DARP Assessment

Are social comparison, competition or cooperation style desktop based screen notifications perceived to be most effective on workers' sedentary behaviour?

Group 3

Group Members:

Group Member 1
Chris Cowl
Group Member 3
Group Member 4

KEYWORDS
Personalised persuasive strategies; perceived persuasiveness; behaviour change; sedentary behaviour; personalised desktop notifications.

ABSTRACT
Sedentary computer-based tasks in the workplace, a reduction in physical jobs and increased attachment to devices mean that people are sitting down more than ever which is associated with numerous illnesses, such as cardiovascular disease. Notifications increase self-awareness which is key to tackling sedentary behaviour. Participants (N = 21) evaluated social comparison, competition and cooperation style notifications using the Perceived Persuasiveness Scale (PPS). Whilst we couldn't find any significant differences in PPS score, interview data uncovered insights about participants' preferences of notification and why they're motivating. Insights on notification preference; competition is preferred, cooperation is similarly preferred and social comparison is demotivating. Insights for reasons for notification choice are that design of the message is important – it's recommended that notifications are trustworthy, believable, calm and private. People respond to competition with others and personalised messages with relevant quantitative information. People were not interested in social comparison and comparing themselves with others.

1 INTRODUCTION AND LITERATURE REVIEW

Sedentary behaviour is a relatively recent issue facing the world's population and has been defined in previous literature as; '*Sedentary behavior is any waking behavior characterized by an energy expenditure ≤1.5 metabolic equivalents (METs), while in a sitting, reclining or lying posture*' (Tremblay et al., 2017; Sedentary Behaviour Research Network, 2012).

Whilst computers have had benefits in the workplace such as increased connectivity, with computers increasingly being intertwined in our lives and daily routines (Cascio & Montealegre, 2016), which is known as ubiquitous computing, it's not surprising that most modern work days involve workers spending large passages of time sitting down at a computer to perform their daily tasks, with previous studies suggesting over 11 hours of total sedentary time, including leisure time, on work days (Parry & Straker, 2013). With trends showing an increase towards low physically demanding jobs, highlighted by research in the US (Brownson et al., 2005), it seems that this is a problem that will remain for some time to come. Indeed, many countries are on a spiral towards less and less physical activity and increased sedentariness as a whole (Ng & Popkin, 2012). As the modern work day is increasingly computer focused and sedentary-based, so too is peoples freetime (Yang et al., 2019). Many people enjoy streaming their favourite shows and films, playing video games with increasing attachment to computer screens and devices (Prince et al., 2020) and generally enjoying many hobbies that include large blocks of sedentary behaviour to relax from the stresses of everyday life.

This, of course, is a worry for countries across the world, and the situation is no different in the UK and US. Unsurprisingly, this trend of increased sedentary behaviour has a strong association risk with many diseases, such as cardiovascular disease, type-II diabetes, stomach and lung cancers (U.S. Department of Health and Human Services, 2018) and decreased number of neurons which influence kidney disease (Kosaki et al., 2021). In the UK, the National Health Service (NHS) suggests cost-of-illness related to prolonged sedentary behaviour was £0.8 billion in 2016-2017. Additionally, 11.6% of total deaths in 2016 in the UK, (69,276 deaths), were correlated with sedentary behaviour

(Heron et al., 2019). A Brazilian study showed that during the Covid-19 Pandemic, which has seemingly accelerated the problem, 70% of people are not meeting government guidelines for minimum activity levels (da Silva et al., 2020).

With decreasing activity levels in the workplace, increased attachment to our devices and with so many negative health implications of sedentary behaviour, we are heading into a perfect storm of conditions that affects the long-term health of not only this generation, but of generations to come. Unfortunately, the situation only gets worse, although having active hobbies such as going to the gym, playing sports, walking or running is undoubtedly good for you, the long-term effects of prolonged sedentary behaviour are independent of how active you are at other times (Bankoski et al., 2011), so it is important that we find ways to break up long periods of sitting or lying down, even if we participate in active hobbies and interests.

Therefore, with computer-based sedentary behaviour seemingly here to stay, both at the workplace and during leisure time, and with the long-term health risks associated with sedentary behaviour understood, it's extremely important that we need to better understand how and what we can do to change sedentary behaviour on-mass.

In recent studies, one way to combat sedentary behaviour was to increase self-awareness of prolonged sitting by office workers and enable them to realise when they have been sat down for long periods of time. For example, it was shown that desktop-based screen notification interventions increase self-awareness of prolonged sedentary behaviour (Luo et al., 2018; Morris et al., 2008). Physical ambient technologies that aren't computer based, such as a physical device that sits on your desk (Brombacher et al., 2020) have also shown that self-awareness is key to interrupt the flow of workers and encourage them to take active breaks from their desks. Whilst notifications are a useful tool to remind people of tasks, research suggests that mobile phone and smartwatch notifications can become lost in a tidal wave of notifications from other apps the user may have installed, and are considered more intrusive compared to computer desktop screen based notifications (Norrie & Murray-Smith, 2015). People seem to prefer being notified of their sedentary behaviour by physical ambient devices, however it was noted that they lack privacy in the workplace (Brombacher et al., 2020), particularly if the office space is open.

It would be very easy to approach the problem with a one size fits all solution. However, great success has been achieved by using a more targeted approach to encourage people to get up and move. Designing notifications that use different mechanisms to entice behaviour change such as social comparison, competition or cooperation messages have been shown to be a very effective way to elicit behaviour change in relation to decreasing participants' sedentary behaviour in many different personality types (Orji et al., 2018).

With this in mind, we are interested in how effective participants perceive these social comparison, competition or cooperation style notifications to be in relation to encouraging them to get up and take active breaks.

Different notification styles may have different success rates in relation to different types of people, but what is also key is timing. The likelihood of any behavioural change is increased by understanding how receptive a person may be to change at that point in time. Therefore, the Transtheoretical Model (TTM) was devised as an integrative model which is used to assess participants likelihood of behavioural change by determining which Stage of Change (SoC) they are currently in (Prochaska & DiClemente, 1983). The stages are; *Precontemplation, Contemplation, Preparation, Action and Maintenance*. We can determine what SoC participants are in with answers from a short questionnaire specifically tailored to sitting avoidance (Han et al., 2015). Although not

the main focus of our study as we haven't controlled for SoC in our sample, it will be interesting to understand which stage of SoC participants are in to enlighten us for future research.

Therefore, the main focus of our study will leverage the benefits of a more targeted approach in terms of notification style and utilise the privacy benefits and ubiquity of desktop computer screens. We will then measure the perceived persuasiveness of these notification styles in order to determine which one has the greatest perceived effect on participants' sedentary behaviour.

Furthermore, we're interested in the deeper insights of participants' preferences and reasons why these preferences are chosen so we can motivate further research in the future to better understand how we can tackle the increase of sedentary behaviour.

Secondary to our main study goals above, we will be aiming to understand what stage in the SoC participants may be in. For example, it is much more likely that our notifications will be perceived as more effective if our participants are in the Preparation stage of SoC rather than the Precontemplation stage. This is however just for exploratory purposes.

2 METHODOLOGY

2.1 Objective

In this paper, we aimed to investigate and compare the perceived persuasiveness of three desktop-based screen notification styles; comparison, competition and cooperation on reducing the sedentary behaviour of workers. We developed three prototypes of desktop notifications for each type of notification respectively and evaluated them by using interviews and surveys to understand the participants' attitudes and opinions towards the interventions.

