# Persuasion in Exercise:
## the Pep to make the extra Step

Master's Project Thesis
Programme: Master Artificial Intelligence, Utrecht University

Yvonne Jolanda Hoven, y.j.hoven@students.uu.nl, 6654185
Hanna J. Hauptman, h.j.hauptmann@uu.nl, Utrecht University
Ana M. D. Ciocarlan, a.m.d.ciocarlan@uu.nl, Utrecht University
Eelco Herder, e.herder@uu.nl, Utrecht University
Thijs Sondag, thijs@weguide.com.au, WeGuide

September 26, 2022

## Abstract

This research investigated which strategies would persuade people to walk more than they would typically do. This research was in cooperation with the IT healthcare company WeGuide, which is interested in getting a more intelligent platform in which they enlarge people's motivation to do their rehabilitation exercises. The exercise under investigation was walking more steps per day, for convenience reasons. The strategies used in the sentences that were sent to the participants to encourage taking more steps, were based on previous literature and research on the topic of persuasion. In addition to demographics, the personal profiles of the participants were deduced according to the `Brainhex` Model, the Susceptibility To Persuasion Scale (`STPS`) and the Trans-Theoretical Model (`TTM`). These profiles however, did not show specific preferences for persuasive strategies. Rather, it appeared from the pre-study that there was a general top three and bottom three along the nine strategies used for this research. The participants for the experiment were distributed into three groups accordingly, each with three persuasive strategies. In contrast to the pre-study, the groups with the less persuasive strategies improved their step-counts more than the group with the most persuasive strategies. `Keywords: persuasion, profiling, gamification, exercise, activity tracking, step-counts, Brainhex, Trans-Theoretical model (TTM), Susceptibility to Persuasion Scale (STPS)`

# Contents

# 1 Introduction

There has been quite some research on persuasion already, in various domains. The art of motivating and encouraging people to do things without using coercion or deception has become an important thing in society. Most persuasion is used for marketing goals, but it has also been used for the greater good, such as health. There are already many apps about health available for phones that use some sort of persuasion, to make the users do the right thing (eating healthier and exercising more). Persuasion has been shown to be effective at changing behaviors, without using deception. This also includes enhanced physical activity (Charity, Nkwo, Orji, and Ebere, 2020). As changing behavior can take a long time, long-term studies may be more applicable here (IJsselsteijn, de Kort, Midden, Eggen, and van den Hoven, 2006). However, there is not much research about the long-term effects yet (Orji and Moffatt, 2018).

There are many different ways to use persuasion and many different techniques of persuasion. For example, gamification can also be seen as persuasion. People are persuaded to play along, as it is made more fun, potentially by a leader-board with your friends' scores on it, to persuade you to do your best even more to win. It has been shown that people differ in their susceptibility to the various kinds of persuasion, due to their different personality characteristics Almutari and Orji (2021); Kaptein, Markopoulos, de Ruyter, and Aarts (2009); Orji, Vassileva, and Mandryk (2014b); Orji, Mandryk, and Vassileva (2014a). One goal of this research is to find some general persuasion profiles that work with most people. This makes it possible to adapt and tailor the persuasion techniques to many different people, which bears the most fruit in people adhering to their tasks. The other goal is to find the best persuasion technique for each of the different personality types.

For this research paper, a few different text formats are used. Here, these formats are explained. Paragraphs in **boldface** are summaries, relating the previous text to this specific research, experiment and/or research questions. Words in **boldface** are important words that are meant to be highlighted and carry the essence of the sentence. Notes about texts are written down in *italics*. Important keywords such as the specific personality profiles used and statistical terms are written in `typewriter font`. The three experimental groups (1, 2 and 3) are colored green, blue and red respectively, which are the same colors as in the figures. Citations of participants are in cyan. Lastly, the research questions and hypotheses are in *violet italics*, as well as the answers to them.

## 1.1 Relevance to Artificial Intelligence

First of all persuasion is a means of changing people's attitudes and/or behavior through communication. This is an important aspect to do research on, especially for (social and cognitive) psychology. Cognitive psychology and Artificial Intelligence (AI) share a similar goal, which is to understand the nature of intelligent behavior. While cognitive psychology tries to grasp cognition by conducting research, testing and building models of the human mind, AI tries to build those processes with advanced technologies through scientific knowledge. Therefore, AI is seen as a multidisciplinary field of science that tries to understand and develop technologies that act similar to human intelligence. This can be done in different levels of complexity, including that the AI can learn even without explicit rules.

Since psychology and mental health are also areas of focus for AI, AI offers a promising approach to assist, simulate or replace certain practices in health treatments. By evolving technology, new types of treatment can be made (e.g. with Augmented or Virtual Reality), or populations can be reached that were previously difficult to reach. These approaches could alleviate stress and lessen the strain on the resources of health specialists.

However, there are still some doubts about having AI take over such tasks. There are ethical issues regarding developing and integrating technologies while keeping clients anonymous, and considerations about the level of human supervision required with the technologies. When people's agency and autonomy are also taken care of, AI can capture and analyze big data sets to create more knowledge and understanding of health treatments.

Another issue is that people tend to trust AI less than humans, even when the AI is better at specific tasks. Therefore, more research needs to be done to increase the trust in AI. When AI is combined with persuasion, the design of the content of messages is generated, augmented and/or changed to increase persuasiveness. The increase in persuasiveness through using AI is promising for research in health especially. There is still ongoing research about the differences between human-made persuasion and persuasion by nonhuman agents.

For this research, the persuasion sentences are human-made, but the sentences that are sent were se-

lected automatically. The sentences have some checks whether they are applicable sentences or not (do not send messages like: 'You did great yesterday with 34 steps!'). Participants that had particular personality profiles or were in a specific experimental group, got messages with specific persuasion strategies based on their profile or group. In such a case, AI can extract answers from a survey with big data, to create a profile per person, to decide which persuasion strategy would work the most on that person specifically. This automated decision-making process would take more time and could be more error-prone if done by hand.

## 1.2 Motivation

Persuasion has become increasingly popular in research. It provides ways to make people change their behavior more easily. For this research, persuasion is used to try and improve the health of people by increasing the number of steps set per day. The goal is to make people exercise and move more to improve their own health. Good health is important for everyone, but since COVID-19 happened, people tend to stay at home more often and work from their home offices, which leads to decreased exercise. To decrease this sedentary behavior and increase their physical activity, an increase in health would be achieved for those people. This is because little exercise can lead to heart diseases and diabetes for example (Oyebode, Maurya, and Orji, 2020c). It is also shown that more exercise could help alleviate depression in people Fog (2009).

Since recent years, persuasion is on the rise and especially in the technology domain. Persuasion could very well be useful as changing behaviors can take quite some time, because they are habitual behaviors (IJsselsteijn et al., 2006; Kelly and Barker, 2016). Artificial Intelligence can also be used to get the best persuasion strategies across people who need that extra motivation to do the right things. Sometimes exercise can be boring, especially when you have to do certain repetitive exercises given by your doctor or physician. The right persuasion technique could help those people to still have the motivation to do their respective exercises, which helps their overall health. In conclusion, using persuasion techniques to exercise more and correctly, is quite a relevant subject.

## 1.3 Problem Statement

The goal of this research was to investigate the potential but also the pitfalls of persuasive strategies and to compare them against one another. More concretely, the goal was to find the most effective and persuasive strategies in the domain of exercise. This was done in a study among (mostly) students, to see which persuasive strategies work best to let people exercise more. The exercise was measured in the number of steps set per day. A ground truth benchmark week for each participant was compared against the weeks that a specific persuasion technique is sent via a message to each participant. After three weeks of receiving persuasive sentences, another week's worth of step-counts was recorded without messages being sent to the participants. This is to investigate the long-term result of receiving persuasive messages. Various different persuasive strategies were used to see which ones had a greater positive impact. For tracking the change in the number of steps, a step-count app was used on mobile phones to count the number of steps per day for each participant.

There has been research on persuasion about differences between nations (Orji, Mandryk, Vassileva, and Gerling, 2013a), sexes and ages, but about persuasion in mobile games, apps and reviews that compare the best working strategies and their pitfalls. For this research, various persuasion techniques were investigated and some were used to check which one(s) are most effective for adhering to exercise (in this case, measured by the number of steps walked per day per participant). Research on (personalized) persuasion has not been analyzed yet in all (sub)domains however. Since there are so many ways to characterize personality types, it gives ample opportunity to cross-reference personality types with all different domains. When there are specific persuasion techniques that improve exercise more or are more effective with some personality types, it could also hold for other domains closely related to either health or exercise as well. This is however not yet proven, as much research still needs to be done. Another interesting aspect is to research whether perceived difference (in a survey) compares to actual persuasion (in an experiment setting), as much research has been done via surveys alone.

When the best persuasion techniques for exercise are found and are personalized to each user, more people will have better adherence to their exercises, which means better health for them overall. Therefore, this research might contribute to giving the optimized persuasion needed for better health for everyone. When everyone has better health, by exercising (or walking) more (especially in these times of COVID-

19), there will be fewer diseases in the world, like diabetes, obesity and high blood pressure. The IT healthcare company WeGuide had also an interest in the possibilities of persuasion for their clients, as most users of their app need to adhere to exercise behaviors as well. Therefore, WeGuide funded the pre-study and the experiment for this research.

A focus group was held to gather the most useful and most influencing persuasion sentences per strategy. The focus group would create an overview of which persuasive strategies and sentences tend to work the best. Then, a pre-study would help match personality profiles to their preferred persuasive strategies. For the profiling of personality types, three common categories of personality profiles were used. The participants were asked which (sentences incorporating the) strategies they found the most persuasive. This was compared to the outcomes of some research on preferred personalized persuasion, to pick the persuasion strategies for the real experiment. This was also used to compare the susceptibility to persuasion of people and whether it differs among people. When the best strategies were decided, they were used to make different conditions for the final experiment. The goal for the participants is then to make a few extra steps per day. There were no specific goals set for any group, as having a goal might influence the results.

The results of the possible change in the number of steps per day were analyzed and reported. In addition to that, the TTM, self-efficacy, attitude and against intention toward walking more steps were compared (between before the experiment and after the experiment). This was done to measure effects in feelings regarding walking, that could not be measured in step-counts alone.

## 1.4 Research Questions

It is known that persuasion can help change behavior in different domains and it is known that personalized persuasion is often better than general persuasion. However, which persuasion has the overall best effect in which domains is still under investigation. The aim of this Master's Thesis research is to answer the following research question: *Does using a specific persuasion technique have an increased effect on top of the effect of using persuasive technology in general, in the domain of exercise measured by the number of steps walked per day, and is there a difference between perceived and actual persuasion?*

To find an answer to these questions, there are a few sub-questions to keep in mind:

- '**Which** persuasive strategies are the **most effective** in increasing physical activity by exercising, measured by the number of steps set per day?'

- 'Is there a **specific** persuasion technique that has a **bigger effect** for each personality type?'

- 'Do people have **differences** in what they **think** is persuasive for them (e.g. in surveys) compared to what is **actually persuasive** for them (e.g. during an actual experiment) and does that depend on their personality profile?'

These questions were answered by comparing strategies one to one but also in groups, whether there is an added or weakened effect when combining strategies. In comparing strategies one to one, the goal was to identify the strategies that perform above the average of the persuasion strategies; looking at how many participants it has an increased effect with and how large that effect is. The effects could be measured in different ways: the increase/decrease in the step-count, how engaged participants were with the experiment, a possible difference perceived in the amount of support that the messages gave, participants' experience with the experiment (like/dislike) and their long-term attitude (possible) changes against persuasion or exercise. For the difference between perceived and actual persuasion, the pre-study will be compared to the outcomes of the experiment. The pre-study entails perceived persuasion of the messages, while the results of the experiment hold the actual change in behavior and thus the actual persuasiveness of the messages.

The hypotheses here are:

- '*The strategies with praise and personalization have a **bigger positive effect** on walking more than strategies like information and punishment*'

- '**Specific persuasion** does have an **added increased effect** in adherence to exercise, compared to general persuasive strategies.'

- '*There is **no difference** between perceived persuasion in surveys, compared to actual persuasion in experiments.*'

What is meant by this is that (1) persuasion does have an increased effect on changing behavior and that (2) specific persuasion strategies in this domain have an even bigger increased effect on changing behavior. The overall best-performing strategies are the strategies that have an increased effect in changing behavior with most personality types. Since people tend to like praise and since personalization is considered more personal advice, these strategies were hypothesized to have a positive influence on the number of steps. Since people tend to not like non-personalized information dumps and negative feedback like punishment, these strategies were hypothesized to have a less positive influence, or even a negative influence, compared to other persuasive strategies.

When these research questions are answered, it has implications for further research. There has been research on persuasion and personalized persuasion, but not every domain has been analyzed yet. Since there are so many ways to characterize personality types, it gives ample opportunity to cross-reference personality types with all different domains. When there are specific persuasion techniques that improve exercise more or are more effective with some personality types, it could also hold for other domains as well. Aside from that, it is important to look at the differences between perceived persuasion and actual persuasion, as much research has only been done in the form of surveys, rather than experiments. When there is a discrepancy between perceived persuasion in surveys and actual persuasion (meaning a change in behavior), previous claims on (personalized) persuasion could be rendered less useful.

For the company WeGuide it also has some implications. WeGuide is a company that uses an app for helping users with their own health. The most used feature is trying to make users keep doing their exercises, for rehabilitation for example. When the best persuasion techniques for exercise are found and are personalized to each user, more people will have better adherence to their exercises, which means better health for them overall. WeGuide is still a company that is mostly known in Australia, but when this adjusted persuasion really helps improve health in Australia, the company might grow, such that people outside Australia will also know the company and will benefit from their app.

## 1.5 Outline of Document

In Section 2, there is relevant literature about persuasion, personalization and technological persuasion in the health domain. This is all to gain an insight into how large this subject is and how and why it is narrowed down to this specific research. In Section 3, the methodology and tactics of the research itself are discussed. This includes a focus group, a pre-study and the experiment with a survey before and after it. The actual implementation of these can be found in Section 4. After that, the results and analysis can be found in Section 5 and Section 6. This is followed by the discussion, conclusion and future work.

# 2 Related Work

In this section, the previous literature on persuasion is introduced. First general persuasion is discussed, and then different types of persuasion are discussed. This is followed by persuasion used in technologies, personalization of persuasion and persuasion in health. These are important to note, since this research is about personalized persuasion, by using technology, to create better health for participants.

## 2.1 Persuasion in General

In this section, it is explained what persuasion is and its importance. To start off, persuasion is part of social influences. People are primarily encouraged and motivated intentionally to follow an attitude, idea or action that is presented symbolically. This is on voluntary basis and thus done without using coercion or deception (Fogg, 2002). Persuasion is important because it can help people convince to do the right thing more easily, which is oftentimes better for themselves, for example for their own health. Some high-level distinctions of persuasion are given below.

There are two types of social influence according to Deutsch and Gerard (1955):

1. Informational influence: This is where a person accepts information because it makes sense, and agrees with the other person.

2. Normative influence: This is where a person accepts information just to fit in and to seek approval from the other person. This influence was also used by Almutari and Orji (2021) and framed it as gathering people with similar goals to help increase change in behavior or reinforcement.

There are three processes of social influence according to Caldini and Goldstein (2004):

1. Compliance: This is a change in behavior in response to a request or under influence of a group, but that person's attitude does not change.

2. Obedience: This is complying with someone's wishes because of that person's authority. An example of obedience research is Milgram's experiment (Milgram, 1963). Participants were instructed to give electric shocks to another person (an accomplice) whenever that person failed to memorize the items of a list, see Subsection 2.1.2.

3. Conformity: This is an actual change in attitude/behavior due to pressure from friends or norms of a group.

Here are two other well-known techniques of compliance:

1. Foot-in-the-door technique (Freedman and Fraser, 1966): This starts with a small request, small enough to be agreed upon. Then it is followed by a large(r) request. In this case, a person will repeatedly accept requests, because of the need for commitment and consistency.

2. Door-in-the-face technique (Cialdini, Vincent, Lewis, Catalan, Wheeler, and Darby, 1975): This starts with a big request, so big that it will be turned down. Then it is followed by a small(er) request. This speaks to our need for reciprocity to the concession of the other person.

**For this research, this general information about persuasion is useful to distinguish different kinds of persuasion that can be used in the experiment later on. It is relevant because this research is about which persuasion works best for whom. A few of these persuasion technologies mentioned were used to find the best-personalized persuasion to get the most exercise out of everyone. For this research, it is more about informational influence than normative influence, because the advice is given as information, rather than social interaction. The door-in-the-face and foot-in-the-door technique were not used, as there was only communication via chat (sending generated messages) between the researcher and the participants. No requests were made. Obedience was not used directly, as the experimenter has no real authority over the participants. However, some messages that were sent to the participants included an official authority figure stating certain things. The idea is that compliance would follow from the experiment by participants setting more steps per day during the experiment. It could also lead to conformity where participants keep walking more, even after the experiment.**

### 2.1.1 Types of Persuasion

There are many ways to look at the various kinds of persuasion techniques and strategies. Researchers do not agree on how many persuasion techniques there are, as it depends on which dimensions or clusters they are put in, or when they are considered individually (Rhoads, 2007). Below, some well-known theories, techniques and strategies are shown in tables, followed by an explanation of which ones were used for this research.

In Table 2.1 some (older) theories on persuasion can be found and Table 2.2 gives an overview of the most recognized persuasive techniques and strategies.

Thus far, these categorizations are used by Orji, Nacke, and Di Marco (2017b) and/or Oyebode, Ganesh, and Orji (2021a). Almutari and Orji (2021) also used competition, cooperation, personalization and rewards, but also a normative influence (see Table 2.3 and the lesser-known strategy list below). In Table 2.3 are the six principles of persuasion/compliance by Cialdini and Griskevicius (2010), which were formerly called the six basic categories of (social) influence:

Furthermore, the seven persuasive strategies that can be used in persuasive tools to make a persuasive system by Fogg (2002) are a bit older and can be found in Table 2.4.

Lastly, here are some lesser-known persuasive strategies.

- Surface credibility: This is about the first impression that you get of a person or company. Whether the object is considered credible without knowing too much about it yet.

- Social learning: The learning that occurs in a social context. No reinforcement or actions are done by the learner, just observations are done, or instructions are followed.

| Name of theory | The thing that persuades people is ... |
|---|---|
| Elaboration Likelihood Model (Petty and Cacioppo, 1986) | Logical sound arguments |
| Yale Model (Aronson, Wilson, and Akert, 2010; Hovland, Janis, and Kelley) | Characteristics of the source or sender (is (s)he credible?), the message itself (facts, arguments, emotions) and the (target) audience (needs, concerns, identify with speaker) |
| Classical Rhetoric (Wróbel, 2015) | Language that appeals to our reason(logos), emotions (pathos), or morality (ethos) |
| Altercasting (Weinstein and Deutschberger, 1963) | Forcing someone into a certain (social) role |
| Social Judgment (Sherif and Hovland, 1961) | Characteristics of the influencer (the idea is evaluated based on their own attitudes) |
| Narrative paradigm (Fisher, 1985) | Narration or storytelling using good reasons |

**Table 2.1: Theories of persuasion**

| Persuasive strategy | Explanation |
|---|---|
| (Social) Comparison | To view and compare their performance with those of others (personal comparison (hot/cold) rather than an objective measure (degrees on thermometer). Social comparison is not about competition. |
| Competition | Competing with others to perform the desired (target) behavior. This is often implemented via a leaderboard. |
| Cooperation | Working together to achieve a shared objective or intended target behavior. Rewards are given for achieving the goals collectively. |
| Customization | Opportunity to adapt a system's content and functionality to people's preferences. |
| Goal setting (and Suggestion) | Setting a clear behavior goal and recommending/suggesting certain actions and tips to achieve that target behavior. |
| Self-Monitoring (and Feedback) | Tracking previous and current states of behavior performance and thereby the progress towards their goals. |
| Simulation | To be able to observe the link between cause and effect of the target behavior. |
| Personalization | Personalized content for users, considered as specific individuals. The system is tailored in features and functionality to user's needs and characteristics. |
| Praise | Motivating users through positive words, images, symbols, or sounds. |
| Punishment | Penalizes the user for not performing the desired behavior or reaching their goal (by removing rewards). |
| Recognition | Public recognition for users who perform their target behavior. |
| Reduction | Reduces the effort that users put in to perform their target behavior. |
| Reminder (and Rehearsal) | Reminding users of their target behavior while using the system. |
| Reward | Giving rewards and credit to users for performing the target behavior and for achieving milestones (by giving badges or points). |

**Table 2.2: Persuasion categories by Orji et al. (2014b) and Oyebode et al. (2020b)**

Since the success of behavior change interventions is depending on the design and implementation of experiments, there is a need for characterizing interventions correctly and using correct analyses. (Michie, van Stralen, and West, 2011; Michie, Atkins, and West, 2014) evaluated older frameworks and created a new framework that was more comprehensible, had more coherence and a clear link to models of behaviors. In total, nineteen frameworks were combined covering nine intervention functions and seven policy categories, see Figure 2.1. At the center of the new framework, there are three essential conditions for behavior change to happen. These are capability, opportunity, and motivation. This is based on criminal law in the United States, where a person can only be found guilty if (s)he has met those three

| Persuasive strategy | Explanation |
| --- | --- |
| Authority (also Trust, Social Role and Obedience) | Authorities are influential people. People are more likely to follow the advice and suggestions of an authority or expert. People trust and obey others with high social roles or who have high expertise. |
| Commitment / Consistency | When committing to a certain position or idea, people comply more with behaviors that are consistent with it afterwards. Also, when one says that (s)he will do something, they more likely to do so, because they like to stay consistent. (In hotels, when there is a note that says that 75% of the guests re-use a towel, more people will re-use their towels https://youtu.be/cFdCzN7RYbw.) |
| Liking (also Similarity and (Social) Normative Influence) | Changing behavior and values to conform to other people's ideas so that they will like you more and you will fit in more. We seek approval and comply more easily to people who we like or people who are more similar to you. |
| Scarcity | When opportunities are limited or scarce, we see them as more valuable and want to secure them. |
| Social proof (also Consensus, Social validation and Third-party endorsements) | Using the beliefs, attitudes and actions of (similar) others to compare it to their own. Behaviors are correct, as others do them and we comply if we think that (similar) others would comply. |
| Reciprocity (and Trust) | When one has provided you with a favor/gift or concession on their half, we will comply more with them later, as we feel a need to repay the favor. |

**Table 2.3: Persuasion categories by Cialdini and Griskevicius (2010)**

| Persuasive strategy | Explanation |
| --- | --- |
| Reduction | Persuading through simplifying. Making a complex task simpler by eliminating steps in the process of achieving the goal. |
| Tunnelling | Guiding the user by reducing "uncertainty". Leading users through a sequence of actions/events, step by step, to encourage specific behavior. |
| Customization / Tailoring | Using personally relevant information to change certain attitudes/behaviors, because people pay more attention to customized information. |
| Suggestion | Finding the best time to remind users to perform the target behavior, as people feel more motivated to perform some behaviors at certain times. |
| Self-Monitoring | A way to track and monitor your progress in achieving the goal. By tracking, one can modify their attitudes/behaviors to achieve that goal. |
| Surveillance (also Social Facilitation and Observation) | People put in more effort and behave differently when they feel watched by others. |
| Conditioning | To reinforce good/target behavior, operant conditioning is used (positive reinforcement). |

**Table 2.4: Persuasion categories by Fogg (2002)**

conditions. Around that, there are nine intervention functions that can address deficits in the three conditions and therefore change them which will minimize the risk of reverting. One of the interventions is persuasion, which is used for the experiment of this research. Lastly, the outer circle has seven policies that can allow and support the interventions to happen. For this research, the communication policy is used to send persuasive messages as an intervention. This will create social and physical opportunities to improve the step-counts of participants. Social opportunity means that the step-counts of participants will be compared, by physical opportunity is meant that sometimes a message would say something about the weather that day. The persuasive messages are also meant to give participants automatic and reflective motivation to walk more than they usually would. Automatic motivation means that messages will provide praise or compassion or information about walking more. The reflective motivation could for

example be triggered by messages that reflect on previous step-count averages of the participant itself. By targeting the opportunity and motivation of participants, the behavior of the participants is expected to change (by setting more steps per day).
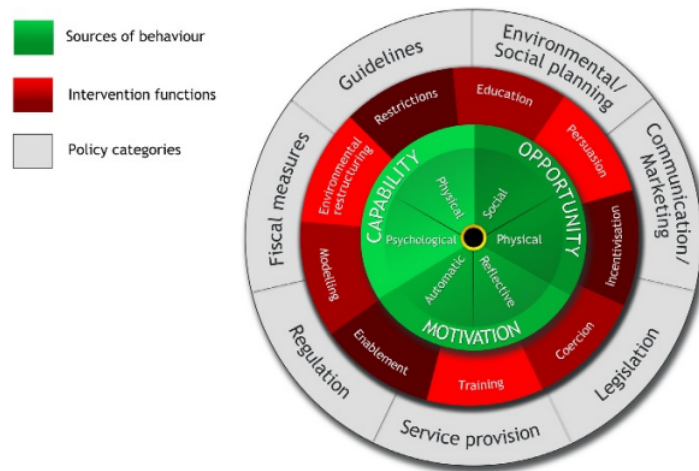


Figure 2.1: The Behavior Change Wheel by Michie et al. (2011, 2014).

For this research, this is relevant because an overview of which persuasive strategies are out there is needed to deduce which ones could and would be used for this research. The goal is then to compare this research to the outcomes of other research. Then, it can be found which strategies have the biggest positive effect on exercise. Knowing which persuasion techniques are used most often and knowing which ones are not can help in choosing the exact persuasive strategies for this research. Self-monitoring is however necessary, because of the nature of this research. While (oldest) seven persuasion types by Fogg (2002) stand at the core of persuasive strategies, the six different persuasion categories by Cialdini (2006) are the base of persuasive strategies now. Since some methods are not independent, they were taken together and considered as one for this research. For example, there are many implementations of Cialdini's strategies, one of them being that social comparison incorporates the liking principle. For this research, various persuasive strategies were evaluated to see which ones work best for whom. Since the number of persuasive strategies does not seem to affect the results (Alhasani, Mulchandani, Oyebode, and Orji, 2020), there is no need to stick to a certain number of persuasion techniques. For this research, the following persuasive strategies were considered for the experiment: authority/obedience, informational influence (elaboration likelihood model and classic rhetoric model), social learning/comparison and competition (social proof, liking, similarity, recognition, gamification), personalization and customization (tailoring), praise (conditioning) and punishment, rewards, reminder/rehearsal and commitment/consistency as consensus. See Subsection 3.6 for more on which persuasion could and could not be used.

### 2.1.2 Ethical Issues

There are various ethical questions and issues involved in research about persuasion. There are design, data and privacy issues as well as issues regarding the safety and well-being of participants. However, in persuasion there is also the issue of intended versus unintended consequences. Some examples are explained below. More ethical issues about personalized persuasion and persuasion in technologies can be found in Subsection 2.3 and Subsection 2.2, respectively.

The most well-known ethical issues arose with Milgram's experiment on obedience (Milgram, 1963, 1974). In that experiment, authority figures pushed participants to give electric shocks to other people. The interpretation of those results, external validity and ethical concerns are still under debate (Caldini and Goldstein, 2004). Burger (2009) mentions that the stress of the patients could not be justified, as well as the potential of long-term harm. Slater, Antley, Davison, Swapp, Guger, Barker, Pistrang, and Sanchez-Vives (2006) made a twist to the original experiment: virtual humans received the shock while participants were in an immersive virtual environment. Participants acted and responded behaviorally

and physiologically as if the virtual human was real, while only slightly more participants stopped the experiment. Differences between participants were visible in empathic concern versus the desire for control, rather than gender (Burger, 2009).

In general, (obedience) research needs to protect the well-being of participants and their stress levels. The challenge is to keep ethical integrity while participants also face the realness of the intended experiment (Cormier, Newman, Nakane, Young, and Durocher, 2013). Most research on obedience was limited because of well-being concerns (Burger, 2009).

Verbeek (2006b) mentioned the external moral questions on persuasive technology. The reliability and trust in the design and the responsibility of the designers might pose a threat to humans, their freedom and even democracy. When technologies have a mediating role in persuasion, designers have a moral activity in engineering such designs. Since technologies give material answers to the question of how to act (by suggesting actions), it raises moral and ethical problems. Thus, technological mediation is important for the ethics around designing technologies (alongside their functionality). Designers should be able to anticipate those mediating roles (Verbeek, 2006a,b). By adding moral reflection to the design process, designers should look at the intended persuasion, the form and method of mediation and the outcomes of that mediation. This could decrease the threat and moral problems of persuasive technologies.

Another possible problem is unintended consequences. While there are ethical guidelines to guide the design and use of persuasive technology, it might happen that people use it such that there are unintended consequences. This is the danger of using persuasive technologies like every other modern technology (Abdullahi, Oyibo, Orji, and Kawu, 2019). Aside from unintended consequences, there are also undesirable consequences. The method of persuasion can be problematic, or one can be persuaded to do something that is not morally justified (Verbeek, 2006b). Persuasion could raise moral questions by threatening human autonomy. Therefore, it has to be designed such that it does not lower the users' autonomy (Vlieghe and De Troyer, 2020; Aldenaini, Alqahtani, Orji, and Sampalli, 2020a; Verbeek, 2006b). It could also be seen as laziness on morality or even anti-democratic as the designers will steer our behavior (indirectly) (Verbeek, 2006b).

**For this research, ethical issues are relevant because different technologies were used to spread the persuasion. In an activity tracker app, participants were able to keep track of their own steps, while a text message system was used to deliver persuasive messages. Participant data needs to stay anonymous throughout and after the experiment. Participants should not be pushed too far in exercising, as it could lead to injuries. Furthermore, there could be counter intuitive effects associated with this experiment. People might push themselves too far in making more steps, which could result in negative side effects. Participants might start running without concern for their joints and hurt themselves. This is another reason why no fixed step-count goals were used in the experiment. Every participant might have their own schedule and capabilities of moving around.**

## 2.2 Persuasion in Technologies

Persuasive strategies can motivate and encourage people to achieve better health by overcoming the boredom of routines. Nowadays, persuasive strategies can also be combined and used with technology. The first one to make the link between persuasiveness and technology was Fogg (2002), who invented the word 'captology', which stood for *computers as persuasive technologies*. Especially gamification or gameful elements can help bring about changes in behavior like initiating healthy behaviors and adhering to them (Kappen and Orji, 2017). Because of the ubiquitous nature of mobile technology, gamification can be used in apps for phones, to make persuasion quickly accessible (Aldenaini et al., 2020a; Aldenaini, Oyebode, Orji, and Sampalli, 2020c; Kappen and Orji, 2017). This is another aspect where technology can help us become healthier.

### 2.2.1 Games

The most common strategies used in persuasive games are self-monitoring, simulation, reminder, reward, reduction, praise, suggestion, recognition, competition, cooperation, personalization and social comparison (Oyebode et al., 2021a). Although persuasive games seem effective, it is hard to design persuasive games to be interesting, compelling and effective. Khaled, Barr, Noble, Fischer, and Biddle (2007) used a game about quitting smoking to come up with five general issues with persuasive game design. Those were contextualized in terms of the persuasive technologies of (Fogg, 2002). The five issues discussed can be found in Table 2.5.

| Issues | Explanation |
| --- | --- |
| Managing player attention | Making them look at information during the game without overloading by making a trade-off between tunnelling and suggestion/self-monitoring |
| Balancing "replay-ability" with reality | Reality shows the long-term linkage of the previous choices, while there are often options to redo (parts of) games nowadays |
| Player control vs. system control | While players want to have full control over their character, some unforeseen outcomes relate to the reality which the game is trying to represent |
| Identity issues | How players perceive the player-avatar relationship |
| Target audience | Customization leads to players believing it is for them, while it might not be tailored to them and their interpretations of the game |

**Table 2.5: The five issues when designing a persuasive game by Khaled et al. (2007).**

While research has shown people respond differently to persuasion based on age, it was not yet clear whether age influences the effectiveness of persuasive games in health. Therefore, Mulchandani and Orji (2021) designed and researched a persuasive game (COVID Pacman-C) to promote COVID-19 precautionary measures. The game was perceived as highly persuasive by the different age groups. There were however significant differences in persuasiveness among different age groups with respect to its ability to motivate them (to adopt the precautionary measures). Competition/comparison, simulation and self-monitoring are the most used strategies in games about behavior change. Younger people tend to be more susceptible to competition, comparison and social learning than older people. Also, younger people are more associated with happiness and social well-being, while older people have higher emotional and psychological well-being and satisfaction in life.

### 2.2.2 Social Pressure

People have a cognitive bias that favors short-term rewards over long-term rewards. This results in neglecting beneficial (long-term) behaviors, which is a cycle that is hard to overcome. New technological approaches like persuasive technologies have been introduced to motivate long-term goals and behaviors by providing short-term goals and rewards. Gamification is one of them. Hamari and Koivisto (2015) did research about 'Working out for likes' with the gamification part being a social network service. The idea was to see how the social influence of a user community could influence behavior positively. More specifically, whether social influence could continue and maintain physical exercise in the context of gamification technologies.

The social influence had four factors. (1) A way to signal and distribute subjective norms (2) A way to communicate recognition by sharing (3) Reciprocal benefits like commenting and giving likes to give feedback (4) Network effects like making social communities. It turned out that social influence, positive recognition and reciprocity all had a positive impact on exercising. Thus, network effects increase the influence of social norms. Moreover, the size of one's social network has an additive influence. When the network is bigger, one can have more benefits from it, as the effects are larger. Therefore, the integration of new users into the community is crucial, just as connecting old and new users. The formation of a strong community helps spread the community's norms more, which helps the users in continuing correct exercise behaviors (Hamari and Koivisto, 2015).

**For this research, a social influence is created through technology. Participants are divided into groups and enter a leader-board in an app. The more participants enter per group, the bigger the social influence and pressure could be for the participants to walk more. Hamari and Koivisto (2015) confirmed that social influence could have a positive influence on exercising, so it might also work for this research.**

### 2.2.3 Conforming to Robots

While there has been research on cultural differences in susceptibility to persuasion (Orji et al., 2014b; Orji, Orji, Oyibo, and Ajah, 2018b; Abdullahi, Orji, Rabiu, and Kawu, 2020), there has not yet been much research about this in the area of computer-mediated communication. Cinnirella and Green (2007) found

an interaction between communication type (computer-mediated communication versus face-to-face) and culture (collectivistic versus individualistic).

Only in the face-to-face conditions, the cultural differences were visible. In face-to-face, the difference was that people from collectivistic cultures conformed more than those from individualistic cultures. For computer-mediated communication, the conformity rates were less than with the face-to-face condition, possibly because of the anonymity acquired by not seeing the other person.

Another research on conformity with non-human agents was done by Cormier et al. (2013). They looked at the conformity behavior while varying the human likeness of the robot/computer and the task type in which they had to conform to the robot. When doing the analytical task, many people conformed to the agent, being either the robot, computer or human. Thus, non-human agents could employ the social conformity effect, just as well. For the social task however, there was a difference. The more the agent was human-like, the more people conformed to it. People thus conform to non-human agents, based on which task they have to do (the context), but it also depends on the individual, the group that (s)he is in and other societal factors.

