planned behavior. Organ Behav. Hum. Decis. Process 50 (2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Aldenaini, N., Alqahtani, F., Orji, R., Sampalli, S., 2020a. Trends in persuasive technologies for physical activity and sedentary behavior: a systematic review. Front. Artif. Intell. 3 (April) https://doi.org/10.3389/frai.2020.00007.

Aldenaini, N., Oyebode, O., Orji, R., Sampalli, S., 2020b. Mobile phone-based persuasive technology for physical activity and sedentary behavior: a systematic review. Front. Comp. Sci. 2 (July), 1–17. https://doi.org/10.3389/fcomp.2020.00019.

Broom, D.R., Flint, S.W., 2018. Gotta catch 'em all: impact of pokémon go on physical activity, sitting time, and perceptions of physical activity and health at baseline and three-month follow-up. Games Health J. 7 (6), 401–408. https://doi.org/10.1089/g4h.2018.0002.

Buckley, J., DeWille, T., Exton, C., Exton, G., Murray, L., 2018. A gamification–motivation design framework for educational software developers. J. Educat. Techn. Sys. 47 (1), 101–127. https://doi.org/10.1177/0047239518783153.

Buunk, A.P., Gibbons, F.X., 2007. Social comparison: the end of a theory and the emergence of a field. Organ. Behav. Hum. Decis. Process 102 (1), 3–21. https://doi.org/10.1016/j.obhdp.2006.09.007.

Chen, Y., Pu, P., 2014. HealthyTogether: exploring social incentives for mobile fitness applications. In: ACM International Conference Proceeding Series, pp. 25–34. https://doi.org/10.1145/2592235.2592240. April 2014.

Chokshi, N.P., Adusumalli, S., Small, D.S., Morris, A., Feingold, J., Ha, Y.P., Lynch, M.D., Rareshide, C.A.L., Hilbert, V., Patel, M.S., 2018. Loss-framed financial incentives and personalized goal-setting to increase physical activity among ischemic heart disease patients using wearable devices: the active reward randomized trial. J. Am. Heart Assoc. 7 (12) https://doi.org/10.1161/JAHA.118.009173.

Chow, S.C., Shao, J., Wang, H., & Lokhnygina, Y. (2017). *Sample Size Calculations in Clinical Research:* 3rd Edition (S.-C. Chow, J. Shao, H. Wang, & Y. Lokhnygina, Eds.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315183084.

Cialdini, R.B., Trost, M.R., 1998. Social influence: social norms, conformity and compliance. In: Gilbert, D.T., Fiske, S.T., Lindzey, G. (Eds.), The Handbook of Social Psychology, 4th editio. McGraw-Hill, pp. 151–192.

Cooper, K., Barton, G.C., 2016. An exploration of physical activity and wellbeing in university employees. Perspect. Public Health 136 (3), 152–160. https://doi.org/10.1177/1757913915593103.

Cotton, V., Patel, M.S., 2019. Gamification use and design in popular health and fitness mobile applications. Am. J. Health Promot. 33 (3), 448–451. https://doi.org/10.1177/0890117118790394.

Dadaczynski, K., Schiemann, S., Backhaus, O., 2017. Promoting physical activity in worksite settings: results of a German pilot study of the online intervention Healingo fit. BMC Public Health 17 (1), 1–9. https://doi.org/10.1186/s12889-017-4697-6.

Deci, E.L., Betley, G., Kahle, J., Abrams, L., Porac, J., 1981. When trying to win. Personal. Soc. Psychol. Bull. 7 (1), 79–83. https://doi.org/10.1177/014616728171012.

Deci, E.L., Eghrari, H., Patrick, B.C., Leone, D.R., 1994. Facilitating internalization: the self-determination theory perspective. J. Pers. 62 (1), 119–142. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x.

Deci, E.L., Koestner, R., Ryan, R.M., 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychol. Bull. 125 (6), 627–668. https://doi.org/10.1037/0033-2909.125.6.627.

Deci, E.L., Ryan, R.M., 1985. Intrinsic Motivation and Self-Determination in Human Behaviour. Plenum.

Deci, E.L., Ryan, R.M., 2008. Self-determination theory: a macrotheory of human motivation, development, and health. Canadian Psychol. /Psychol. Canadienne 49 (3), 182–185. https://doi.org/10.1037/a0012801.