2.2 Hypothesis

H_O: There are no differences between the perceived persuasiveness of the three notification styles: social comparison, cooperation and competition.

H_1: There are differences between the perceived persuasiveness of the three notification styles: social comparison, cooperation and competition.

2.3 Notification Content Development

We developed the content of the notifications used for this study based on a collection of ten storyboards created by previous researchers (Orji et al., 2018). The storyboards depicted ten different persuasive messages for behaviour change respectively, in which we adopted the social comparison, cooperation and competition notifications in the storyboards as our references. We evaluated and refined the messages iteratively following the pilot studies and discussions with our colleagues. At the start of the interview, to make the notifications more realistic for the participants, we asked them to give us the first names only of people who they study or work with in order for us to personalise the notification to increase ecological validity.

Fig. 1. shows the notification styles and in brackets, where the personalised friends names were.

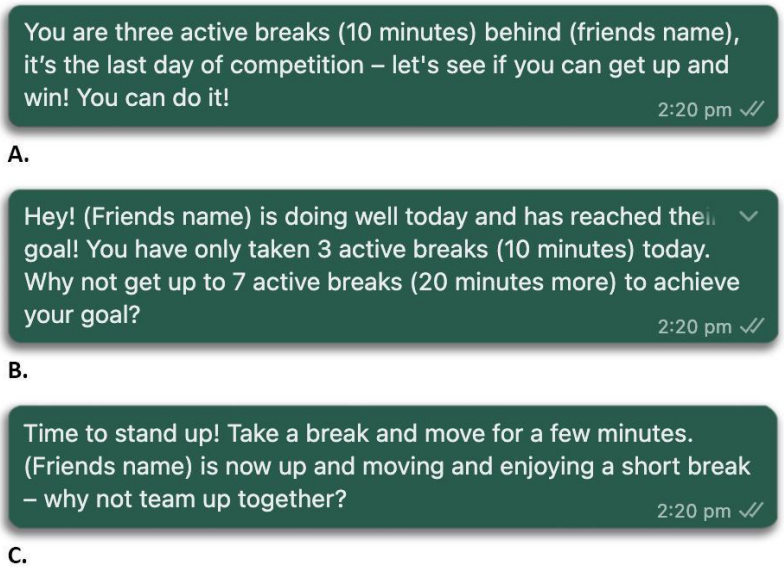

Fig. 1. Screenshots of the (A) competition, (B) social comparison and (C) cooperation notifications.

2.4 Participants

We recruited participants through our social networks by using the snowball sampling method. Participants were either invited in class or via WhatsApp. Participants were informed about the goals of the study, the duration of their participation and the fact they couldn't be compensated with a short spiel including these details. This sampling method was at the convenience of the researchers and efficient within the limitation of time and cost in this study. The participants were mostly either Undergraduate or Master students studying at the University of York and the University of Leeds. Two participants were in full-time employment.

This was a within-participants study, in which participants were shown three prototypes one by one. To eliminate ordering bias, we divided the participants into three groups to randomise the order of prototypes shown to the participants. Before the main studies, we conducted a pilot study to test the feasibility and validity of our study.

A total of 21 responses were gathered in this study. The age of the participants ranged from 22 to 49 years old and mean at 26.71 years old. There were 10 males and 11 females. 3 participants are currently studying for their undergraduate degrees, while 19 participants are studying postgraduate degrees and 2 for other educational levels.

2.5 Questionnaire measures

Our survey, administered through Qualtrics, had two main sections. The first section contained demographic-related questions, including participants' age, gender and educational levels. In the second section, we presented a survey to participants to collect objective data to validate our interview results. We used the Perceived Persuasive Scale (PPS) developed by Thomas et al., (2019). The Perceived Persuasiveness Scale was validated in the area of using messages to intervene healthy eating behaviour and email security habits. The construct of the survey was developed based on the model of persuasive technologies. The survey consisted of three factors: effectiveness, quality and capability. Each factor consists of 3 items. In the factor *effectiveness*, for example, the participants were asked "*The messages will cause change in my behavior*". In the factor *quality*, for example, the participants were asked "*I believe this message is true*". In the factor *capability*, for example, the

participants were asked "*This message has the potential to inspire users*". The scales were measured by using participant agreement with a 7-point Likert scale ranging from "1 = Strongly disagree" to "7 = Strongly agree". A full set of the questions in the PPS is available in Appendix A.

Next, as a secondary aim of the study, we wanted to measure how receptive our participants were to behaviour change in relation to their sedentary behaviour. To do this the Transtheoretical Model (TTM) was devised as an integrative model which is used to assess participants likelihood of behavioural change, originally developed to assess behavioural change in smokers (Prochaska & DiClemente, 1983). We used a TTM short questionnaire '*Stages of Motivational Readiness to Avoid Sitting Time*' (Han et al., 2015), that was specifically developed and validated to measure where participants are in the five stages of readiness to avoid sitting time. The stages are Precontemplation, Contemplation, Preparation, Action, and Maintenance. The survey consists of two questions in order to ascertain which stage the participants are in. Validation of the questionnaire has indicated strong concurrent ($\chi 2 = 25.0$, $p < 0.001$) and construct ($p < 0.01$) validities and strong internal consistency ($k = 0.62$) and test-retest reliabilities (Cronbach's alpha ranging from 0.72 to 0.88). The questions in the TTM short questionnaire are available in Appendix B.

2.6 Interview questions

After answering the survey, participants were also asked a series of open-ended questions to provide further insights on their preference of the three notification styles. Participants were also asked about their daily activity levels, motivations for increasing their physical activity, attitudes towards desktop-based notifications and opinions on the three notifications prototypes we designed. All interview questions could be found in Appendix C.

2.7 Data collection procedures

Before the data collection, we gained ethical approval from the University of York. Participants were then recruited via our personal contacts. We gathered participants' names and phone numbers for contacting purposes. After consolidating a list of participants, we divided the participants into three groups, and we randomly shuffled the orders of notifications shown to participants by R studio code in each group. The orders of notifications were:

Group A: Social Comparison, Cooperation, Competition
Group B: Cooperation, Competition, Social Comparison
Group C: Competition, Social Comparison, Cooperation

Afterwards, we arranged study times with the participants. Shortly after, we sent participants information sheets and instructions for preparing for the study electronically. The participants' information sheet is included in the submission folder. The information sheet ensured there would be no harm to participants and it was a no compensation research study. The instruction sheet described the process of the study. More importantly, the instruction manual detailed the steps to download WhatsApp on desktop and made sure the participants had downloaded the WhatsApp desktop application, which would be a tool used for showing participants the three notifications, before the study. WhatsApp was chosen because it is a common messaging application which the majority of participants would likely already have. Also, WhatsApp desktop notifications closely resemble the appearance and nature of desktop notifications sent by other applications.