Hertz and Wiese (2018) did similar research with non-human agents and an analytical versus social trial. The outcome was similar, such that people conformed to analytical tasks, but complied more with social trials when the agent was more human-like. Here the idea was that people judge their conformity rates on the agents' expertise and competence on the specific task. This is to say that people judge computers to do calculations correctly, but prefer (to comply with) people when judging about society's principles.

**For this research, conforming to non-humans is relevant because the participants will think that the persuasive messages will come from the system itself, a non-human, rather than from the experimenter. As this research is about the number of steps per day, its base is a numerical issue and not a social issue. Therefore, there is no need for concern about lower conformity rates to the non-human system. Also, the persuasion might be less coming from computer-mediated communication (a text message) rather than from normal face-to-face communication, so that could also influence the results.** *Note: During the experiment, participants sometimes responded to the persuasive messages in such a way that it was clear that they did think that the experimenter made the messages rather than an automatic bot. This was confirmed after the experiment.*

### 2.2.4 Ethics in Persuasive Technologies

New ethical problems arise when persuasion is used in combination with technologies. One such problem has to do with similarity or affiliation. When persuasion is given by someone or something (a computer or an app) similar in personality/affiliation to the one person receiving the persuasion, the effect of persuasion can be enhanced. When computers are perceived as they have psychological concepts (motivate through conveying emotions, negotiating and reaching agreements), they can apply social pressure onto people. Therefore, ethical questions arise when one implements those psychological cues into computers. There are two sides to this story. The first is that it is unethical to deliberately implement computers with psychology concepts, as it is misleading for its users. The other side is that people will always infer some psychology from computers, so it is bad practice to avoid implementing some psychological concepts into computers. Fogg (2002) states that designers must be aware of the ethical implications that follow from implementing psychological cues into computers, but still use some appropriate cues.

An area closely related to that of ethics of persuasion in computers, is the ethics of games and gamification. There are ethical issues about a game itself, playing that game and about the corresponding game designers. Players reflect on and relate to games and their characters. Therefore, the designers must make ethical considerations to make the game ethically possible so that players cannot do unethical things in the game. For more information about ethics in game design, see Sicart (2005).

## 2.3 Personalization of Persuasion profiles

It was shown multiple times that different people react differently to the same persuasive strategies. This is because there is a relation between people's personality traits and their susceptibility to persuasion (Abdullahi et al., 2020; Alqahtani, Meier, and Orji, 2021a; Orji et al., 2017b). Features that motivate one person might demotivate another. Because of this, there have been multiple attempts to categorize the different personality types of people, in order to assign the best performing persuasive strategies

for them. Thus, the traditional one-fits-all approach to persuasion made room for the personalized and tailored approach, to treat each user differently (Abdullahi et al., 2020; Alqahtani et al., 2021a).

Ironically, personalization is actually a general-purpose strategy, while its aim is adapting to specific individuals. A way of using personalization as a general strategy is for example to include usernames in messages, or to consider color preferences based on culture or age (Orji et al., 2014b). There are many dimensions in tailoring persuasion however. The player type or personality type could be taken into account and one could tailor it to an individual or to a group of people. It is not yet clear which tailoring fully maximizes persuasion (Kappen and Orji, 2017).

To come up with personality types to personalize the persuasion, one can either use implicit or explicit profiling. For explicit profiling, one uses prior research by asking questions about one's personality (traits), attitudes, preferences, behaviors and beliefs (Alqahtani et al., 2021a). There are multiple scales made and used for this aspect, like the STPS (see the seventh and last item in the list hereafter) and the 'preference for consistency scale'. However, the person being interviewed knows that their answers will influence the persuasion that is coming after that.

The alternative, implicit profiling, is used so that the person for who the persuasion will be adapted is not aware that the persuasion will be personalized. This is because persuasion is adapted through interactions with the user. To do this, one can look at the actual behavior and responses of that person and other objective measures. The plus side here is that the user is not interrupted or disturbed in their experience, but the downside might be that there are design and ethical challenges involved (Kaptein and Eckles, 2010; Kaptein, Markopoulos, De Ruyter, and Aarts, 2015; Verbeek, 2006b; Fogg, 2002; IJsselsteijn et al., 2006; Berdichevsky and Neuenschwander, 1999), see Subsection 2.1.2.

Alqahtani et al. (2021a) did persuasive research on the Big-Five (Goldberg, 1990) personality traits, see Table 2.6. It appeared that persuasive features could motivate one person while demotivating another. Conscientious people were motivated by relaxation audio, encouragement, suggestions and trusted information but not by customization. People who are open to experience were motivated by relaxation exercises and audios, self-monitoring and social support. Neurotic people were slightly motivated by relaxation exercises and audios, social support, rewards and a clear privacy policy. This shows that personality plays a significant role in persuasive features in technology, interface design and gamification. So, to appeal to a wide audience, apps should have relaxation built in along with social support, self-monitoring and distraction tools and a clear privacy policy.

**Personalizing of persuasion is an important aspect of this research, because oftentimes personalized persuasion has a bigger effect than general persuasion. The aim of this research was to find not only the overall best persuasion strategy for increasing exercise, but also the most effective persuasion per personality profile. However, because of the results of the pre-study, the goal was adjusted to find specific persuasive strategies that have a bigger positive effect on increasing exercise than other persuasive strategies. Because of this, the 'personalization' of this research is shifted from being based upon personality profiles towards the general domain of increasing physical activity.**

| Personality Trait | They have a tendency to. . . |
|---|---|
| Agreeableness | . . . be considerate, cooperative, tolerant, friendly, caring and helpful |
| Extraversion | . . . be outgoing, expressive and seek need opportunities |
| Conscientiousness | . . . to be self-disciplined, plan actively, organized and dependable |
| Neuroticism | . . . be nervous fearful, sensitive, distrustful and emotionally unstable |
| Openness to experiences | . . . be curious, imaginative, hold unconventional values and be creative |

Table 2.6: **Different personality traits identified as the Big-Five by Goldberg (1990).**

### 2.3.1 Personalizing Gamified Systems

Gamification and persuasive games are effective at changing behavior, but there has not been much research yet on personalizing in gamified systems and the long-term effects of exercise promoting systems (Zhao, Arya, Orji, and Chan, 2020; Orji and Moffatt, 2018). Zhao et al. (2020) did an experiment on the

long-term effects of a personalized exercise recommending system. It was shown that both gamification and personalization had positive effects on motivation, engagement and satisfaction with exercise over time. The increase in games that are designed to change behavior and attitudes is however limited: they are often not based on what changes behavior and players are often treated as a monolithic group (Orji et al., 2013a, 2017b).

Orji et al. (2013a) did an experiment that tailored healthy eating persuasion based on seven different models, each model representing one of the gamer types identified by `Brainhex`. There were some differences between the gamer types and their susceptibility to kinds of persuasion. The finding was that persuasive games that are based on behavior theories were more successful and tailored persuasion was also more effective. The best persuasion to use for a broad group of players was self-efficacy, as it had a positive or no effect on all player types. The worst one was perceived barriers; this was seen as a negative or no effect on all player types.

Orji et al. (2014a) continued to research personalization based on gamer types. Again, they found that while some people were motivated by one strategy, others were demotivated by it. So they made suggestions, based on the most commonly used persuasion technologies and seven gamer types identified by `Brainhex`. The persuasion technologies were: simulation, reward, comparison, competition and cooperation. The seven gamer types were: achiever, conqueror, daredevil, mastermind, seeker, socializer and survivor. It was concluded for each gamer type what the best and worst persuasive technologies were. This outcome can be seen in Table 2.7. Orji et al. (2014b) also did this research with ten persuasive technologies and the same seven gamer types. In addition to simulation, reward, comparison, competition and cooperation there were customization, self-monitoring, suggestion, personalization and praise. These ten are the most commonly used persuasive technologies in persuasive health games. There were no sub-gamer types yet involved. Reward and Praise had the least effect overall, but a persuasion that is beneficial to almost all gamer types (except for Achiever and Daredevil) is Competition/Comparison.

| Gamer type | Best PT | Worst PT |
|---|---|---|
| Achiever | Cooperation and Reward *Self-Monitoring/Suggestion* | |
| Conqueror | Competition/Comparison and Simulation *Personalization and Self-Monitoring/Suggestion* | |
| Daredevil | Simulation | Competition/Comparison *Self-Monitoring/Suggestion* |
| Mastermind | Competition/Comparison and Simulation *Customization and Personalization and Self-Monitoring/Suggestion* | |
| Seeker | Competition/Comparison *Customization and Personalization and Praise* | |
| Socializer | Competition/Comparison and Cooperation | *Customization and Praise and Self-Monitoring/Suggestion* |
| Survivor | Competition/Comparison | Cooperation and Reward *Customization and Self-Monitoring/Suggestion* |

**Table 2.7: The best and worst persuasive technology (PT) for each gamer type from `Brainhex` (Nacke et al., 2013) by Orji et al. (2014a). The one in *italics* are added from Orji et al. (2014b)**

Later on, Orji et al. (2017b) also found out that people's personality traits have an effect on the perceived persuasiveness of strategies. Careful people can be motivated by goal setting, simulation, self-monitoring and feedback. People who are open to experience can be demotivated by rewards, competition, comparison and cooperation. Extraversion and agreeableness are traits that respond the most to persuasive strategies, the neuroticism trait is not affected and openness is negatively affected by persuasion strategies. Again, tailored persuasion in game settings turned out to be more effective than general persuasion in games. Personalization, simulation, goal setting, suggestion, self-monitoring and feedback were the most persuasive strategies while punishment and customization were the least persuasive. However, when designing, designers should also include some so-called fun persuasive strategies to make it less

boring and reinforce good behavior more, like reward, competition and comparison.

**As such, gamification is a great way to motivate people to change behavior. For this thesis, the use of gamification lies in the social comparison and competition component of the persuasive messages. By comparing participants through a leader-board, participants might get competitive, which could link to a gamification aspect. However, since participants will not open the step-counter app during the experiment to see the data of other participants, gamification is not a real and integrated fixed part of this research.**

### 2.3.2   Personality Type Categorizations

Here follows a list of some personality type categorizations, which are commonly referred to. Most of them use profiles based on gaming characteristics, but some can generalize out of the scope of (online/offline) gaming. The last two are explained in more detail, as they are the most relevant ones for this research.

1. Myers-Briggs Type Indicator (Myers, 1962): Myers used four axes to categorize personality types into sixteen types. Each axis represents two opposite psychological types. The axes are Introversion/Extroversion, Sensing/Intuition, Thinking/Feeling and Judgement/Perception.

2. Demographic Game Design Model (Bateman and Boon, 2005): They adapted the previous categorization so that it fits with gamer types (conqueror, manager, wanderer and participant) and it can be used beyond multiplayer games.

3. Bartle's four gamer types (Bartle, 1996): This categorizes into achiever, explorer, socializer and killer, but the problem is that those types are not mutually exclusive and it is not empirically based, so they cannot be validated.

4. Yee's player types (Yee, 2005, 2006): Yee looked at multiplayer games online about role-playing. They came up with 5 different reasons to play: achievement, relationship, immersion, escapism and manipulation. These are however not generalizable beyond multiplayer games.

5. Big-Five (Goldberg, 1990): This is about five personality traits: openness, conscientiousness, extraversion, agreeableness and neuroticism. This had inconsistent and low results for predicting gamers' satisfaction.

6. `Brainhex` (Nacke et al., 2013): This personality type categorization is validated much and is based on neurobiology. It identifies seven gamer types of players that are not mutually exclusive, so one can have a dominant gamer type and another gamer sub-type. The types correspond to gameplay elements found in games. Its categories are daredevil, achiever, survivor, mastermind, seeker, socializer and conqueror.

7. Kaptein's Susceptibility to Persuasion Scale (`STPS`) (Kaptein et al., 2009): It is shown that people differ in their motivations and beliefs and therefore differ in responding to certain persuasive strategies. Some people respond better to persuasion from authority figures than other people, while even different people respond worse to authorities than to other persuasive strategies. Using a strategy that does not correspond with the person itself (in fact, it may be counterproductive and lead to adverse reactions), is called a contra-tailored strategy. This scale can be used to tailor persuasive strategies to people correctly. The susceptibility is based on the six principles of compliance of Cialdini and Griskevicius (2010).

The `STPS` is based on a questionnaire of twelve questions and relates to the six principles of compliance of Cialdini (Cialdini and Griskevicius, 2010), see Table 2.8, taken from Kaptein et al. (2009).

The `Brainhex` model is an important well-known model to identify different personality types. This model uses an online questionnaire of about 30 questions to divide people into seven different gamer types (and some demographic questions on top of that). Although it is a rather new model, it is less based on intuition and more based on neurological foundations and has been validated a lot already, so it is reliable (Orji et al., 2014a). In Table 2.9, the seven gamer types of the `Brainhex` model are explained.

**For this research, this is relevant because each participant's personality profile needs to be specified before personalized persuasive strategies can be connected to specific profiles. Only the last three (out of the seven) personality type categorizations are still used by many researchers: the Big-Five, `Brainhex` and the STPS. The Big-Five is used the least of**

| Principle of compliance | Questions in questionnaire |
| --- | --- |
| Reciprocation | 1. When a family member does me a favor, I am very inclined to return this favor.<br>2. I always pay back a favor |
| Scarcity | 1. I believe rare products (scarce) are more valuable than mass products.<br>2. When my favorite shop is about to close, I would visit it since it is my last chance. |
| Authority | 1. I always follow advice from my general practitioner.<br>2. When a professor tells me something I tend to believe it is true. |
| Commitment | 1. Whenever I commit to an appointment I do as I told.<br>2. I try to do everything I have promised to do. |
| Consensus | 1. If someone from my social network notifies me about a good book, I tend to read it.<br>2. When I am in a new situation I look at others to see what I should do. |
| Liking | 1. I accept advice from my social network.<br>2. When I like someone, I am more inclined to believe him or her. |

Table 2.8: The questions for the STPS from Kaptein et al. (2009).

| Gamer types | Explanation |
| --- | --- |
| Achievers | Achievers are goal-oriented and motivated by the reward of achieving long-term goals. Therefore, an achiever often gets satisfaction from completing tasks and collecting things (e.g., points). |
| Conquerors | Conquerors are challenge oriented. They enjoy struggling against impossibly difficult foes before eventually achieving victory and beating other players. |
| Daredevils | Daredevils are excited by the thrill of taking risks and enjoy playing on the edge. They enjoy rushing around at high speeds while still in control. |
| Masterminds | Masterminds enjoy solving puzzles, devising strategies to overcome puzzles that defy several solutions and making efficient decisions. |
| Seekers | Seekers enjoy exploring things and discovering new situations. They are curious, have sustained interest and love sense-simulating activities. |
| Socializers | Socializers enjoy interacting with others. For instance, they like talking, helping and hanging around with people they trust. |
| Survivors | Survivors love the experience associated with terrifying scenes and enjoy the excitement of escaping from terrifying situations. |

Table 2.9: The seven gamer types from Brainhex explained (Nacke et al., 2013; Orji et al., 2014a)

them. For this thesis, the Brainhex categorization was chosen, as it has more questions and therefore more accurate in dividing categories, while it could still be filled in quickly by the participants. Brainhex was also shown to be effective and is already validated. More information about how the Brainhex was used in research can be found in the next subsection. As a backup, since the Brainhex questions tend to be on the gamer side, the STPS was also used to put participants into personality profiles.

### 2.3.3 `Brainhex`

As described, one particular way to personalize persuasion is by using the `Brainhex` profile. To get some more insight on how this personalization technique works, an example research will follow from Orji et al. (2018b). It uses `Brainhex` to find different personalities and then finds congruent persuasive strategies. An example of the `Brainhex` survey can be found in Appendix A.1.

The following research was about combining the seven gamer types of `Brainhex` with the six determinants of health behavior by the Health Belief Model (Rosenstock, 1966) and took place in Africa. The Health Belief Model is one of the oldest and most widely used models for promoting healthy behavior. The seven gamer types were: achiever, conqueror, daredevil, mastermind, seeker, socializer and survivor. The six determinants were: perceived susceptibility, perceived severity, perceived benefit, perceived barrier, cue to action and self-efficacy. It was discovered that Africans have more differences than similarities with previous research on the general population by Orji et al. (2013a). Also, achievers, seekers and survivors were the most motivated gamer types by the determinants of the Health Belief Model overall (Orji et al., 2018b).

Achievers were most motivated by perceived susceptibility, so what they stand to loose while portraying the risks that are associated with unhealthy behaviors. Daredevils however, were more motivated by perceived benefits, so what they stand to gain, suggesting benefits and self-efficacy. Lastly, socializers were more motivated by self-efficacy and severity. To motivate the majority of gamer types for behavior change, the severity of the consequences of unhealthy behavior should be promoted. Another determinant that worked for most gamer types is self-efficacy. So, one should focus on building the confidence of people to promote healthy behaviors. This can be done by implementing ownership, loyalty, pride and repeating simple actions and cascading information. Lastly, urgent optimism could be an effective approach, as long as players can believe that they will succeed. Determinants that were not effective are barriers and susceptibility, as they had negative or no effect on healthy behaviors (Orji et al., 2018b).

Other examples of papers that discuss the `Brainhex` profile and persuasive strategies are that of Orji et al. (2014a,b). The best and worst strategies for each profile can be found in Table 2.7. The strategies included: simulation, reward, comparison, competition, cooperation, praise, customization, self-monitoring, suggestion and personalization.

**`Brainhex` is an important aspect of this research, as it is used as one of the three personality profile categorizations. It was chosen because it is a widely used personality profile that is also validated, such that much previous research could be compared to this research. However, since the questions are rather game-oriented, other personality profile categorizations were also used.**

### 2.3.4 Domains

Besides personalizing persuasion, domains could also matter when providing persuasion to people. There was a paper published (recently, 2022) about research on whether there were within-domain differences and within-personality differences for `STPS` (Ndulue, Oyebode, Iyer, Ganesh, and S. I. Ahmed, 2022). In this paper, they used the Big-Five model and used five different persuasive strategies: competition, cooperation, reward, personalization and normative influence. The domains were healthy eating and quitting smoking. The results showed that indeed persuasion effectively differs between domains and between personalities of participants. More specifically, the reward and competition strategies were effective in promoting healthy eating behaviors. However, those were not effective for conscientious people who tried to stop smoking.

Kaptein, De Ruyter, Markopoulos, and Aarts (2012) and Orji et al. (2017b) also showed that the preference and effectiveness of persuasive features can be domain-dependent. Alqahtani et al. (2021a) showed domain dependency by relating the Big-Five personality traits to seven different domains. The comparative analysis showed that relationships between personality traits and persuasive features varied between domains. This means that a persuasive feature that works for one individual for one task, might not work for a different task in another domain. Although healthy eating, stopping with smoking and physical activity are in the health domain, they differ in their subdomain. Therefore, they can still differ in the effectiveness of the persuasion techniques (Alqahtani et al., 2021a; Ndulue et al., 2022). Alqahtani et al. (2021a) compared twelve persuasive features against seven different domains. Certain features were positively associated with some domains, while negatively or not associated with others. Self-monitoring and personalization were most often positively associated with extraversion and agreeableness, just as

in the mental health domain. The mental health domain was however often negatively associated with openness.

**This domain-dependency of persuasion is relevant for this research, as it is now known that there is no universal best-performing persuasive strategy. Not for domains and neither for individuals. While eating healthy and stopping with smoking are health-related, they are different subdomains and have differences in effectiveness of persuasion techniques. In this research the focus is on finding the best-performing persuasion strategy in the subdomain of exercise in health.**

### 2.3.5 Ethics in Personalized Persuasion

With personalization of persuasion, new ethical concerns are raised regarding tailoring and data privacy (Barton, 2016; ). Tailoring the technology to users (feedback, notifications and progress updates) should not turn into surveillance (Aldenaini et al., 2020a). A trade-off needs to be made between the collection of data and the intrusion of privacy users. Especially for health-related persuasion, the user should be able to control how their data is tracked and used (Vlieghe and De Troyer, 2020; Aldenaini et al., 2020a). Bessenyei, Suruliraj, Bagnell, McGrath, Wozney, Huguet, Elger, Meier, and Orji (2021) suggest that users allow specific terms on data sharing, rather than giving general consent. *Note: It should not be a trick to get more participants, but it should include strong criteria and acceptable risks).* The lack of openness about collecting and using (health) data raises ethical concerns, especially when marketing gains power about healthy living. The privacy, confidentiality and safety of data are not guaranteed in health apps and therefore better and more ethical guidelines are needed (Oyebode, Alqahtani, and Orji, 2020a). This also holds for unwanted advertisements in those apps.

Data linking/sharing and recruiting participants can also raise ethical challenges themselves (Lang, 2011). Data processing and algorithms bring concerns about biases, discrimination, fairness, accountability, transparency and explainability (Shin, 2020; Rieger, Shaheen, Sierra, Theune, and Tintarev, 2022). When designing AIs, fairness needs to be addressed in order to have no unfair biases and discriminatory results (Shin, 2020). A failure to adapt the influence strategies of an app to a person, such that a person increases instead of decreases undesired behavior, is considered unethical as well (Kaptein et al., 2012).

## 2.4 Persuasion in Health

This section discusses previous research on persuasion, in the domain of health. There are many areas of research in the health domain. Some of them are discussed below in the subsections. Examples are health apps, papers about well-being, increasing physical activity, and the use of text-message interventions.

There is growing evidence that persuasion is effective at promoting healthy behavior (Alqahtani et al., 2021a) and wellness (Charity et al., 2020). Examples are eating healthy, stopping with smoking, drinking less alcohol and being more physically active. There was also a paper that concerned personalized persuasion for care for babies of expecting mothers in Nigeria (Charity et al., 2020). Changing behaviors can be hard, so persuasion could be used to make healthy living easier. When better health is achieved, with for example more physical activity, there will be fewer diseases like diabetes, obesity and high blood pressure. That is why persuasion in the health domain is so important. Since health problems are nowadays more lifestyle-related, there has been a shift from treatment-and-description-centered healthcare system towards a patient-centered prevention-focused system. Therefore, the new system emphasizes the promotion of healthy behaviors (Orji et al., 2018b). There are however many ways to implement persuasion in health to motivate people to initiate, adhere and maintain healthy behaviors (Kappen and Orji, 2017). A few examples are reward, competition and comparison, and personalization.

It is known that there are many ways to implement persuasive strategies to motivate people to start and maintain healthy behaviors. To assess the usability of persuasive strategies in health apps, sentiment analysis can be used on reviews of such apps. Oyebode et al. (2020a) used machine learning to classify sentiment analysis on reviews of health apps. The negative themes had these categories: usability issues, content issues, ethical issues, customer support issues and billing issues. The positive themes included: aesthetically pleasing interface, app stability, customizability, high-quality content, content variation/diversity, personalized content, privacy and security and low subscription cost. This led to some design suggestions for persuasive apps. First of all, the app needs to be usable, so it should be nice to look at and simple to use and navigate and the app should have stable performance. Apps should be high in quality, complete, easily accessible and they should (eventually) help improve the health of users. To achieve this, features should be included in the app that supports users in their tasks, while not

interfering with the daily lives or tasks of users. Finally, there should be fair and low prices for the apps while maintaining high privacy and security of data.

### 2.4.1 Changing Behavior

Kelly and Barker (2016) found six reasons why changing behavior is so difficult. One reason is that people think of behaviors as common sense, obvious and self-evident. However, the context in which behavior occurs and the psychological, political, economic and social context matters too. The second problem is that people often think that the key to changing behavior is just getting the message across. It is however, not just about the right packaging of the message. The third one is that it is not just knowledge and information that drive behavior. It is not the added knowledge of medical practitioners that makes you change your behavior.

Another problem is that people think that they only act rational, sensible and logical. When health educators give evidence and tell people what is right for them and what they need to do, people often will not do it (yet). The opposite is also not true, that people always act irrational. When an asthma patient refuses to stop smoking, (s)he might just seem foolish and addicted. However, it might seem more rational when considering their past experiences. Everyone has their own reasons to keep having their 'functional' behavior. Lastly, the final problem with changing behavior is that humans think that it is possible to predict accurately. In the best models, there is still a lot of variance in how individuals will act to specific stimuli in certain situations. This is mainly because of the automatic responses of individuals and the short time frames in which it will happen.

The big problem is that the focus is more on the prediction of behavior than to understand what led to it. Prediction is simpler, but it is less effective and accurate than finding the cause. The idea is that predicting and supporting behavior is not obvious or common sense. It requires a deep understanding of motivations and social and economic pressure. When this is acquired, people will be better supported in their behavior changes. Those individual changes in behavior together will advance the health of the population (Kelly and Barker, 2016).

**Since the goal of this research is to change the (walking) behavior of participants, it is important to note and consider the possible pitfalls of changing behavior. Just getting knowledge about how walking more often is better for one's health would not necessarily increase the step-counts of participants. To get an idea of the motivation acquired from the persuasive messages, participants will be asked to rate the top three and bottom three sentences that they got during the experiment. Participants that agree with a small interview, will also be asked about the influence of the messages they got during the experiment.**

### 2.4.2 Well-being

Abdullahi et al. (2020) did some research on personalized persuasion on Africans regarding their subjective well-being. Based on personality traits, they developed design guidelines for personalizing persuasion. By following the guidelines, health applications will engage users more and enhance the desired behavior change. Extroverted people had higher scores for subjective well-being overall. People with neuroticism however, seemed to be more negatively associated with subjective well-being. Furthermore, extroverted people could benefit from psychological well-being. Conscious people could benefit from psychological, emotional and social well-being. Neurotic people could benefit most from emotional and psychological well-being. People who agree much, benefit the most from social and emotional well-being, satisfaction with life and happiness to promote their subjective well-being and thus their overall health. For people with openness, satisfaction with life is the best way to improve their subjective well-being.

Alhasani, Alkhawaji, and Orji (2021) also pointed out that persuasive technologies are effective and increasingly used in various health systems. They focused on research that reduced stress and therefore anxiety. To do this, they looked at the Persuasive System Design framework of Oinas-Kukkonen (2009), which was built upon the seven strategies of Fogg (2002). This Persuasive System Design framework involves 28 strategies to design content and functionality for persuasive systems. There are four categories, based on which kind of support they provide: primary task, dialogue, system credibility and social support, see Figure 2.2. In the research of Alhasani et al. (2021), they mapped three categories of organization to corresponding persuasive strategies. Using those strategies for interventions could improve the perception control of time of students and therefore reduce stress and anxiety. With a preference for task organizations, the persuasive strategies should be goal setting, reduction and self-monitoring. With

a preference for space organizations, the persuasion should include suggestions. Finally, for the preference for being organized and its benefits, the persuasive strategies should be rewards, social learning and reminders.
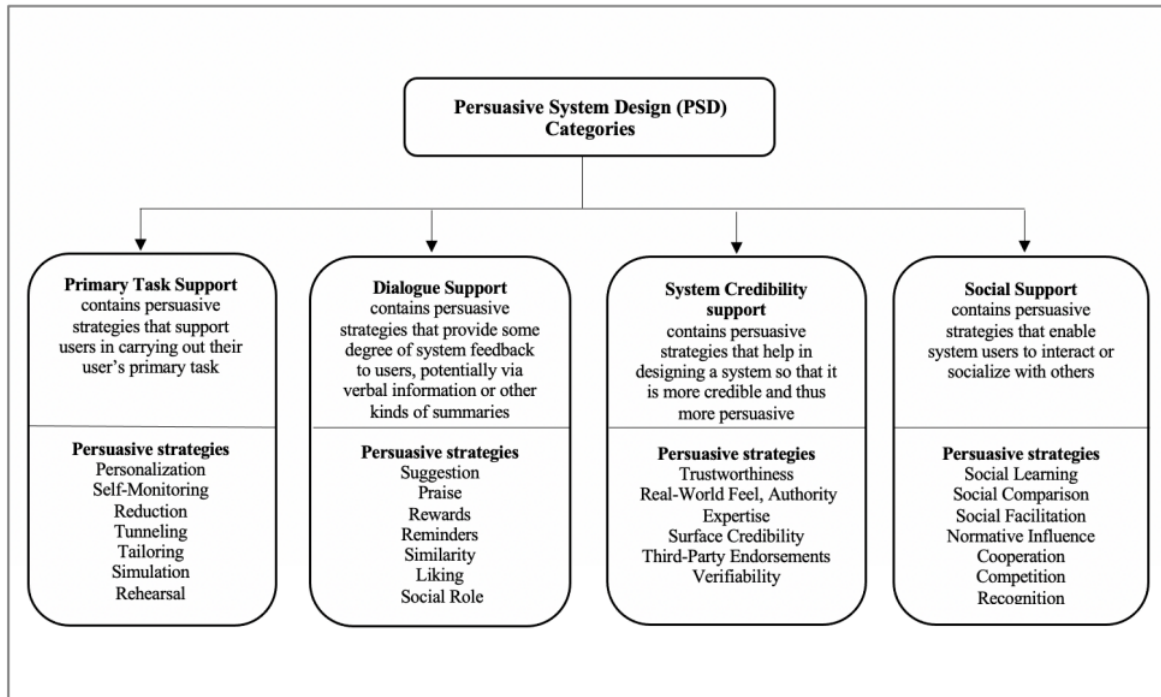


**Figure 2.2: Framework for persuasive system design by Alhasani et al. (2021).**

Alhasani et al. (2020) did a systematic review of stress management apps and the persuasive strategies used in them. Personalization was the most used strategy overall, followed by self-monitoring, simulation and tailoring. Rehearsal was the least used strategy. Technological advancements have made it possible to have mobile health apps. Those apps promote stress regulation and stress management and improve users' well-being over time (Weber, Lorenz, and Hemmings, 2019; Ly, Asplund, and Andersson, 2014). Some even prefer mobile health apps to have face-to-face help from psychiatrists (Apolinário-Hagen, Hennemann, Fritsche, Drüge, and Breil, 2019). The apps can detect stress levels through heart rates or skin conductance, but can also have relaxation features. To achieve that change in behavior, persuasive strategies are needed that target the individual user.

In anxiety, depression and sleep-focused apps, personalization is again the most frequently used persuasive strategy, followed by self-monitoring (Alhasani et al., 2020). However, for mood disorder apps, self-monitoring is more used than personalization. For anger, fear, panic and worry-oriented apps, personalization is used equally much as self-monitoring. Tunnelling and rehearsal are the least implemented strategies in anxiety, depression and stress management apps. Tailoring, simulation and reduction were not used in mental health apps outside stress management apps, so that distinguishes them from the other mental health apps. Overall, personalization, self-monitoring and simulation are the most used persuasive strategies in mental health apps, while rehearsal and reduction the least.

Out of the reviewed apps, personalization was often operationalized as customization, like customizing themes, sounds and timers. This was also often the case when no tailoring strategy was implemented. The other apps personalized the app content by utilizing an assessment approach or letting users choose from a list of topics. So all apps used some kind of personalized persuasion, as it often has more benefits than general strategies for stress management apps (Torbica and Sad, 2004). Self-monitoring and rehearsal were the most employed strategies after personalization. Self-monitoring is also often used in mental health apps (Cohen, Edmunds, Brodman, Benjamin, and Kendall, 2013; Bakker, Kazantzis, Rickwood, and Rickard, 2016) and in health and wellness apps (Orji and Moffatt, 2018). Most reviewed apps used manual tracking however, which can be quite difficult when suffering from serious mental issues (Orji, L., Oyibo, Orji, Blustein, and Shahid, 2018a).

Tunnelling and reduction were rarely employed in stress management apps. This is surprising, as they

can reduce efforts to achieve the right behaviors and thus could increase the likelihood that that behavior will occur. Similarly, rehearsal is used the least while it can encourage and motivate users to perform the target behavior.

Furthermore, Alhasani et al. (2020) found no significant relationship between the number of strategies used in an app and its effectiveness based on its ratings. This means that the number of persuasive strategies employed does not affect the effectiveness of the app, neither positively nor negatively.

**In this research, the well-being of participants is considered in the sense that they can opt out of the experiment, whenever they want to, without giving a reason for it. After the experiment, participants were asked how they felt about the experiment and the messages that they got, to assess their well-being. Not a single participant quit the experiment before it was over. Most participants did like the experiment and its messages.**

### 2.4.3 Self-Monitoring

Self-monitoring and self-evaluation of behavior have both positive and negative implications for motivating health and wellness (Orji et al., 2018a). The reasons why self-monitoring may not be effective include that self-monitoring can be tedious, boring and not fun to do. This, along with finding it oppressive or pushy, can discourage users from using it. The other thing is that self-monitoring could lead to health disorders (eating disorders) which could lead to mental problems (depression). In all, to motivate users, one needs other strategies as well.

Self-monitoring is however mainly a positive aspect in changing behaviors for health and wellness. First of all, it raises people's consciousness and awareness about their health and wellness. This makes them reflect on their own behaviors and take responsibility for it. Furthermore, it reveals their real and concrete behavior (whether it is problematic or not) and it promotes competition between people. It engages users in monitoring their progress and performance. By reflecting on behavior, it can guide people to take control of their health and well-being by making informed decisions (Orji et al., 2018a).

**One of the aspects of the experiment of this research, is self-monitoring. Participants could see their current step-count throughout the day on their phone and/or app. It is important to note that this can have a positive as well as a negative influence on the experiment. The self-monitoring in this experiment is automatic (participants can see it easily) and not required. In addition to that, it takes place in the setting of changing behavior in the health domain. Taken together, this would mean that the self-monitoring aspect would be mainly positive and raise people's consciousness and awareness of their own step-count.**

### 2.4.4 Socially-Oriented Persuasion

The three most common persuasive strategies in health domains are socially-oriented: competition, comparison and cooperation. These are effective as they can use social influence to influence and change behavior, so they are widely used across health domains. Although they all use social influence to motivate behavior, they differ in their operationalization of the principle. Hence, there are different strengths and weaknesses attached to these strategies in motivating or demotivating people. Those need to be considered when employing them. There is still a knowledge gap around how and why socially-oriented strategies influence behavior either positively or negatively. Orji, Oyibo, Lomotey, and Orji (2019) developed design guidelines for operationalizing these three strategies in persuasive health intervention to increase its success. By knowing the strengths and weaknesses of each strategy, the appropriate strategy can be chosen and the manner in which it will be operationalized, so that its strengths are amplified.

Here follows some strengths and weaknesses of each socially-oriented strategy. Competition allows users to compete with one another as motivation to perform the desired behavior. The plus side is that it makes people committed and a sense of accomplishment. However, it could trivialize the importance of that behavior and annoy or discourage people as well. Comparison allows viewing and comparing the performance of behavior, without direct interaction and without winning or losing. It allows subtle and empowering peer pressure and the opportunity of role models. However, it could also encourage body shaming, health disorders and reduce self-esteem, just as competition. Cooperation allows collaboration between people to achieve a shared behavior goal. This is often accompanied by a shared reward of some sort. It provides mutual support, group encouragement and interactions. Unfortunately, it could cause stress, tension and pressure and make people anxious as well (Orji et al., 2019).