Deterding, S., 2019. Gamification in management: between choice architecture and humanistic design. J. Manage. Inq. 28 (2), 131–136. https://doi.org/10.1177/1056492618790912.

Deterding, S., Dixon, D., Khaled, R., Nacke, L., Sicart, M., O'Hara, K, 2011. Gamification: using game design elements in non-game contexts. In: Proceedings of the 2011 Annual Conference Extended Abstracts on Human Factors in Computing Systems (CHI 2011), pp. 2425–2428. https://doi.org/10.1145/1979742.1979575.

Direito, A., Jiang, Y., Whittaker, R., Maddison, R., 2015. Apps for IMproving FITness and increasing physical activity among young people: the AIMFIT pragmatic randomized controlled trial. J. Med. Internet Res. 17 (8), 1–13. https://doi.org/10.2196/jmir.4568.

Edney, S.M., Olds, T.S., Ryan, J.C., Vandelanotte, C., Plotnikoff, R.C., Curtis, R.G., Maher, C.A., 2020. A social networking and gamified app to increase physical activity: cluster RCT. Am. J. Prev. Med. 58 (2), e51–e62. https://doi.org/10.1016/j.amepre.2019.09.009.

Ellis, P.D., 2010. The Essential Guide to Effect Sizes. The Essential Guide to Effect Sizes. Cambridge University Press. https://doi.org/10.1017/cbo9780511761676.

Epstein, J.A., Harackiewicz, J.M., 1992. Winning is not enough: the effects of competition and achievement orientation on intrinsic interest. Personal. Soc. Psychol. Bull. 18 (2), 128–138. https://doi.org/10.1177/0146167292182003.

Farrow, M., Lutteroth, C., Rouse, P.C., Bilzon, J.L.J., 2019. Virtual-reality exergaming improves performance during high-intensity interval training. Eur. J. Sport Sci 19 (6), 719–727. https://doi.org/10.1080/17461391.2018.1542459.

Festinger, L., 1954. A theory of social comparison processes. Human Relat. 7, 117–140. https://doi.org/10.1177/001872675400700202.

Fiedler, J., Eckert, T., Burchartz, A., Woll, A., Wunsch, K., 2021. Comparison of self-reported and device-based measured physical activity using measures of stability, reliability, and validity in adults and children. Sensors 21 (8), 2672. https://doi.org/10.3390/s21082672.

Fogg, B.J., 2002. Persuasive technology. Ubiquity 2002 (December), 2. https://doi.org/10.1145/764008.763957.

Garde, A., Umedaly, A., Abulnaga, S.M., Junker, A., Chanoine, J.P., Johnson, M., Ansermino, J.M., Dumont, G.A., 2016. Evaluation of a novel mobile exergame in a school-based environment. Cyberpsych., Behav., Social Network. 19 (3), 186–192. https://doi.org/10.1089/cyber.2015.0281.

Geelan, B., Zulkifly, A., Smith, A., Cauchi-Saunders, A., De Salas, K., Lewis, I., 2016. Augmented exergaming: increasing exercise duration in novices. In: *Proceedings of the 28th Australian Computer-Human Interaction Conference, OzCHI 2016, May 2019*, pp. 542–550. https://doi.org/10.1145/3010915.3010940.

Gremaud, A.L., Carr, L.J., Simmering, J.E., Evans, N.J., Cremer, J.F., Segre, A.M., Polgreen, L.A., Polgreen, P.M., 2018. Gamifying accelerometer use increases physical activity levels of sedentary office workers. J. Am. Heart Assoc. 7 (13), 1–12. https://doi.org/10.1161/JAHA.117.007735.

Hair, J., Black, W., Babin, B., Anderson, R., 2010. Multivariate Data Analysis, 7th Ed. Pearson Education.

Hamari, J., Hassan, L., Dias, A., 2018. Gamification, quantified-self or social networking? Matching users' goals with motivational technology. User Model User-adapt. Interact 28 (1), 35–74. https://doi.org/10.1007/s11257-018-9200-2.

Hamari, J., Koivisto, J., 2013. Social motivations to use gamification: an empirical study of gamifying exercise. In: ECIS 2013 Proceedings. http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1328&context=ecis2013_cr.