Due to the limitation during the Covid restriction, the study was done online on Whatsapp and Zoom was used for troubleshooting purposes. Before the start of the study, we made sure the participants had WhatsApp installed on their desktop computers. The entire study was conducted via text on WhatsApp. At the beginning of the study, we sent the informed consent form, which is included in the submission folder, to participants via Qualtrics. we greeted the participants and sent a message with their participant number to conduct a test run on their WhatsApp desktop notifications functionality. If needed, we asked participants to share screens with us by using Zoom, so we could help them to solve any technical difficulties. This was to ensure participants' WhatsApp and desktop notifications were both enabled. After that, participants were allowed to stop sharing their screens. Second, we asked participants to provide the first name of two of their friends to help us personalise the notifications. Third, we sent a link for participants to sign the informed consent sheet and fill in the demographic, PPS and TTM questionnaires. The survey was to be used throughout the first part of the interview in order for participants to record their perceived persuasiveness scores via the PPS survey after each exposure to each notification. Fourth, was the core part of the study. We conducted a Wizard-of-Oz study where we sent participants a WhatsApp notification, which was either social comparison, competition or cooperation, according to the group that the participants were in. As previously mentioned, after exposing each participant to a notification, we asked participants to return to the previously sent survey and record their PPS scores. After they saw all three styles of notification and rated each notification's persuasiveness, we then conducted a semi-structured interview via WhatsApp, which asked which notification they preferred and why that might be. We also asked questions to explore their opinions towards their background activity levels and notifications in general. The interviews lasted approximately 45 minutes.

The interviews were done through instant messaging apps as suggested by Braun & Clarke (2013, 239-243). Firstly, This was because participants could participate in their own chosen locations. Second, participants would have a greater sense of control because they had time to think and reflect on their responses. Third, it facilitated the participation of shy people and people who lack confidence in communicating their answers in face-to-face communication. Fourth, it levels the traditional power dynamic between interviewer and interviewee, with the purpose of gaining deeper insight and richer data because interviewees feel more at ease. For the researchers, there was no need for an observer in the interview and no need for transcription post interview, which saved time and resources. Because of the extended time frame of instant messaging interviews, the researchers could formulate, prompt and ask follow-up questions in their own time and whenever necessary during the process. Once qualitative interview data was collected, it was immediately anonymised and collated in a spreadsheet to ensure no ethical issues could arise later on in the reporting of findings phase of the study.

The data from the interviews that specifically relates to the research question was analysed using content analysis. This enabled us to understand clearly the preference of notification style each participant had and more importantly, what motivated their preference. It also had the potential to inform us of why participants disliked other notification styles. The content analysis was performed inductively where we used codes that emerged from the data, by one researcher initially, with a second researcher checking the draft code dictionary for potentially missed codes or ambiguity. Codes emerged on a phrase by phrase basis, however, if participants reiterated the same point later in the paragraph, it was only counted as one count as answers were generally only a short paragraph or two. We felt that counting the code once was consistent and fair. However, two or more different

codes could be applied to a participant's answer. Once the code dictionary was agreed, both researchers then recounted the codes that appeared in the data. Lastly, code counts were compared and assessed for inter-coder reliability.

To address the research question explicitly, the combination of exposure to the three notification styles, the elicitation of notification persuasiveness from participants using the Perceived Persuasiveness Scale (PPS) survey via an online survey, coupled with semi-structured interviews specifically relating to participants notification preference and reasons why, we have been logical and rigorous in our approach.

3 RESULTS

3.1 Survey Data Analysis

3.1.1 *Stages of Motivational Readiness to Avoid Sitting Time.*

From the TTM survey of the Stages of Motivational Readiness to Avoid Sitting Time, participants answered two questions in order for us to understand which Stage of Change (SoC) they were currently in. As shown in Fig. 3., most participants were in the Preparation stage (N = 8, 38.1%), showing that many people were seriously considering increasing their activity levels and had taken steps, albeit unsuccessfully. The second most popular SoC was Contemplation (N = 5, 23.8%), in which participants had not yet committed to taking action but plan to in the next six months. The third most popular SoC was Maintenance (N = 4, 19.0%), where participants have been engaged in a more active lifestyle for more than six months. The last two SoC are Precontemplation (N = 2, 9.5%), where participants just aren't interested in an active lifestyle and have no intention of changing that, and Action (N = 2, 9.5%), in which participants had taken action to increase physical activity for less than six months, were both equally represented in terms of participant frequency. Percentages don't add up to 100% due to rounding.

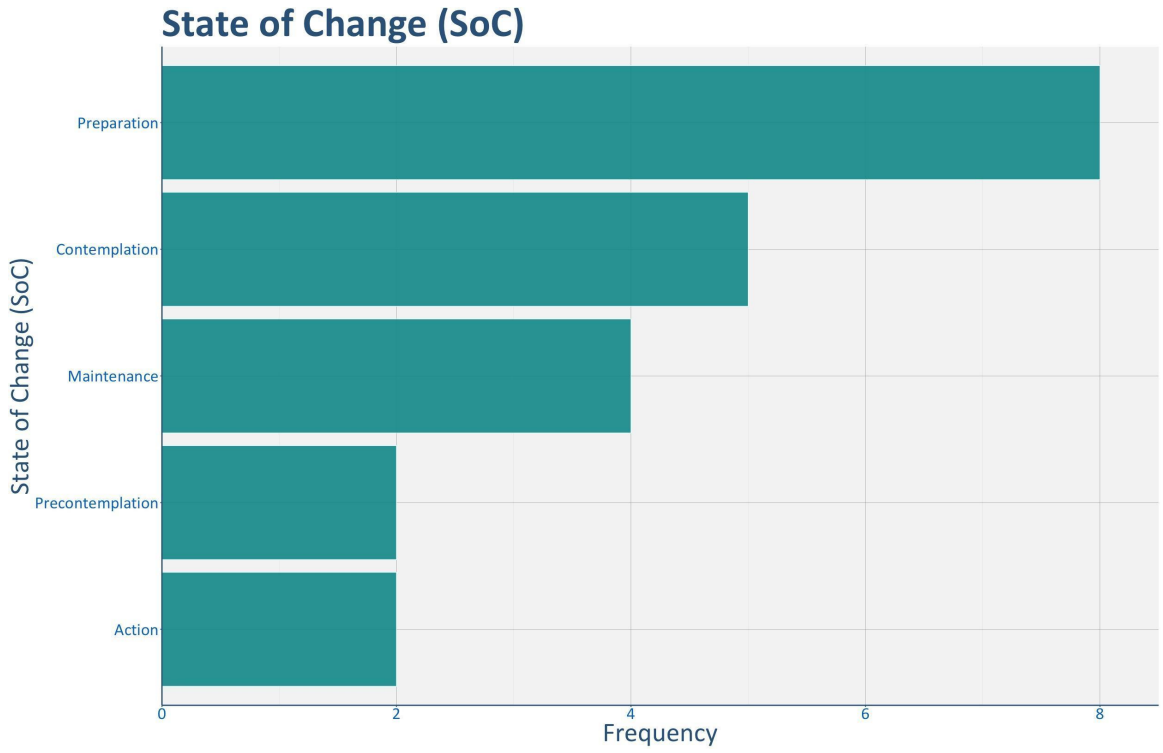

Fig. 3. Number of participants in each State of Change (SoC)

3.1.2 *Perceived Persuasiveness Scale (PPS).*

Every participant responded to all three PPS of notification designs (N = 21). The mean PPS score of Social comparison design is the highest (M = 47.33, SD = 7.03) followed by Cooperation design (M = 46.86, SD = 10.20), and Competition design. (M = 44.86, SD = 7.48) The PPS score of Cooperation design has the widest range at a minimum of 24 and a maximum of 61 from the total score at 63 (Table 1).

Table 1. Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Social Comparison design	21	47.333	7.0309	33.0	57.0
Cooperation design	21	46.8571	10.19944	24.00	61.00
Competition design	21	44.8571	7.48522	32.00	58.00

For testing our hypothesis, we used a nonparametric method on SPSS to analyse due to the limited sample size. We computed the total Perceived Persuasiveness Scale (PPS) score for each type of notification and then conducted a Friedman's Test to test whether participants perceive the difference in persuasiveness score between each notification type. However, the result found that Friedman's test was not statistically able to detect a significant difference in the mean ranks between the three different notification types ($\chi 2(2) = 1.76, p = 0.414$), so we didn't need to analyse post-hoc pairwise comparisons (Table 2).