**One of the persuasive strategies used in this experiment is competition/comparison, which is socially-oriented persuasion. It is important to know that competition could have a positive as well as a negative influence on participants. While the sentences were more of a comparison between participants than a win-or-lose situation, participants might get competitive because of the messages. What was not mentioned yet however, is the difference between competition between anonymous participants and between friends. Participants might become more competitive when they know the other participants or are friends of theirs.**

### 2.4.5 Social Influences

The health of people is in danger if they do not exercise enough, so Morton, Layton, and Muraven (2019) came up with a way that people learn from behavior of others. This is called social influence. Previous research on social support for exercising was primarily about having an exercise partner and encouragement to exercise (Chogahara, 1999; Sallis, Grossman, Pinski, Patterson, and Nader, 1987). Chogahara (1999) and Morton et al. (2019) found that people who emphasized the importance of exercising regularly and people who gave information on how to do exercises were perceived as helpful. In addition to that, Morton et al. (2019) found that people who set a positive example in exercise and people who pressured others to exercise were perceived as helpful as well.

Research has favored social influence processes that were indirect, subtle and non-conscious. Social psychology has also focused more on goal activation without awareness and automaticity (Caldini and Goldstein, 2004).

The social influence on exercising can come from anyone, from significant others to strangers, but it works best if the social exchanges are about the importance and benefits of exercise. It can be advice or guidance based on their own exercise or bringing a positive attitude towards exercise. Since a person's social world impacts perceptions and consequences (Schunk and Zimmerman, 1996), seeing others exercise regularly could increase the likelihood of that behavior by the person itself (Morton et al., 2019). Social networks are also important in altering social norms for example in the case of obesity and weight loss (Napolitano, Hayes, Bennett, Ives, and Foster, 2013).

It is hard for any individual to stand up against group pressure. As one allows to be influenced, there is a distortion in judgement and a loss of authenticity. Normative social influences when feeling like belonging or wanting to belong to a group can be very strong (Deutsch and Gerard, 1955).

Overall, social modelling with discussing benefits of exercising (information and motivation) and observing other who exercise regularly is necessary to incorporate regular exercise into one's own lifestyle (Lau, Quadrel, and Hartman, 1990; Morton et al., 2019).

**For this research however, text messages are used which are sent to one participant at a time and participants will not look at their step-count app to look for other participants' step-counts. So, they cannot compare or compete against each other based on their step-count in real-time. Cooperation is therefore also not included, as each participant is evaluated on itself without a created social group in which they could cooperate with one another. This is however an interesting follow-up study of this research, whether an active social group or platform would have an increased positive effect on the number of steps of the participants.**

### 2.4.6 Trans-Theoretical Model (TTM)

There was another study based on the TTM of health behavior change (Prochaska, 2013). The theory describes the six stages of change that people go through in their behavior. The first is when people do not intend to change their behavior, as they lack awareness or have given up (precontemplation). The second (contemplation) is when people do want to change in the coming future, but not immediately weighing the pros and cons. The third stage (preparation/determination) is when there are prepared to start changing their behavior, but still have some concerns. The fourth stage (action) is when the behavior change is observable and the threat of regression is the highest. The fifth stage (maintenance) is when the behavior change seems done, but relapses might still happen in certain challenging situations. The last stage (termination) is unfortunately not for everyone, as it is the state where one is confident that (s)he will never relapse again into the old behavior.

The transition from precontemplation to contemplation, is a switched amount of pros and cons. In the early stages the cons seem to be overwhelming, but later on, the pros keep adding up. So, to

make the transition easier, people could list all pros and then be challenged to double that amount (Prochaska, 2013). The most promising outcomes were found with computer-based individualized and interactive interventions and personalized counsellors. It was also shown that people who work on one health behavior change can be as effective as people who work on multiple health changes at the same time. This is great news for people with multiple high health risks.

Grimes, Kantroo, and Grinter (2010) and Orji, Vassileva, and Mandryk (2013b) also chose to use this TTM of behavior theory to frame their results. It is the most frequently used model for health-related behavior interventions, as it is useful for characterizing the processes of change that make the transition to healthier lifestyles easier.

Oyebode, Ndulue, Mulchandani, Adib, Alhasani, and Orji (2021b) looked at how persuasion could be tailored to people in each stage of change. It turned out that the different stages of change did have an impact on how individuals responded to persuasive strategies. The persuasive strategies used were self-monitoring, suggestion, reminder, social role and praise. People in the precontemplation stage benefited from the self-monitoring strategy, as it raised their self-awareness and consciousness, which is the core thing in that stage of change. People in the action stage can be motivated more by suggestions, reminders and social roles. They allow people to learn healthy substitute behaviors, give cues and encouragement and let people engage in encouraging relationships. This is in line with the processes of counter-conditioning, stimulus control and helping relationships. These results show that tailoring to people's stages of change is useful. By being informed of these differences, it will not happen that persuasion is targeted at a wrong stage.

**For this research, the six stages of change are used to divide participants evenly among the three experimental groups. It is also used as a comparison before and after the experiment, whether participants increased or decreased the stage of change they were in. The hope is that (at least) the stage of maintenance is achieved by many of the participants when the persuasive text messages stop and the experiment ends. This would mean that there will be long-term effects of the persuasive messages to exercise more.**

### 2.4.7 Health Apps

Here will follow some examples of health-related apps, as this research is also about health and uses an app to track steps.

Oyebode et al. (2021b) used a focus group to select persuasive features that came across as the most motivating and influencing ones. The app was therefore more user-centered than designer-driven. This was done for an app called SleepFit. Users could track their own sleep patterns and factors that negatively affected their sleep and receive personalized interventions to improve their sleep. SleepFit uses social influence through a community feature. This is similar to normative peer pressure influence, which is one of the five most strategies that is used in health apps (Oyebode et al., 2021b).

Oyebode, Ndulue, Alhasani, and Orji (2020d) reviewed 80 health apps, in order to find the most employed persuasive strategies in health apps. Four health domains were used: (1) physical activity and fitness, (2) diets, (3) emotional and mental health and (4) health assessment and healthcare. Personalization, surface-credibility, trustworthiness and self-monitoring were the most used strategies, respectively. Commitment/consistency, reduction, reminders, real-world feel, expertise, liking, tunnelling and suggestion were also often used. The least dominant strategies used were: rehearsal, praise, third-party endorsements and cooperation. The conclusion was that surprisingly, using more strategies is not necessarily better. When fewer strategies were used, the persuasion was still effective and less complex (Orji, Mandryk, and Vassileva, 2017a; Oyebode et al., 2020d).

In a later study, text mining with natural language processing was used to get persuasive strategies out of health apps (Oyebode and Orji, 2020a). There, self-monitoring was the most employed persuasive strategy, followed by personalization and tailoring, simulation and rehearsal. Reduction and especially tunnelling were used the least.

Alqahtani, Khalifah, Oyebode, and Orji (2019) also looked at the most used persuasion in mobile apps. There were between one and ten strategies used per app. Self-monitoring, personalization and reminders were the most frequently used persuasive strategies. When coming up with design suggestions for mental health apps, Alqahtani, Winn, and Orji (2021b) found that self-tracking was the most implemented persuasive strategy was self-tracking.

Even though there are many mental health apps around and they can be easily downloaded on any smartphone, there is no guarantee that they will be used to improve mental illnesses. Many people do

not use the app anymore after a few times. The problem is supposed to be usability, so Alqahtani and Orji (2019) researched usability issues with mental health apps. Usability issues could be classified into six different categories: (1) bugs, (2) poor user interface design, (3) data loss, (4) battery and memory usage issues, (5) lack of guidance and explanation and (6) internet connectivity issues. Even though apps designed for mental health issues should be intuitive and simple to use, they often have difficult instruction, navigation and no good orientation. This could often be due to focusing on the functionalities rather than usability of the app. Therefore, mental health apps should have a usability evaluation before putting them online. While the core of mental health apps is to track data to realize and understand patterns of behavior, some apps suffer from data loss, by which they lose the trust of their users. The suggestion is to include an easy-to-use guide for the app, update the app regularly to fix issues, include some adaptations to the app (font size), an option to back up data safely and have an offline mode, so that can still access it in times of need without having an internet connection.

**For this research, this is relevant because, in order to be successful, it is necessary to look at what research already has been done before. Personalization and self-monitoring are often used in health apps and that also holds for this research. While participants can monitor their step-counts, they receive personalized persuasive messages. Since rehearsal, praise, cooperation, reduction and tunnelling were not used so often, there is less research to compare it to, when this research would use those tactics. Therefore, and because of irrelevant strategies, some of those tactics (reduction , tunnelling and cooperation) are not implemented in this research. Praise and rehearsal (in the form of reminders) however are still being used, as to test as many strategies as possible in this research.**

### 2.4.8 Physical Activity Versus Sedentary Behavior

Physical inactivity and sedentary behavior are risk factors for people's health. It could cause cancer, diabetes, heart disease or even a stroke (Oyebode et al., 2020c). In an attempt to increase people's health, Oyebode et al. (2020c) build a game that employs twelve commonly used persuasive strategies in health games. These are personalization, self-monitoring, reduction, simulation (all used for primary support), praise, suggestion, reminders, rewards (all for dialogue support), competition, cooperation, recognition and social comparison (all for social support). There will also come an updated version of the game with suggestions, reminders, reduction, social comparison and cooperation.

Oyebode also used gamification through an app to get people to exercise more (Oyebode et al., 2020c; Oyebode and Orji, 2020b; Oyebode et al., 2021a). The goal was to increase physical activity and thereby decrease sedentary behavior. The means was a metaphorical tree that represented participants' health. This concept draws from the emotional connection of people with nature, where we nurture plants and trees. When a participant did not have much physical activity, the tree would look like it was withering and dying, while when a participant did have much physical activity, the tree would start flourishing growing beautiful leaves. This allowed for visualizing the progress of people in real-time. Strategies were tailored to the personalities of the players and that personalization was effective in increasing physical activity (Oyebode et al., 2020c). Goal setting but also the option to compete with one another on the leaderboard reinforced and encouraged the right behaviors. It also made it easier to do the behavior and made people more committed and focused on healthy behavior.

Similar research was done by Lin, Mamykina, Lindtner, Delajoux, and Strub (2006). To promote physical activity, a social computer game called Fish'n'Steps was developed. The daily step-count was presented as the activity and growth of a fish in a fish tank. By allowing to see another person's fish in their own tank, cooperation and competition were also possible. This app increased the activity levels of participants, even after the excitement of a new app subsided after a couple of weeks.

Orji and Moffatt (2018) did research on trends of persuasive technologies in health and wellness. The persuasive technologies were effective as 92% of the reviewed studies did have a positive outcome in health and wellness. The majority focused on behavior and/or attitude change, as conceptualized by Fogg (2002), but there were other behavior-related and psychological outcomes as well. An example is reinforcing and strengthening existing behavior (increase the step-count), rather than changing behavior (quit smoking). This shows that persuasion can do much more than stated in its initial conceptualization.

However, there are differences in effectiveness of persuasion in health and wellness. When it is targeted at smoking, substance abuse and disease, it is less effective at changing behavior. When it is targeted at dental health, sexual behaviors, eating and physical activity however, persuasion is more successful in promoting behavioral changes. This difference in effectiveness could also be due to the operationalization

and appropriateness of the persuasive strategies. The length and depth of evaluation could also make have an influence (Orji and Moffatt, 2018).

Persuasive techniques could be implemented differently across technological platforms. Aldenaini, Orji, and Sampalli (2020b) compared the personalization technique in different persuasive technologies on their effectiveness, motivation, educational content, recommendations, personal goals and feedback. Persuasion for promoting physical activity and decreasing sedentary activity was more effective and promising when it was personalized.

Aldenaini et al. (2020a) also researched the most effective approaches for persuasion for physical activity. 79% of those reviewed studies have successful outcomes. Having persuasion on a mobile phone along with activity trackers and/or sensors was the most effective approach. The second most effective approach was with just the mobile phone and its own sensors. Combining mobile phones with games was ranked third. The top ten most effective strategies that can be used to persuade from most effective to least effective were: tracking/self-monitoring, personalization, goal setting, reminders, other social support strategies, praise and reduction, social competition, suggestion, social comparison and tunnelling and social cooperation.

In another physical activity research, Aldenaini et al. (2020c) focused on the age differences in previous research. It became clear that (young) adults were most targeted, as they would have more mobile and handheld devices, sensors and activity trackers than older people or young people. Mobile phones are always with you and thus a clever way to track your activity with all of their integrated sensors (Almutari and Orji, 2019; Aldenaini et al., 2020a; Oyebode et al., 2020c, 2021a). The most effective strategies here (in decreasing order) were: tracking/self-monitoring, reminders, personalization, goal setting, rewards and a set of social support strategies.

To increase physical activity, it is helpful to use trackers to track physical activity. Mobile phones, which use persuasion for physical activity, use suggestions in a timely manner to motivate people to be more physically active. Another often used and effective way to increase physical activity is socially-oriented persuasion strategies. Competition, social comparison and cooperation were effective at promoting physical activity in 97% of the reviewed studies (Almutari and Orji, 2019).

Almutari and Orji (2021) suggest that we could use self-monitoring, loss-framed appeal and simulation as strategies to let people show the risks of their physical inactivity. This is important, because it might not affect you alone, but also your family as you have close ties to them. While knowing the benefit of healthy behaviors, people often still need a nudge to take appropriate actions.

Mobile apps that encourage physical activity were often based on designer's intuition rather than using knowledge of the effectiveness of persuasive strategies. That is why Oyebode et al. (2021a) did a study with user-centered methodology. First participants assessed the perceived effectiveness of persuasive strategies. The best strategies were then employed inside a game. After that, a field study to evaluate the game. That game did improve participants' physical activity and was considered fun, engaging and easy to use.

**For this research, improving health through persuasion is relevant because it is about increasing physical behavior and decreasing sedentary behavior (walking more steps per day). Personalizing persuasion seems to be effective here as well, so that is evaluated further in this research. Self-monitoring, tracking and personalization are used, as is often the case with persuasion on physical activity. The socially-oriented persuasion strategies are not used although they seem to work. This is because the privacy of each participant is taken into account. While there is a kind of leaderboard per day to see for all participants, it is not encouraged to look it up (as it is also not up-to-date most of the time), and it is anonymous (participants either use only their first name or their participant ID). There are some messages with comparisons however, as to say that a participant was in the top 20% of the day before.**

### 2.4.9 Text Messages

For this research, text messages are sent to the participants with the persuasive message to walk more steps. There is already evidence that text message interventions can help in creating short-term changes in behavior. The details of the intervention and the long-term effects are still under investigation (Orji et al., 2014b; Orji and Moffatt, 2018).

While personalization is officially a general-purpose strategy, it can be used in a way such that text messages can be adapted to individuals and groups. It could use individual names or account for certain

group preferences like preferred color palettes (Orji et al., 2014b).

Armanasco, Miller, Fjeldsoe, and Marshall (2017) found preliminary evidence that positive change in preventive health behaviors can be maintained after interventions have stopped. The theoretical basis and length of the intervention can increase that effect on health behaviors.

It was also shown that emotions, that come with a message, can have an impact. Emotion biases one's expectation of events matching that emotion and expectation facilitates positive attitudes toward persuasive strategies (DeSteno, Petty, Rucker, Wegener, and Braverman, 2004; IJsselsteijn et al., 2006).

Kaptein et al. (2012), after finding out the effect of contra-tailored strategies, concluded that persuasive text messages can be effective, but depend on the right strategy for the right participant. Hirsh, Kang, and Bodenhausen (2012) also examined tailoring persuasive messages to personalities, but to the Big-Five Model (also known as Five-Factor-Model by Goldberg (1990)).

By adapting the content of texts to a person's characteristics, the content will be more relevant and less redundant. Examples could be socio-demographics or an intention to change. By adding personal information to a text, the persuasion of the text is increased. Dijkstra (2006) and IJsselsteijn et al. (2006) found that double the number of participants would stop smoking when their personal information was incorporated in the text messages.

Another research on persuasion also used text messages as intervention by Smids (2017). Personalization of text messages was done based on personality traits. Personality profiles were based upon the STPS of Kaptein et al. (2012) to six of the principles of Cialdini (2006): reciprocity, authority, scarcity, consensus (social proof), liking and commitment/consistency of behavior. The goal was to increase the water intake of participants. This research did not have the expected results, possibly because of the season, the name of the study, the questionnaire used, or because the control condition did also have explicit goals (Smids, 2017).

**This current research study is based on this last research by Smids (2017). For this research, the goal is to increase the step-count per participant but the company for which the research is done, is the same. Also, a different questionnaire is used, as this research focuses on Brainhex profiling (Nacke et al., 2013) rather than solely on the STPS (Kaptein et al., 2012). The season should have a decreased effect on this study (compared to the previous one). While water intake (of the previous study) depends on the weather (more water intake on hot days in the summer), walking (this study) can be done both inside and outside the house independent of the current season. However, when it is spring or summer, people tend to go on a walk a bit more. The idea is that participants are encouraged to walk more than they would normally do. Another difference is that there are no specific goals for participants to reach during the experiment including the control condition. This is to prevent influences on step-count from just 'having a goal'. The monitoring effect would probably still have an influence, but participants were informed that the text messages were sent at random by a bot, rather than by the experimenter, to decrease that effect slightly. *Note: Participants did act like the persuasive messages were from the experimenter itself in responding to them.***

### 2.4.10 Timing of Messages

There has not been research yet about the timing of messages and/or reminders. There are some websites that claim that between 8-12 AM is the best time to send messages (although they differ in exactness) and some claim that around 1 PM is also a suitable time (as it is lunchtime and people will check their mail). The idea is that one should send messages at the beginning of the working day, so people can apply the information from within the messages on that same day. There are also some papers about the timing of sending an email and those conclusions are similar.

For this research, we had mostly students. Students generally have classes from 9 AM onwards, so the aim is to send the messages before 9 AM. Some participants might not wake up around that time, but later on the day. Those participants would still see that they have received a message when they wake up and start their day. By sending the messages early in the morning, the messages should be encouraging for the day to come.

In the blog on Moosend (moosend.com/blog/best-time-to-send-an-email), they suggest that marketing emails should be sent between 8 and 9 AM. This is due to working hours. At the start of the day, people tend to open and answer their emails more, than when they are working. At nighttime, people will not open and answer their emails, but they might do that when they wake up and see an email. This

is also the theory of the blog on hunter.io/blog/best-time-to-send-an-email/#what-is-the-best-time-to-send-an-email. According to a study from Yesware university, emails will get higher reply rates around 10-11 AM and 1 PM (because of the lunch break time) emailanalytics.com/best-time-to-send-an-email yesware.com/blog/best-time-to-send-email/. Lastly, a blog on messente.com/blog/most-recent/best-time-to-send-text-messages suggests that you adjust the timing of messages to their content and the receivers. When the message is valuable for the start of the day, sent them between 8 and 12 AM. If messages are more for the end of a working day, sent it between 5 and 9 PM. The different time zones should be accounted for as well.

In the study of Czerwinski, Cutrell, and Horvitz (2000), the disruptiveness of messages was researched. Relevant messages were considered less disruptive than irrelevant messages, but otherwise, the timing is also relevant. When messages were queued or when messages were sent early in the task (before the user was engaged in that task), there was less disruption as well. So, messages should be early in the process rather than during evaluation and planning of execution phases (Czerwinski et al., 2000; Miyata and Norman, 1986). This could be a sign that messages in the early morning are effective as well, since people are not yet in the middle of another process.

**For this experiment, it was decided that participants would get the message early in the morning. That way, it could remember them to make enough steps that day while they still had time to plan their day. When the persuasive messages are sent during the day, participants might not have time to read them during busy hours and when a message is sent in the evening, they do not have time anymore to make more steps that day and they might forget about the message again after a good night sleep. That is why the messages are sent in the morning. The time was chosen to be 7 AM. This would ensure that most participants got (or already had) the message while waking up, to not intervene with other tasks and to make sure that participants had the whole day to plan for more steps on that day.**

## 2.5 Measuring Attitude against Walking

Both Lin et al. (2006); Orji et al. (2013b) used Prochaska's TTM (Grimley, Prochaska, Velicer, Vlais, and DiClemente, 1994) to evaluate their studies. This was done to see if there was a change in attitude toward healthy eating before and after the study. The survey before the experiment and after the experiment of Orji et al. (2013b) had questions about their thoughts about healthy diets, health concerns, nutrition knowledge and their attitude towards healthy eating. By comparing those surveys, it became evident that participants had more knowledge about healthy eating and were able to use that outside of the game in real life and changed their attitude towards healthy eating as a consequence. Lin et al. (2006) used five distinct levels to divide participants into the six stages of change of the TTM (numbers four and five were taken together). The first level includes participants that had no established exercise routines and had no intention to take such actions in the next six months or in the near future. The second level included participants that had no routine yet, but were developing plans to do so in the next six months. The third level included participants that had intentions to take action in the next 30 days and did some initial steps in that direction already (like joining gyms). The level that combines stages four and five included participants that had an established exercise routine (in the past) but had not made a habit out of it. The final and sixth level included participants that followed their exercise routines as strong habits and did not fear or think about relapsing into behavior with less exercise.

Orji et al. (2017a) used a different model to assess the attitude of participants toward healthy eating, namely the Theory of Planned Behavior by Ajzen (2006). Three scales were used: seven items to assess the attitude towards healthy eating, seven items for the intention of eating healthy and seven items about self-efficacy towards healthy eating, see Figure 2.3 and Figure 2.4. Participants also answered how likely it was for the ten persuasive strategies to influence their eating decisions. The persuasive strategies were depicted through storyboards, which were verified and easy to understand. The feedback of participants was gathered through a validated scale for measuring perceived persuasiveness (Drozd, Lehto, and Oinas-Kukkonen, 2012). In the study of Drozd et al. (2012), the persuasive system design categories in the model of Oinas-Kukkonen (2009) had a significant impact on perceived persuasiveness.

**Aside from actual differences in step-counts of the participants, there should also be a difference measured in people's psychology. The feelings against walking more cannot be measured by absolute differences in step-count alone. For this research, a combination of Prochaska's TTM (six stages of change) and the Ajzen Theory of Planned Behavior model**

## Self-efficacy

**32. Please rate your level of agreement with the following statements in a 7-point scale. Choosing 1-Strongly Disagree to 7-Strongly Agree.**

|  | 1-Strongly disagree |  |  |  |  |  | 7-Strongly agree |
| --- | --- | --- | --- | --- | --- | --- | --- |
| If I want, I could easily eat healthily within the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I have control over whether or not I eat healthily | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Whether or not I eat healthy diets in the next week is entirely up to me | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I believe I have the ability to eat healthy diet next week | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I am confident that I could eat healthily within the next two weeks if I want | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

## Attitude

| Eating healthy food in the next two weeks would be: |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Unimportant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Important |
| Useless | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Useful |
| Worthless | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Valuable |
| Unenjoyable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Enjoyable |
| Harmful | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Beneficial |
| Unpleasant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Pleasant |
| Bad | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Good |

Figure 2.3: Self-efficacy and attitude questionnaire of the Ajzen Theory of Planned Behavior model by Orji et al. (2017a)

**Intention to Eat Healthy**

| | 1-Extremely unlikely | | | | | | 7-Extremely likely |
|---|---|---|---|---|---|---|---|
| I intend to avoid eating junk and fast food during the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I intend to eat more vegetable during the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I will try to eat only a healthy amount of fat during the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I intend to eat more fruits during the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I intend to eat healthily during the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I will try to eat healthily during the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| I plan to eat healthily during the next two weeks | 1 | 2 | 3 | 4 | 5 | 6 | 7 |

**Figure 2.4: Intention questionnaire of the Ajzen Theory of Planned Behavior model by Orji et al. (2017a)**

(intention, attitude and self-efficacy) is used (edited from healthy eating to walking more steps). This is done to measure differences (if any) in attitude towards walking and taking enough steps per day, aside from numerical and percentile differences in step-counts (if any).

# 3 Methodology

In this section, every aspect of leading up to the experiment is described. First, the focus group to come up with the right messages is discussed. Then the pre-study of matching strategies to personality profiles is discussed, followed by the set-up of the real experiment and the post-survey. The actual implementation of each of these can be found in Section 4.

## 3.1 Focus Group

After choosing which persuasion techniques are applicable and are worth investigating for this specific research, specific messages should be generated that belong to the different persuasion techniques. To make sure that those sentences are persuasive in an intended way and to come up with more reasonable messages to send to participants during the experiment, a focus group was held.

The focus group should be similar to the participants in the actual experiment (to ensure similar reactions to the persuasive sentences). Using few people for this focus group would result in too little input, but using many people could distort the flow of talking and listening. The aim was to come up with as many sentences as possible that corresponded to the persuasive strategies to let participants set more steps per day.

Before diving into the nine persuasive strategies, all participants should give consent and be told about the aim of the research in general and the aim of the focus group specifically. After that, an example should be given so that participants can grasp the subject better. Then the strategies can be discussed. For each of the strategies, some sentences (three up to six) should be prepared beforehand to give some extra guidance for the participants of the focus group. This is also to make participants understand the strategies better.

After discussing the strategies, participants of the focus group need to fill in a survey, which is discussed in Subsection 4.2. Participants should be thanked properly for their participation and time before leaving the meeting. The actual implementation of the focus group can be found in Subsection 4.1.

## 3.2 Pre-study

A pre-study was done to check whether the persuasive messages and strategies from the focus group have a connection with some of the persuasion profiles (`Brainhex`, `STPS` and/or `TTM`).

To get as many participants as possible, while also getting a variety of people, a suitable platform should be used (possibly online). The survey would have questions to place each participant in certain personality profiles (`Brainhex`, `STPS` and/or `TTM`), along with some demographic questions. For each persuasive strategy, there should be three sentences (incorporating that strategy) that participants have to rate. This is needed to ensure that participants unknowingly rate the strategy behind the sentence, rather than the sentence in itself, which might not have been as representative of the strategy.

With the data of personality profiles and their persuasion scores on the sentences, the pre-study can help deduce which `Brainhex` profiles are more susceptible to which persuasive strategies. This can also be compared to the outcomes of previous research on the link between `Brainhex` profiles and persuasive strategies (Subsection 2.3.3 and Table 2.7). This is also to confirm that the sentences (that are used as persuasive strategies in the actual experiment) are correctly correlated to their respective persuasive strategies. The results of the pre-study can then be used to create different experimental groups for the experiment. The link to the survey for the pre-study and the implementation of the pre-study can be found in Subsection 4.2.

**The pre-study is used to see which persuasive strategies have the most influence on walking more steps in general as well as for each personality profile individually. This is however only a partial answer to the research questions, as there could still be a difference between perceived persuasion and the effects of actual persuasion in the experiment setting.**

## 3.3 Study Design

For this research, a `between-subjects design` was used, as each participant enters the experiment once, while being in different experimental groups. Since the pre-study showed that there was no personalization possible, while there was a clear top and bottom three out of the nine persuasion strategies, participants were divided into three groups. This division of participants into three groups was one of the `independent variables` for this research. Other `independent variables` include but do not exhaust the susceptibility of participants to persuasion in general and participants' susceptibility to the specific persuasion strategies.

To account for some `confounding variables`, all participants had to fill in a survey before being placed in one of the three groups. Participants were divided into three experimental groups based on their answers to the survey. This way, it was possible to divide males, females and others evenly among the three groups. Not only the genders were taken into account, but also the age (young, middle or old), the main `Brainhex` profile and the place in the `TTM` were considered when dividing participants evenly among the three groups.

Other `confounding variables` could not be accounted for. These include how much participants walk in general on their own. Perhaps they do not have spare time to walk more, or have sustained an injury or gotten ill so that they cannot walk more steps. Also, the general daily activities of participants are considered `confounding variables`. For example, people who have work that includes sitting behind a desk will walk less than people who work in hospitals or work while standing. Aside from injuries and short illnesses, this will not have a great impact on the research however. This is because participants are compared against their-selves, rather than to each other in their step-counts. Another example of a `confounding variable` is the amount of holidays participants have during the experiment. While holidays might interfere with the step-counts and the intervention of persuasive messages, they could not be avoided as the holidays were already planned and participants have the right to do so.

The main point is to see whether people can improve their own step-counts during this experiment. Therefore, a benchmark week is included in the experiment, where no persuasive messages are sent yet. That first week is to measure and set a benchmark on step-counts per participant. The `dependent variable` is then the number of steps set per day per participant compared to their benchmark week, which is a discrete measure per day. The comparisons also involve differences (increase/decrease) in

percentages. This is done because people could have big relative differences in step-counts between weeks, while not being a big absolute difference. People that normally walk 5,000 steps a day and later walk 6,000 steps a day have a more significant increase than someone that walked 10,000 and later on walked 11,000 steps.

The controlled variables include the messages and the experimental groups. The plain-text messages were randomly selected per experimental group from a group of messages that were verified through a focus group. The messages were all sent via WhatsApp and were sent at the same time (7 AM). The three different experimental groups got different persuasive messages per day. Participants were evenly distributed among these three experimental groups considering age, gender, `Brainhex` profiles and `TTM` profiles. The first group got the top three persuasive strategies (according to the pre-study), the third group got the bottom three strategies (according to the pre-study), while the second group got the middle three strategies. This way, the groups also have logical numbers and colors. Group 1 is expected to come out on top, while group 3 is expected to come in third place. Green is often associated with good things (e.g. balance, abundance and green traffic lights), while red is often associated with bad things, when compared to green (e.g. anger, danger and red traffic lights).

For the significance tests, there are two ways of looking at the data. One is comparing the groups against each other with `ANOVA`'s or `linear models`, to see whether there was a difference between the conditions of the experimental groups. The other is comparing participants to themselves during the experiment. All this data is numerical data, as they are the step-counts recorded of each participant during the experiment. Participants can be compared between their first benchmark week, the persuasive messages weeks, and the final fifth week without messages with several `paired T-tests`. The standard value of $\alpha = 0.05$ was used.

## 3.4 Survey

For the real experiment, participants also needed to answer questions about their attitude towards walking and exercising. This needed to be filled out after the experiment as well, to notice if there are any differences in attitude towards walking (in case the step-counts do not differ, but the intention to go walking has grown). The questionnaires of two papers are combined for this (Lin et al., 2006; Orji et al., 2017a). This is to ensure a thorough study on whether there were any changes in actual walking or thinking about walking. The survey was held online, to avoid waste of paper and setting up a meeting with every participant. Qualtrics (survey.uu.nl) was used for this, as it is a safe and private way to store data of participants, and is often used by the University of Utrecht for research projects. The links to the surveys before and after the experiment can be found in Subsection 4.3.

**In the survey, personality profiles are calculated for each participant as well as some attitudes against walking. To answer the research questions, not only the absolute step-counts of each participant are taken into account, but also these attitudes. More specifically, the place in the `TTM`, the attitude against walking, the intention to walk more and self-efficacy of participants are compared before and after the experiment. This is to measure the psychological influence and likeability of the persuasive strategies. In the survey, participants were also asked to enter the most and least liked/persuasive messages that they for during the experiment, for the same reason.**

## 3.5 Experiment

For the actual experiment, the people from the focus group are not allowed to participate again, as they would have inside information about the persuasive sentences and could recognize the sentences from before. Having seen the sentences before could influence their behavior.

A step-counter should be used to track the steps of participants. So, in the first week, participants were instructed to install an activity-tracker app. Most of its notifications had to be deselected to control for confounding messages from the app itself. Participants also filled out a survey before participating in the experiment (see Subsection 3.4). In that survey, their `Brainhex` profile was deduced as well as their `STPS` and their place in the `TTM`, as that was used later on in for the analysis. No messages were sent in this first week, but the number of steps per participant per day was noted down as a benchmark.

Reliable results will come when the longitude of the study is long enough, but the available time for this research was restricted. (Aldenaini et al., 2020a) found that 61% of the studies on persuasive technology that they reviewed, took between one week and three months. Due to the limited time and

the general duration of other similar longitude studies, the idea of one benchmark week, three weeks of persuasive messages and one week of the after-effect was executed.

The measure in weeks is necessary (as opposed to measuring days) because people behave and walk differently during working days and days off. People might also work only on certain/few days per week. Comparing days (instead of weeks), could influence the results while it has nothing to do with the impact of the persuasive messages. Walking during work also differs per person, as one could have a job sitting behind a desk versus one that is a running coach or athlete. This is another reason why participants are not compared to one another, as it would be unfair. Participants are compared to themselves only, as to see whether they improved their own walking behavior, by setting more steps per day than in their benchmark week at the beginning.

In week two to four, participants get daily persuasive messages via a chosen platform. Participants are told that these are automatic messages from a bot, to lessen the surveillance aspect, as opposed to the experimenter personally typing the messages specifically for the participants. To send all persuasive messages to participants in a timely manner, some message scheduler could be used.

The messages were sent early in the morning, so participants could see the persuasive message at the start of their day. This is based on some research and some intuition. Participants will have all day to pick a time to take some extra steps if they receive the message in the early morning. When this message is done in the evening, participants will not have time anymore to take more steps and they might forget about that message the next morning.

After the fourth week, the persuasive messages stopped, but the number of steps per participant per day was still recorded. This was to see whether persuasion had a (long-term) effect even after the persuasion itself has stopped. To analyze the data gathered, several things could be important: Do people walk more during the week or at the weekend, what was the longest streak of walking more than (for example) 5,000 steps, was there a decline or improvement in the number of steps, are there differences in steps taken by age or gender, did people walk more steps in week two to four than in week one?

Further implementation of the experiment and its surveys can be found in Subsection 4.3.

**Relating back to the research questions and hypotheses, the experiment itself allows comparing the absolute and relative differences in step-counts of the participants individually and between groups. This implicitly compares the persuasiveness of the different strategies per group. When one group has higher step-counts or increases it's step-count more than other groups, the persuasive strategies in that group might be more persuasive than those in other groups. The last research question can be answered by comparing the results of the pre-study with the results of the experiment.**

## 3.6 Possible Strategies for this research

Here each strategy found in the previous literature will be divided into three parts; strategies that are used in the messages for this experiment, strategies that might play a role, and strategies that are not used. The main inspiration for selecting the strategies came from Table 2.7, since the main aim was to look at `Brainhex` profiles.

The following strategies **do** play a role in this experiment:

- Authority/obedience: The authority strategy could be an ethical issue, when stating that a world health organization suggests a certain number of steps per day. However, since the ethical issues were not addressed in the focus group, the authority strategy was used in the experiment. An example of such a sentence would be that some (specific) health researchers state that you should walk more than a certain number of steps per day.

- Informational influence/elaboration likelihood model/classical rhetoric: This is similar to the authority strategy, but it excludes the authority figure(s). There are more general statements like: 'It is good for health that you walk a lot every day, rather than sitting around.'