Hamari, J., Koivisto, J., 2015a. Why do people use gamification services? Int. J. Inf. Manage. 35 (4), 419–431. https://doi.org/10.1016/j.ijinfomgt.2015.04.006.

Hamari, J., Koivisto, J., 2015b. Working out for likes": an empirical study on social influence in exercise gamification. Comput. Human Behav. 50, 333–347. https://doi.org/10.1016/j.chb.2015.04.018.

Hamari, J., Koivisto, J., Sarsa, H., 2014. Does gamification work? - a literature review of empirical studies on gamification. In: Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 3025–3034.

Harris, M.A., 2018. Beat the street: a pilot evaluation of a community-wide gamification-based physical activity intervention. Games Health J. 7 (3), 208–212. https://doi.org/10.1089/g4h.2017.0179.

Hedges, L.V., Olkin, I., 1985. Statistical Methods for Meta-Analysis. Elsevier. https://doi.org/10.1016/C2009-0-03396-0.

Höchsmann, C., Müller, O., Ambühl, M., Klenk, C., Königstein, K., Infanger, D., Walz, S. P., Schmidt-Trucksäss, A., 2019. Novel smartphone game improves physical activity behavior in type 2 diabetes. Am. J. Prev. Med. 57 (1), 41–50. https://doi.org/10.1016/j.amepre.2019.02.017.

Huotari, K., Hamari, J., 2012. Defining gamification - a service marketing perspective. In: Proceeding of the 16th International Academic MindTrek Conference, pp. 17–22. https://doi.org/10.1145/2393132.2393137.

Huotari, K., Hamari, J., 2017. A definition for gamification: anchoring gamification in the service marketing literature. Electron. Markets 27 (1), 21–31. https://doi.org/10.1007/s12525-015-0212-z.

Johnson, D., Deterding, S., Kuhn, K.A., Staneva, A., Stoyanov, S., Hides, L., 2016. Gamification for health and wellbeing: a systematic review of the literature. Intern. Intervent. 6, 89–106. https://doi.org/10.1016/j.invent.2016.10.002.

Johnson, D.W., 2003. Social interdependence: interrelationships among theory, research, and practice. Am. Psychol. 58 (11), 934–945. https://doi.org/10.1037/0003-066X.58.11.934.

Johnson, D.W., Johnson, R.T., 1989. Cooperation and Competition: Theory and Research. Interaction Book Company Edina.

Johnson, D.W., Johnson, R.T., 2005. New developments in social interdependence theory. Genet. Soc. Gen. Psychol. Monogr. 131 (4), 285–358.

Kaczmarek, L.D., Misiak, M., Behnke, M., Dziekan, M., Guzik, P., 2017. The Pikachu effect: social and health gaming motivations lead to greater benefits of Pokémon GO