Table 2. Friedman's Test

Ranks

	Mean Rank
Social Comparison design	2.17
Cooperation design	2.05
Competition design	1.79

Test Statistics

N	21
Chi-Square	1.763
df	2
Asymp. Sig.	0.414

3.2 Interview Data Analysis

3.2.1 *Content Analysis and Inter-coder Reliability.*

The data from the interviews that specifically relate to the research question were analysed using content analysis. The content analysis was performed inductively, we used codes that emerged from the data, by one researcher initially, with a second researcher checking. Codes emerged on a phrase by phrase basis, however, if participants reiterated the same point later in the paragraph, it was only counted as one count as answers were generally only a short paragraph or two. Two or more different codes could be applied to a participant's answer. Once the code dictionary was agreed, both researchers then recounted the codes that appeared in the data. Lastly, code counts were compared and assessed for inter-coder reliability (Table. 3).

Table 3. Inter-coder Reliability

	Agreement (%)	Cohen's Kappa (K)
Notification Preference	100	1
Reason for Preference	94	0.939

3.2.2 *Participants Notification Preference.*

Looking at Fig. 4, two notification styles were clearly preferred by participants. Overall, Competition was preferred (N = 9, 42.9%), with Cooperation just behind (N = 8, 38.1%). Next was Social Comparison (N = 3, 14.3%) and lastly, one participant could not decide which notification they preferred (N = 1, 4.7%).

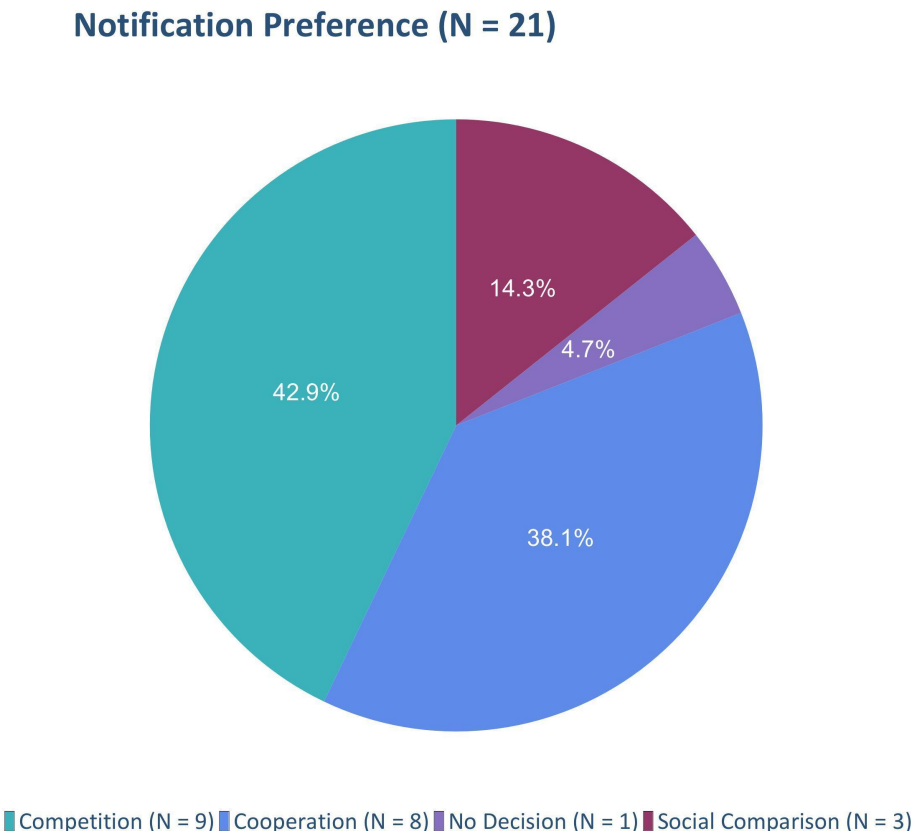

Fig. 4. Participants notification preference

3.2.3 *Reason for Participants Notification Preference.*

The insights behind participants' notification preference in Fig. 5 were coded into 13 categories; Design of message (N = 10, 20.8%); Health benefits (N = 6, 12.5%); Competition with others (N = 6, 12.5%); Specific information (N = 4, 8.3%); Dislike comparison (N = 4, 8.3%); Dislike social (N = 4, 8.3%); Comparison with others (N = 3, 6.3%); Social benefits (N = 3, 6.3%); Dislike competition (N = 3, 6.3%); Goal setting (N = 2, 4.2%); Anxiety (N = 1, 2.1%); Precontemplation (N = 1, 2.1%) and Continue with change (N = 1, 2.1%). Percentages don't add up to 100% due to rounding.

These insights enabled us to understand clearly the preference of notification style each participant had and more importantly, what motivated their preference. It also had the potential to inform us of why participants disliked other notification styles.