- Social learning/social comparison/competition/liking/similarity/recognition/gamification: Comparison and competition were taken together as one strategy, just as in the original papers by Orji et al. (2014a,b). It also includes the liking/similarity principle, as the sentences are about other participants in the same experiment and in the same group. This means that those participants 'are on the same boat' as themselves and that might create recognition for those other participants. This merged strategy compares step-counts of participants within one group to create a sentence that

invites the participants to be competitive. This is achieved by implying other participants' step-counts (that could be verified by opening the app), with an example being, 'You were at the top 5 participants yesterday, good job'.

- Personalization/customization/tailoring: The names of participants are used in the messages to make it more personal, and sometimes the type of weather of that specific day is used.

- Praise/punishment/(operant)conditioning: Participants get praise in the form of sentences starting with sentences like 'Well done!' or 'Good job yesterday!'.

- Reward: Rewards are similar to praise, as it means someone did a good job, but the focus is more on the achievement than on the praise itself. These sentences include 'You achieved', 'You unlocked' or 'You accomplished'. An example is, 'You were 465 steps above average for the past 7 days, you achieved a silver trophy'.

- Reminder/rehearsal: Although reminders are more technical and could be seen as a daily message (compared to informational influence, which is about the content), reminders are implemented as a separate strategy, with an example being: 'Please do not forget to set a good amount of steps today'.

- Consensus/social proof: Here the data of other participants is taken into account as well, just as with the competition strategy, but the aim is different. Instead of competing on a leader-board, sentences are more focused on facts. These sentences include percentiles and/or 'On average'. An example is: 'About 35% of participants have a step-count higher than 5000 on average'.

- Tips/suggestions: There are messages with tips in them, which could be seen as suggestions. An example is: 'You can get off public transport one stop earlier to walk more'.

The following strategies **may** play a role in this experiment:

- Self-monitoring/feedback: The notification on the mobile phone that displays the current step-count could be seen as self-monitoring. Sometimes participants had to open the app to synchronize, then they would also see their own step-count of the day and possibly of the last few days (depending on what they did when opening the app).

- Compliance/conformity: This is what is expected to be a side issue resulting from the experiment; a change in behavior and/or attitude.

- Scarcity: Perhaps participants feel more inclined to win at the social comparison part in the final week(s).

- Commitment/consistency: We cannot be sure whether participants will commit to walking more steps per day (and therefore stay consistent with that commitment).

- Timing of messages: All participants get their message at the same time in the morning around the time that they wake up or even before that. This way, participants can read the message without having other ideas, plans or tasks in their heads yet.

- Surface credibility: There might be a first impression of the experiment or experimenter, that might influence the experiment.

- Surveillance: There has been an attempt made to avoid surveillance in this experiment by stating that the messages were generated and sent by a bot rather than a person watching over their step-counts.

- Liking: Similarly to the previous one, the attempt to avoid this influence was made by stating that the messages are automatically generated (by a bot). Since the messages come from a bot, rather than from the experimenter, the possibility that participants comply with the message (just because they like the messenger) is reduced.

The following strategies are **not** used or not possible for this experiment:

- Tunneling: There are no encouragements, leading or step-by-step guidance along the way.

- Reduction: Tasks are not simplified or broken into steps, as there are no such tasks.

- Reciprocity: There are no concessions or favors involved between people.

- Simulation: The experiment is set in the real world, so there is no simulation of a world. There is also no clear cause-and-effect linkage in the experiment between walking more and having better health.

- Goal setting: There are no specific goals for participants, except for generally increasing their step-count.

- Customization: Participants are refrained from adapting their app, and they also cannot influence the kind of messages they are getting or the timing of those messages.

- Cooperation: There is no communication between participants to increase their step-counts, as the participants do not know each other, as they stay anonymous.

- Normative influence: There is no gathering of people with similar goals, as participants will not know about a goal, since there is no goal set for this experiment.

- Narrative paradigm: There are no stories, just short sentences that are sent to the participants.

- Altercasting: There will not be social roles involved, as participants are anonymous and they all have the same role, being a participant that gets some daily messages.

# 4 Implementation

Here, the actual implementation of the focus group, the pre-study and the experiment can be found. The overview of the process until the experiment (creating strategies and sentences) can be seen in Figure 4.1. The overview during the experiment (selecting strategies and sentences) can be seen in Figure 4.2.

## 4.1 Focus group

In the focus group, mainly students participated, as participants were recruited via friends and via the university of Utrecht via Teams. There were six participants present during the focus group (three males and three females of which one was older than a student). Using fewer people would result in too little input, but more could distort the flow of talking and listening.

To come up with the exact persuasive strategies for the focus group to evaluate, some strategies were not fit to be part of the experiment and were left out of consideration. For example, the foot-in-the-door technique as well as the door-in-the face technique were considered not usable, as it works with making requests, while for the experiment, the participants get only messages via WhatsApp. This is also the reason why reciprocity/trust and liking were not applicable, as the messages were automatically generated, which was also told to the participants. Cooperation was also not considered usable, as participants would not work together to achieve some goal, every participant was considered individually. Similarly, simulation, surface credibility, social learning, scarcity, reduction/tunnelling and goal setting (or suggestion) were not applicable for this experiment. Participants should not have specific goals to reach, as that might have had an effect on the step-counts of participants. The self-monitoring and surveillance strategies might have been implicit for this research, as participants could see their own step-count and they knew that those scores would be checked-upon and used for this research.

There also were some strategies that were closely related to other strategies that were not considered for the focus group. For example, normative influence were considered close to informational influence, while compliance and obedience were considered close to authority. Conformity, commitment/consistency and social proof were considered close to consensus. Recognition was considered close to praise, customization/tailoring was considered close to personalization, and rehearsal was considered close to reminders.

The strategies that were discussed in the focus group were: authority, informational influence, comparison/competition, personalization, consensus, praise, rewards, punishment and reminders. These strategies were derived from different previous studies. Other strategies used by previous studies mentioned in Section 2 were not fit to be used for this experiment or duplicated of the used strategies. The sentences that were presented to the focus group can be found in Appendix A.4.

Before diving into these nine persuasive strategies, all participants gave consent and were told about the aim of the research in general and the aim of the focus group specifically. After that, an example was given to grasp the subject better; *'A friend of you will walk to campus or take the bus. What would you say to this friend to let him/her walk more steps?'*. This was later on combined with the information that the friend has an app which counts the steps of that person. After that, feedback was gathered on the same situation, except that the friend could also see the step-counts of other friends as well.

Finally, the nine strategies were discussed. For each of the strategies, it was told what was meant by it, and participants had to come up with some example sentences. For each of the strategies, some sentences (three up to six) were prepared beforehand to give some extra guidance to the participants of the focus group. After discussing all nine strategies, participants of the focus group also had to fill in a survey, which is discussed in Subsection 4.2. After that, the participants were properly thanked for their participation and time and they could leave the meeting.

The set-up for the focus group can be found in Appendix A.3. The overview of the process of creating sentences for the experiment, can be seen in Figure 4.1.

## 4.2 Pre-Study

As a pre-study, the nine persuasive strategies were evaluated by 124 participants. The gender was evenly distributed between males and females with a few categorized as 'other'. Most participants were young (18-27 years), with a few middle-aged and even less old participants (53+). There were 25 nationalities recorded, of which 9 were well represented. The gender, age and nationality distributions of the pre-study can be found in Appendix A.1. The survey for the pre-study can be found in Appendix A.5 and on https://forms.gle/5uXZ4rpJm2csnjXW6.

The platform Prolific (https://www.prolific.co/) was used for this pre-study, to get as many participants as possible, while also getting a variety of people. For each participant, there were `Brainhex` questions, `STPS` questions and `TTM` questions to place each participant into certain personality profiles. In addition to that, the age and gender of the participants were asked for demographics. Participants were also asked to rate sentences (with persuasive strategies) on persuasiveness to walk more steps. Out of each of the nine selected persuasive strategies (by the focus group), three sentences were selected (the same ones for every participant), which represented that strategy the most. This was done to ensure that participants unknowingly rated the strategy behind the sentence, rather than the sentence itself, which might not have been as representative of the strategy. This means that there were 27 sentences to be judged upon by the participants. Participants of the pre-study had to select whether they loved/liked/disliked/hated the sentence or whether they had a neutral feeling about it. *Note: Reminders and tips were not used for this survey, as they are more general persuasive strategies rather than having to do with personal preferences. Reminders and tips were however used for the actual experiment as persuasive strategies.*

Since the results of the pre-study showed there was no personalization possible, while there were three clear top as well as bottom strategies (see Subsection 5.1), it was decided to divide participants among three groups for the experiment. The number of participants was 43, so each of the three experimental groups got either fourteen or fifteen participants.

## 4.3 Experiment

To recruit as many people as possible for the experiment, the research was shared on Facebook and Teams groups of Utrecht University. In total 47 participants were recruited, of which 43 fully participated in the experiment.

While the pre-study showed no connection between personal profiles and strategies, there was however a clear distinction between the top three and bottom three strategies. Therefore, participants were divided into three groups: a group with the three top-ranked persuasive strategies, a group with the bottom three ranked strategies, and a group with the other and thus middle three strategies. Each group received randomly assigned persuasive sentences (taken from the chosen three strategies) in the second, third and fourth week (but not the fifth). There was no room for an actual control condition and a contra-tailored persuasion condition, as there were too few participants. In a sense, it can still be called personalized persuasion, as the persuasion strategies are considered the best or worst for the participants in group 1 or group 3 respectively. It just happens to be the same top and bottom three for almost everyone. The core idea is to compare and see whether 'personalized' persuasion has an increased effect on top of (general) persuasion, of which we already know that it has a positive effect on changing behavior.

Rather than an activity tracker, an app was used and installed on participants' phones that tracks the number of steps set per day. It does so by using the built-in step-counter of the mobile phones. This is to avoid confounding (extra) step-count results from suddenly having an extra and cool gadget that shows increasing numbers in real-time. Mobile phones can be used as every student has one nowadays. The issue of women not having their phones in their pocket (as they are too small to hold them), as opposed to men, is not a problem for this research. This is because each participant was compared to him or herself during the experiment. If one does not take their mobile phone to the bathroom in their pocket, for example, they will neither do so in the week thereafter. That way the actual difference in step-counts is the only thing that is measured and matters most.

The app that was used is **Pacer**. This is an app on which you can log in either through Facebook, email, or Google account. It can synchronize the steps either via the phone itself or via Google Fit. The app can be found on https://play.google.com/store/search?q=pacer &c=apps&hl=nl&gl=US for Android phones and on https://apps.apple.com/nl/app/pacer-pedometer-bmi-tracker/id600446812 for Apple phones. Participants were instructed to install the app properly with its allowances and without reminders/ popups/messages and told to ignore that app after that. This is to avoid having gamification mechanics and popup messages influencing this research.

In Pacer, the experimenter can keep track of participants' step-counts per day, by inviting them into a group. There were three groups, each with three different persuasive strategy messages. This is to avoid having participants from one group increasing/decreasing their steps as a result of seeing the step-counts from another group. This is also to ensure that participants knew that the messages about the leaderboard were true, as they could verify it by checking the app. This is however not mentioned or encouraged for the participants. The app also gave a continuing message that shows the number of steps of that current day including the number of steps from the previous day. This is for participants to see whether they have beaten themselves yet, without having to open the app. *Note: This was however only possible for Android phones, but not for Apple phones yet.* Therefore, a question in the post-survey was added, concerning which device was used for the experiment.

Participants had to fill in a survey before the experiment and after the experiment. The link to the survey before the experiment is: https://forms.gle/8EDtSrmujvLeBpvS6 and can be found in Appendix A.7. The link to the survey after the experiment is: https://survey.uu.nl/jfe/form/SV_0IXBmEf0YFIc5Ey and can be found in Appendix A.8. *Note: for this survey, the nationality and kind of mobile phone used (Android/Apple) was added as a question, as they also turned out to be valuable information.*

During the experiment, participants should forget about the app (thus not use it, because it could interfere with the results) and go on with their lives, as they would normally do. Meanwhile, the participants get daily 'personalized' persuasion messages on **WhatsApp**. Participants were told that these are automatic messages from a bot, to lessen the surveillance aspect, as opposed to the experimenter personally typing the messages specifically for the participants. The messages were sent from a **WhatsApp Business** account, rather than a personal account. That way, the idea that a bot is sending those messages is more reliable (especially when participants know the experimenter and her phone number).

A Google Chrome extension was used, which is called **Blueticks**, along with the web version of WhatsApp that is called **WhatsApp Web**. That way, messages can be assigned and scheduled in the evening (when checking everyone's step-counts) to be sent in the morning at 7 AM. This is to ensure that participants get their message in a timely matter when or before they get up and start their day. Two other apps were also investigated, namely WhatsApp Scheduler and SKEDit. These however require the screen of a phone to be on at the specified time (7 AM), rather than a web browser on a laptop/desktop. With a computer, the messages were sent faster automatically, than with a phone that needs to open up each contact individually.

The way of selecting which persuasive strategy to send to which participant on which day, can be seen in Figure 4.2. The strategy was depended on a randomized number and a few other factors. These factors were: which day it was, which strategies were already used that week, and how often each strategy was already used in the experiment. The sentences were also selected via a randomized number, but it was checked whether that sentence made sense for each participant individually. This way, there would be no messages like: 'You did great yesterday with 34 steps!'. Rather, a new randomized number would be created and remembered for other participants that also could not get that specific sentence on that day.

**Participants' step-counts were recorded in weeks one through five, while persuasive messages were sent in weeks two through four. These messages were persuasive strategies, based on which experimental group a participant was in. The survey before the experiment was**

compared to the survey after the experiment, and the step-counts per participant and per group were compared to each other. The second hypothesis was that specific persuasion would have an added increased effect on walking more. This would mean that participants in group 1 (and perhaps group 2) would improve their walking behavior more than in group 3, as the pre-study showed that people think that the persuasive strategies in group 1 (and group 2) are more persuasive to walk more steps than in group 3. When the group, with persuasive strategies that were considered the most useful and persuasive in the pre-study, has the highest increase in step-counts, the third hypothesis could be confirmed, that there is no difference between perceived and actual persuasion. To test the first and second hypotheses, comparisons of step-counts within participants were made. Each participant was looked at on its own, so that different live styles were not compared against one another. The goal was to improve the number of steps for each participant individually.

# 5    Results

In this section, the results of the experiment are discussed. This is divided into three parts. The first part is about the results from the pre-study. Another part is comparing the survey answers of participants before the experiment with the answers after the experiment. This includes demographic, personal profiles, and attitude against walking. The last part is about comparing the actual step-count data of participants during the experiment within and between the three experimental groups. Some figures and tables were left out of the results, but were included in Appendix A.9 for convenience. To answer the research questions and to test for a general best persuasion strategy, the attitude against walking and the actual step-counts are relevant. To see if there were strategies that worked better for certain profiles (e.g for `Brainhex`, `STPS` and `TTM`), the step-counts need to be connected to the profiles of the survey. Finally, to test the third hypothesis whether perceived persuasion is similar to actual persuasion, the step-counts of the experimental groups are compared to the results of the pre-study.

## 5.1    Pre-Study

For the pre-study, personality profiles of the participants were gathered along with their preferences for specific persuasive strategies. There were 124 participants in total, which all got a `Brainhex` profile, `STPS` profile and `TTM` profile by answering the survey. Participants had to judge whether a sentence would help them walk more than they would do usually. There were three sentences per strategy and nine strategies, so they had to judge 27 sentences. The pre-study aims to answer part of the research question about which strategies are perceived as the most persuasive. The results of this pre-study are later on compared to the actual step-counts of the experiment, to answer the research question about the similarity between perceived and actual persuasion.

### 5.1.1    Comparing Profiles to Strategies

Since some personality profiles were less represented, there has been an attempt to reduce the seven `Brainhex` profiles to either six, five or four profiles. The same was done for the `STPS` profiles (six to five, four and three). Since the `TTM` profiles were more evenly divided among the participants, it was only reduced from six to five and four strategies in total. The means of reduction was done by pretending the lowest represented profiles were never profiles, to begin with. This way participants who did fall into that profile, would fall into another and more represented profile. As a result, there is more data per profile. An example is that the survivor profile was removed from the `Brainhex` profiles, and participants from that profile therefore fell into another profile, like daredevil.

This however, did not change the result of comparing the personal profiles to the available persuasive strategies. The scores of participants, which gave only small differences between profiles or strategies ($\pm$ 1 or 0), were also normalized to ensure similar data entries for every participant in the pre-study (such that all scores were between -2 and +2). This however did also not affect the results.

The results of matching personality profiles to persuasive strategies can be seen in Figure 5.1 for `Brainhex` profiles, Figure 5.2 for `TTM` profiles and Figure 5.3 for `STPS` profiles. The results showed that, for almost every personality profile (either `Brainhex`, `STPS` or `TTM`), the three most likes persuasive strategies were praise, information and rewards. In the few cases that this was not the case, there was too few data per profile and/or at least two out of the three strategies were still the most preferred ones.
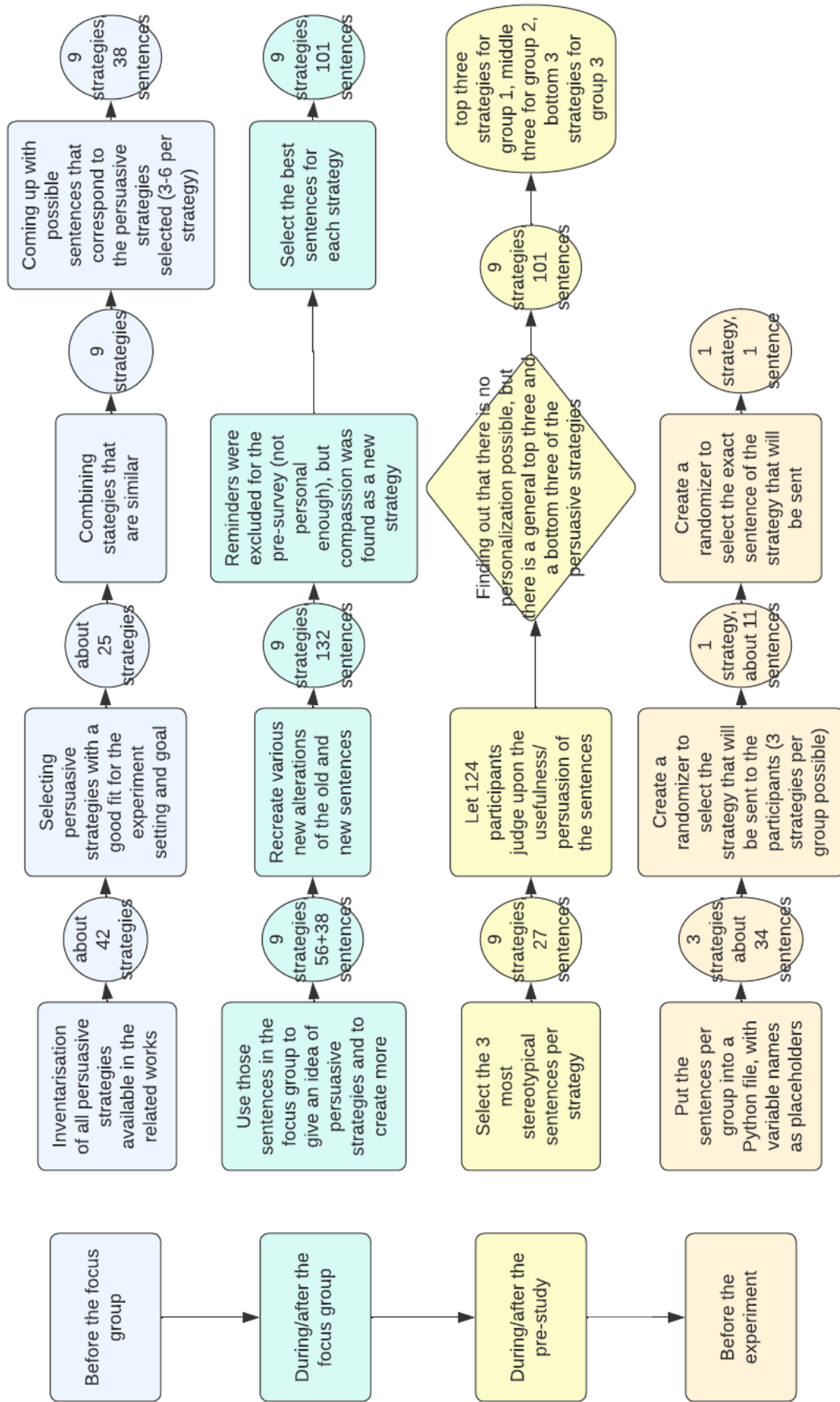
Figure 4.1: The overview of the process of creating sentences for the experiment, starting from previous literature, via the focus group towards the pre-study.
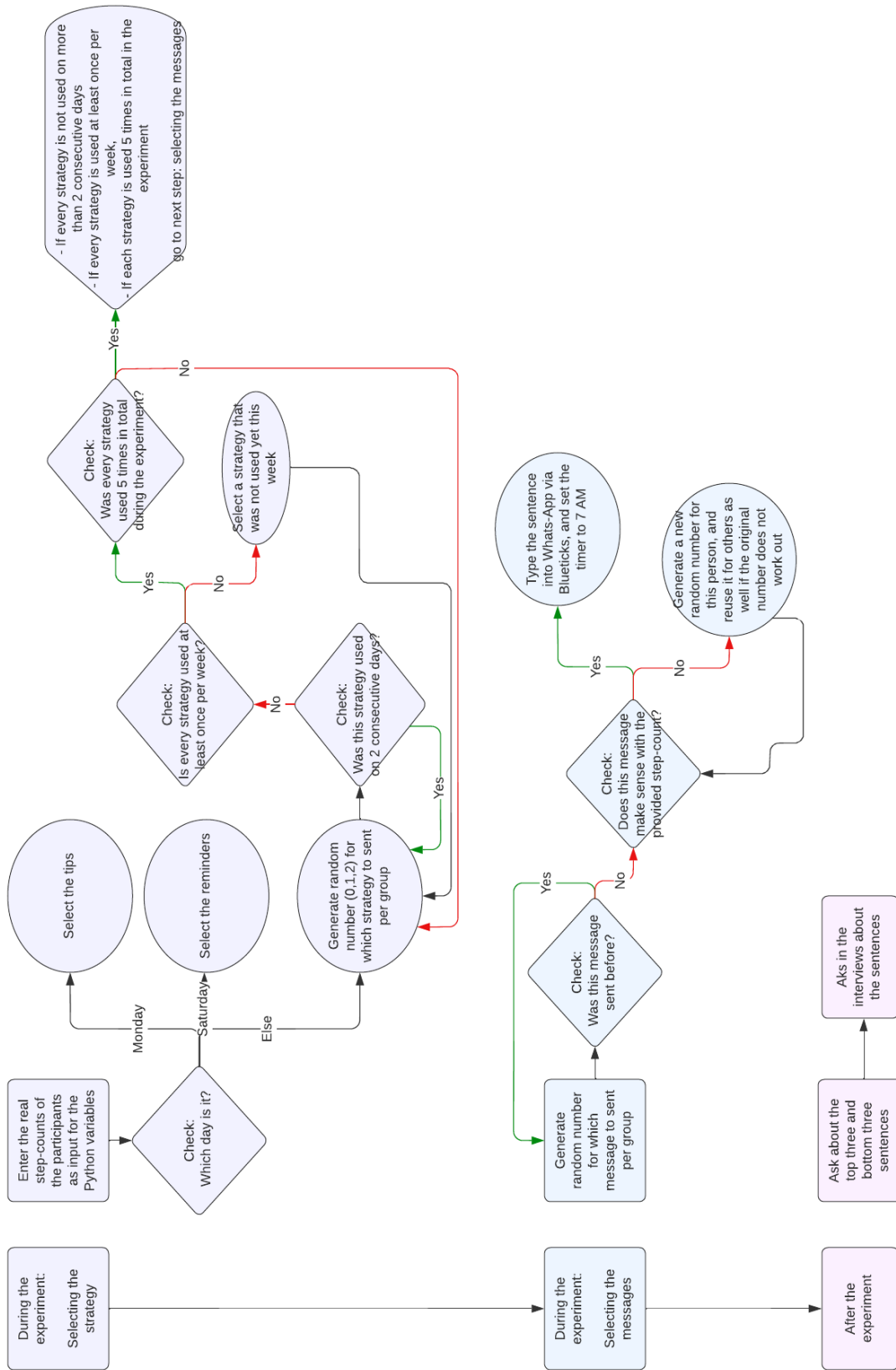
Figure 4.2: The overview of the process of selecting sentences to send during the experiment.

There was also a clear bottom three strategies that were liked the least for almost each personality profile. There were: competition, consensus and punishment. This leaves the strategies of personalization, compassion and authority in the middle.

Unsurprisingly, the same results (top and bottom three) showed up when the total scores of each strategy were taken into account for all participants together. For each strategy, there were three sentences that received a scoring from the participants. When taken together, praise, information and rewards scored a total of 337, 305 and 281 points respectively. personalization, compassion and authority scored a total of 233, 226 and 190 points respectively. Finally, the bottom three strategies, competition, consensus and punishment had a total score of 150, 34 and 34 points.

For the individual scores of each strategy, see Table 5.1. Some sentences might have been more representative of the strategy than others. For example, the first personalization sentence had a lower score than the other two. For other strategies, there was often one sentence with a lower or higher score than the other two sentences of that strategy. When these values were considered as outliers however, the ranking of the nine strategies did not change. Praise, information and rewards still came out on top, and competition, consensus and punishment were still liked the least.

**According to the pre-study, the personality profile did not matter when selecting the most persuasive strategy. This was later on compared to the step-counts of each experimental group along with their personality profiles, to see if perceived persuasion was similar to actual persuasion.**

| Strategy | Sentence 1 | Sentence 2 | Sentence 3 | Total score | Average | Times in top 3 |
|---|---|---|---|---|---|---|
| Praise | 92 | 113 | 132 | 337 | 0.92 | 60 |
| Information | 74 | 111 | 120 | 305 | 0.83 | 64 |
| Rewards | 91 | 117 | 73 | 281 | 0.77 | 43 |
| Personalization | 46 | 94 | 93 | 233 | 0.64 | 29 |
| Compassion | 96 | 47 | 83 | 226 | 0.62 | 39 |
| Authority | 67 | 62 | 61 | 190 | 0.52 | 45 |
| Competition | 22 | 60 | 68 | 150 | 0.41 | 21 |
| Punishment | 36 | 41 | -43 | 34 | 0.09 | 12 |
| Consensus | 28 | 8 | -2 | 34 | 0.09 | 11 |

**Table 5.1: Individual, total and average scores of each strategy from the pre-study, along with how often a certain strategy was chosen to be in the top three strategies.**

## 5.2 Survey

Below, some aspects of the survey are discussed and compared. The demographic aspects are gender, age and nationality. The personal profiles of `Brainhex`, `STPS` and the `TTM` are also included. The other aspects are ways to inspect how a person feels about walking: self-efficacy, attitude, intention and the `TTM`. *Note: the `TTM` is used twice, once as a profile and once as an attitude.* These variables are used to try and deduce the answer to which strategy is the most persuasive for each profile, while controlling for demographics.

### 5.2.1 Gender

Gender identification was one part of the survey, as it could be an `independent variable` for this research as stated in the problem statement. Males or females could react differently to a single persuasion strategy message. This could consequently influence the number of steps set per day, and thus the difference between the number of steps set and between that person's normal number of steps. Furthermore, males tend to have their phones in their pockets at all times, while females tend to keep the phone more on the table or in a purse, so that fewer steps are registered.

There were 43 participants in total, of which there were 34 females, 8 males and 1 participant who identified as 'other'. The males were evenly distributed among the three experimental groups; 3 in group 1, 3 in group 2 and 2 in group 3. This was to ensure that possible gender differences were not due to the condition of the experimental groups. The distribution of participants' gender among the three experimental groups can also be seen in pie charts in the middle part of Figure 5.4. *Note: since there was only one participant in the 'other' category, not much can be said about that category.*
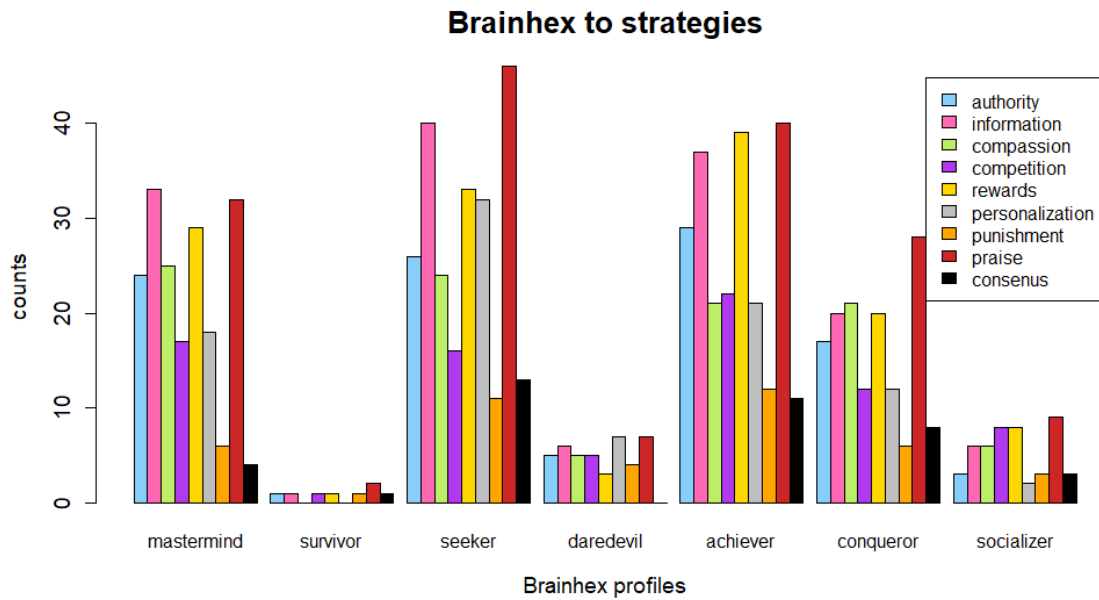
**Figure 5.1:** Comparing the `Brainhex` profiles to the nine persuasive strategies. It shows how often a particular `Brainhex` profile was correlated with a certain persuasive strategy.
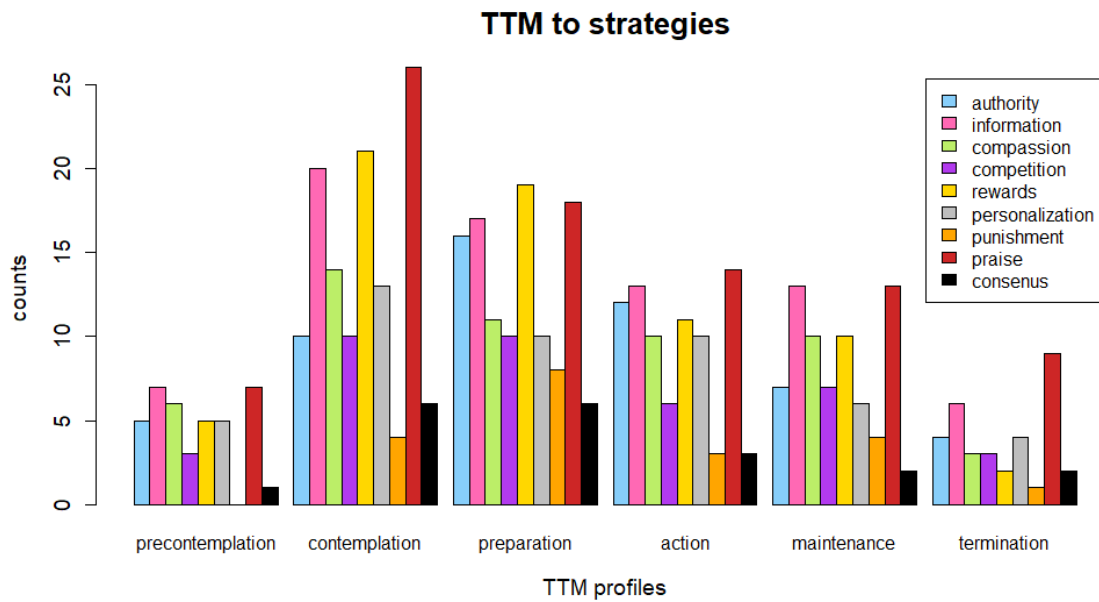


**Figure 5.2:** Comparing the `TTM` to the nine persuasive strategies. It shows how often a particular TT profile was correlated with a certain persuasive strategy.

The means and standard deviations of the different genders in the experiment can be found in Table 5.2.

**There were a few male participants. Females and males did seem to walk a similar number of steps, but females did walk more (6880) than males did (6466).**

### 5.2.2 Age

Age was another part of the survey, as it also could be an `independent variable` for this research. Young people could react differently to the same persuasive message received as elderly people. This

**Figure 5.3: Comparing the STPS profiles to the nine persuasive strategies. It shows how often a particular STPS profile was correlated with a certain persuasive strategy.**
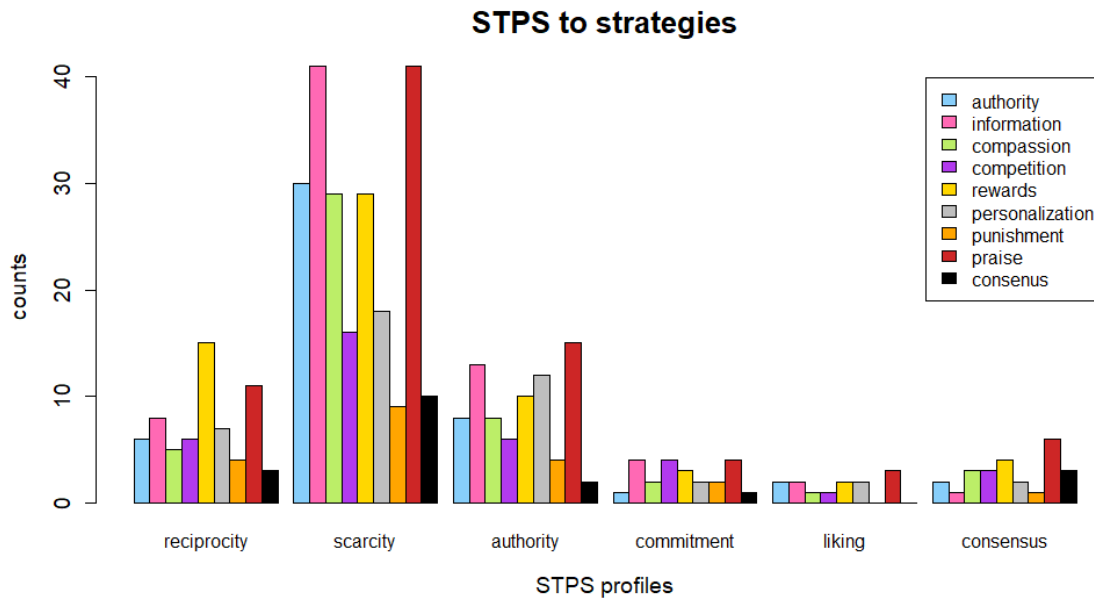
| Gender | Mean | SD | Amount of people |
|--------|------|------|------------------|
| Other | 4343 | 2899 | 1 |
| Male | 6466 | 4958 | 9 |
| Female | 6880 | 5838 | 33 |

**Table 5.2: Table with an overview with means and standard deviation of the step-counts per gender category.**

could consequently influence the number of steps set per day as well as the difference in number of steps because of the persuasive message.