use. Comput. Human Behav. 75, 356–363. https://doi.org/10.1016/j.chb.2017.05.031.
Kelman, H.C., 1958. Compliance, identification, and internalization three processes of attitude change. J. Conflict Resol. 2 (1), 51–60. https://doi.org/10.1177/002200275800200106.
Koivisto, J., Hamari, J., 2014. Demographic differences in perceived benefits from gamification. Comput. Human Behav. 35, 179–188. https://doi.org/10.1016/j.chb.2014.03.007.
Koivisto, J., Hamari, J., 2019a. Gamification of physical activity: a systematic literature review of comparison studies. In: 3rd International GamiFIN Conference, GamiFIN 2019. CEUR-WS, pp. 106–117.
Koivisto, J., Hamari, J., 2019b. The rise of motivational information systems: a review of gamification research. Int. J. Inf. Manage. 45 (December 2018), 191–210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013.
Kurtzman, G.W., Day, S.C., Small, D.S., Lynch, M., Zhu, J., Wang, W., Rareshide, C.A.L., Patel, M.S., 2018. Social incentives and gamification to promote weight loss: the lose it randomized, controlled trial. J. Gen. Intern. Med. 33 (10), 1669–1675. https://doi.org/10.1007/s11606-018-4552-1.
Landers, R.N., Collmus, A.B., Williams, H., 2019. The greatest battle is within ourselves: an experiment on the effects of competition alone on task performance. Internat. J. Human Comp. Stud. 127 (September 2017), 51–61. https://doi.org/10.1016/j.ijhcs.2018.09.011.
Latham, G.P., Locke, E.A., 1991. Self-regulation through goal setting. Organ. Behav. Hum. Decis. Process. 50 (2), 212–247. https://doi.org/10.1016/0749-5978(91)90021-K.
Lenhard, W., Lenhard, A., 2016. Computation of different effect sizes. Psychometrica. https://www.psychometrica.de/effect_size.html.
Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H.B., 2006. Fish'n'Steps: encouraging physical activity with an interactive computer game. In: Dourish, P., Friday, A. (Eds.), UbiComp 2006: Ubiquitous Computing. Springer, Berlin Heidelberg, pp. 261–278.
Liu, D., Li, X., Santhanam, R., 2013. Digital games and beyond: what happens when players compete? MIS Quarterly 37 (1), 111–124. https://doi.org/10.25300/MISQ/2013/37.1.05.
Liu, D., Santhanam, R., Webster, J., 2017. Toward meaningful engagement: a framework for design & research of gamified information systems. MIS Quarterly 41 (4), 1011–1034.
Maher, C., Ferguson, M., Vandelanotte, C., Plotnikoff, R., De Bourdeaudhuij, I., Thomas, S., Nelson-Field, K., Olds, T., 2015. A web-based, social networking physical activity intervention for insufficiently active adults delivered via Facebook app: randomized controlled trial. J. Med. Internet Res. 17 (7) https://doi.org/10.2196/jmir.4086.
Martin, A.J., Dowson, M., 2009. Interpersonal relationships, motivation, engagement, and achievement: yields for theory, current issues, and educational practice. Rev. Educ. Res. 79 (1), 327–365. https://doi.org/10.3102/0034654308325583.
Mazéas, A., Duclos, M., Pereira, B., Chalabaev, A., Mazeas, A., Duclos, M., Pereira, B., Chalabaev, A., 2022. Evaluating the effectiveness of gamification on physical activity: systematic review and meta-analysis of randomized controlled trials. J. Med. Internet Res. 24 (1), e26779. https://doi.org/10.2196/26779.
McAuley, E., Duncan, T., Tammen, V.V., 1989. Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: a confirmatory factor analysis. Res. Q. Exerc. Sport 60 (1), 48–58. https://doi.org/10.1080/02701367.1989.10607413.
Mo, D., Xiang, M., Luo, M., Dong, Y., Fang, Y., Zhang, S., Zhang, Z., Liang, H., 2019. Using gamification and social incentives to increase physical activity and related social cognition among undergraduate students in Shanghai, China. Int. J. Environ. Res. Public Health 16 (5), 1–17. https://doi.org/10.3390/ijerph16050858.
Morschheuser, B., Hamari, J., Maedche, A., 2019. Cooperation or competition – when do people contribute more? a field experiment on gamification of crowdsourcing. Int. J. Hum. Comput. Stud. 127 (July 2019), 7–24. https://doi.org/10.1016/j.ijhcs.2018.10.001.
Morschheuser, B., Maedche, A., Walter, D., 2017. Designing cooperative gamification: conceptualization and prototypical implementation. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 2410–2421. https://doi.org/10.1145/2998181.2998272.