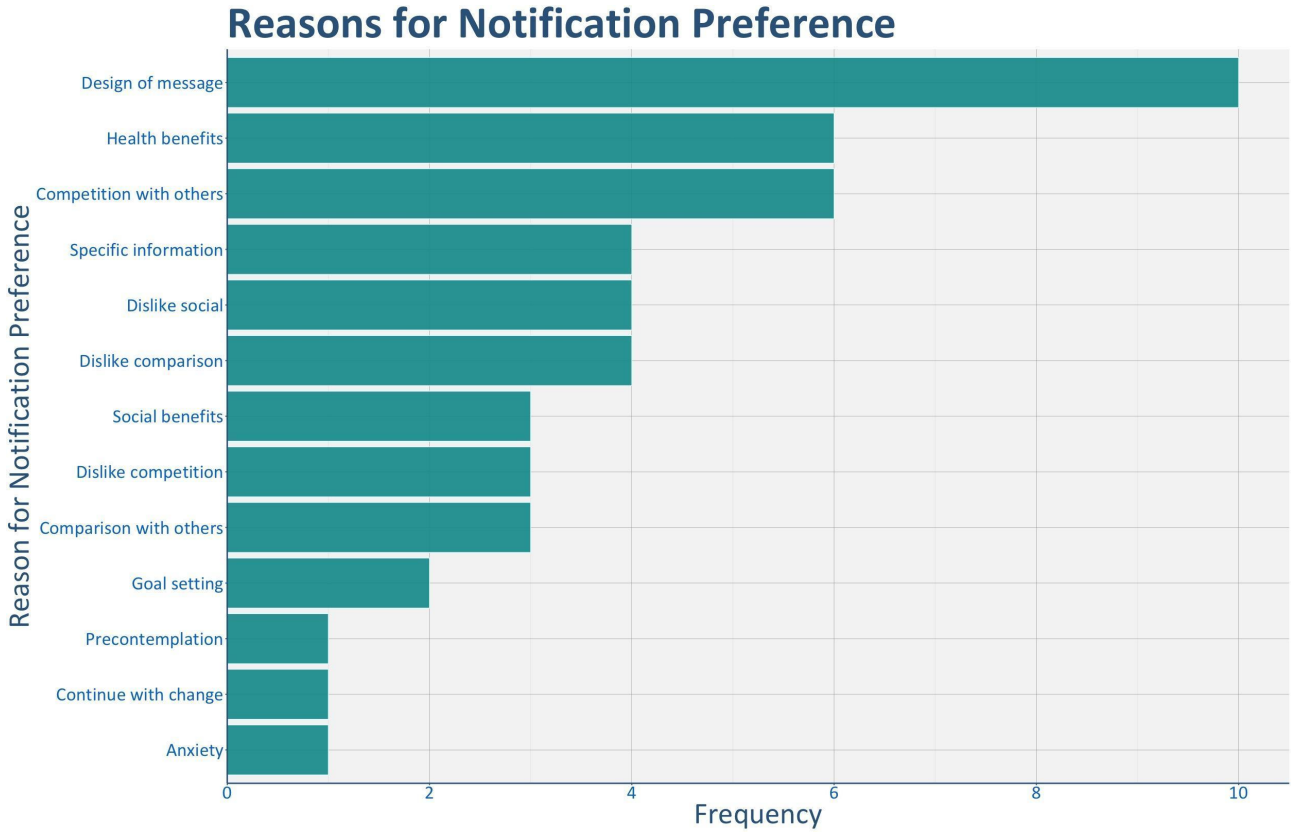

Fig. 5. Participants reasons for notification preference

Design of Message (N = 10, 20.8%)

The design of the message had the most impact on why participants chose the notification style they did. This category included the requirements that the message design creates trust, uses calm tone, is believable, private or just simply reminds. This is interesting because this category doesn't explicitly relate to competition, cooperation or social comparison, it is more about how you deliver the message style.

> *"The first one I think its tone makes me feel less stressed and make me want to get up and do some exercise."* (P9)

“The third one, because the time gap it mentions for active breaks seems trustworthy[…] I might be motivated by this.” (P18)

Competition with Others (N = 6, 12.5%)
As the competition notification style was the most preferred, we would expect to see this as one of the most popular reasons for notification choice. This category includes any comments where competition was referenced as a positive factor and reason to get up and move. Competition seems to be a major factor in participants' motivations to change behaviour.

> *"But when I see the second message, I might get up and exercise because I like to compete against others and it makes me feel great."* (P5)

> *"I feel because it is a competition and it gives me a goal I would need to get up and move in order to win. I think it would push me to do something."* (P16)

Health Benefits (N = 6, 12.5%)
Health benefits affected participants' notification choices. This category included any comments made relating to the reasons why they are persuaded to take notice of any of the notification styles. Again, this category doesn't explicitly relate to competition, cooperation or social comparison, but it does inform us that people are concerned about their health and are possibly more committed to getting up and taking active breaks. Also, some people had existing health issues and seem to be more likely to be further on in the State of Change (SoC) through necessity.

> *"I wouldn't want to risk developing chronic aches and pains just because I was gaming for extended periods with minimal exercise."* (P11)

> *"I have back pain. So if I end up sitting for too long, I'm unable to stand or walk for a while."* (P12)

Specific Information (N = 4, 8.3%)
Participants seemed to be motivated by quantitative figures included in the message, such as the specific amount of time in minutes or the amount of active breaks they are behind their friends. Again this doesn't directly relate to the notification styles of competition, cooperation or social comparison, but it indicates that participants find specific information relating to them very motivational.

> *"The competition notice will motivate me more, because I can clearly see how much time I have left to work on and make a more specific plan of action for myself."* (P2)

> *"I think the second one is the best because it informed me with accurate data, which would be more convincing."* (P1)

Dislike Social (N = 4, 8.3%)
The first four reasons for participants' notification preference were all positive reasons relating to their choices, this is the first category where participants expressed what they don't like and why. This category included comments directly relating to the dislike of social elements of the notification such as encouraging them to meet with others.

"I think the idea of taking breaks with others or competing with them will not work for me." (P20)

"I will ignore it because I don't want to join my friends." (P3)

Dislike Comparison (N = 4, 8.3%)

Again we see that participants know what they don't find motivational, this is the second instance of a category where participants expressed what they don't like and why. This category included comments directly relating to the dislike of comparing elements of the notification, such as comparing their activity levels with the activity levels of their friends.

"For the first one and the second one I will turn off it directly, cause i'm not care about other people." (P7)

"For the first desktop notifications: I will ignore it because I don't like to compare myself." (P3)

4 DISCUSSION

4.1 Limitations That Could Have Contributed to Lack of Significance in Survey's Results

This study does not find any significance between the perceived persuasiveness of the three types of notification styles. There are several possibilities contributing to this result. One is that the participants were in different Stages of Change (SoC). In our survey, there were possibly too many participants that had no reason to be persuaded by any of the notifications, but for very different reasons. Participants in Precontemplation (N = 2) were unlikely to want to change in terms of reducing their sedentary behaviour simply because they didn't want to change no matter what notification they were exposed to. Participants in the Action stage (N= 2) had already made changes to their lives and successfully continued with the changes for less than six months and had no reason to need to be persuaded to change anything. Equally, participants in the Maintenance stage (N = 4) had succeeded in making positive health choices in relation to getting up and moving for even longer at over six months, again meaning they had no reason to be persuaded to change their behaviour as they had already done so a long time ago. So collectively, participants across Precontemplation, Action and Maintenance stages account for 38.1% of our sample size (N=8). Considering the sample size was small to begin with, it's a lot to lose to people who are likely to not find any of the notifications persuasive and therefore detrimental to our ability to be able to find significant differences between the persuasiveness of the three notification styles. So essentially this was a confound.

It's also possible that a Type II Error could have occurred, meaning a difference in persuasion levels was there, however, we didn't have enough participants to detect the significant difference in persuasion levels between the three notification groups. The notifications could have persuaded participants in some small way but we will only find out if we increased the sample size.

Another reason could be due to the Wizard of Oz study design, in order to make the study work, we could only realistically gather a small sample size to enable us to execute data collection in the time allocated. So whilst trying to conduct a study that was more ecologically valid, exposing participants to digital messages rather than just showing participants images of the notifications, we had to sacrifice sample size. This does not only reduce the external validity, but again also limits the

probability of finding any significant results. Furthermore, the Wizard of Oz nature of the study can have had an effect on the results. Participants may not be persuaded by the fictitious notifications, the ubiquitousness of WhatsApp which made it such a good choice for convenience for participants and therefore the study, may also have been its downfall. Participants might have been so comfortable with WhatApp that they may not have been convinced of the seriousness of the study. Participants may not have been able to separate the notification exposure from the casual chat, which we purposely induced to make participants feel at ease in order to elicit richer qualitative data. Therefore, the study may have still seemed like a casual chat.

Perhaps the inconsistencies of exposure time for the participants could have confounded the experiment. We weren't consistent with the time we interviewed participants and what the participants had done prior to the notification exposure, therefore some participants may have been more open to being persuaded having just engaged in large blocks of sedentary behaviour, and some could have just come in from being active, and therefore been less open to being persuaded.

As participants were only exposed to the notifications in a short amount of time, it's likely that receiving one notification in each style was nowhere near enough to persuade participants anyway, even if participants did find them subconsciously persuasive. Long-term exposure is likely needed to elicit behaviour change (Gabrielli et al., 2014). Participants were also put in a situation which required them to imagine their reactions to the notifications as if they had been sitting and working for a long period of time. As this study did not measure the participants' actual reactions, this could have an effect on the answers given in the interview and PPS survey.