There were 43 participants in total, of which 30 were between 18 and 27 years old, 8 were between 28 and 52 years old and 5 people were older than 52. The ages were also evenly distributed among the groups, to ensure that possible age differences were not due to the condition of experimental groups. Group 1 had 14 people in total of which 10 were young people, 2 were older people and 2 were in-between. Group 2 had 15 people in total of which 11 young ones, 1 older one and 3 in-between. Group 3 had 14 people in total of which 9 young ones, 2 older ones and 3 in-between. The distribution of participants' age among the three experimental groups can also be seen in pie charts in the top part of Figure 5.4.

The means and standard deviations of the different ages in the experiment can be found in Table 5.3.

**Young people did walk more on average (7088) than older people (6064) which in turn did walk more than middle ages people on average (5830).**

| Age | Mean | SD | Amount of people |
|--------|------|------|------------------|
| Young | 7088 | 5754 | 30 |
| Middle | 5830 | 5857 | 8 |
| Old | 6064 | 4048 | 5 |

**Table 5.3: Table with an overview with means and standard deviation of the step-counts per age category.**

### 5.2.3 Nationality

As nationality can influence the results, even when people have the same `Brainhex` profile (Orji et al., 2013a), the nationality of the participants was also recorded. There were eight nationalities recorded, of

which Dutch was the most represented with 35 (out of 43) participants. In Appendix A.2, the distribution of nationalities of the participants is shown along with their respective experimental groups.

Since nationality was only considered after the experiment, and because there were few other nationalities besides Dutch, the different nationalities were not evenly distributed among the experimental groups. The distribution of participants' nationality among the three experimental groups can be seen in pie charts in the bottom part of Figure 5.4.

The means and standard deviations of the different nationalities in the experiment can be found in Table 5.4.

**Foreigners did walk more on average (8624) than Dutch people (6303), but their standard deviation was higher as well. There were also fewer foreigners than Dutch people.**

| Nationality | Mean | SD | Amount of people |
|---|---|---|---|
| Dutch | 6303 | 4875 | 35 |
| Other | 8624 | 7868 | 8 |

**Table 5.4: Table with an overview with means and standard deviation of the step-counts by Dutch people and by foreigners.**
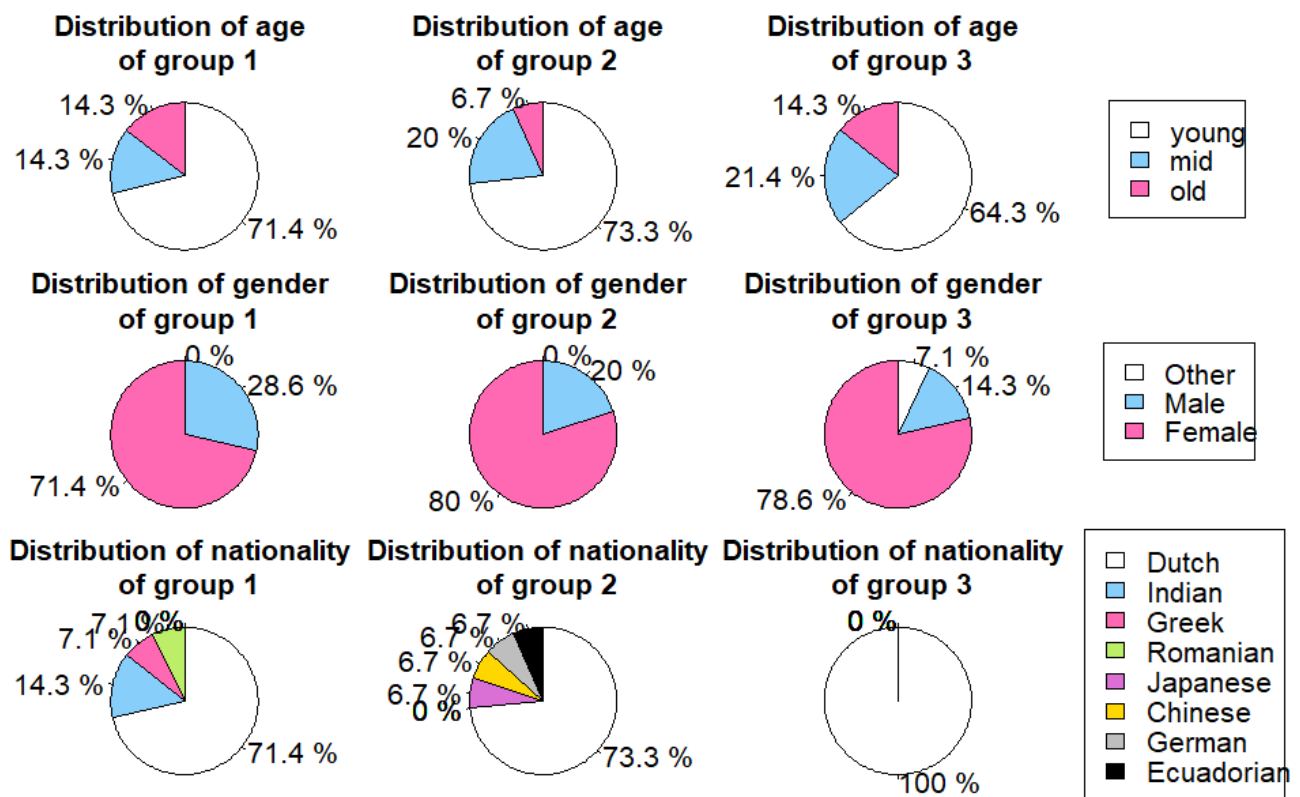


**Figure 5.4: Distribution of age, gender and nationality among the three experimental groups.**

### 5.2.4   Brainhex

The distribution of `Brainhex` (main) profiles was also considered when assigning the participants to their groups. This was also done because of the possibility of different outcomes when `independent variables` interact with each other (e.g. the `Brainhex` profile interacting with the persuasion messages received). `Brainhex` is one of the three relevant personality profiles in this research for testing and answering the research question about specific strategies that might work better for specific profiles.

There were 8 people with an achiever (main) profile, 6 conquerors, 3 daredevils, 10 masterminds, 12 seekers and 4 socializers in total. The exact distributions can be found in Appendix A.3. The ones in brackets are the sub-profiles. The `Brainhex` sub-profiles were not considered when dividing participants
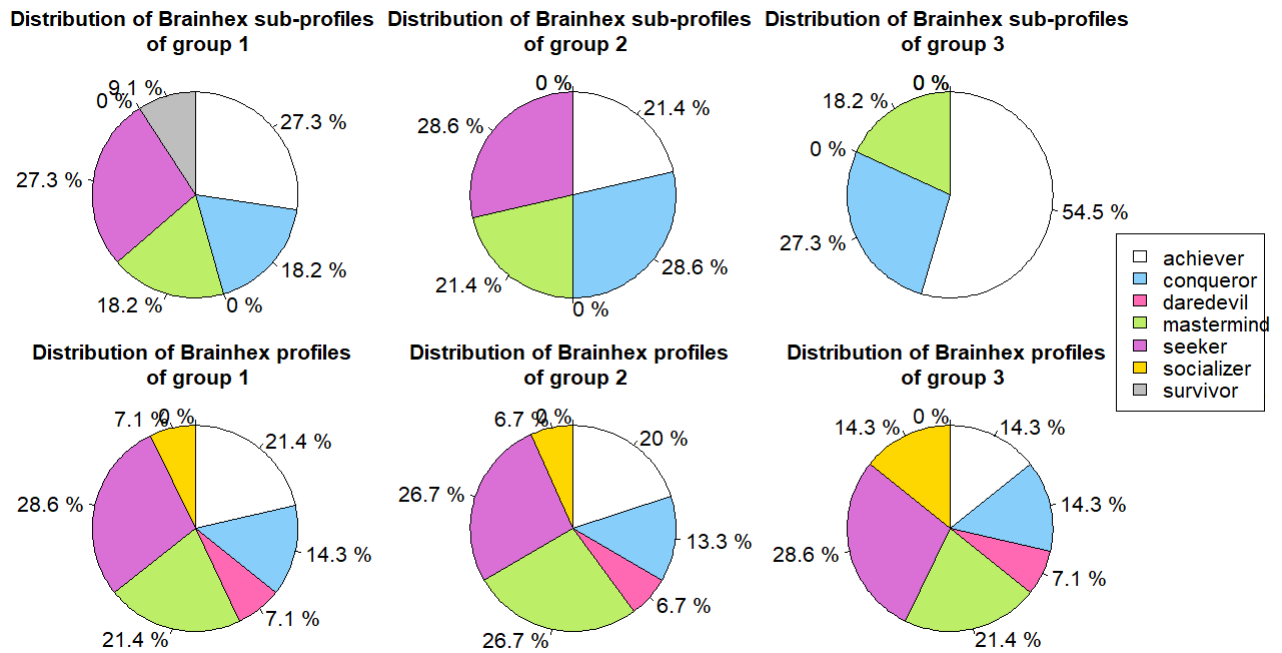
**Figure 5.5:** Distribution of main and sub- `Brainhex` profiles among the three experimental groups.
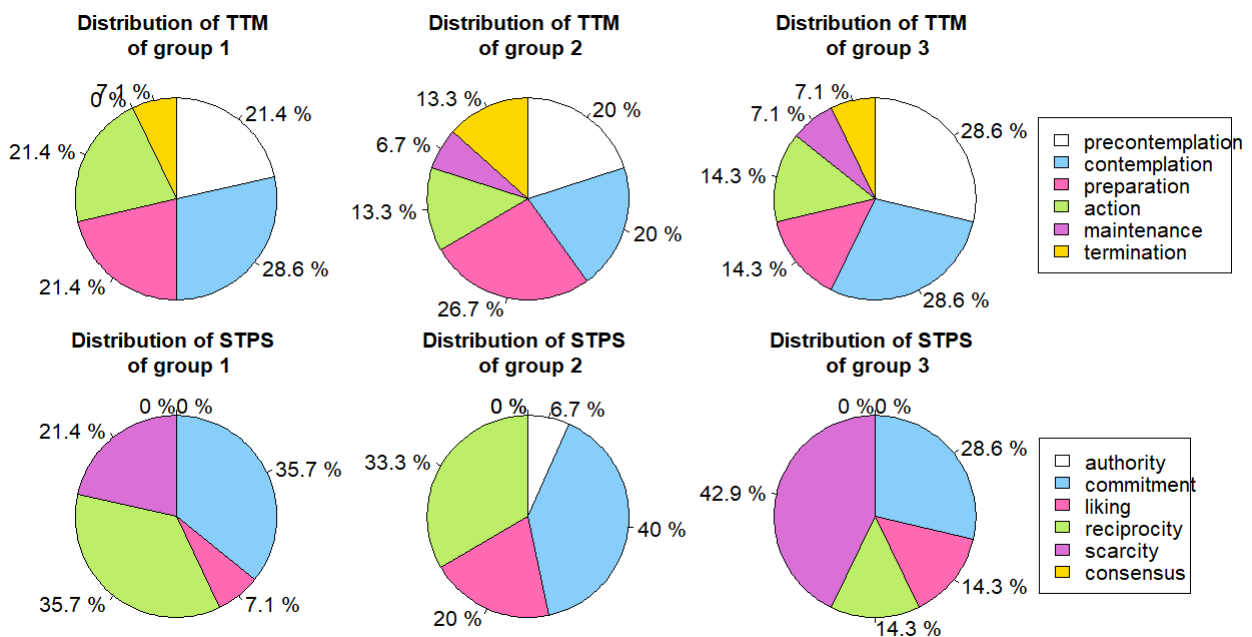


**Figure 5.6:** Distribution of `TTM` and `STPS` among the three experimental groups.

into groups. This is because the main profiles bears more meaning and there were already many other variables to look out for. The distribution of participants' `Brainhex` main and sub-profile among the three experimental groups can be seen in pie charts in the top and bottom parts of Figure 5.5, respectively. There were no participants with survivor as their main `Brainhex` profile.

The means and standard deviations of the different main `Brainhex` profiles in the experiment can be found in Table 5.5.

**Participants who had the conqueror profile seemed to walk the least on average (4985), while participants with the profile mastermind seemed to walk the most (8252) on average.**

| Brainhex profile | Mean | SD | Amount of people |
|---|---|---|---|
| Achiever | 6519 | 4915 | 8 |
| Conqueror | 4985 | 4077 | 6 |
| Daredevil | 5580 | 5085 | 3 |
| Mastermind | 8252 | 7211 | 10 |
| Seeker | 6845 | 5434 | 12 |
| Socializer | 6531 | 4287 | 4 |
| Survivor | - | - | 0 |

**Table 5.5: Table with an overview with means and standard deviation of the step-counts per (main) `Brainhex` profile.**

### 5.2.5 Susceptibility To Persuasion Scale (STPS)

The distribution of the six Cialdini principles (Cialdini, 2006; Cialdini and Griskevicius, 2010) was not considered when dividing participants into groups, as there were enough other variables to look out for, which were considered more important. It was noted down however, to later see if there were any interesting conclusions on the effect between this personality feature and the different persuasion strategies. This is also used to answer the research question about specific strategies that might work better on some profiles.

It turned out that only scarcity and reciprocity were not evenly distributed. The exact distribution of the STPS among the three experimental groups can be found in Appendix A.3. The distribution of participants' STPS profiles among the three experimental groups can also be seen in pie charts in the bottom part of Figure 5.6. *Note: since there was only one participant with authority as their STPS profile, not much can be said about that profile.* There were no participants with consensus as their STPS profile.

The means and standard deviations of the different STPS profiles in the experiment can be found in Table 5.6.

**Participants who had commitment as their STPS profile walked less on average (5713) than other participants. Participants who had reciprocity as their STPS profile walked more on average than other participants. No other differences were visible.**

| STPS | Mean | SD | Amount of people |
|---|---|---|---|
| Authority | 9404 | 4529 | 1 |
| Commitment | 5713 | 4758 | 15 |
| Liking | 7161 | 6691 | 6 |
| Reciprocity | 7341 | 5875 | 12 |
| Scarcity | 7048 | 5696 | 9 |
| Consensus | - | - | 0 |

**Table 5.6: Table with an overview with means and standard deviation of the step-counts per STPS profile.**

### 5.2.6 Trans-Theoretical model (TTM)

The TTM profile also contributes to the research question of which specific strategy is considered more persuasive for certain profiles. For convenience, the stages in the places in the TTM are given an ordered number. The first precontemplation stage gets the value zero and it goes up until the last termination stage, which gets the value five. This way, the progress of participants can be seen more clearly. The differences between before and after the experiment in the TTM can be seen in Figure 5.7 and Table 5.8. The last one is also used for psychological changes during the experiment, see Subsection 5.2.7.

In total, the 43 participants degraded 10 steps in the TTM during the experiment. Group 1 actually gained a step upwards in the TTM, while group 2 had lost 4 steps and group 3 had lost 7 steps in the TTM. The exact distribution of the TTM among the three experimental groups, before and after the experiment, can be found in Appendix A.3. The distribution of participants' TTM profiles among the three experimental groups before the experiment can also be seen in pie charts in the top part of Figure 5.6.

The means and standard deviations of the different TTM profiles in the experiment can be found in Table 5.7. Participants who had preparation as their TTM profile walked less on average (5063) than other
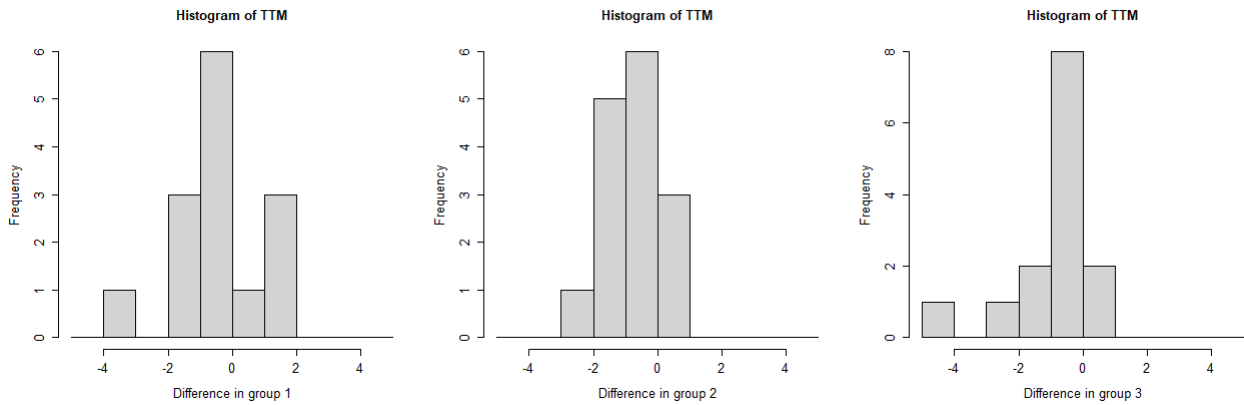
**Figure 5.7: Frequency of differences between TTM before and after the experiment. The TTMs are converted into numbers for convenience. So, if one goes from termination (5) to precontemplation (0), the difference is recorded as -5.**

participants. This is contrary to expectations, as it would be expected to find the lowest number of steps in the lowest TTM profile (precontemplation). Participants who had termination as their TTM profile walked more on average (9811) than other participants. This was expected, as the last profile of the TTM is supposed to be the profile that walks the most. *Note: since there were only two participants with preparation as their TTM profile, not much can be said about that profile.*

**Participants in the fifth and sixth stage of TTM had higher step-counts than participants in stages one and two, as expected. There was not as visible for the third and fourth stage. This was either due to the high standard deviations or the low amount of people per stage of the TTM.**

| TTM | Mean | SD | Amount of people |
|---|---|---|---|
| Precontemplation | 6065 | 4954 | 7 |
| Contemplation | 6792 | 5281 | 11 |
| Preparation | 5063 | 4001 | 2 |
| Action | 7371 | 7242 | 10 |
| Maintenance | 8911 | 4932 | 9 |
| Termination | 9811 | 6578 | 4 |

**Table 5.7: Table with an overview with means and standard deviation of the step-counts per TTM profile.**

### 5.2.7 Psychological Changes

In this Subsection, the differences before and after the experiment are discussed for self-efficacy, attitude, intention and the TTM. The differences can be seen in Table 5.8. The differences are compared between groups and between nationalities (Dutch or others). For the exact frequencies of differences in self-efficacy, attitude, intention and TTM, see Figure 5.8. The self-efficacy, attitude and intentions will be compared before and after the experiment, to see if differences have been recorded. This is needed to answer part of the research questions about which strategy would be the most persuasive. When one experimental group would have a (bigger) increase than other groups, the persuasive strategies of that group would be more persuasive.

**Self-efficacy:** There was a drop in self-efficacy of 60 points in total after the experiment, compared to before the experiment. In group 1 the drop was 41 points, in group 2 it was 36 points but in group 3 it was an increase of 17 points. The difference in self-efficacy was the most negative category for group 2, while it was the most positive category for group 3.

**Attitude:** There was an increase in attitude of 17 points in total after the experiment, compared to before the experiment. The increase for group 1 consisted of 14 points, in group 2 it was 26 points, but for group 3 it was a decrease of 23 points. The difference in attitude was the most positive category for
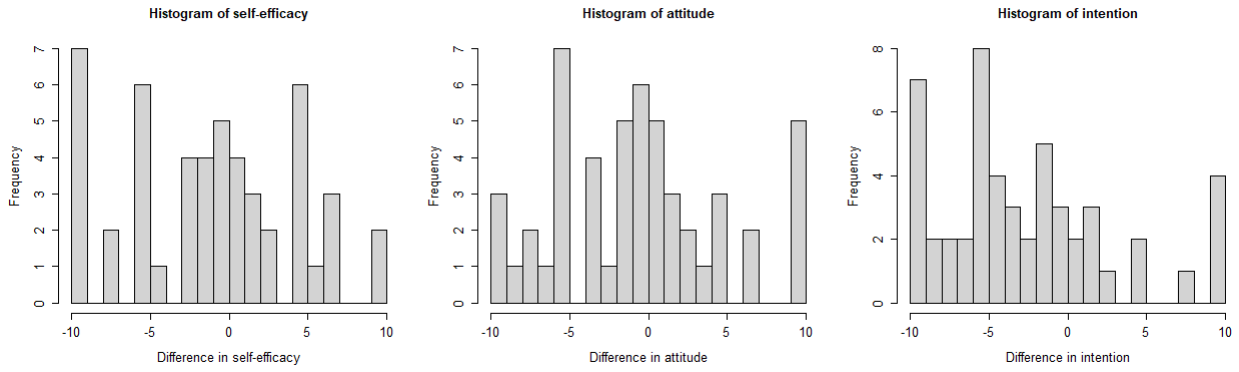
**Figure 5.8: Histogram of the difference between before and after the experiment of self-efficacy, attitude and intention. Differences higher than 10 (+/-) were recorded as 10 (+/-).**

group 1 as well as for group 2.

**Intention:** There was a drop in intention of 83 points in total after the experiment, compared to before the experiment. In group 1, the drop was 44 points, in group 2 it was an increase of 3 points, and for group 3 it was again a drop of 42 points. The difference in intention was the most negative category for both group 1 and group 3.

**TTM:** There was a drop of 10 places in total after the experiment, compared to before the experiment. In group 1, the increase was 1 point, in group 2 it was a decrease of 4 points, and for group 3 it was a drop of 7 points. These differences are mostly negative, contrary to expectations, but are proportionally within expectations. Group 1 had the most positive effect, while group 3 had the most negative effect.

In Appendix A.5, all information can be found regarding all participants ID's, their nationality, their self-efficacy, their attitude against walking, their intention to walk more and their place in the TTM before and after the experiment. The difference (which was shown here already) is also noted in that table but between brackets.

**Most psychological feelings towards walking were decreased during the experiment. Only the attitude towards walking increased a bit.**

| Group | Nationality | Self-efficacy | Attitude | Intention | TTM |
|--------|-------------|---------------|-----------|------------|------------|
| Group1 | Dutch | -33 (-3) | +37 (+4) | -42 (-4) | ±0(±0) |
| Group1 | Other | -8 (-2) | -23 (-6) | -2 (-1) | +1 (±0) |
| Group1 | All | -41 (-3) | +14 (+1) | -44 (-3) | +1 (±0) |
| Group 2 | Dutch | -4 (±0) | +15 (+1) | +18 (+2) | -2 (±0) |
| Group 2 | Other | -32 (-8) | +11 (+3) | -15 (-4) | -2 (-1) |
| Group 2 | All | -36 (-2) | +26 (+2) | +3 (±0) | -4 (±0) |
| Group 3 | Dutch | +17 (+1) | -23 (-2) | -42 (-3) | -7 (-1) |
| Group 3 | Other | - | - | - | - |
| Group 3 | All | +17 (+1) | -23 (-2) | -42 (-3) | -7 (-1) |
| All | Dutch | -20 (-1) | +29 (+1) | -66 (-2) | -9 (±0) |
| All | Other | -40 (-5) | -12 (-1) | -17 (-2) | -1 (±0) |
| All | All | -60 (-1) | +17 (±0) | -83 (-2) | -10 (±0) |

**Table 5.8: Table with another overview of differences in self-efficacy, attitude, intention and the TTM of the participants before and after the experiment. The number between brackets is the averaged difference in score per person, rounded to a whole number.**

### 5.2.8   Best and Worst Sentences

In the survey after the experiment, participants were asked to give their top three and bottom three sentences (out of the 21 sentences during the experiment) and to rate them accordingly. This was used to answer the research questions about the most persuasive strategy (which strategy is mentioned the most) and about the relation between perceived and actual persuasion (by comparing these results to

those of the pre-study). Four (out of 43) participants had difficulty in deciding which sentences were more persuasive and useful or less persuasive for them personally, and chose to fill in no sentences and no ranking. The ranking ranged from 'big negative impact' (0) to 'big positive impact' (+5) in five steps. Each participant could only enter sentences that they had received during the experiment, so group 1 entered different sentences than group 2 or group 3. The tips and reminders were sent to each participant however. First, the general persuasive strategies are considered and shown in Table 5.9 and after that, the individual sentences are considered.

For group 1, the entries (aside from reminders and tips) consisted of praise, information and rewards. Only praise and rewards had a small positive effect on walking more, as they had an average of around four. For group 2, the entries (aside from reminders and tips) consisted out of personalization, compassion and authority. Compassion and authority almost had a small positive effect. For group 3, the entries (aside from reminders and tips) consisted of competition, consensus and punishment. This was the most interesting group, as competition had a small positive effect, while consensus had no effect, but punishment had a small negative effect on walking more according to the participants after the experiment. The reminders and tips, that were sent to all participant groups, were mentioned the most, but had no effect according to the participants.

When entries with a score of three, which stands for no effect, were not considered when taking the averages of each strategy, a few different results were acquired. Compassion, authority and consensus were then considered to have a small positive effect rather than no effect, while competition went from a small positive effect to a small negative effect.

| Group | Strategy | Times mentioned | Average score | Effect on walking more |
|-------|----------|-----------------|---------------|------------------------|
| Group1 | Praise | 8 | 3.88 (4.4 +) | small positive effect |
| Group1 | Information | 10 | 3.15 (3.4 ±) | no effect |
| Group1 | Rewards | 14 | 3.93 (4.08 +) | small positive effect |
| Group 2 | Personalization | 19 | 3.11 (3.14 ±) | no effect |
| Group 2 | Compassion | 13 | 3.46 (3.6 +) | no effect |
| Group 2 | Authority | 22 | 3.45 (3.77 +) | no effect |
| Group 3 | Competition | 28 | 3.63 (2.4 -) | small positive effect (small negative) |
| Group 3 | Consensus | 13 | 3.15 (3.55 +) | no effect |
| Group 3 | Punishment | 15 | 2.46 (2.25 -) | small negative effect |
| All | Reminders | 36 | 3.19 (3.56 +) | no effect |
| All | Tips | 47 | 3.23 (3.69 +) | no effect |

**Table 5.9: Table of each strategy and how often it was mentioned and it's average score in the survey after the experiment. Scores were between 0 (big negative effect) and 5 (big positive effect). Reminders and tips were sent to all experimental groups. Scores between brackets were acquired when scores of three were not taken into account.**

When considering individual sentences, there are a lot of sentences which have an average score of around three, which would mean that they have no effect on walking less or more. Other sentences are mentioned only one or two times in total. The sentences that were mentioned three or more times can be found in Appendix A.6 and Appendix A.7.

The highest scoring sentences were found in the personalizing (4.0), reminders (3.83) and competition (4.67 and 3.75) strategies [*]. The lowest scoring sentences were for the personalization (2.0) and punishment (2.0) strategies [†]. The sentences that were used most often were from the reminders (6 times) and authority (7 times) strategies [‡]. There was one sentence that had very diverse scores per participant. This was a sentence from the rewards strategy: '*If you manage to get more steps today than your highest record*

[*] 'Hi ##, you have increased your step-count with ## since you began the experiment, keep it going! (Yesterday you walked ## steps.)', 'Do you remember that you reached ## steps during the last 7 days? Let us try to reach this number again today!', 'You walked more than the average (5582) of the participants in your group yesterday. (You walked ## steps yesterday.)' and 'Someone was only ## steps behind you yesterday, to keep ahead, try to walk at least ## steps today. (Yesterday you walked ## steps.)'

[†] 'Hello ##, you probably have some time left today to walk around the block. (Yesterday you walked ## steps.)' and 'Your step-count of yesterday was not so good (##), aim higher.'

[‡] 'Do you remember that you reached ## steps during the last 7 days? Let us try to reach this number again today!' and 'A study conducted by Amanda Paluch, PhD, assistant professor in the department of kinesiology at the University of Massachusetts Amherst found that taking at least 7.000 steps per day was associated with a 50-70% lower chance of early death among middle-aged adults.'

*(##), you may treat yourself with something nice.*'. It had scored a five for two times but also a one and a two, which resulted in an average of 3.25, which stands for no effect.

The results of the sentences of the survey after the experiment can also be compared to the results of the sentences of the pre-study. In the pre-study, 124 participants could rate 27 sentences in total on their persuasiveness. Those sentences were implementations of the nine persuasive strategies, with reminders and tips excluded. The results of the pre-study can be found in Table 5.10. The scale of the pre-study was slightly different from the scale of the sentences after the experiment, but the essence was the same. In the pre-study, the scale went from -2 up to +2, while in the survey after the experiment went from 0 up to +5.

In the pre-study, almost every strategy would have had a small positive effect on walking more according to the participants. Only competition, punishment and consensus would have no effect, which were exactly the three strategies for group 3 of the experiment. While in the pre-study almost every sentence would have had a small positive effect, the experiment survey showed that almost every strategy had no effect. However, when scores of the value three of the experiment survey were excluded from calculating the averages, the tables do look more similar. There are still three differences. (1) Information and personalization had no effect in the experiment, while they did have a small positive influence in the pre-study. (2) Consensus had a positive influence in the experiment, but no effect in the pre-study. (3) Lastly, competition and punishment had a small negative influence in the experiment but no effect in the pre-study.

**So, in general, the experiment survey showed fewer positive effects, or even more negative effects than the pre-study, with consensus being the only exception.**

| Group | Strategy | Times in top 3 | Average score | Effect on walking more |
|-------|----------|----------------|---------------|------------------------|
| Group 1 | Praise | 60 | 0.92 | It would help |
| Group 1 | Information | 64 | 0.83 | It would help |
| Group 1 | Rewards | 43 | 0.77 | It would help |
| Group 2 | Personalization | 29 | 0.64 | It would help |
| Group 2 | Compassion | 39 | 0.62 | It would help |
| Group 2 | Authority | 45 | 0.52 | It would help |
| Group 3 | Competition | 21 | 0.41 | Neutral |
| Group 3 | Punishment | 12 | 0.09 | Neutral |
| Group 3 | Consensus | 11 | 0.09 | Neutral |

**Table 5.10: Table of each strategy and how often it ended up in the top three strategies per participant and it's average score. Scores were between -2 ('It would not help at all') and +2 ('It would help a lot!'). Reminders and tips were not considered in the pre-study.**

## 5.3 Holidays

This research was aimed at finding the most persuasive strategies and the comparison between perceived and actual persuasion. Holidays were included in the results to have an overview of this variable and to see whether it had effects on the step-counts of participants on their own.

Out of the 43 participants, 24 went on some holiday during the experiment. The number of holidays differed from 0 to 24, out of the 35 experiment days. Only holidays of six or more days were considered for the results, which includes 20 participants. Fewer days are not representative of holiday behavior and there is not enough data otherwise. For these 20 participants (eight in group 1, six in group 2 and ten in group 3), the box plots of normal days and holidays are shown in Figure 5.9. The means are included as well. In Appendix A.6 a similar box plot is shown but the results are separated per group. While there is not much data per group, the groups do not seem to deviate much from each other, except that group 2 has a larger fourth quartile. While the bottom part of the box plots look similar, from the third quartile upwards (including the mean) the step-counts of holidays are higher than that of the normal days. The normal days have a third quartile of 7924 steps, while the holidays have a third quartile of 12381 steps. The biggest difference can be seen in the fourth quartile, holidays have much more variation and higher values than normal days. The mean of the normal days is 6212 with a standard deviation of 3330, while for holidays the mean is higher, with 9776 and a standard deviation (that is also higher) of 5917.

Another comparison of normal days versus holidays can be made between week one and week five. These are the most important weeks to look at, since they represent the before and after effects of this

experiment. In the first week, 6 out of 43 participants went on holiday. Two times a single day, and four times the full seven days. In the last week, 12 out of 43 participants went on holiday. One time a holiday of one, three, and five holidays and four times a holiday of six days and five times a holiday of the full seven days. This means that in week one, 29 out of 301 (7 days times 43 participants) days were holidays, while for week five there were 68 holidays, which is double the amount of week one.

This shows that holidays could indeed influence the results, as holidays have generally higher step-counts than normal days. For group 3, this change is the highest, as there were more people on holiday than in group 1 and group 2.
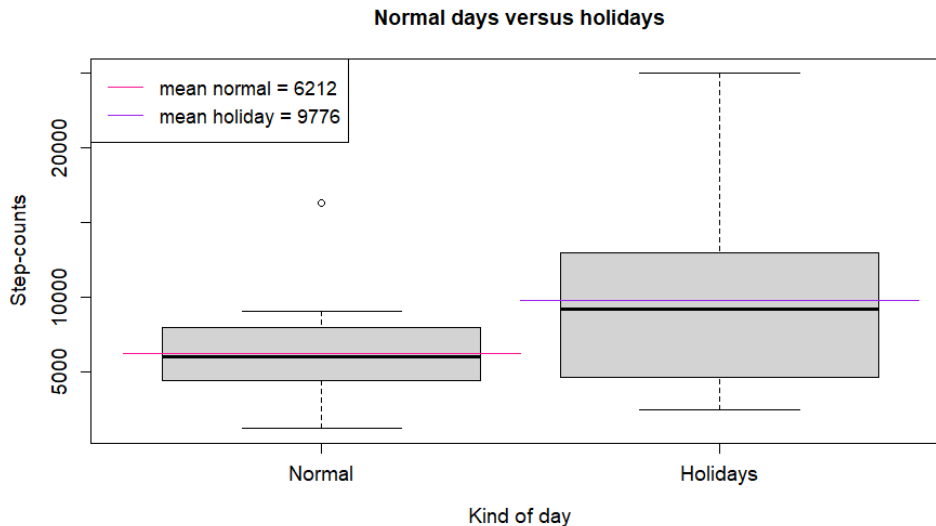


**Figure 5.9: Box plots of normal days versus holidays. Only holidays with at least six days were considered.**

## 5.4 Interviews

Some participants agreed to have some discussion about the experiment after the experiment was over. This included five people from group 1, four people from group 2 and three people from group 3. They all got the same five questions about the likability, clarity, usefulness of the experiment and its messages and the long-term effects. This is to be able to see if there were interesting aspects to the experiment, that could not be found when looking at the step-count data and survey answers alone. Below, the answers are presented to these questions, per group. In Appendix A.9.1, the full and exact answers of the participants can be found to these questions. Here, only the general gist is presented of each question and group.

**(1) Did you like participating in this experiment?**
**Group 1:** Participants liked the experiment, but gave advice on using activity trackers, because they are not forgotten and are more accurate and synchronize better. Furthermore, a participant liked keeping track of the step-count and wanted to walk even more the next day. Another participant noted that the group was working hard on increasing their step-counts, while not having an influence on his/her own step-count. A citation from a participant [**translated to English**]: "(3) I enjoyed participating! It's nice to keep track of your number of steps in a day. You try to walk more on the next day!"

**Group 2:** Participants liked to participate, as it was considered fun and did not take much time. One participant noted that it was great that (s)he got some daily activity as part of the experiment. Another participant became happy when (s)he reached his/her own goals. A note was made on the duration of the survey (it was quite long) and that a Dutch version of the survey would have been nice.