Nacke, L.E., Deterding, S., 2017. The maturing of gamification research. Comput. Human Behav. 71 (June), 450–454. https://doi.org/10.1016/j.chb.2016.11.062.
Neupane, A., Hansen, D., Fails, J.A., Sharma, A., 2021. The role of steps and game elements in gamified fitness tracker apps: a systematic review. In: Multimodal Technologies and Interaction, 5, p. 5. https://doi.org/10.3390/mti5020005.
Neupane, A., Hansen, D., Sharma, A., Fails, J.A., Neupane, B., Beutler, J., 2020. A review of gamified fitness tracker apps and future directions. In: Proceedings of the Annual Symposium on Computer-Human Interaction in Play, pp. 522–533. https://doi.org/10.1145/3410404.3414258.
Oinas-Kukkonen, H., 2013. A foundation for the study of behavior change support systems. Pers. Ubiquitous Comput. 17 (6), 1223–1235. https://doi.org/10.1007/s00779-012-0591-5.
Patel, M.S., Benjamin, E.J., Volpp, K.G., Fox, C.S., Small, D.S., Massaro, J.M., Lee, J.J., Hilbert, V., Valentino, M., Taylor, D.H., Manders, E.S., Mutalik, K., Zhu, J., Wang, W., Murabito, J.M, 2017. Effect of a game-based intervention designed to enhance social incentives to increase physical activity among families: the BE FIT randomized clinical trial. JAMA Intern. Med. 177 (11), 1586–1593. https://doi.org/10.1001/jamainternmed.2017.3458.
Patel, M.S., Small, D.S., Harrison, J.D., Fortunato, M.P., Oon, A.L., Rareshide, C.A.L., Reh, G., Szwartz, G., Guszcza, J., Steier, D., Kalra, P., Hilbert, V., 2019. Effectiveness of behaviorally designed gamification interventions with social incentives for increasing physical activity among overweight and obese adults across the united states: the step up randomized clinical trial. JAMA Intern Med 179 (12), 1624–1632. https://doi.org/10.1001/jamainternmed.2019.3505.
Prince, S.A., Adamo, K.B., Hamel, M., Hardt, J., Connor Gorber, S., Tremblay, M., 2008. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Intern. J. Behav. Nutrit. Phys. Act. 5 (1), 56. https://doi.org/10.1186/1479-5868-5-56.
Rapp, A., 2015. A qualitative investigation of gamification: motivational factors in online gamified services and applications. Int. J. Technol. Human Interact. 11 (1), 67–82. https://doi.org/10.4018/ijthi.2015010105.
Rapp, A., Hopfgartner, F., Hamari, J., Linehan, C., Cena, F., 2019. Strengthening gamification studies: current trends and future opportunities of gamification research. Internat. J. Human Comp. Stud. 127 (November 2018), 1–6. https://doi.org/10.1016/j.ijhcs.2018.11.007.
Reeve, J., Deci, E.L., 1996. Elements of the competitive situation that affect intrinsic motivation. Personal. Soc. Psychol. Bull. 22 (1), 24–33. https://doi.org/10.1177/0146167296221003.
Rigby, C.S., 2014. Gamification and motivation. The Gameful world: Approaches, issues, Applications. MIT Press, pp. 113–138.
Roberts, C., Torgerson, D.J., 1999. Understanding controlled trials. Baseline imbalance in randomised controlled trials. Br Med J 318 (7203), 185. https://doi.org/10.1136/bmj.319.7203.185.
Ryan, R.M., 1982. Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. J. Pers. Soc. Psychol. 43 (3), 450–461. https://doi.org/10.1037/0022-3514.43.3.450.
Ryan, R.M., Deci, E.L., 2000a. Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp. Educ. Psychol. 25 (1), 54–67. https://doi.org/10.1006/ceps.1999.1020.
Ryan, R.M., Deci, E.L., 2000b. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 55 (1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68.
Ryan, R.M., Rigby, C.S., Przybylski, A., 2006. The motivational pull of video games: a self-determination theory approach. Motiv. Emot. 30 (4), 344–360. https://doi.org/10.1007/s11031-006-9051-8.
Santhanam, R., Liu, D., Shen, W.C.M., 2016. Gamification of technology-mediated training: not all competitions are the same. Inf. Syst. Res. 27 (2), 453–465. https://doi.org/10.1287/isre.2016.0630.
Schmidt-Kraepelin, M., Thiebes, S., Tran, M.C., Sunyaev, A., 2018. What's in the game? developing a taxonomy of gamification concepts for health apps. In: *Proceedings of the 51st Hawaii International Conference on System Sciences, October 2017*. https://doi.org/10.24251/hicss.2018.150.