The measurement scales we used perhaps could have been more appropriate and valid. The Perceived Persuasiveness Scale (PPS) uses three dimensions with nine scale items in total, each with a 7-point Likert scale. In order to measure the Transtheoretical Model (TTM) of Stages of Motivational Readiness to Avoid Sitting Time, the scale only has one dimension with two scale items. For both scales we can say that the measures are too brief, inadequate and ethically questionable (DeVellis & Thorpe, 2022). Additionally, this version of Perceived Persuasiveness Scale (PPS) by Thomas et al., (2019) was newly developed from analysing the previous studies about perceived persuasiveness measurements from 2013 to 2018 and then narrowed down to 34 scale items related to persuasiveness. After that they developed the new scale analysing Factor Analysis by recruiting 92 participants to rate 34 scale items. Nevertheless, this PPS scale had a flaw in validity of Factor Analysis, especially in sample size. According to Subjects-to-Items ratio (STI), the number of participants and items should be around 5:1 (Hair at al., 2014).

What is also a plausible explanation for the lack of significance from our quantitative analysis is that the persuasiveness of the notification styles have similar persuasion levels. It's possible that the mechanisms of competition, cooperation and social comparison are so similar and appeal to a broad range of personality types that we may not find significance, even in a future study.

4.2 Useful Insights from Our Interview Data

4.2.1 *Notification Preferences.*

From our interview data, we can see that competition was the most preferred notification style (N = 9, 42.9%), closely followed by Cooperation (N = 8, 38.1%). This showed us that there were two clear preferred notification choices. However, if participants were given more choice of notification styles it's possible that a different outcome may occur. All we can determine is between competition, cooperation and social comparison, competition and cooperation were most popular in our study.

What is also interesting is that Social Comparison scored so low (N = 3, 14.3%) and Dislike Comparison scored relatively high (N = 4, 8.3%) for notification preference reasons. This could be because people feel shamed into getting up and perhaps feel pressured to change their behaviour rather than motivated in a positive way. This concurs with previous research (Orji et al., 2019) that argues that comparisons with others can make people feel anxious, feel like they are being body shamed and feel like their privacy is being eroded due to the sharing of their daily activities with friends and vice versa. *"[…] but if you keep getting notifications about the performance of other friends, it seems to be a little stressful."* (P9).

4.2.2 *Notification Preference Reasons.*

The most popular reason for notification preference related to the actual design of the message (N = 10, 20.8%). We expected reasons simply relating to the style of the notification, for example, enjoyment or dislike of competition so it's interesting to understand how the particular notification makes people feel as well as the mechanism behind the notification. Trust, tone of message, believability and privacy are all dimensions of the Design of Message category and useful insights for further research. For example, *"The first one I think its tone makes me feel less stressed and make me want to get up and do some exercise."* (P9). Therefore, building notification messages that convey these traits should be paramount for future notification systems attempting to successfully reduce people's sedentary behaviour. This finding agrees with recent research on push notifications that notification content and how you speak to people is also critical to notification behaviour change effectiveness (Bidargaddi et al., 2018).

As Competition (N = 6, 12.5%) was joint second for reasons for notification preference, it's clear people love beating their friends; *"I feel because it is a competition and it gives me a goal I would need to get up and move in order to win."* (P16), are concerned if they are falling behind others; *"Competing with others because I don't want to be behind others."* (P4).

Participants also found the Specific Information category motivating; *"I think the second one is the best because it informed me with accurate data, which would be more convincing."* (P1), which shows that participants appreciated the clarity of having specific information relating to time or how many active breaks they are behind their friends in their notification choice. This is similar to previous research findings suggesting participants' responses to notifications increased when they were exposed to personalised notification messages (Compernolle et al., 2021; Orji et al., 2017).

Disliking the notification style was also a common narrative running through participants' comments. Disliked Social (N = 4, 8.3%) was a common category and showed us that people actually didn't respond to the social benefits of getting together and enjoying a break with friends; *"I'm more concerned about my health than catching up my friends."* (P1). We found this surprising as we thought this could be a main driver especially in participants who didn't like being active in general. As mentioned previously, Dislike Comparison (N = 4, 8.3%) was also a popular reason for being turned off by the notification styles; *"I will ignore it because I don't like to compare myself."* (P3).

Although not a notification preference reason, it's perhaps worth noting that from the interview data that was not relating directly to the research question, most people we interviewed don't understand the health risks from sedentary behaviour are independent of whether you have active hobbies or not. They think that if you have active hobbies you are ok from a health perspective. The health risks come from sitting or lying down even if you have just engaged in an active activity (Bankoski et al., 2011).

4.3 Further Study

Future studies can go in many directions, however from our experience of this study's findings, we propose the following improvements in further research to increase the validity of research to gain richer data. With the following suggested improvements in mind, we would also suggest two future research directions.

To increase the construct validity, we have three recommendations for the following researchers. First, we suggest using a different psychometric scale, other than the Perceived Persuasiveness Scale survey we used in this study, to measure the perceived persuasiveness of the notifications. The scale in future studies should be chosen carefully, ensuring that the Subjects-to-Items ratio (STI) in the development of the survey should be around 5:1 (Hair at al., 2014). Also, when choosing the survey, researchers should make sure that the number of items in each construct should be more than three (Hair at al., 2014). Second, there was no control group in this study, so there was no baseline for researchers to compare the three persuasiveness scores of the three notifications. We recommend adding a control group to help future researchers to compare the persuasiveness scores obtained from the experimental groups to the persuasiveness scores obtained from the control group which receive notification without application of any persuasive techniques. Third, we suggest using fitness trackers or apparatus such as smartwatches or pedometers to track the participants' level of sedentary behaviour. In this study, the sitting time of the participants was self-reported data, the participants in this study only need to report their daily sitting time approximately from their perception, so the data could be biassed because of the participants' self-serving bias or social desirability bias. Also, the researchers could not know if the participants would change their behaviours in real-life after they receive the persuasive notifications in this study. By using a tracker, researchers could measure the sitting time before, during and after the participants are exposed to the persuasive notifications. Hence, the persuasive effect of the notifications could be measured quantitatively and objectively. This could help researchers understand if the notifications' design can really affect participants' sedentary behaviour.

To increase the internal validity, we also have three suggestions for future enhancements. First, it would be beneficial to separate the participants into different groups. In this study, we found that the participants were in different States of Change (SoC), which might affect the total scores of the perceived persuasiveness of notifications. This could also be the reason that we could not find significant results in quantitative survey results in this study. The variability of the participants' willingness or intention to change confounded the perceived persuasive scores collected. For future study, it would be interesting to separate people who are in the Contemplation and Preparation stages and who are in the Precontemplation stage into two groups. This will need a participants' pre-screen before study begins. People who are in the Contemplation and Preparation stages understand the benefits of physical activities, have been mentally preparing to change and have not yet taken any action. They are the group at a tipping point in their lives in terms of their receptiveness to change. People who are in the Precontemplation stage are very resistant to behavioural change at this point in their lives and may need a totally different approach to get them to the Contemplation and Preparation stages. Therefore, future researchers might get interesting results of the perceived persuasiveness from the two different groups of participants. Second, in this study, the participants were divided into three groups, which had three orders of the notification styles. This was to limit the number of experimental groups and number of scripts administered by the researchers, reducing the chance of researchers errors. Ideally, a fully randomised study would have six experimental groups for the three notifications types. The partial randomised experimental

groups in this study discounted its internal validity. For future study, it would be beneficial to increase the experimental groups which fully randomise the order of notifications types. Third, a future study could control for what times participants were exposed to the notifications and what activity they did before they were exposed to notifications, to ensure confounding conditions are minimised.