**Group 3:** Participants considered the experiment fun. It was interesting for participants to see how many steps were set per day. A citation from a participant [**translated to English**]: "Yes, I did like to participate! I know that I set a lot of steps in my daily life, so I liked the confirmation of that.".
**(2) Were all the instructions clear for this experiment?**
**Group 1:** Participants found the experiment clear and the instructions were easy to follow. Sometimes a

sentence had to be read twice, because it was not in their native language. A citation from a participant [**translated to English**]: "I had to read it twice sometimes, my English has improved because of it. It was very clear though.".

**Group 2:** There were no problems, in general it was all very clear.

**Group 3:** Instructions were clear, but sometimes it was a bit unpractical, since iPhones worked differently than Android phones.

**(3) Did the experiment help to set more steps or become more aware of your step-count?**
**Group 1:** In general, the answer was 'not really'. One participant noted that (s)he already exercises enough and does not look at step-counts, even on a Fitbit. It was also mentioned that the weather was nice to set higher step-counts and that it would be lower if there was more rain during the experiment.

**Group 2:** Participants did not feel like the experiment increased their step-counts. This was either caused by frustration and demotivation of the not accurate step-counts, or that participants already walked much. One participant was quite positive however, and noticed that (s)he still watches his/her step-count (while never doing so before the experiment). It was nice to see and know the step-count and (s)he felt obliged to walk more when the step-count was lower. One mentioned that it was unfortunate that there was no display of the step-count on the home screen (iPhone), as it would have helped more in increasing the step-count. A citation from a participant: "My walking has definitely increased after I installed the app and now too. It's higher than the usual.".

**Group 3:** There was no clear put answer, it ranged from 'not really' to becoming aware of the step-count (but also forgetting about it sometimes).

**(4) Did the messages have an influence on you?**
**Group 1:** While participants did read the messages, it was said that they did not have an influence. This was either due to knowing them from another app ('Ommetje'), or considering that they already walk/exercise enough or have no specific goal to reach. The messages were considered fun and informative and felt like an extra stimulant however.

**Group 2:** There was a lot of variation. Some read the messages and did nothing with them, some enjoyed the messages, and some disliked them even. On the one side, messages could make someone competitive, but on the other side, they could come across as nagging. One participant noted that not the content of the message, but rather the cue of the message made him/her stand on point with the step-counts. A citation from a participant [**translated to English**]: "Personally, the messages were not really pleasant, it came across like my mom who nagged that I had to do something."

**Group 3:** On one side, the messages were considered fun and gave a stimulant to do better on the next day. On the other side, the competitive messages might have worked (better) if participants knew the other participants or when they were close friends with them. A citation from a participant [**translated to English**]: "I became very competitive because of the messages."

**(5) Would this kind of experiment work for you in the long-term with taking more steps per day than usual?**
**Group 1:** In general, the answer was no. This had to do with participants either already having increased their step-counts during COVID-19 or having a lot of exercise already in general.

**Group 2:** While one participant said it did not work for him/her, (s)he mentioned that (s)he could imagine it working out for other participants. The other participants did note that they increased their step-counts after the experiment. This was often due to seeing the step-count, rather than the messages.

**Group 3:** Here participants were also positive about the effects in the long-term. The awareness it created about one's own step-count made participants walk more than before. One participant even mentioned that (s)he was going to use such an app again to gain even more awareness of his/her step-count. A citation from a participant [**translated to English**]: "I think the awareness about the number of steps does work for the long-term, I notice that I watch my step-counts more than I did before (just because of some number on my home screen)."

**In general, participants liked participating in the experiment. The messages were not necessarily creating an influence on the step-counts, but the cue of the messages and the display of the step-count did. This is an interesting result, that could not be found when looking at step-counts and the surveys alone. Some suggestions were made as well, for example regarding the use of an activity tracker rather than a mobile phone. This was because the phones were often forgotten and their step-counts seemed inaccurate.**

## 5.5 Weather

The weather was also recorded during the experiment. This was to check for any irregularities concerning the weather influence on the number of steps set per day. While this research aims to answer which strategies are the most persuasive for walking more, and whether there is a difference between perceived and actual persuasion, the weather could have an influence on its own. Just in case the weather was indeed a variable with influences on its own, it was recorded as well, so see whether it had any influence on the step-counts of participants.

The data about the weather was gathered from https://www.wetterzentrale.de/weatherdata.php?station=260. The temperature (maximum, average and minimum), the rain (millimetres in total and the number of hours that it rained on that day), and the strength of the wind (in meters per second) were recorded from 'De Bilt' which lies close to Utrecht. For the exact numbers, see Appendix A.4. This place was chosen to record because most of the participants were students from Utrecht University. Another reason to choose a place near Utrecht (as Utrecht itself was no option on this site), was that it is also in the center (and thus some average) of the Netherlands, where the experiment took place. *Note: two participants were in India during the experiment and some participants might have had a holiday in another place than the Netherlands.*

The coldest day had a maximum temperature of 19.9 and an average temperature of 16.2 degrees Celsius. It also rained on that day. The hottest day had a maximum temperature of 35.5 and an average temperature of 26.9 degrees Celsius, with no rain. During the experiment, there was rain on about 9 out of 35 days, with only 31.6 millimetres of rainfall in total. This is rather typical weather of the summer in the Netherlands, although it might have been a bit dry, as several countries including the Netherlands feared a drought a few days after the experiment.

On the coldest day, participants walked between 56 and 12770 steps with a median of 4526 and an average of 4338 (this was also the lowest average recorded). On the hottest day, participants walked between 118 and 20742 steps with a median of 4582 and an average of 5368.

The lowest median step-count was 3415 on the seventh day of the experiment with a maximum temperature of 21.2 an average temperature of 16.7 degrees Celsius and with 0.6 mm of rain.

The highest median step-count was 7193, which was on the fifth day of the experiment. The highest average step-count recorded was 7905, and was on the same day. This day had a maximum temperature of 20.5 and an average of 16.2 with no rainfall.

**In general, it was a normal, hot and dry summer in the Netherlands during the experiment.**

## 5.6 Step-count data

In this Subsection, the results of the step-count data are considered. First, an overview of individual participants is given, followed by the comparison of experimental groups. This is needed to answer the research questions about the most persuasive strategy and the difference between perceived and actual persuasion. The most persuasive strategies would be found in the group that increased its step-count the most during the experiment. The difference between perceived and actual persuasion could be inferred from comparing the results of the pre-study with the step-counts of the experiment. This is to see whether the most perceived persuasive strategies were also used in the experimental group that had the highest increase in step-counts.

Since this research is about the long-term effects of persuasion, the first and last week of the experiment are the most relevant weeks to look at. The first week represents the normal routine of each participant, while the fifth and last week shows the step-count after having received persuasion. The weeks in-between are the weeks with daily messages of persuasion, which is interesting when looking at the immediate effects of persuasion only. For the results, only the step-counts that were sent per screenshot were considered. This is because these were more reliable that the step-counts used during the experiment, due to connectivity issues. The goal here is to see whether some groups had higher or lower step-counts than other groups. If group 1, has higher step-counts than group 2 and especially group 3, the three hypotheses could be verified: Perceived persuasion is similar to actual persuasion and there are strategies that work better than others (e.g. praise works better than punishment).

### 5.6.1 Step-counts per Participant

Here, step-counts will be compared between participants. This is to answer the research question of which strategy is the most persuasive for walking more. When many participants in one group have increased their step-count (more than in other groups), the most persuasive strategy can be found in that group.

In Appendix A.4, a line plot is shown with the step-counts of each participant, divided by groups in different colors. Some participants had step-counts above 20,000 steps, those participant lines are shown again but separate in Appendix A.5. In Figure 5.10, the box plots of step-counts of each participant are displayed for the first benchmark week. The participants are partitioned into three experimental groups, and the combined means per group are shown. It is visible that group 1 had higher average step-counts to begin with, in the first benchmark week than group 2, which in turn had higher step-counts than group 3. The variation in group 1 seems to be the highest, as participants 1 and 10 had relatively high step-counts.

In Figure 5.11, the box plots of step-counts of each participant are displayed for the fifth and final week of the experiment. Again, the participants are partitioned into the three experimental groups, and the combined means per group are shown. Here it can be seen that group 2 had a higher average step-count than the other groups. The variation seems bigger than in the first benchmark week, with participants 2, 13 and 23 having relatively high step-counts. Another interesting insight here is that most participants that tend to have few step-counts, also have low deviations, while participants with more step-counts, also have higher deviations in their step-count.
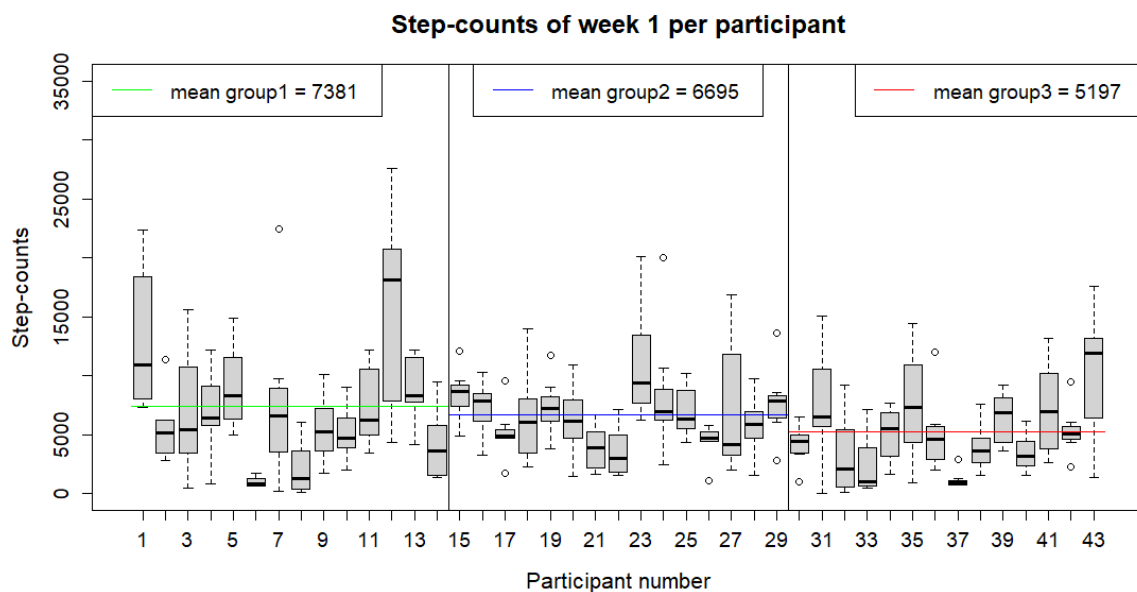


**Figure 5.10: Box plots with means included of week 1.**

By comparing Figure 5.10 to Figure 5.11, the most noticeable changes are that participant 1, 12, 27, 38 and 41 have decreased their step-counts, while participants 2, 13, 15, 16, 23 and 42 have increased their step-count. In group 1, seven people have decreased and seven people have increased their average step-count per week. In group 2 six people have decreased and nine people have increased their average step-count per week. In group 3 three people have decreased and eleven people have increased their average step-count per week. This would mean that group 3 has had the most people that increased their step-counts, more than group 2 and especially more than group 1.

**This has two implications: (1) The strategies used in group 3 were not less persuasive and useful than in group 1. They were more persuasive than the ones in group 1, since more people increased their average step-count in group 3 than in group 1. (2) There is a difference in perceived persuasion in surveys compared to actual persuasion that relates to changes in behavior. While the most positive effect was expected to be found in group 1, it was found in group 3. So, while information, praise and rewards came out on top with perceived persuasion, it seems that competition, punishment and consensus had a bigger**
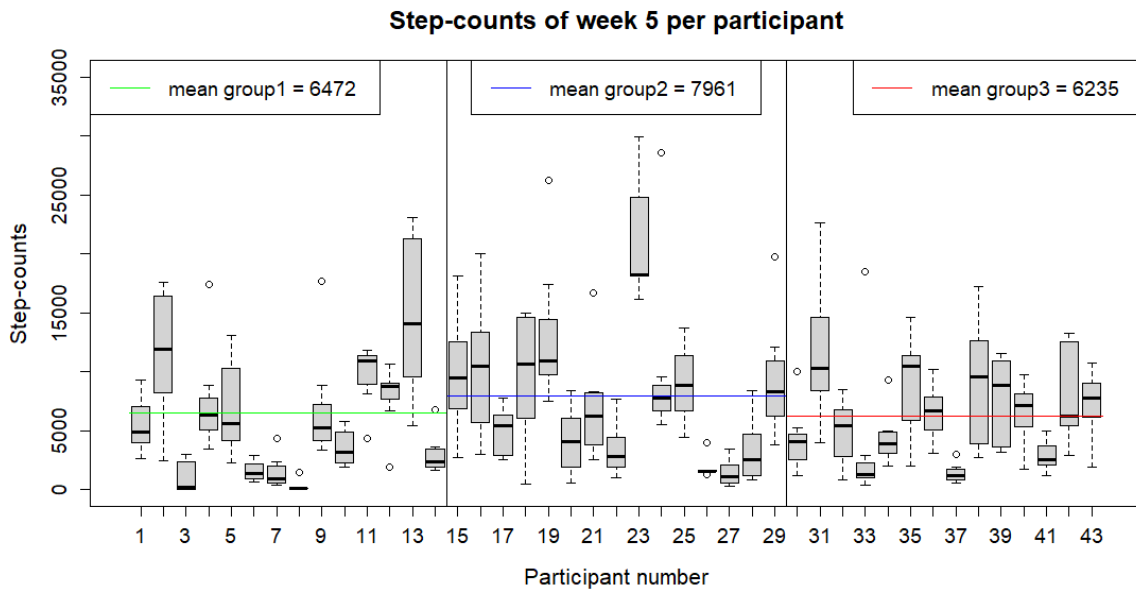
**Figure 5.11: Box plots with means included of week 5.**

**positive impact on the step-counts.**

### 5.6.2 Step-counts per Group

Here, the step-counts will be compared between groups, rather than between participants. This is done to answer the research questions regarding the best persuasive strategy and the difference between perceived and actual persuasion. The group that has the most increase in step-counts, would have been the group with the most persuasive strategies. The increase in step-count per group is also compared to the results of the pre-study in Subsection 6.8. This is done to see whether the most perceived persuasion strategies are similar to the strategies that were actually persuasive during the experiment.

In Figure 5.12, the box plots of the step-counts of all three groups are presented for each of the five weeks of the experiment. In Appendix A.10, a similar box plot is shown, but for every day of the experiment, rather than weeks. While the weeks within each group can be compared, the differences between groups can be seen as well. In general, the average step-counts per week lie between 2,500 and 10,000 steps, but there are outliers of above 20,000 steps on average per week. In weeks two, four and five group 2 had the highest step-counts. For weeks one and three, group 1 has the highest step-counts. Group 1 seems to alternate between decreasing and increasing their average step-count per week, resulting in a lower step-count for week five than week one. Group 2 seems very stable in their step-count, but it increases from week four to week 5. Group 3 had a slight increase in their step-counts in weeks two and three, but lost that increase in week four, to increase it again in the fifth week.

In Figure 5.13, the groups are separated in Figure 5.12, to be able to see the differences within groups more clearly. The means per week are added as well as a line in this figure, but the means can also be found in Table 5.11. Aside from group 1, group 2 and group 3 seem to have an overall increase in step-counts during the experiment. Group 1 was alternating the decrease and increase in their step-count means, which resulted in a lower average for week five than in week one. Group 2 continuously increased their mean step-counts, except in week four. Group 3 was similar to group 2, but has a barely noticeable decrease in means in week three as well. Week four seems to be a bad week, as every participant group had a decrease in the average step-count of that week.

This could be due to the last and final week of the persuasive messages. This would mean that the longer persuasive messages are sent, the less effect they would have on increasing the step-count. Another possibility is that the weather was worse in week four, or that fewer people were on holiday in week four. There were 62 out of 301 (7x43) holidays in week four, which was similar to week five, so that is not necessarily why week four had a lower average than the other weeks. However, the second day of the fourth week did have the second lowest median and mean of the whole experiment, so that might have
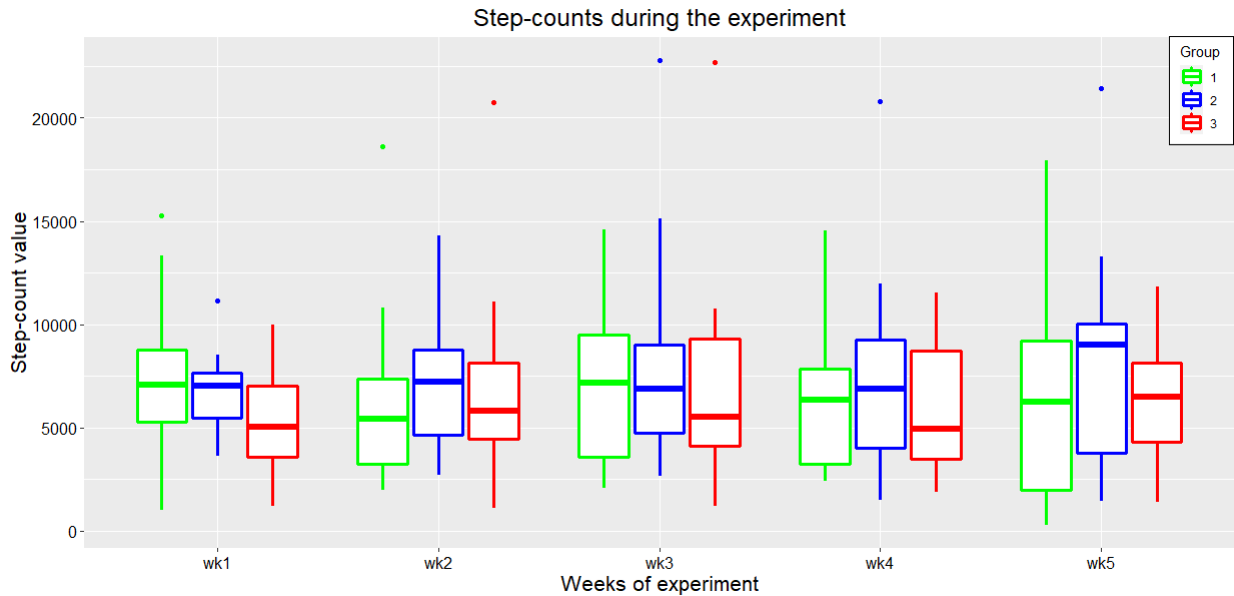
**Figure 5.12: Box plots of each week of each group during the whole experiment.**

been another reason why the average of week four is lower than that of the other weeks.
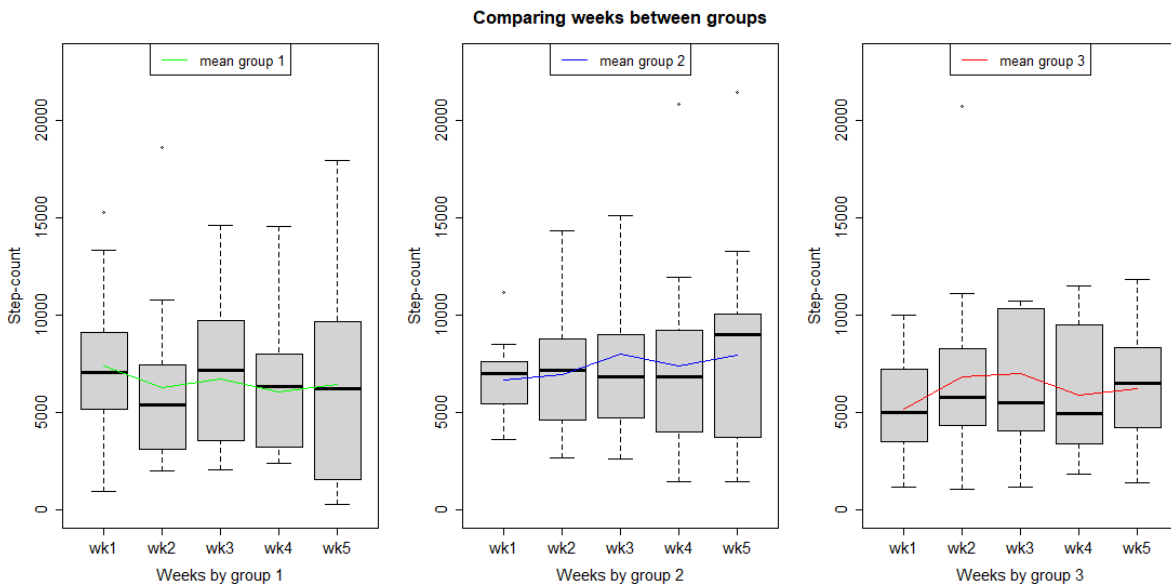


**Figure 5.13: Comparing the step-counts from the three groups by weeks with means included.**

| Mean step-count | week 1 | week 2 | week 3 | week 4 | week 5 |
|---|---|---|---|---|---|
| Group 1 | 7381 | 6269 (-) | 6423 (+) | 6297 (-) | 6472 (+) |
| Group 2 | 6695 | 6955 (+) | 8027 (+) | 7304 (-) | 7961 (+) |
| Group 3 | 5197 | 6854 (+) | 6840 (±) | 5874 (-) | 6235 (+) |

**Table 5.11: Table with an overview with means of the step-counts per group per week. The sign between brackets shows whether it was an increase or decrease.**

Since this research is mainly about comparing the step-counts from the benchmark week to the long-term effect in the last week, Figure 5.14 displays the box plots of the step-counts in week five minus the step-counts of week one, per group. The means per group per week are included in both Figure 5.14 and

Table 5.11. An average, week five had 465 steps more than week one. The median and mean of group 1 lie below zero, meaning that participants in group 1 did generally not increase their step-counts from week one compared to week five. Group 2 and group 3 have a similar median and mean, that are above zero. This means that group 2 and 3 did increase their step-count from week one compared to week five. Group 2 however did have more variation in their step-counts. There are two possible outliers here. Participant 23, which was mentioned already in Subsection 5.6.1, did have an increase of 10,284 steps. Participant 41, which was also mentioned in Subsection 5.6.1, did have a decrease of 4,324 steps.

   **The third hypothesis was that perceived persuasiveness is similar to actual persuasion. If this was true, then group 1, with strategies that have the most perceived persuasiveness would have increased their step-count more than the other groups. At the same time, group 3, with the least perceived persuasive strategies would have the lowest increase in their step-counts, or even a decrease. This however not the case, when looking at the data of the experiment. While group 1 had a decrease in step-counts, group 2 and group 3 had an increase in step-counts. Thus, the three strategies that were perceived as the least persuasive strategies did have a more positive impact on the step-counts than the strategies that were perceived as the most persuasive. In fact, the strategies that were perceived as the most persuasive, did have a negative influence on the step-counts.**
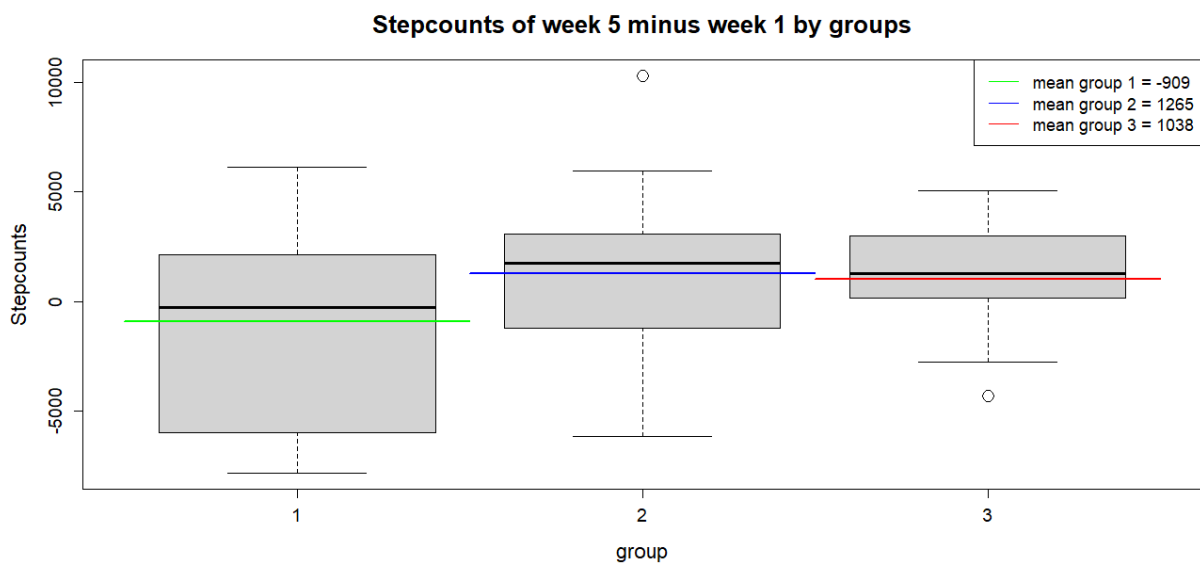


**Figure 5.14: Step-counts of participants of week 5 minus those in week 1, separated by group number.**

# 6 Statistical Analysis

In this Subsection, the analysis of the results of the step-counts, the theory of behavior, the TTM and the messages are discussed. Various statistics are used to compare the results between groups. The results from before and after the experiment are also analyzed. This is to see whether the research questions can be answered and whether the hypotheses were true. The hypotheses were that perceived persuasion was equal to actual persuasion, and that there are specific persuasive strategies that work better in the context of this experiment. To check these hypotheses, there need to be significant differences in the step-counts. Below the significance tests are used on the step-counts.

## 6.1 Step-Counts

First of all, a `Shapiro test` was done to check for a normally distributed population. For each group, a Shapiro test was done on the step-counts of the first week of the experiment. For group 1, the `p-value` was 0.560, for group 2 it was 0.719 and for group 3 it was 0.974. These values are all above 0.05, so we accept the null hypothesis that the step-counts of the participants of each group were normally distributed.

Furthermore, when a Shapiro test for normality was done on all participants together for week one, the `p-value` was 0.07, which shows that all step-counts of all participants together also behaved like a normal distribution.

When the strategies that were perceived to be the most persuasive (group 1), would have the highest increase in step-counts, an answer to one of the research questions would be that perceived persuasion is similar to actual persuasion. The other research questions would be answered by finding the group with the highest increase in step-count, and relating those participants back to their personality profiles. The first one answers the question of the best persuasive strategies for walking more, and the second one answers whether there are certain strategies that tend to work better for specific profiles.

First, comparisons are made between weeks by all participants together. Then, comparisons are made between groups, which are followed by comparisons within groups. Finally, some analysis is done on the variables of the experiment and on holidays as well.

### 6.1.1 Comparison with all Participants

When considering all participants of the experiment, there was almost a significant difference between week one and week three through a `T-test` with a `p-value` of 0.083. The mean of week one was 6431 and the mean of week three was 7118, so there was a difference (an increase) of 687 steps, which was almost significant.

The comparison of all participants between week three and week four was also almost significant, with a `p-value` of 0.072. The mean of week four was 6511, so there was a difference (a decrease) in the number of steps towards week four that was almost significant. In Subsection 5.6.2, there was already a discussion of why week four could have lower step-counts. One explanation was that the day with the second lowest median and mean were included in week four. The other explanation was that persuasive messages have less of an effect when given for longer periods of time. The actual difference (decrease) was however not significant according to the `T-test`.

**When all participants were taken together, there was no significant difference found when comparing the experiment weeks to each other. This means that the research questions of the most persuasive strategy and the difference between perceived and actual persuasion cannot be answered, as there are no differences between the strategies.**

### 6.1.2 Comparison Between Groups

For the comparisons between groups, a normal `T-test` was used. There were two significant differences in week one. The comparison between group 1 and group 3 had a `p-value` of 0.005. The mean of group 1 was 7381, while the mean of group 3 was 5197, so there was a significant difference of 2184, such that group 1 had done more steps in week one than group 3. The comparison between group 2 and group 3 had a `p-value` of 0.006. The mean of group 2 was 6695, so there was a significant difference of 1498, such that group 2 had a higher step-count than group 3. This means that group 1 had walked the most in week one, while group 3 had walked the least in week one.

Furthermore, there was almost a significant difference in week three between group 1 and group 2, with a `p-value` of 0.051. Group 2 had walked an average of 7499 that week, while group 1 walked 6383 that week, which is less. This is unexpected behavior, since group 1 was expected to receive the most persuasive messages, which would lead to the highest step-count. This difference was however almost significant.

In week four another difference was found between group 2 and group 3, which was also almost significant with a `p-value` of 0.069. Group 2 had an average of 7301, while group 3 had an average of 5874, which is lower. This was an expected result, as group 3 would have fewer persuasive messages than group 2 to increase the step-count. This difference was however almost significant.

Finally, a real significant difference was found in the final fifth week when comparing group 2 to group 3. This `p-value` was 0.031. The mean of group 2 was 7961, while the mean of group 3 was 6235, which is less. This is in line with the expectations, since group 2 was supposed to have messages that were more persuasive than those of group 3.

**For the comparison between groups, it was found that for week one, group 1 had walked the most, while group 3 walked the least. Since this was the benchmark week, not much can be said about it. The only significant difference in week five was that group 2 had a higher step-count than group 3. This is partly in line with the third hypothesis that perceived persuasion is similar to actual persuasion. It was expected that group 2 had a bigger increase**

in step-count than group 3, since the perceived persuasiveness of its strategies was greater. However, since group 1 had no significant difference (increase) compared to group 2, the answer to the research question is ambiguous.

### 6.1.3 Comparison Within Groups

For the comparisons within groups, a `paired sample T-test` was used. The data of each day within a week was used for this analysis, rather than the average step-count per week. There were no significant differences within group 1 during the experiment. For group 2 and group 3 however, there were multiple significant differences.

For example, when comparing week one against week five in group 2, the p-value was 0.042. The means were 6695 and 9761 respectively, so there was a significant difference of 3066 steps, such that week five had a higher step-count than week one. This is a positive outcome, as it shows that the intervention of sending persuasive messages did help in increasing the step-count of group 2.

For group 2, that was not the only significant increase during the experiment. Another significant difference (increase) was found when comparing week two against week three (p-value = 0.039, mean week two = 6955, mean week three = 8027, increase) and when comparing week one against week three (p-value = 0.029, mean week one = 6955, mean week three = 8027, increase). For group 3 there were also some significant differences. When comparing week one against week two the p-value was 0.011. The mean of week one was 5197 and the mean of week two was 6854. This means that there was a significant difference (increase) for group 3 in week two compared to week one. This outcome shows that the intervention of sending persuasive messages did help in increasing the step-count of group 3, even though it was the group that was expected to have the least increase. Another significant difference (increase) was found in group 3 when comparing week one against week three (p-value=0.009, mean week one = 5197, mean week three = 6840).

Finally, there were some **differences that were almost significant** for group 2 and group 3. For group 2, it was found when comparing weeks one and two combined against weeks four and five combined (p-value = 0.058, mean week one+two = 6825, mean week four+five = 7632, increase). Another difference that was almost significant for group 2 was the comparison between week two and week five (p-value = 0.099, mean week two = 6955, mean week five = 7961, increase). For group 3, it was the comparison between week one and week five (p-value = 0.057, mean week two = 5197, mean week five = 6235, increase).

**Since there were no significant differences within group 1, part of the research question regarding differences in perceived and actual persuasion can already be answered negatively. The group with the highest perceived persuasion strategies, did not significantly increase its step-count, so perceived persuasion is not necessarily the same as actual persuasion. Group 2 had three significant differences (increases), while group 3 had two such differences. This is however in favor of the third hypothesis that perceived and actual persuasion are similar. This is because group 2 had more significant increases than group 3, which had the least persuasive strategies according to the pre-study.**

## 6.2 Possible Outliers

Eliminating certain participants from the results was considered. Participants who had more than three 'outliers' in their box plots were considered for this, see Appendix A.7, A.8 and A.9. However, when those participants were taken out of the results, the statistical tests gave almost the exact same results. Instead of group 2 having a significant difference between week two and week three, a new difference was found between week one and week four that was almost significant (p-value = 0.060, mean week one = 6913, mean week four = 8176, increase).

**While eliminating participants with odd step-counts was considered, it did not affect the results.**

## 6.3 Relative differences

The step-counts were compared between weeks in percentages of increase and decrease. The exact differences in percentages can be found in Table 6.1. When looking at the average differences of all participants

| Comparison | Group 1 (±%) | Group 2 (±%) | Group 3 (±%) |
|---|---|---|---|
| Week 1 vs 2 | -0.2 (-15.1) | +3.0 (+3.9) | +**36.1** (+31.9) ✓ |
| Week 1 vs 3 | +11.0 (-8.6) | +16.6 (+19.6) ✓ | +**37.1** (+34.5) ✓ |
| Week 1 vs 4 | -0.2 (-17.7) | +4.2 (+10.7) | +17.9 (+13.2) |
| Week 1 vs 5 | -7.4 (-12.3) | +14.1 (+18.9) ✓ | +**29.2** (+20.0) |
| Week 2 vs 3 | +**25.8** (+7.6) | +10.5 (+15.1) ✓ | +11.0 (+2.0) |
| Week 2 vs 4 | +16.3 (-3.2) | -0.01 (+6.5) | +5.3 (-14.2) |
| Week 2 vs 5 | +14.4 (+3.2) | +10.7 (+14.5) | +**24.9** (-9.0) |
| Week 3 vs 4 | -4.2 (-10.0) | -7.5 (-7.5) | +1.0 (-15.8) |
| Week 3 vs 5 | -7.2 (-4.0) | +4.6 (-0.6) | +13.3 (-10.8) |
| Week 4 vs 5 | -6.1 (+6.6) | +14.2 (+7.4) | +17.0 (+6.0) |

**Table 6.1: Relative differences between weeks in percentages. The numbers without brackets are the averages of each participant difference. The numbers with brackets are the differences between the means of the group. The ✓ means that the difference was found significant in the analysis with the paired sample T-test.**

together, there are a few big differences in percentages noticeable. The differences bigger than ± 20% are highlighted.

There was one difference above 20% for group 1, which was between weeks two and three. There were no such differences found in group 2. In group 3 however, there were four such big differences found in total (out of ten comparisons). This shows that group 3 had the most relative differences between their weeks in the experiment. These relative differences were positive, which means that group 3 had the most relative increase in step-counts compared to the other groups.

This was an unexpected result, as participants in group 3 were expected to have the least increase (or even a decrease) in their step-counts, since they got the messages with the persuasive strategies that were perceived as the least persuasive.

**Group 3 had more relative increases in their step-count than the other groups. This is in contrast to the third hypothesis that perceived persuasion is similar to actual persuasion. Group 3 with the least persuasive strategies (according to the pre-study) had more relative increase in step-counts than group 1 with the most persuasive strategies (according to the pre-study).**

## 6.4 Variables

For each week, an `ANOVA` and `Linear Regression Model` were tested upon the average step-count data per week to see if any of the variables had an effect on the step-count. The average step-count data per week was used for this, rather than the data of every day within a week, because of the nature of these tests. This was done in order to answer the research question whether there were specific strategies that were more persuasive for certain profiles. The available variables were: experimental group number, age (young, middle or old), gender (male, female or other), main `Brainhex` profile, sub-profile of `Brainhex`, place in the `TTM` and profile in the `STPS`.