Schmidt-Kraepelin, M., Warsinsky, S., Thiebes, S., Sunyaev, A., 2020. The role of gamification in health behavior change: a review of theory-driven studies. In: *Proceedings of the Annual Hawaii International Conference on System Sciences, 2020-Janua*, pp. 1256–1265. https://doi.org/10.24251/hicss.2020.155.
Seaborn, K., Fels, D.I., 2015. Gamification in theory and action: a survey. Int. J. Hum. Comput. Stud. 74, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006.
Sigmundová, D., Vašícková, J., Stelzer, J., Řepka, E., 2013. The influence of monitoring interval on data measurement: an analysis of step counts of university students. Int. J. Environ. Res. Public Health 10 (2), 515–527. https://doi.org/10.3390/ijerph10020515.
Stanne, M.B., Johnson, D.W., Johnson, R.T., 1999. Does competition enhance or inhibit motor performance: a meta-analysis. Psychol. Bull. 125 (1), 133–154. https://doi.org/10.1037/0033-2909.125.1.133.
Tam, M.K., Cheung, S.Y., 2019. Validation of consumer wearable activity tracker as step measurement in free-living conditions. Finnish J. EHealth EWelfare 11 (1–2), 68–75. https://doi.org/10.23996/fjhw.76673.
Tauer, J.M., Harackiewicz, J.M., 2004. The effects of cooperation and competition on intrinsic motivation and performance. J. Pers. Soc. Psychol. 86 (6), 849–861. https://doi.org/10.1037/0022-3514.86.6.849.
Thorsteinsen, K., Vittersø, J., Svendsen, G.B., 2014. Increasing physical activity efficiently: an experimental pilot study of a website and mobile phone intervention. Int J. Telemed. Appl. 2014 https://doi.org/10.1155/2014/746232.
Tu, R., Hsieh, P., Feng, W., 2019. Walking for fun or for "likes"? the impacts of different gamification orientations of fitness apps on consumers' physical activities. Sport Manage. Rev. 22 (5), 682–693. https://doi.org/10.1016/j.smr.2018.10.005.
Tudor-Locke, C., Bassett, D.R., 2004. How Many Steps/Day Are Enough? Sports Med. 34 (1), 1–8. https://doi.org/10.2165/00007256-200434010-00001.
Tudor-Locke, C., Burkett, L., Reis, J., Ainsworth, B., Macera, C., WilsonILSON, D., 2005. How many days of pedometer monitoring predict weekly physical activity in adults? Prev. Med. 40 (3), 293–298. https://doi.org/10.1016/j.ypmed.2004.06.003.
Twisk, J., Bosman, L., Hoekstra, T., Rijnhart, J., Welten, M., Heymans, M., 2018. Different ways to estimate treatment effects in randomised controlled trials. Contemp. Clinical Trials Commun. 10 (November 2017), 80–85. https://doi.org/10.1016/j.conctc.2018.03.008.
Vansteenkiste, M., Niemiec, C.P., Soenens, B., 2010. The Development of the Five Mini-Theories of Self-Determination theory: an Historical overview, Emerging trends, and Future Directions, pp. 105–165. https://doi.org/10.1108/S0749-7423(2010)000016A007.
Vansteenkiste, M., Simons, J., Lens, W., Sheldon, K.M., Deci, E.L., 2004. Motivating learning, performance, and persistence: the synergistic effects of intrinsic goal contents and autonomy-supportive contexts. J. Pers. Soc. Psychol. 87, 246–260. https://doi.org/10.1037/0022-3514.87.2.246. American Psychological Association.
Walz, S.P., Deterding, S., 2015. The Gameful world: Approaches, issues, Applications. MIT Press.
WHO, 2022. WHO EUROPEAN REGIONAL OBESITY REPORT 2022.

Wu, J., Lu, X., 2013. Effects of extrinsic and intrinsic motivators on using utilitarian, hedonic, and dual-purposed information systems: a meta-analysis. J. Assoc. for Info. Sys. 14 (3), 153–191. https://doi.org/10.17705/1jais.00325.

Wu, Y., Kankanhalli, A., Huang, K., 2015. Gamification in Fitness Apps: how do leaderboards influence exercise?. In: Thirty Sixth International Conference on Information Systems.

Xie, J., Wen, D., Liang, L., Jia, Y., Gao, L., Lei, J., 2018. Evaluating the validity of current mainstream wearable devices in fitness tracking under various physical activities: comparative study. JMIR Mhealth Uhealth 6 (4). https://doi.org/10.2196/mhealth.9754.

Yang, X., Ma, L., Zhao, X., Kankanhalli, A., 2020. Factors influencing user ' s adherence to physical activity applications : a scoping literature review and future directions. Int. J. Med. Inform. 134, 104039 https://doi.org/10.1016/j.ijmedinf.2019.104039.

Zuckerman, O., Gal-Oz, A, 2014. Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. Pers. Ubiquitous Comput. 18 (7), 1705–1719. https://doi.org/10.1007/s00779-014-0783-2.