To increase the ecological validity of future study, we suggest two improvements. First, further research could be a longitudinal study which exposes the participants to the three styles of messages for a longer period of time. This not only can ensure that the participants will receive the messages well, but also make sure that the participants have experienced persuasion. Furthermore, a long-term exposure of the notifications makes sure that the research resembles how a user will receive notifications from real-world applications. Second, we would suggest that future researchers employ a specific notification push system or develop their own notification systems to deliver the notifications, so participants can gain the experience of receiving a notification from a desktop's application rather than receiving a notification from WhatsApp Desktop. This measure can ensure the formal quality of the notifications, resembling the real-world experiences of using a desktop application. Through the above mentioned improvements, this could increase the ecological validity of the study.

For future research directions, we would suggest fully separating the quantitative and qualitative parts of the study into two individual parts.

The first research direction is survey-based research. In this study, the researchers needed to be present with the participants to send the participants the WhatsApp notifications and then conducted the surveys and interviews with the participants. Doing both the surveys and interviews at the same time limited the number of participants who could be involved in the survey. As we were constrained to analyse the qualitative interview data, we would only get a small sample size. Therefore, we would suggest doing a full-scale quantitative study in the future, which would be researched with a larger sample size of at least 100 participants, hence increasing the external validity of the research. Future research could be a within-subject study that shows participants the images of the three notifications on the survey forms and ask participants to rate the perceived persuasiveness of the notifications. This not only can increase the power of the study, lowering the risk of committing Type 2 errors, but also can increase the generalizability of the study. Moreover, there is no need for the researchers to send the messages manually to participants, it could also reduce the time cost of researchers being present in the study.

The second research direction is an in-person study. Our study was conducted on WhatsApp and at the participants' choices of locations and times. The researchers could not make sure the participants were exposed to the notifications after they had sat and worked on a laptop for a period of time. Also, the researchers could not see the facial expressions of the participants. The researchers could not make sure if the participants paid attention to the notifications. Hence, the persuasiveness rated by the participants were hypothetical. A further study, conducted in-person, would involve participants being exposed to the same interventions every hour after working on a work task replicating the working day and environment. This design would be beneficial for design recommendations for use in the real world. Conducting an in-person study can increase the external and ecological validity of the research. Moreover, conducting in-person interviews can help researchers to interpret the emotion and attitude of the participants towards the intervention through their facial expressions and tones of voices.

5 CONCLUSIONS

Understanding the notification styles that are most persuasive in order to combat the worrying increase in sedentary behaviour (Parry & Straker, 2013) is extremely important for human health across the globe. With many workforces becoming more reliant on sedentary computer-based tasks (Cascio & Montealegre, 2016), with jobs seemingly decreasing in physical nature (Brownson et al., 2005), people becoming increasingly attached to their devices in their free time (Yang et al., 2019) and sedentary behaviour correlated with many long term negative health impacts (U.S. Department of Health and Human Services, 2018), for example in the UK, 11.6% of total deaths were correlated with sedentary behaviour (Heron et al., 2019), it's very clear that any sort of understanding and insights we can gather in order to help change people's behaviour on-mass, are extremely worthwhile. Through our qualitative research we have shown that people prefer competition and cooperation style notifications. We have found that social comparison style notifications aren't popular in our study, which mirrors previous research that comparison with others can make people feel anxious, body shame them and generally either not motivate them or motivate them in an unenjoyable way (Orji et al., 2019). We also discovered that special care must be taken when designing notification messages, as this is just important as the notification style itself. The following must be considered; the tone of the message, creating trust, believability, privacy and being specific in regards to quantifying times and active breaks in order to fully maximise the potential of notification message styles.

Although we found no difference in persuasion levels between competition, cooperation and social comparison notifications, this could be due to the limitations of the study. Having a small sample size further reduced by having over 38% in States of Change (SoC) where participants are unlikely to need or want to be motivated to change behaviour and therefore possibly creating a type II error, the Wizard of Oz study perhaps not being believable to participants, not using the most valid scales to measure persuasiveness and SoC and only exposing participants to one notification of each style once, we can therefore only inform and motivate further research to address these limitations to further understand which one, if any, of competition, cooperation or social comparison desktop screen-based notifications is more effective at inspiring people to get up and take regular active breaks for the benefit of the health of humankind.

REFERENCES

Bankoski, A., Harris, T. B., McClain, J. J., Brychta, R. J., Caserotti, P., Chen, K. Y., Berrigan, D., Troiano, R. P., & Koster, A. (2011). Sedentary activity associated with metabolic syndrome independent of physical activity. *Diabetes Care, 34*(2), 497–503. https://doi.org/10.2337/dc10-0987

Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners.* SAGE Publications Ltd.

Bidargaddi, N., Pituch, T., Maaieh, H., Short, C., & Strecher, V. (2018). Predicting which type of push notification content motivates users to engage in a self-monitoring app. *Preventive Medicine Reports, 11,* 267-273. https://doi.org/10.1016/j.pmedr.2018.07.004.

Brombacher, H., Ren, X., Vos, S., & Lallemand, C. (2020). *Visualizing computer-based activity on ambient displays to reduce sedentary behavior at work*. In 32nd Australian Conference on Human-Computer Interaction (OzCHI '20). Association for Computing Machinery, New York, NY, USA. 760–764. https://doi.org/10.1145/3441000.3441022

Brownson, R. C., Boehmer, T. K., & Luke, D. A. (2005). Declining rates of physical activity in the United States: What are the contributors?. *Annual review of public health, 26*, 421–443. https://doi.org/10.1146/annurev.publhealth.26.021304.144437

Cascio, W. F., & Montealegre, R. (2016). How technology is changing work and organizations. *Annual Review Of Organizational Psychology And Organizational Behavior, 3*, 349-375. https://doi.org/10.1146/annurev-orgpsych-041015-062352.

Compernolle, S., Van Dyck, D., Cardon, G., & Brondeel, R. (2021). Exploring Breaks in Sedentary Behavior of Older Adults Immediately After Receiving Personalized Haptic Feedback: Intervention Study. *JMIR mHealth and uHealth, 9*(5), e26387. https://doi.org/10.2196/26387

DeVellis, R. F. & Thorpe, C. T. (2022). *Scale Development. Theory and Applications* (5th Ed.). SAGE Publications, Inc.

Gabrielli, S., Forbes, P., Jylhä, A., Wells, S., Sirén, M., Hemminki, S., Nurmi, P., Maimone, R., Masthoff, J., Jacucci, G., (2014). Design challenges in motivating change for sustainable urban mobility. *Computers in Human Behavior, 41*, 416-423. https://doi.org/10.1016/j.chb.2014.05.026.

Hair, J. F., Black, W. C., Babin, B. J. & Anderson, R.E. (2014). *Multivariate Data Analysis*. (7th Ed.). Pearson Education, Upper Saddle River.