For week one, the group variable had a `p-value` of 0.147, but group 3 was almost significant in that model with a `p-value` of 0.057. There were no significant differences found for age, gender and the `STPS` profile.

For week five compared to week two, the sub-profile of `Brainhex` had a significant `p-value` of 0.038. This was mostly due to the low `p-value` of the profile conqueror, which was 0.004. Here, conqueror has the lowest estimate. A similar result was found for week four compared to week two. Here, the was `p-value` of 0.019, also due to the low `p-value` of the profile conqueror, which was 0.001. Again, conqueror had the lowest estimate out of all the profiles, but this time it might have been due to two low outliers. Another significance was found for the sub-profile conqueror in the comparison of week two to week one. The `p-value` was 0.019, due to the low `p-value` of conqueror, which was 0.001. This time however, the sub-profile conqueror had higher estimates than the other profiles. Since the results differ and since it is only a sub-personality profile, the significance is seen as a coincidence.

For the `TTM`, a significance was found when comparing week four to week three, with a `p-value` of 0.035. Since the `TTM` is actually an ordered factor, a quadratic model had significance with a `p-value` of 0.028. When `TTM` was not considered to be ordered, the `p-value` of the text was 0.035, with maintenance

having the highest estimate and a `p-value` of 0.012. This is not expected behavior of the `TTM`, as it is an ordered factor and thus should have a somewhat linear effect. Besides that, the effect was only found in one of the comparisons between weeks, so it is also not a robust effect.

Since the main focus was on comparing week five to week one, an attempt was made to model the difference between week five and week one by the different variables available. When entering the model with all variables, the step-wise AIC only removed age and gender as variables, while going from an AIC of 713 to 710. When a model was entered of week five, based upon all other variables, the step AIC removed the variables: group, gender and main `Brainhex` profile. The variables age, `TTM`, `STPS` and sub-profile of `Brainhex` remained, while the AIC value dropped from 734 to 726. These models are still rather complicated, while the variables that were removed were expected to be more useful (especially group) than the ones that remained. This is visible the most on removing the main profile while keeping the sub-profile of `Brainhex`.

Finally, when the **step-counts per week** were taken as the explanatory variables for the step-counts in weeks, all `paired sample T-Tests` gave significant `p-values` (between 0.006 and $5.155e-13$). This is to be expected, as the step-counts of each week are highly dependant on the step-counts of the previous weeks, as they entail the same (and all) participants each time. Since the data comes from the same participants, no drastic changes were expected.

**Since many variables were explored for this research, it is no surprise that some of them were not significant. The sub-profile of `Brainhex` and the TTM did seem to be significant. For `Brainhex` it was however mainly about one sub-profile that sometimes had fewer steps and sometimes had more steps than the other profiles. Therefore, these significant findings are disregarded. For the `TTM`, a quadratic model was significant, or the 'maintenance' category had higher results than the other categories. This is unexpected behavior, as the `TTM` is meant to be an ordered (linear) factor. Therefore, this result was also disregarded. This means that no variable had significance for the step-counts of participants. The research question regarding specific persuasive strategies for specific profiles can therefore be answered negatively. No profile seemed to have a preference for any specific persuasive strategy.**

## 6.5 Holidays

The holidays are subjected to a significance test, as holidays might have significant differences compared to normal days. This is important to know as holidays could interfere with the step-count differences due to persuasion. For comparing the normal days to the holidays of the same participants, a `paired sample T-test` was used. Holidays with fewer than six days were not considered, as there would be too few data to compare to. Holidays between 6 and 21 days were considered for this test statistic. *Note: One participant had a holiday of 24 days, but in the last three days the connection was lost, so there were three days with 0 as a step-count. These three days were truncated from the results.*

The participants that went on holidays that were considered for this analysis, were about equally distributed among the three experimental groups (eight in group 1, six in group 2 and ten in group 3). The difference in means was 3564, with normal days having a mean of 6212 and holidays having a mean of 9776 steps. The `paired sample T-test` gave a `p-value` of 0.001.

**Since the `p-value` was below 0.05, when comparing normal days to holidays, the null hypothesis can be rejected. The difference in the means between the two groups is not equal to zero. This means that holidays do have a significant influence on the results of the experiment, in contrast to the other (independent) variables. This effect could have an influence on answering the research questions. When higher step-counts for a week are found, the cause could either be the persuasive messages or the number of holidays in that week. The highest effect of holidays could be found in group 3 (ten people), followed by group 1 (eight people), as there were only six people that went on holidays longer than six days in group 2.**

## 6.6 Overview of Step-Count Analysis

In summary, there were several significant results. The results from the variables were peculiar and were disregarded. A summary of all other significant findings in the step-counts can be found in Table 6.2. The step-counts on holidays were significantly different (higher) that the step-counts of normal days. The

| Where | Benchmark | Comparison | P-value | Mean 1 | Mean 2 | Difference |
|---|---|---|---|---|---|---|
| Week 1 | Group 1 | Group 3 | 0.005 | 7381 | 5197 | +2184 |
| Week 1 | Group 2 | Group 3 | 0.006 | 6695 | 5197 | +1498 |
| Week 5 | Group 2 | Group 3 | 0.031 | 7961 | 6235 | +1726 |
| Group 2 | Week 1 | Week 5 | 0.042 | 6695 | 9761 | +3066 |
| Group 2 | Week 2 | Week 3 | 0.039 | 6955 | 8027 | +1072 |
| Group 2 | Week 1 | Week 3 | 0.029 | 6695 | 8027 | +1332 |
| Group 3 | Week 1 | Week 2 | 0.011 | 5197 | 6854 | +1657 |
| Group 3 | Week 1 | Week 3 | 0.009 | 5197 | 6840 | +1643 |
| All groups | Normal days | Holidays | 0.001 | 6212 | 9776 | +3564 |

Table 6.2: Summary of all significant differences found in the step-counts.

most insightful results were that group 1 did not have a significant difference in their step-counts during the experiment, while group 2 and group 3 did have a significant increase in their step-counts during the experiment. Group 3 also had the most relative increase in step-counts during the experiment, while group 2 had a higher step-count in the last week than group 3. The effect of holidays could be the highest in group 3, followed by group 1. While group 3 did have a significant increase, group 1 did not have those. This contrasts with the idea of holidays having an influence on step-counts. Even when holidays did have an effect, they would have the least effect in group 2, which also had significant increases in step-counts. Therefore, it can be said that at least group 2 did have significant increases in step-counts, which were not due to holidays, and possibly group 3 as well.

**Regarding the research questions, the following things can be said. (1) Since no variable was significant for the step-count, no specific strategies were more persuasive for certain personality profiles. (2) Since group 2 and group 3 did significantly increase their step-counts, the most persuasive strategies for increasing exercise can be found in group 2 and group 3. (3) Since group 3 (and group 2) had a higher increase in step-counts than group 1 (with the highest perceived persuasion strategies), the perceived persuasion differs from the actual persuasion. Hence, the answers to the research questions are: (1) '*The persuasive strategies that are most effective in increasing physical activity by exercising can be found in these three to six strategies: personalization, compassion, authority, and perhaps competition, consensus and punishment.*', (2) '*No, there is no specific persuasion technique that has a bigger effect on a personality type*'. For research question three, the top and bottom three strategies in the final survey were also considered along with the step-count differences, see Subsection 6.8.**

## 6.7 Surveys

For this analysis, `paired sample T-tests` were used on the differences in theory of behavior and `TTM`, see Appendix A.5. Differences before and after the experiment were analyzed, as well as differences between nationalities. The analysis was done for self-efficacy, attitude, intention and `TTM`.

When comparing Dutch people to foreigners, no significant differences were found at all. When comparing the results between the three experimental groups, the only difference found was a difference before the experiment began between group 1 and group 3. There were no other significant differences between groups and no significant differences within groups (before compared to after the experiment). In conclusion, there were no significant differences in the theory of planned behavior and `TTM`.

**Therefore, this part of the survey will not be able to answer the research questions about the best persuasive strategies or about the comparison of perceived and actual persuasion.**

## 6.8 Type of Messages

In this section, an attempt is made to answer the research question regarding the difference between perceived and actual persuasion. This is done by comparing the results of the pre-study with the final survey of the experiment and the step-counts of the experiment. For the pre-study, participants had to score sentences that had implicitly embedded persuasion strategies. In the final survey of the experiment, participants could freely enter 6 of the 21 messages that they got during the experiment and rate them.

In the pre-study, it was shown that almost all 124 participants highly preferred the persuasive strategies of information, praise and rewards. The least preferred strategies were competition, consensus and punishment. The groups for the experiment were made distributed among the top three, bottom three and the remaining three strategies accordingly.

The final survey showed that rewards, praise and competition were good persuasive strategies and in that order of magnitude. Punishment was found the least persuasive, along with personalization. This was followed by information and consensus, which had received the same score.

Rewards, praise and consensus remained in the same place among the nine persuasive strategies as before. Competition went upwards from the bottom three to the top three, while information dropped from the top three to the bottom three persuasive strategies. This shows that perceived persuasion in surveys is not the same as actual persuasion used in experiments.

While group 1 had the most perceived persuasive strategies, it had no significant increase in step-counts during the experiment. Group 2 and group 3 however did have significant increases in step-counts, while having strategies that were perceived as less persuasive. This also shows a discrepancy between what people think is persuasive (in surveys) and what is actually persuasive for them in an experiment.

**The comparison between the pre-study and final survey as well as the comparison between the pre-study and the step-counts of the experiment show a discrepancy in the most and least persuasive strategies. This would mean that there is a difference between perceived persuasion in surveys and actual persuasion in experiments. Hence, the research question would be answered with: '*Yes, there is a difference between perceived and actual persuasiveness*' This is however not a clear-cut answer, as the nationalities differed between the pre-study and the actual experiment.**

# 7 Discussion

In this section, the research itself is evaluated. Limitations are addressed and results are related to previous research. The following aspects are discussed: The app that was used to track steps, the means of sending messages, the content of the messages and the surveys. Furthermore, some remarks of participants are mentioned along with some unforeseen circumstances and a discussion of the uncontrollable variables.

## 7.1 The Pacer app

The step-counter app that was used for this experiment (Pacer), was useful such that participants could be put into different groups. It was also easy to make an account without needing to give much personal information. Since the experimenter could look up the step-count data of every participant per day in the app, no step-count data needed to be sent.

In contrast, participants had to send their step-count data (per screenshot) to the experimenter, as it became clear that the Pacer app was not always as accurate or up to date. During the experiment, participants sometimes noted that the step-count of their Pacer app was lower than their smartwatch, activity tracker or even their phone or Google Fit. When two people would walk together for the same distance, the step-counts could differ between hundreds or even thousands of steps. Having battery saving mode on could also lead to fewer steps being recorded.

Another problem was the synchronization of the app. Also, without a steady WiFi connection, the app could not synchronize at all, even at night. This was sometimes the case for people on a holiday or on a festival. The messages sent to those participants were something like: 'You did great yesterday, with 0 steps' followed by another message that their step-count was probably not synchronized well. Participants were asked then to open up the app to synchronize (by opening the app with a WiFi connection it would immediately synchronize), and then close the app again.

So, although it was an easy-to-use app, it was not as accurate as preferred. For a follow-up study, it would be advisable to run the experiment with another app, or maybe an activity tracker.

## 7.2 Blueticks

The chrome extension Blueticks used to send the WhatsApp messages to the participants worked well. Individual messages between the participant and the experimenter could be scheduled to be sent at specific times. There was also an option to send a message every day at the same time. This was used to send the daily messages, while changing the content of the message each day. Aside from individual

messages, there was also an option to create groups of participants and templates of messages to be sent. By upgrading the account, campaigns were unlocked. This means that one can select a group and a template and select when those messages need to be sent. This was primarily used before the experiment officially began, to distribute participant IDs, consent forms and the link to the first survey.

Privacy-wise, it is however not clear whether Blueticks has access to the messages sent via Blueticks. Therefore, the WhatsApp Business account and Blueticks account were deleted after the experiment, along with the messages sent. In all, Blueticks was a nice addition to this research.

## 7.3 Content of the Messages

The messages that were chosen to use for the experiment might not be as representative of their respective strategies. There were some sentences that could also belong to another persuasive strategy. For example, some sentences with praise or reward might mention the average of the group in some way, while it could also be seen as a competition/comparison strategy.

Another possible issue is that the information strategy could look similar to the authority strategy, as the only difference is that some authority states a fact, rather than solely a fact being given.

The difference between praise and rewards was that rewards had either 'achieved' or 'unlocked' or some mention of a trophy in the sentence, besides it being a positive/praiseworthy sentence. This distinction could be good enough, but that could also be investigated in another research.

Lastly, after the experiment, it became clear that for group 2, the participants did receive one and the same message two times in total during the experiment. This could happen due to the randomization factor in sending persuasive messages. None of the participants mentioned it however, until one noticed it when filling in the final survey. This was at the part where participants had to come up with the top 3 and bottom 3 sentences of the experiment and rate them accordingly.

## 7.4 Survey

In general, there were no problems with the surveys. Sometimes a participant could not seem to open the survey, but could do so at a later moment. There were a few cases (3) that a survey was received which was filled in partially, even though there were questions which required an answer to end the survey successfully. These participants were contacted to fill in the survey once again.

For the survey itself, some people preferred it to be shorter and/or available in Dutch. Some questions could have been made more clear: 'walking more steps' could be seen as a bit vague, while it was meant to mean taking more steps per day than you would do on a normal or regular day. A more profound unclear part of the survey, appeared to be the part about attitude against walking more. There was seven times the same sentence ('Walking more steps in the next two weeks would be') but with different scales for the answer possibilities (e.g. useless/useful and unimportant/important). Some people did not notice the scale differences at first glance.

For the data and content of the survey, there were some slight issues as well. Some participants did not answer all sanity check questions correctly, but missed only one out of three. This could be due to a miss-click or looking at the wrong column or row. Those participants were asked whether they could remember some of the sanity checks (they did notice them). Therefore, those participants were not excluded. In hindsight, one participant clicked wrongly on the gender identification, but the males versus females were still evenly distributed among the three experimental groups.

As mentioned before, the nationality question and kind of phone used (Android/Apple) were only added to the final survey, as they were not expected to have an impact before.

Lastly, participants found it hard to judge the top three and bottom three sentences of the 21 sentences during the experiment. Some said neither sentence had an effect, some required an excel sheet where they could look up their messages (without the interference from chats in-between).

## 7.5 Remarks of Participants

Out of the 43 participants, 15 showed interest in talking about the experiment, after it was done, of which 12 actually replied. Those participants liked the experiment in a general sense. There were some issues however, regarding forgetting the phone when walking or the difference between iPhones and Androids in the interface of the app. There was also the issue of making the experiment and its survey in Dutch, as many participants were native Dutch people. Most participants said that the messages itself did not have an influence on their step-count. Exceptions were that the messages came across as nagging or that the

messages made someone really competitive. The most impact and increase on step-counts was said to be from the increased awareness of participants' own step-count, which was visible through the (Pacer) app. So, having a visible step-count to look at could very well increase one's awareness of one's step-count as well as the actual increase in step-counts.

## 7.6 Unforeseen circumstances

The goal of this research was to personalize persuasion messages. To do this, personalities (either via `Brainhex`, STPS or TTM) were to be matched with their corresponding best and worst persuasive strategies [§]. To do this, a pre-study was held to see which personality profiles preferred which persuasive strategies. According to the results of the pre-study, there was however no personalization possible. The pre-study did show that there was a general preference for three strategies (out of the nine presented), and that there was also a bottom three. Therefore, the 'personalization' was changed into a more general direction. People who want to walk more steps than they would normally, prefer the strategies: praise, information and rewards. The least preferred strategies are competition, consensus and punishment. The other three strategies in the middle were: compassion, authority and personalization. Reminders and tips were sent to each experimental group.

It is unsure whether the problem with personalization for this research was a coincidence, or would be true in general for this specific area of persuasion in health exercise with walking more steps than usual. More research would be needed to see if there is indeed a general preference for persuasion for walking more, or whether there can be personal profiles and thus personalized persuasion.

## 7.7 Uncontrollable variables

Here some variables are discussed that could not be controlled during the experiment. These include the holidays of participants, getting sick, and the influences of the weather. While holidays could and did have a positive influence of step-counts, illnesses could have negative influences, while the weather could influence step-counts in both ways.

### 7.7.1 Holidays

Because the experiment took place in the summer (the end of June till the end of July), many participants went on holidays. The number of holidays per participant differed from 0 up and to 24 (out of 35 experiment) days. [¶] When participants were on holiday, they generally walked more steps per day than they normally would. This difference in steps was also higher than the persuasive messages could achieve. While holidays had an influence of increasing the step-count on average by 3564, the increase due to persuasion lay between 1072 and 3066 steps.

This could mean that more holidays per person per year could lead to more steps being set per day, which could lead to better health overall for those people. It is not yet known why participants walk more during holidays. This could be due to the fact that they want to make the most out of their holidays. Another possible explanation could be that people generally want to walk more, but cannot do so because of other obligations like work. For this research, there was no data collected in other seasons than the summer. Therefore, it cannot be known whether, for example, winter holidays also have an increased effect on the number of steps set per day.

On the other hand, some participants did work in a supermarket at either the beginning or end of the experiment, which also increased their step-count. This is due to the way of working. By working in a supermarket, you would walk and stand more than when relaxing at home.

### 7.7.2 Illnesses

Some participants either had COVID at the start of the experiment or got COVID during the experiment. This did affect them to walk less than they would normally. Others just felt ill on some days or had trouble walking because of either long-term illness (e.g. arthritis) or a short-term injury.

---

[§]Out of a total of nine persuasive strategies with two more general strategies for every experimental group: reminders and tips

[¶]The one participant with 24 holidays, had lost connection with the app on the last 3 holidays, so it displayed 0 as a step-count.

### 7.7.3 Weather

The weather during the experiment was also measured. On sunny days, people would be more eager to go out for a walk than when it rains by common sense. This was also visible in the data. However, there were also people who noted down the weather as being too hot to walk comfortably outside. There was one day, which was recorded as the hottest day of the year (in the Netherlands). This was the 19th of July with 35.5 degrees Celsius. On that day, only the average of group 2 was above 6000. The median of group 3 was even below 4000 steps. Therefore, warmer days are not guaranteed to lead to higher step-counts, as it could also be too hot to walk around comfortably and without getting a sunburn or heatstroke.

## 7.8 Relating Results to Previous Literature

In this subsection, this research is reflected upon by comparing it to previous research.

There were no visible differences in the persuasiveness of the strategies per personality type, either in the pre-study or in the experiment. Therefore, no comparisons can be made based upon previous research on personalized persuasion like in Table 2.7 for the `Brainhex` profiles. Although research showed that differences in nationality could lead to different results (Almutari and Orji, 2021; Orji et al., 2013a), most participants were Dutch (81%), so that had no big impact on this research. There were also no differences in persuasiveness in age and place in the `TTM` and `STPS`. Therefore, no comparisons can be made on this either. Rather, the individual persuasive strategies used in this experiment will be compared to their effects in previous research.

Although **self-monitoring** was not a persuasion strategy used for the messages, it was part of this research. Self-monitoring is also one of the most used strategies, especially in the domain of health (Aldenaini et al., 2020a; Alqahtani et al., 2019; Oyebode and Orji, 2020a). Self-monitoring is often the most persuasive strategy (Aldenaini et al., 2020a,c; Oyebode et al., 2020b), or the second most (Aldenaini et al., 2020a; Alqahtani et al., 2021a; Orji et al., 2014b), or the third most persuasive strategy (Orji et al., 2017b; Oyebode et al., 2021a,b). In general, most tracking research was done with younger people, as they already have phones with integrated tracking (Aldenaini et al., 2020c). According to Oyebode et al. (2021b), people in the precontemplation stage of the `TTM` would benefit the most from self-monitoring, as it could raise the core things (self-awareness and consciousness) in that stage of change. There were no differences found within the `TTM` for this research however. Through the interview with participants, it also became clear that self-monitoring and thus seeing your step-count on your home screen was very effective at persuading you to set more steps per day.

Information, compassion, consensus, authority and punishment were rather new strategies to be used for research on persuasion, especially in the domain of exercise and health. Authority and punishment did have a result to compare this research to. People did have the least preference for punishment in the research by (Orji et al., 2017b). Authority did come in fourth place with the research of (Oyebode et al., 2020b). Therefore, more research needs to be done on these strategies to be able to compare the results of this research.

**Praise** had rather different outcomes in previous research. Sometimes it was one of the least persuasive strategies (Alqahtani et al., 2021a; Orji et al., 2014b; Oyebode et al., 2020b), sometimes it was somewhere along the middle (Aldenaini et al., 2020a,c; Oyebode et al., 2021a), but sometimes it was the second most persuasive strategy found (Oyebode et al., 2021b). Alqahtani et al. (2021b) propose to use praise anyways, as people prefer and expect to have in a (health) app.

In this research, praise was considered to be the second most persuasive strategy in the pre-study, as well as in the final survey. The experimental group with this strategy however (group 1), did not have a (significant) increase in step-counts. So even though people think they like praise, it seems as if it is not as effective.

**Rewards** also had mixed results in previous research. Sometimes it was one of the least persuasive strategies (Alqahtani et al., 2021a; Orji et al., 2014b; Oyebode et al., 2020b), and sometimes it was the fourth most persuasive strategy or in the middle (Aldenaini et al., 2020a,c; Oyebode et al., 2021a). Rewards could promote health by incentivizing users and by making the behavior fun and appealing. However, the focus might be on the extrinsic reward and thus trivialize and undermine the benefits of the behavior or they are seen as the only benefit of adopting healthy behaviors (Kappen and Orji, 2017; Orji et al., 2014b). Orji et al. (2018b) propose to use (extrinsic) rewards anyways, as changing behaviors for better health can take quite long and the intrinsic reward will take a while.

In this research, rewards were considered to be the third most persuasive strategy in the pre-study, and the most persuasive strategy in the final survey. The experimental group with this strategy however

(group 1), did not have a (significant) increase in step-counts. So even though people think they like rewards, it seems as if it is not as effective.

**Personalization** had a slightly more stable result. In most of the previous research, it ended up somewhere in the middle of persuasive strategies (Orji et al., 2014b; Oyebode et al., 2020b, 2021a). In the research of Alqahtani et al. (2021a), it was one of the least persuasive strategies, but in the research of Aldenaini et al. (2020a,c), it was the second most persuasive strategy.

In this research, personalization came in fourth place in the pre-study, but it was the second least persuasive strategy in the final survey. The experimental group with this strategy however (group 2), did have a (significant) increase in step-counts. So even though people think they do not prefer personalization, it seems that it is effective.

**Competition** had slightly better results in previous research. While it was one of the least persuasive strategies or along the middle in some research (Aldenaini et al., 2020c; Oyebode et al., 2020b, 2021a), it was the top best strategy in other research (Almutari and Orji, 2019; Orji et al., 2014b,a). Competition and comparison can motivate behavior change by making it fun and interesting and letting people commit to the behavior. On the other hand, it could also lower self-esteem and confidence and threaten one's privacy (Kappen and Orji, 2017).

In this research, competition was the third least persuasive strategy in the pre-study, but it was the third most persuasive strategy in the final survey. The experimental group with this strategy (group 3), did have a (significant) increase in step-counts. So even though people in the pre-study did not think that they preferred competition, participants from the experiment did think it was persuasive. The experiment also showed that competition was a persuasive strategy, as the step-count increased with this strategy. In the interviews, it became known that participants would have found it even more persuasive, if they knew the other people that they were competing with.

**Reminders** were considered quite persuasive in previous research. It either came in first place (Aldenaini et al., 2020a), third place (Aldenaini et al., 2020c; Oyebode et al., 2020b) or fourth place (Oyebode et al., 2021a,b). According to Oyebode et al. (2021b), people in the actions stage of the TTM would benefit most from reminders and suggestions. There were no differences found within the TTM for this research however.

In this research, reminders were not considered for the pre-study but were used in all experimental groups. In the final survey, it was the seventh most persuasive strategy (out of eleven). According to its average score in the final survey, it had no effect on walking more or less. This difference in persuasiveness could be due to the difference in the implementation of reminders. In previous research, it was more used as notifications, than actual text messages incorporating reminders.

**Tips** were compared to the 'suggestion' strategy, as they encapsulate the same meaning, see Table 2.2. While it was the eighth persuasive strategy in one research (Aldenaini et al., 2020c), it was the most or second most persuasive strategy in many types of researches (Orji et al., 2014b; Oyebode et al., 2020b, 2021a,b). According to Oyebode et al. (2021b), people in the actions stage of the TTM would benefit most from reminders and suggestions. There were no differences found within the TTM for this research however.

In this research, tips/suggestions were not considered for the pre-study but were used in all experimental groups. In the final survey, it was the sixth most persuasive strategy (out of eleven). According to its average score in the final survey, it had no effect on walking more or less. This difference in persuasiveness could be due to a difference in implementation between tips and suggestions. While tips convey rather neutral information, suggestions might be more tailored and/or push people more to do actions.

**In general, previous research had mixed results on how persuasive each strategy was. For this research, there were also discrepancies between the results of the pre-study, the step-counts and the final survey. This could be due to nationality differences or domain differences. Even though some strategies were not very effective at persuading people (like praise), researchers advice to include it anyways, as people expect to receive them and like them.**

# 8    Conclusion

This research seemed to conclude that there is no personalized persuasion possible for walking more steps per day. By slightly changing the experiment set-up, different persuasive strategies were tested in three different experimental groups. The most preferred strategies were used in group 1, the least preferred in group 3 and the others in group 2. It was expected to see the most improvement in group 1 (preferred persuasive strategies) and the least improvement in group 3 (least preferred strategies). However, it seemed

that people have differences in thinking about what would help them change their behavior and what seems to be actually improving their behavior into walking more steps per day (compared to now/before).

While group 1 had the strategies that were perceived as the most persuasive, it was the group with the least increase (none) in step-counts. Group 3 had the most relative improvement of all groups, while it was expected to have the least increase (or even a decrease in the number of steps). Holidays did have a significant difference compared to normal days. Group 2 had the least number of holidays, while it had the most significant differences (increases) between weeks during the experiment. Therefore, the strategies used in group 2 and group 3 could be considered the most persuasive in reality.

It is interesting to note that while praise and rewards were perceived as most persuasive in both the pre-study and in the final survey, they are not the actual most persuasive strategies. While competition was considered not so persuasive in the pre-study, it was considered persuasive in the final survey and in the increase in step-counts. Participants noted that it would be even more persuasive if they knew the other people who they were competing with. The differences between the pre-study and the final survey could either be due to nationality differences or differences between perceived and actual persuasion.

From the interviews, it became clear that an activity tracker would be preferred over a mobile phone, as it tends to be forgotten less while being more accurate. Participants told that the messages were not persuasive per se, but the cue of them, along with the visible current step-count on their phone were persuasive. Participants walked more steps per day than before because of that, even in the long-term.

So, there was a discrepancy between the results of the pre-study, the results of the final survey, and the differences in increase in step-count among the three experimental groups. Therefore, more research would be needed to make an accurate identification of which strategies would benefit most people for getting more exercise and whether personalized persuasion is an option in this domain.

# 9 Future Work

In this Section, some suggestions for future research can be found. These suggestions are made based on the experiences of this research. The main areas of concern are the app for tracking steps and the scale of the research (number of participants as well as the longitude of the study). Some suggestions are made about persuasive messages and the domain of research.

## 9.1 App for Measuring Step-Counts

First of all, another app than the **Pacer** app could be used to run this experiment. This is advisable, as the synchronization of the app could have been better. Some steps were not taken into account, or the step-count was inaccurate because of a lack of synchronization. It would be advisable to use an app that either has a leader-board display, or has an export button for the step-count data, so that participants do not need to send screenshots of their data. Furthermore, the app should be the same for Android and iPhone users, because otherwise only participants with one kind of phone could participate in the experiment, to avoid confounding variables. A notification or widget displaying the step-count is also advisable, as participants noted that as one of the positive impacts on their step-count, as it made them more aware of their step-counts.

## 9.2 Scale of the Research and Experiment

Another issue with this research might have been the scale. This research was done with 43 participants in the summer, for five weeks. While 43 participants might be enough to measure differences, larger research is needed. This research is needed to ascertain the possibilities of persuasion in health, as it is such an important topic and could be widely profitable for everyone if it shows promising results. Another possibility to increase participants for such research is to include surveys and messages in the native language of each participant. Sometimes a participant left the experiment beforehand, just because (s)he had difficulties in understanding English.

There was only one week of measuring step-counts as a benchmark. This could be enough, but there is always a possibility that that particular week was not a regular week for some of the participants. Therefore, the study should include more than one benchmark week. Also, to really study the long-term effects of the research, longer research would be needed than one week of measuring step-counts after the persuasive messaging weeks. Lastly, this research could also be done in another season, like fall or spring or even winter, to see if there are differences in this kind of research per season. While people may like

to go out in the summer, it may also be the case in other seasons. For example, when there is snow in the winter or falling leaves in the fall or plants starting to bloom in the spring.

## 9.3   Messaging

For this research, WhatsApp was used to distribute persuasive messages in the morning, however other possibilities could also be investigated. For example, the messages could come through an app, or via face-to-face communication. The content of the messages could also be different. The messages used for this research were mostly validated through a focus group of seven people, but some messages might have been ambiguous as to what strategy they belonged to. In the future, a pre-study could be used to see if each message is persuasive and whether it can be deduced as to which strategy it belonged to. Another aspect that could be changed regarding the messages, is the language that is used. Most participants for this research were not native English speakers and preferred to have the survey as well as the messages in their own native language (Dutch). This might then also relate to the scale of the research.

## 9.4   Definition of Exercise

This research could also be created entirely different. For now, the focus was on setting more steps per day than regularly as the exercise. There are plenty of other ways to do research about exercising more with persuasion. For example, more closely related to the company WeGuide, research could be done on re-validation exercises. The amount of exercise and the time spent on them could be evaluated, regarding (personalized) persuasive strategies being tested on the participants. This could be in the form of messaging (as with this research), but it could also be face-to-face or via an app for example.

# References

Creating persuasive technologies: An eight-step design process. *Persuasive '09: Proceedings of the 4th International Conference on Persuasive Technology*, 350:44, 2009. doi: 10.1145/1541948.1542005.

A. M. Abdullahi, K. Oyibo, R. Orji, and A. A. Kawu. The influence of age, gender, and cognitive ability on the susceptibility to persuasive strategies. *Information*, 10(11), 2019. ISSN 2078-2489. doi: 10.3390/info10110352. URL https://www.mdpi.com/2078-2489/10/11/352.

A. M. Abdullahi, R. Orji, A. M. Rabiu, and A. A. Kawu. Personality and subjective well-being: Towards personalized persuasive interventions for health and well-being. *Online journal of public health informatics*, 12(1), 2020. doi: 10.5210/ojphi.v12i1.10335.

I. Ajzen. Constructing a tpb questionnaire: Conceptual and methodological considerations. Technical report, University of Massachusetts, 2006.

N. Aldenaini, F. Alqahtani, R. Orji, and S. Sampalli. Trends in persuasive technologies for physical activity and sedentary behavior: A systematic review. *Frontiers In Artificial Intelligence*, 3, 2020a. doi: 10.3389/frai.2020.00007.

N. Aldenaini, R. Orji, and S. Sampalli. How effective is personalization in persuasive interventions for reducing sedentary behavior and promoting physical activity: A systematic review. In *persuasive*, 2020b.

N. Aldenaini, O. Oyebode, R. Orji, and S. Sampalli. Mobile phone-based persuasive technology for physical activity and sedentary behavior: A systematic review. *Frontiers in Computer Science*, 2, 2020c. doi: 10.3389/fcomp.2020.00019.

M. Alhasani, D. Mulchandani, O. Oyebode, and R. Orji. A systematic review of persuasive strategies in stress management apps. In *Behavior Change Support Systems at Persuasive*, 2020.

M. Alhasani, A. Alkhawaji, and R. Orji. Mental health and time management behavior among students during covid-19 pandemic: Towards persuasive technology design. *medRxiv*, 2021. doi: 10.1101/2021.10.01.21264409. URL https://www.medrxiv.org/content/early/2021/10/01/2021.10.01.21264409.

N. Almutari and R. Orji. How effective are social influence strategies in persuasive apps for promoting physical activity? a systematic review. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, page 167–172, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450367110. doi: 10.1145/3314183.3323855. URL https://doi.org/10.1145/3314183.3323855.

N. Almutari and R. Orji. *Culture and Health Belief Model: Exploring the Determinants of Physical Activity Among Saudi Adults and the Moderating Effects of Age and Gender*, page 138–146. Association for Computing Machinery, New York, NY, USA, 2021. URL 10.1145/3450613.3456826.

F. Alqahtani and R. Orji. Usability issues in mental health applications. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*, UMAP'19 Adjunct, page 343–348, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450367110. doi: 10.1145/3314183.3323676. URL https://doi.org/10.1145/3314183.3323676.

F. Alqahtani, G. A. Khalifah, O. Oyebode, and R. Orji. Apps for mental health: An evaluation of behavior change strategies and recommendations for future development. *Frontiers in Artificial Intelligence*, 2:30, 2019. ISSN 2624-8212. doi: 10.3389/frai.2019.00030. URL https://www.frontiersin.org/article/10.3389/frai.2019.00030.

F. Alqahtani, S. Meier, and R. Orji. Personality-based approach for tailoring persuasive mental health applications. *User Modeling and User-Adapted Interaction*, 07 2021a. doi: 10.1007/s11257-021-09289-5.

F. Alqahtani, A. Winn, and R. Orji. Co-designing a mobile app to improve mental health and well-being: Focus group study. *JMIR Form Res*, 5(2):e18172, Feb 2021b. ISSN 2561-326X. doi: 10.2196/18172. URL https://doi.org/10.2196/18172.

J. Apolinário-Hagen, S. Hennemann, L. Fritsche, M. Drüge, and B. Breil. Determinant factors of public acceptance of stress management apps. *Journal of Medical Internet Research*, 2019. doi: 10.2196/15373.

A. A. Armanasco, Y. D. Miller, B. S. Fjeldsoe, and A. L. Marshall. Preventive health behavior change text message interventions: A meta-analysis. *American Journal of Preventive Medicine*, 3(52):391–402, 2017.

E. Aronson, T.D. Wilson, and R.M. Akert. *Social Psychology*. Pearson, 2010. URL https://books.google.nl/books?id=1wOmQAAACAAJ.

D. Bakker, N. Kazantzis, D. Rickwood, and N. Rickard. Mental health smartphone apps: Review and evidence-based recommendations for future developments. *JMIR Mental Health*, 3, 2016. doi: 10.2196/mental.4984.