Han, H., Gabriel, K. P., & Kohl, H. W. (2015). Evaluations of validity and reliability of a Transtheoretical Model for sedentary behavior among college students. *American Journal of Health Behavior, 39*(5), 601-609. https://doi.org/10.5993/AJHB.39.5.2

Heron, L., O'Neill, C., McAneney, H., Kee, F., & Tully, M. A. (2019). Direct healthcare costs of sedentary behaviour in the UK. *Journal of Epidemiology and Community Health, 73*(7), 625–629. https://doi.org/10.1136/jech-2018-211758

Kosaki, K., Takahashi, K., Matsui, M., Yoshioka, M., Mori, S., Nishitani, N., Shibata, A., Saito, C., Kuro-o, M., Yamagata, K., Oka, K., & Maeda, S. (2021). Sedentary behavior and estimated nephron number in middle-aged and older adults with or without chronic kidney disease. *Experimental Gerontology, 154*. 111531. https://doi.org/10.1016/j.exger.2021.111531

Luo, Y., Lee, B., Wohn, D. Y., Rebar, A. L., Conroy, D. E., & Choe, E. K. (2018). *Time for Break: Understanding information workers' sedentary behavior through a break prompting system.* Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY. Paper 127, 1–14. https://doi.org/10.1145/3173574.3173701

Morris, D., Brush, A. B., & Meyers, B. R. (2008). SuperBreak: Using interactivity to enhance ergonomic typing breaks. *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*. https://doi.org/10.1145/1357054.1357337

Ng, S. W., & Popkin, B. M. (2012). Time use and physical activity: a shift away from movement across the globe. *Obesity reviews: an official journal of the International Association for the Study of Obesity, 13*(8), 659–680. https://doi.org/10.1111/j.1467-789X.2011.00982.x

Norrie, L., & Murray-Smith, R. (2015). *Impact of smartphone notification display choice in a typing task*. In Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '15). Association for Computing Machinery, New York, NY, USA. 1094–1099. https://doi.org/10.1145/2786567.2794335

Orji, R., Mandryk, R. L., & Vassileva, J. (2017). Improving the efficacy of games for change using personalization models. *ACM Transactions on Computer-Human Interaction, 24*(5), 1-22. https://doi.org/10.1145/3119929

Orji, R., Oyibo, K., Lomotey, R. K., & Orji, F. A. (2019). Socially-driven persuasive health intervention design: Competition, social comparison, and cooperation. *Health informatics journal, 25*(4), 1451–1484. https://doi.org/10.1177/1460458218766570

Orji, R., Tondello, G. F., & Nacke, L. W. (2018). *Personalizing persuasive strategies in gameful systems to gamification user types*. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY. Paper 435, 1–14. https://doi.org/10.1145/3173574.3174009

Parry, S., & Straker, L. (2013). The contribution of office work to sedentary behavior associated risk. *BMC public health, 13*. 296. https://doi.org/10.1186/1471-2458-13-296.

Prince, S. A., Melvin, A., Roberts, K. C., Butler, G. P., & Thompson, W. (2020). Sedentary behaviour surveillance in Canada: trends, challenges and lessons learned. *The International Journal Of Behavioral Nutrition And Physical Activity, 17*(1), 34. https://doi.org/10.1186/s12966-020-00925-8

Prochaska, J. O., & DiClemente, C. C. (1983). Stages and processes of self-change of smoking: Toward an integrative model of change. *Journal of Consulting and Clinical Psychology, 51*(3), 390–395. https://doi.org/10.1037/0022-006X.51.3.390

Sedentary Behaviour Research Network. (2012). Letter to the editor: Standardized use of the terms "sedentary" and "sedentary behaviours". *Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition et Metabolisme, 37*(3), 540–542. https://doi.org/10.1139/h2012-024

da Silva, J. M., Verlengia, R., de Oliveira, J. J., Ribeiro, A. G. S. V., Barbosa, C. G. R., Stotzer, U. S., & Crisp, A. H. (2020). Associations between sociodemographic factors and physical activity and sedentary behaviors in adults with chronic diseases during COVID-19 pandemic. *Sports Medicine and Health Science*, *2*(4), 216–220. https://doi.org/10.1016/j.smhs.2020.11.002

Thomas, R. J., Masthoff, J., & Oren, N. (2019). Can I influence you? Development of a Scale to Measure Perceived Persuasiveness and Two Studies Showing the Use of the Scale. *Frontiers in Artificial Intelligence, 2*(November). https://doi.org/10.3389/frai.2019.00024

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S., Altenburg, T. M., Chinapaw, M., & SBRN Terminology Consensus Project Participants (2017). Sedentary Behavior Research Network (SBRN) - Terminology consensus project process and outcome. *The international journal of behavioral nutrition and physical activity, 14*(1), 75. https://doi.org/10.1186/s12966-017-0525-8

U.S. Department of Health and Human Services. (2018). 2018 Physical Activity Guidelines Advisory Committee Scientific Report. Retrieved February, 2018 from Office of Disease Prevention and Health Promotion, Department of Health and Human Services: https://health.gov/our-work/nutrition-physical-activity/physical-activity-guidelines/current-guidelines/scientific-report

Yang, L., Cao, C., Kantor, E. D., Nguyen, L. H., Zheng, X., Park, Y., Giovannucci, E. L., Matthews, C. E., Colditz, G. A., & Cao, Y. (2019). Trends in sedentary behavior among the US population, 2001-2016. *JAMA, 321*(16), 1587–1597. https://doi.org/10.1001/jama.2019.3636

APPENDIX A

Perceived Persuasive Scale (PPS)

Effectiveness

1. This message will cause changes in my behavior.
2. This message causes me to make some changes in my behavior.
3. After viewing this message, I will make changes in my attitude.

Trustworthy

1. This message is accurate.
2. This message is trustworthy.
3. I believe this message is true.

Capability

1. This message has the potential to change user behavior.
2. This message has the potential to influence user behavior.
3. This message has the potential to inspire users.

APPENDIX B

Transtheoretical Model's questionnaire for 'Stages of Motivational Readiness to Avoid Sitting Time'

Q1. Do you achieve sufficient levels of physical activity on most days?

Yes/No

Q2. Do you think you are currently sitting most of the day?

1. Yes, and I do not intend to avoid my sitting time
2. Yes, but I intend to avoid my sitting time within the next 6 months
3. Yes, but I intend to avoid my sitting time within the next 30 days or sometimes do some movements such as interruption (break) of prolonged sitting to reduce sitting time
4. No, I am not sitting. I began doing frequent movements to interrupt or avoid prolonged sitting time within the last 6 months
5. No, I am not sitting. I began doing frequent and regular movements to avoid or break prolonged sitting time more than 6 months ago.

APPENDIX C
Interview Questions

Background activity questions

1. How long would you say you sat down on an average day?
2. What are the main activities in your daily routine that make you sit down for long periods of time?
3. What are the important factors that influence you to sit down?/Why do you sit down instead of standing up or being active?
4. Do you think that having active hobbies counteracts the effects of sitting or lying down for long periods of time? Why is that?

General notification questions

1. What do you think of desktop notifications as a reminder tool?
2. Do you usually turn on desktop notifications or have them off?
    a. What makes you turn on/off desktop notifications?
    b. Why is that?
3. What positive (or useful) things do you associate with desktop notifications?
4. What negative (or unuseful) things do you associate with desktop notifications?
5. When you are studying or working, what are usually your first thoughts when a notification appears on your desktop?
    a. Do you want to read it, read it and action it or just ignore it?
    b. What makes you ignore a notification?
    c. Does it depend on the relevance of the notification?
    d. Or something else?

Specific intervention questions

1. Let's return to the three notifications I sent you earlier. Imagine you receive these notifications after_____(fill in the participants' habits)____, how would you respond to them? Please don't rush your answers, have a good think before you reply.
2. Which notification would persuade you to move more? Why is that?
3. Apart from these three notification designs, do you have any other design or message in mind that can persuade you to take more active breaks? Or do you have any persuasive notifications you have got before that you can share?

Final question to wrap up the interview

1. Any other thoughts about sedentary behaviour, active breaks or notifications?

Interviewers are encouraged to ask further questions if necessary.