R. Bartle. Hearts, clubs, diamonds, spades: Players who suit muds. *Journal of MUD research*, 1(1):19, 1996.

H. Barton. Persuasion and compliance in cyberspace. pages 111–122, 2016.

C. Bateman and R. Boon. 21st century game design (game development series). 2005.

D. Berdichevsky and E. Neuenschwander. Toward an ethics of persuasive technology. *Communications of the ACM*, 42(5):51–58, 1999. doi: 10.1145/301353.301410.

K. Bessenyei, B. Suruliraj, A. Bagnell, P. McGrath, L. Wozney, A. Huguet, B. S. Elger, S. Meier, and R. Orji. Comfortability with the passive collection of smartphone data for monitoring of mental health: An online survey. *Computers in Human Behavior Reports*, 4:100134, 2021. ISSN 2451-9588. doi: https://doi.org/10.1016/j.chbr.2021.100134.

J. Burger. Replicating Milgram: Would people still obey today? *American Psychologist*, 64(1):1–11, 2009.

R. B. Caldini and N. J. Goldstein. Social influence: Compliance and conformity. *Annual Review of Psychology*, 55:591–621, 2004.

M. Charity, M. Nkwo, R. Orji, and E. Ebere. Personalized persuasive technology for maternal healthcare in Nigeria. 04 2020.

M. Chogahara. A multidimensional scale for assessing positive and negative social influences on physical activity in older adults. *Journal of Gerontology*, 54:S356–S367, 1999. doi: 10.1093/geronb/54B.6.S356.

R. B. Cialdini. Influence: the psychology of persuasion. 2006.

R. B. Cialdini and V. Griskevicius. Social influence. *Advanced social psychology: The state of the science*, pages 385–417, 2010.

R. B. Cialdini, J. E. Vincent, S. K. Lewis, J. Catalan, D. Wheeler, and B. L. Darby. Reciprocal concessions procedure for inducing compliance: The door-in-the-face technique. *Journal of Personality and Social Psychology*, 31(2):206–215, 1975. doi: 10.1037/h0076284.

M. Cinnirella and B. Green. Does 'cyber-conformity' vary cross-culturally? exploring the effect of culture and communication medium on social conformity. *Computers in Human Behavior*, 23(4):2011–2025, 2007.

J.S. Cohen, J.M. Edmunds, D.M. Brodman, C.L. Benjamin, and P.C. Kendall. Using self-monitoring: Implementation of collaborative empiricism in cognitive-behavioral therapy. *Cognitive and Behavioral Practice*, 20:419–428, 2013. doi: 10.1016/j.cbpra.2012.06.002.

D. Cormier, G. Newman, M. Nakane, J. E. Young, and S. Durocher. Would you do as a robot commands? an obedience study for human-robot interaction. *International Conference on Human-Agent Interaction, Sapporo, Japan*, 2013.

M. Czerwinski, E. Cutrell, and E. Horvitz. Instant messaging: Effects of relevance and timing. In *People and computers XIV: Proceedings of HCI*, volume 2, pages 71–76, 2000.

D. DeSteno, R.E. Petty, D.D. Rucker, D.T. Wegener, and J. Braverman. Discrete emotions and persuasion: The role of emotion-induced expectancies. *Journal of Personality and Social Psychology*, 86:43–56, 2004.

M. Deutsch and H. B. Gerard. A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, 51(3):629–636, 1955.

A. Dijkstra. Technology adds new principles to persuasive psychology: Evidence from health education. volume 3962, pages 16–26, 2006. doi: 10.1007/11755494$_4$.

F. Drozd, T. Lehto, and H. Oinas-Kukkonen. Exploring perceived persuasiveness of a behavior change support system: A structural model. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, page 157–168, 2012. doi: 10.1007/978-3-642-31037-9$_1$4.

W. R. Fisher. The narrative paradigm: An elaboration. *Communications Monographs*, 52(4):347–367, 1985.

B. J. Fogg. Persuasive technology: Using computers to change what we think and do. *Ubiquity*, 2002(5):32, 2002. doi: 10.1145/764008.763957.

J. L. Freedman and S. C. Fraser. Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, 4(2):195–202, 1966. doi: 10.1037/h0023552.

L. R. Goldberg. An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, 59(6):1216–1229, 1990.

A. Grimes, V. Kantroo, and R.E. Grinter. Let's play!: mobile health games for adults. *international conference on Ubiquitous computing*, page 241–250, 2010.

D. Grimley, J.O. Prochaska, W.F. Velicer, L.M. Vlais, and C.C. DiClemente. The transtheoretical model of change. in t.m. brinthaupt  r.p. lipka, changing the self: Philosophies, techniques, and experiences. page 201–227, 1994.

J. Hamari and J. Koivisto. "working out for likes": An empirical study on social influence in exercise gamification. 50, 2015. doi: 10.1016/j.chb.2015.04.018.

N. Hertz and E. Wiese. Under pressure: Examining social conformity with computer and robot groups. *Human Facrtors*, 60(8):1207–1218, 2018.

J. B. Hirsh, S. K. Kang, and G. V. Bodenhausen. Personalized persuasion: tailoring persuasive appeals to recipients' personality traits. *Psychological Science*, 23(6):578–581, 2012.

C .I. Hovland, I. L. Janis, and H. H. Kelley. Communication and persuasion: Psychological studies of opinion change. *American Political Science Review*, 48(2):600–600.

W. IJsselsteijn, Y. de Kort, C. Midden, B. Eggen, and E. van den Hoven. Persuasive technology for human well-being: setting the scene. *Persuasive technology*, pages 1–5, 2006.

D. L. Kappen and R. Orji. Gamified and persuasive systems as behavior change agents for health and wellness. *XRDS*, 24(1):52–55, sep 2017. ISSN 1528-4972. doi: 10.1145/3123750. URL https://doi.org/10.1145/3123750.

M. Kaptein, B. De Ruyter, P. Markopoulos, and E. Aarts. Adaptive persuasive systems: A study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems Association for Computing Machinery*, 2012.

M. Kaptein, P. Markopoulos, B. De Ruyter, and E. Aarts. Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles. *International Journal of Human-Computer Studies*, 77:38–51, 2015.

M. C. Kaptein and D. Eckles. Selecting effective means to any end: Futures and ethics of persuasion profiling. *Persuasive Technology*, pages 82–93, 2010.

M. C. Kaptein, P. Markopoulos, B. de Ruyter, and E. Aarts. Can you be persuaded? individual differences in susceptibility to persuasion. *IFIP conference on human-computer interaction*, pages 115–118, 2009.

M. P. Kelly and M. Barker. Why is changing health-related behavior so difficult? *Public health*, 136:109–116, 2016.

R. Khaled, P. Barr, J. Noble, R. Fischer, and R. Biddle. Fine tuning the persuasion in persuasive games. *Lecture Notes in Computer Science*, 4744, 2007. doi: $10.1007/978$-$3$-$540$-$77006$-$0_5$.

T. Lang. Advancing global health research through digital technology and sharing data. *Science*, 331(6018): 714–717, 2011.

R. R. Lau, M. J. Quadrel, and K. A. Hartman. Development and change of young adults' preventive health beliefs and behavior: Influence from parents and peers. *Journal of Health and Social Behavior*, 31: 240–259, 1990.

J.J. Lin, L. Mamykina, S. Lindtner, G. Delajoux, and H.B. Strub. Fish'n'steps: Encouraging physical activity with an interactive computer game. in: Ubiquitous computing. *Lecture Notes in Computer Science*, 4206, 2006. doi: $10.1007/11853565_16$.

K.H. Ly, K. Asplund, and G. Andersson. Stress management for middle managers via an acceptance and commitment-based smartphone application: A randomized controlled trial. *Internet Interventions*, page 95–101, 2014. doi: 10.1016/j.invent.2014.06.003.

S. Michie, M.M. van Stralen, and R. West. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science*, 6, 2011. doi: 10.1186/1748-5908-6-42.

S. Michie, L. Atkins, and R. West. *The Behaviour Change Wheel: A Guide to Designing Interventions*. Silverback Publishing, 2014. URL https://books.google.nl/books?id=am6ytAEACAAJ.

S. Milgram. Behavioral study of obedience. *Abnormal and Social Psychology*, 67(4):371–378, 1963. doi: 10.1037/h0040525.

S. Milgram. Obedience to authority: An experimental view. 1974.

Y. Miyata and D. Norman. Psychological issues in support of multiple activities, chapter in user-centered design. pages 265–284, 1986.

L. Morton, R. Layton, and M. Muraven. Social modeling of exercise inventory: A measure of exercise-specific social influence. *Journal of Applied Social Psychology*, 2019. doi: 10.1111/jasp.12575.

D. Mulchandani and R. Orji. Age and the persuasiveness of a game to promote the adoption of covid-19 precautionary measures. In *2021 IEEE 9th International Conference on Serious Games and Applications for Health(SeGAH)*, pages 1–9, 2021. doi: 10.1109/SEGAH52098.2021.9551909.

I. B. Myers. The Myers-Briggs type indicator: Manual. 1962. doi: 10.1037/14404-000.

L. Nacke, C. Bateman, and R. Mandryk. Brainhex: A neurobiological gamer typology survey. *Entertainment Computing*, 5, 2013. doi: 10.1016/j.entcom.2013.06.002.

M. A. Napolitano, S. Hayes, G. G. Bennett, A. K. Ives, and G. D. Foster. Using Facebook and text messaging to deliver a weight loss program to college students. *Obesity*, 21(1):25–31, 2013.

C. Ndulue, O. Oyebode, R. S. Iyer, A. Ganesh, and R. Orji S. I. Ahmed. Personality-targeted persuasive gamified systems: Exploring the impact of application domain on the effectiveness of behavior change strategies. *User Modeling and User-Adapted Interaction*, 2022.

H. Oinas-Kukkonen. Persuasive systems design: Key issues, process model, and system features. *Communications of the Association for Information Systems*, 24(1), 2009. doi: 10.17705/1CAIS.02428.

R. Orji and K. Moffatt. Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health Informatics Journal*, 24(1):66–91, 2018. doi: 10.1177/1460458216650979. URL https://doi.org/10.1177/1460458216650979. PMID: 27245673.

R. Orji, R. L. Mandryk, J. Vassileva, and K. M. Gerling. Tailoring persuasive health games to gamer type. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 2467–2476, New York, NY, USA, 2013a. Association for Computing Machinery. ISBN 9781450318990. doi: 10.1145/2470654.2481341.

R. Orji, J. Vassileva, and R.L. Mandryk. Lunchtime: a slow-casual game for long-term dietary behavior change. *Personal and Ubiquitous Computing*, 17:1211–1221, 2013b. doi: 10.1007/s00779-012-0590-6.

R. Orji, R. Mandryk, and J. Vassileva. Selecting effective strategies for tailoring persuasive health games to gamer types. pages 1–4, 05 2014a.

R. Orji, J. Vassileva, and R. L. Mandryk. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. *User Modeling and User-Adapted Interaction*, 24(5):453–498, 2014b. doi: 10.1007/s11257-014-9149-8.

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

R. Orji, L. E. Nacke, and C. Di Marco. Towards personality-driven persuasive health games and gamified systems. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 1015–1027, New York, NY, USA, 2017b. Association for Computing Machinery. ISBN 9781450346559. doi: 10.1145/3025453.3025577. URL https://doi.org/10.1145/3025453.3025577.

R. Orji, R. L., K. Oyibo, F. Orji, J. Blustein, and S. Shahid. Tracking feels oppressive and 'punishy': Exploring the costs and benefits of self-monitoring for health and wellness. *DIGITAL HEALTH*, 4:2055207618797554, 2018a. doi: 10.1177/2055207618797554. URL https://doi.org/10.1177/2055207618797554. PMID: 30202544.

R. Orji, F. Orji, K. Oyibo, and I. A. Ajah. Personalizing health theories in persuasive game interventions to gamer types: An African perspective. In *Proceedings of the Second African Conference for Human Computer Interaction: Thriving Communities*, AfriCHI '18, New York, NY, USA, 2018b. Association for Computing Machinery. ISBN 9781450365581. doi: 10.1145/3283458.3283467.

R. Orji, K. Oyibo, R. K. Lomotey, and F. A. Orji. Socially-driven persuasive health intervention design: Competition, social comparison, and cooperation. *Health Informatics Journal*, 25(4):1451–1484, 2019. doi: 10.1177/1460458218766570. URL https://doi.org/10.1177/1460458218766570. PMID: 29801426.

O. Oyebode and R. Orji. Deconstructing persuasive strategies in mental health apps based on user reviews using natural language processing. In *Behavior Change Support Systems at Persuasive*, 2020a.

O. Oyebode and R. Orji. A persuasive mobile game for reducing sedentary behaviour and promoting physical activity. 04 2020b.

O. Oyebode, F. Alqahtani, and R. Orji. Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. *IEEE Access*, 8:111141–111158, 2020a. doi: 10.1109/ACCESS.2020.3002176.

O. Oyebode, B. Graham-Kalio, and R. Orji. Hearthealth: A persuasive mobile app for mitigating the risk of ischemic heart disease. In *International Conference on Persuasive Technology*, pages 126–138. Springer, 2020b.

O. Oyebode, D. Maurya, and R. Orji. Nourish your tree! developing a persuasive exergame for promoting physical activity among adults. In *2020 IEEE 8th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–7, 2020c. doi: 10.1109/SeGAH49190.2020.9201637.

O. Oyebode, C. Ndulue, M. Alhasani, and R. Orji. Persuasive mobile apps for health and wellness: a comparative systematic review. In *International Conference on Persuasive Technology*, pages 163–181. Springer, 2020d.

O. Oyebode, A. Ganesh, and R. Orji. Treecare: Development and evaluation of a persuasive mobile game for promoting physical activity. In *IEEE Conference on Games*, 2021a. doi: 10.1109/CoG52621.2021.9619035.

O. Oyebode, C. Ndulue, D. Mulchandani, A. A. Z. Adib, M. Alhasani, and R. Orji. *Tailoring Persuasive and Behaviour Change Systems Based on Stages of Change and Motivation*. Association for Computing Machinery, New York, NY, USA, 2021b. ISBN 9781450380966. URL `https://doi.org/10.1145/3411764.3445619`.

R. E. Petty and J. T. Cacioppo. The elaboration likelihood model of persuasion. In *Communication and persuasion*, pages 1–24. Springer, 1986.

J. O. Prochaska. Transtheoretical model of behavior change. *In Encyclopedia of behavioral medicine*, pages 1997–2000, 2013.

K. Rhoads. How many influence, persuasion, compliance tactics & strategies are there? 2007. doi: workingpsychology.com/numbertactics.html.

A. Rieger, Q. Shaheen, C. Sierra, M. Theune, and N. Tintarev. Towards healthy engagement with online debates: An investigation of debate summaries and personalized persuasive suggestions. page 192–199, 2022. doi: 10.1145/3511047.3537692.

I.M. Rosenstock. Why people use health services. *The Milbank Memorial Fund quarterly*, 44(3):94–127, 1966.

J. F. Sallis, R. M. Grossman, R. B. Pinski, T. L. Patterson, and P. R. Nader. The development of scales to measure social support for diet and exercise behaviors. *Preventive Medicine*, 16:825–836, 1987.

D. H. Schunk and B. J. Zimmerman. Modeling and self-efficacy influences on children's development of self-regulation. page 154–180, 1996.

M. Sherif and C. I. Hovland. Social judgment: Assimilation and contrast effects in communication and attitude change. 1961.

D. Shin. User perceptions of algorithmic decisions in the personalized ai system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting Electronic Media*, 64 (4):541–565, 2020. doi: 10.1080/08838151.2020.1843357.

M. Sicart. The ethics of computer game design. *Proceedings of Digital Games Research Association*, 2005.

M. Slater, A. Antley, A. Davison, D. Swapp, C. Guger, C. Barker, N. Pistrang, and M. V. Sanchez-Vives. A virtual reprise of the Stanley Milgram obedience experiments. *Plos One*, 1:1–10, 2006. doi: 10.1371/journal.pone.0000039.

L. Smids. Changing behavior one drop at a time: Changing daily water intake using tailored persuasive strategies. Master's thesis, Radboud university, Netherlands, 2017.

A. Torbica and N. Sad. Influence of different factors on user motivation for e-health. Master's thesis, University of Twente, Netherlands, 2004.

P. P. Verbeek. Materializing morality: Design ethics and technological mediation. *Science, Technology and Human Values*, 31(3):361–380, 2006a.

P. P. Verbeek. Persuasive technology and moral responsibility toward an ethical framework for persuasive technologies. *Persuasive*, 6:1–15, 2006b.

J. Vlieghe and O. De Troyer. Report d3: Literature study on persuasive techniques and technology. 2020. URL `wise.vub.ac.be/ tickle/wp-content/uploads/2015/12/Report-D3`$_f inal.pdf$.

S. Weber, C. Lorenz, and N. Hemmings. Improving stress and positive mental health at work via an app-based intervention: A large-scale multi-center randomized control trial. *Frontiers in Psychology*, 10, 2019. doi: 10.3389/fpsyg.2019.02745.

E. A. Weinstein and P. Deutschberger. Some dimensions of altercasting. *Sociometry*, pages 454–466, 1963.

S. Wróbel. Logos, ethos, pathos. classical rhetoric revisited. *Polish Sociological Review*, 191(3):401–421, 2015.

N. Yee. Motivations of play in mmorpgs. *International Conference of Digital Games Research Association on Changing Views: Worlds in Play*, 2005.

N. Yee. Motivations for play in online games. *CyberPsychology & behavior*, 9(6):772–775, 2006.

Z. Zhao, A. Arya, R. Orji, and G. Chan. Effects of a personalized fitness recommender system using gamification and continuous player modeling: System design and long-term validation study. *JMIR Serious Games*, 8, 2020. doi: 10.2196/19968.

# A    Appendix

Here the questionnaire from `Brainhex` can be found in Figure A.1, A.2 and A.3 and Section A.1, which can also be found at http://survey.ihobo.com/BrainHex/index.php. The set-up for the focus group can also be found here, along with the sentences that the participants had to judge upon. The link to the (preliminary) survey for the pre-study is: https://forms.gle/5uXZ4rpJm2csnjXW6, but a simplified version is also included in A.5. The link to the (preliminary) survey for the real experiment is: https://forms.gle/8EDtSrmujvLeBpvS6, but a simplified version is also included in A.7.

## A.1    Survey about `Brainhex` profiles

The `Brainhex` survey consist of Figure A.1, Subsection A.2, Figure A.2 and Figure A.3, and can be found on http://survey.ihobo.com/BrainHex/.

## A.2    Questions from the original `Brainhex`

This is a representation of page number 2 from the questionnaire on: **http://survey.ihobo.com/BrainHex/**
***The possible answers were: I love it! / I like it. / It's okay. / I dislike it. / I hate it! The scoring of those answers was: +1 +0.5 0 -1 -2***
Exploring to see what you can find.
Frantically escaping from a terrifying foe.
Working out how to crack a challenging puzzle.
The struggle to defeat a difficult boss.
Playing in a group, online or in the same room.
Responding quickly to an exciting situation.
Picking up every single collectible in an area.
Looking around just to enjoy the scenery.
Being in control at high speed.
Devising a promising strategy when deciding what to try next.
Feeling relief when you escape to a safe area.
Taking on a strong opponent when playing against a human player in a versus match.
Talking with other players, online or in the same room.
Finding what you need to complete a collection.
Hanging from a high ledge.
Wondering what's behind a locked door.
Feeling scared, terrified or disturbed.
Working out what to do on your own.
Completing a punishing challenge after failing many times.
Co-operating with strangers.
Getting 100% (completing everything in a game).

## Part 1: About You

**Year of birth:** [1996 ▾]

**Gender:**
- ◉ Female
- ○ Male

**Geographical Territory:**
- ○ North America
- ○ Southern or Central America
- ◉ Western Europe or UK
- ○ Eastern Europe or Russia
- ○ South Asia (incl. China, India and Japan)
- ○ Africa
- ○ Middle East
- ○ Australasia
- ○ Other (please specify) [        ]

**I typically play computer or videogames**
- ○ Every day
- ◉ Every week
- ○ Occasionally
- ○ Rarely
- ○ Never

**I would consider myself**
- ○ Hardcore gamer
- ○ something between a Hardcore and a Casual gamer
- ◉ Casual gamer
- ○ I have no idea!

**I work in:**
- ◉ a non-videogames related industry (or I don't work/am a student)
- ○ videogame development
- ○ videogame publishing
- ○ videogame retail
- ○ videogame press
- ○ videogames in some other context (e.g. research)

**I prefer the following way of playing games:**
- ○ Single player alone
- ○ Single player with other people helping or pad-passing
- ○ Multiplayer, in the same room
- ◉ Multiplayer, over the internet
- ○ Team play or Clan play over the internet
- ○ Virtual worlds or MMORPGs

**My attitude to videogame stories is:**
- ○ Stories are very important to my enjoyment of videogames
- ◉ Stories can help me enjoy a videogame
- ○ Stories are not important to me in videogames
- ○ I prefer videogames without stories
- ○ I don't play videogames

**Name three games that exemplify what you enjoy about games (these don't have to be videogames - any game you enjoy counts):**

[animal crossing]
[kirby]
[mario]

**I live with, and/or like living with:**
- ○ a cat or cats
- ○ a dog or dogs
- ◉ both cats and dogs
- ○ neither

*Optional:*
**If you know your Myers-Briggs Type, please select it here:**
[- ▾]

*Case Study Volunteers Only:*
Email address : [                    ]

[ Onto Page 2 of 4 ]

Figure A.1: Part 1 of the `Brainhex` questionnaire.

**Part 3: Rate**

**Instructions**

Please read through the following seven statements.
Place them in order of preference by choosing the highest number for the most preferred statement to the lowest number for the least preferred. Please choose each number only once.

Please list the seven statements in order of preference, choosing each number only once:

6 — A moment of jaw-dropping wonder or beauty.
1 (worst) — An experience of primeval terror that blows your mind.
2 — A moment of breathtaking speed or vertigo.
5 — The moment when the solution to a difficult puzzle clicks in your mind.
3 — A moment of hard-fought victory.
4 — A moment when you feel an intense sense of unity with another player.
7 (best) — A moment of completeness that you have strived for.

See your results!

Last Updated: 27 February 2014
BrainHex is the latest player satisfaction model from International Hobo Ltd
©1999 - 2014 International Hobo.

Figure A.2: Part 3 of the `Brainhex` questionnaire.

## Results

**Your BrainHex Class is Achiever.**

**Your BrainHex Sub-Class is Achiever-Seeker.**

*You like collecting anything you can collect or doing everything you possibly can as well as finding strange and wonderful things or finding familiar things.*

Each BrainHex Class also has an Exception, which describes what you dislike about playing games. Your Exceptions are:

» No Punishment: You dislike struggling to overcome seemingly impossible challenges, and repeating the same task over and over again.
» No Pressure: You dislike being asked to perform under pressure, preferring to take your time so you can make the right decision.
» No Fear: You do not enjoy feeling afraid, preferring to feel safe or in control.

Learn more about your classes and exceptions at BrainHex.com.

Your scores for each of the classes in this test were as follows:

Achiever: 20
Seeker: 15
Socialiser: 12
Mastermind: 11
Conqueror: 0
Daredevil: -2
Survivor: -6

Go to BrainHex.com to learn more about this player model, and the neurobiological research behind it.

Feel free to take a copy of your BrainHex icon and display it anywhere you wish! Simply right click and choose "save as". All we ask is you provide a link to BrainHex.com anywhere you use our images.

Thanks for taking part in the BrainHex survey!

Figure A.3: An example result of the `Brainhex` questionnaire.

## A.3   Set-Up for the Focus Group

Hello and welcome everyone to this focus group for my master thesis. First of all, I want to thank you for your time and interest to help me for my thesis, it is much appreciated. Did everyone fill in the questionnaire of Brainhex beforehand, and send the outcome to me personally? Feel free to introduce yourself by stating how far you are in your studies.

I will give you some insight in my thesis, so that you better understand the circumstances and aim of this focus group. My thesis is about persuasion in the health domain. In this case, health is measured the number of steps set per day per participant. Many people can sit almost all day for studying or working from home, which is not good for their health. Enough physical activity is necessary for having good health. By sending daily persuasive text messages to participants, the hope is to increase the number of steps set per day per participant.

The aim here today is to find out which sentences come off as most persuasive for you, and to generate more sentences like that, so that I can use them for my experiment. This will take about 1 hour of your time. First, I will lay out an example story to get a grasp on the subject. Then we will go through 9 different strategies for persuasive text messages one by one. After discussing 5/9 strategies, we will have a small break.

Please stay humble to one another during this focus group. By continuing this meeting, you give consent to incorporate your collective ideas on the persuasive sentences into my thesis, and nothing else. To do so, this meeting will be recorded in audio only. You will remain anonymous. There are no right or wrong answers to any of the questions.

Any questions so far? [this should stay within 5-10 minutes]

Now I want you to imagine a student, just like you, let's name her Eliza for now. It is Monday morning and the sun is shining, but it is a bit chilly outside. Eliza wakes up to go to campus to have her lectures for that day. She lives a bit close to the campus. She is thinking about walking to campus or taking the bus.

What would persuade her to walk rather than taking the bus?

Thank you for your input. Now we add some information to the story: Eliza has an app installed that tracks the number of steps set per day. She can also see the total number of steps of yesterday in her notifications.

Are there any new insights on what could be persuasive text messages that can persuade her to go walking?

Thank you for your input. Now we add some other information to the story: Eliza can also see the number of steps of her friends in that app. Both of today (Monday) and of yesterday (Sunday).

Are there any new insights on what could be persuasive text messages that can persuade her to go walking? [this should have taken another 10 minutes]

1. Now we will look at the first persuasive strategy, authority. This means that the message is some advice from an authority figure, like someone with high social status or an expert on the matter.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved?

2. Now we will look at the persuasive strategy of informational influence. These messages include information that makes sense to us or has sound arguments and thus appeals to our reason. General knowledge of the subject of health for example.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved?

3. Now we will look at the persuasive strategy of comparison and competition. This includes a social context in which people live. A well know example is a leaderboard, where you can compare your scores to that of another, and can compete against one another.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved?

4. Now we will look at the persuasive strategy of personalization. This includes tailoring messages to specific individuals. Personal information or characteristics could be used for this, like one's name.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved?

5. Now we will look at the persuasive strategy of consensus. The idea here is that people will look at others how they act and behave, to create a baseline on what is good and acceptable behavior and what is not. They base their own behaviors on that of others, even if they have not observed it themselves. An example being notes that say that 75% of people do reuse their towels in a hotel.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved?

6. Now we will look at the persuasive strategy of praise. This includes positive words and images to reinforce the target behavior of walking more steps a day.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved?

7. Now we will look at the persuasive strategy of rewards. This includes giving credit for performing the target behavior or achieving milestones. This would be walking a lot every day or taking more than 7K steps a day or a streak of more than 5K steps for 5 days in a row for example.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved? [5 minutes]

8. Now we will look at the persuasive strategy of punishment, which means penalizing the unwanted behavior like taking few steps a day.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved? [5 minutes]

9. Now we will look at the final persuasive strategy or reminders, which are actual reminders of performing the target behavior, which is taking more steps a day.

Could you think about what kinds of sentences would be persuasive or not for students in general? What is good about the sentences provided, and what could be improved? [5 minutes]

Lastly, I would like to know which of the 9 persuasive strategies are considered more persuasive by each of you individually. Aim to make a top 3.

That was it for the focus group today. Thank you for participating.

If you have anything to add or ask, feel free to do so, otherwise, have a nice day!

## A.4   Sentences of the Focus Group

• Authority: This means that the message is some advice from an authority figure, like someone with high social status or an expert on the matter.

1. The world health guidelines state that you should walk more than you did yesterday.

2. According to the world health guidelines, it is recommended to walk 7000 steps a day.

3. Health researchers state that one has to set at least 5000 steps a day to prevent bad health.

4. By walking more than 6K steps a day, you have less change at health risks, researchers have found.

• Informational influence: These messages include information that makes sense to us or has sound arguments and thus appeals to our reason. General knowledge of the subject of health for example.

1. It is good for health that you walk a lot every day, rather than sitting around.

2. Exercise is good for your health.

3. Too much sitting still is not a healthy habit.

4. It is better to walk than to sit during the day.

• social comparison/competition: This includes a social context in which people live. A well know example is a leaderboard, where you can compare your scores to that of another, and can compete against one another.

1. You were 5th among our 15 participants yesterday (good job).

2. You were at the bottom 5 people yesterday (try to keep up).

3. You walked as much as the average of the participants in your group yesterday.

4. There were 3 participants better than you yesterday.

5. 70% of the participants in this study walked more than you yesterday, try to beat them today!

6. graph of leaderboard

• Personalization: This includes tailoring messages to specific individuals. Personal information or characteristics could be used for this, like one's name.

1. Hello [name], you were doing great yesterday.

2. Good morning [name], you were not doing great yesterday.

3. Hi [name], you were better yesterday than the day before that.

4. Hello [name], you can do this.

5. Hello [name], you did great yesterday with a total amount of 2435 steps.

• Consensus: The idea here is that people will look at others how they act and behave, to create a baseline on what is good and acceptable behavior and what is not. They base their own behaviors on that of others, even if they have not observed it themselves. An example being notes that say that 75% of people do reuse their towels in a hotel.

1. 75% of participants has a step count higher than 5000 on average.

2. 100% of the participants has done more than 1000 steps yesterday.

3. 60% of participants had a 5-day streak of more than 4000 steps a day.

4. All participants stayed above 3K steps yesterday.

5. On average, participants walked 5K yesterday.

• Praise: This includes positive words and images to reinforce the target behavior of walking more steps a day.

1. Good job yesterday.

2. Try to aim as high as yesterday.

3. That was going great yesterday.

4. smiley emoji

• Reward: This includes giving credit for performing the target behavior or achieving milestones. This would be walking a lot every day or taking more than 7K steps a day or a streak of more than 5K steps for 5 days in a row for example.

1. You unlocked the achievement of 10K steps for a day.

2. you unlocked the streak of more than 5K for 5 days in a row.

3. picture of a badge

• Punishment: This means penalizing the unwanted behavior like taking few steps a day.

1. Too bad, you have not walked a lot yesterday.

2. Unfortunately, you were one of the worst performing participants yesterday.

3. sad face smiley

4. sad gif

• Reminder: These are actual reminders of performing the target behavior, which is taking more steps a day.

1. Do not forget to exercise well today as well.

2. Let's aim for another personal record today.

3. Do not forget to walk enough today.

## A.5 Survey of the Pre-Study

Survey for step-count thesis

(**Q1**) Which participant number are you? This is to ensure your anonymity. (You can use your Prolific ID if you have one)

**This was a mandatory open text answer box.**

(**Q2**) Choose from a scale between "I love it!" for experiences you enjoy through "It's okay" to "I hate it!" for experiences you would rather avoid. If you have difficulty answering them, try imagining a work-related scenario or a gaming (boardgames or computer) scenario. The feeling that you would get in such a situation matters, do you like or dislike that feeling?

**The scales for the following statements were: I love it! / I like it. / It's okay. / I dislike it. / I hate it!**

- Exploring to see what you can find.
- Frantically escaping from a terrifying foe.
- Working out how to crack a challenging puzzle.
- The struggle to defeat a difficult boss.
- Playing in a group, online or in the same room.
- Responding quickly to an exciting situation.
- Picking up every single collectible in an area.
- Looking around just to enjoy the scenery.
- Being in control at high speed.
- Devising a promising strategy when deciding what to try next.
- Feeling relief when you escape to a safe area.
- Taking on a strong opponent when playing against a human player in a versus match.
- Talking with other players, online or in the same room.
- Finding what you need to complete a collection.
- Hanging from a high ledge.
- Wondering what's behind a locked door.
- Feeling scared, terrified or disturbed.
- Working out what to do on your own.
- Completing a punishing challenge after failing many times.
- Co-operating with strangers.
- Getting 100% (completing everything in a game).

(**Q3**) Please read through the following seven statements. Place them in order of preference by choosing the highest number for the most preferred statement to the lowest number for the least preferred. Please choose each number only once. So the WORST one is 1, and the BEST/preferred one is 7.

**The scales for the following statements where the numbers 1 (worst) up to 7 (best) could only be chosen once.**

- A moment of jaw-dropping wonder of beauty.
- An experience of primeval terror that blows your mind.
- A moment of breathtaking speed or vertigo.
- The moment when a solution to a difficult puzzle clicks in your mind.
- A moment of hard-fought victory.
- A moment when you feel an intense sense of unity with another player.
- A moment of completeness that you have strived for.

(**Q4**) Please select how much you agree or disagree with the following statements.

**Questions from Susceptibility to Persuasion Scale**

*The possible answers were: Strongly Agree / Agree / Somewhat Agree / Neutral / Somewhat Disagree / Disagree / Strongly Disagree. The scoring of those answers*

*was: +2 +1 +0.5 0 -0.5 -1 -2*
- When a family member does me a favour, I am very inclined to return this favour.
- I believe rare products (scarce) are more valuable than mass products.
- I always follow advice from my general practitioner.
- Whenever I commit to an appointment I do as I told.
- I always pay back a favour.
- Please answer "Agree" here. **This was one of the sanity checks in the survey.**
- When my favourite shop is about to close, I would visit it since it is my last chance.
- I accept advice from my social network.
- When I am in a new situation I look at others to see what I should do.
- When a professor tells me something I tend to believe it is true.
- When I like someone, I am more inclined to believe him/her.
- I try to do everything I have promised to do.
- If someone from my social network notifies me about a good book, I tend to read it.

(**Q5**) Which of these following sentences would have the most positive influence on you? Meaning the sentence that would make you change your behavior. Please select for each message whether it would make you set extra steps per day. (1/3)

**Questions about persuasive strategies via example sentences per strategy**

*The possible answers were: It would help a lot! / It would help / Neutral / It would not help / It would not help at all. The scoring of those answers was: +2 +1 0 -1 -2*

The world health organization advises to set at least 7.000 to 10.000 steps per day. Make sure that you get this.

It is good for health that you walk a lot every day, rather than sitting around.

The more often you walk, the easier it gets. If one day it doesn't work out, that is fine as well.

5 participants in this study walked more than you yesterday, try to beat them today!

Please select "It would help a lot!" **This was one of the sanity checks in the survey.**

You have consistently reached more than 5000 steps, you're a master walker now!

Hey there Tim, if you are heading somewhere today, maybe you can go by foot?

Walking more can prevent or manage various conditions, including heart disease, stroke, high blood pressure, cancer and type 2 diabetes.

You unlocked the achievement of 10.000 steps for a day!

It's easy to get started: All you need is a pair of comfortable shoes and to get yourself out the door.

(**Q6**) Which of these following sentences would have the most positive influence on you? Meaning the sentence that would make you change your behavior. Please select for each message whether it would make you set extra steps per day. (2/3)

**Questions about persuasive strategies via example sentences per strategy**

*The possible answers were: It would help a lot! / It would help / Neutral / It would not help / It would not help at all. The scoring of those answers was: +2 +1 0 -1 -2*

You only walked 3400 steps yesterday, you can do better than that! About 72% of the participants has done more than 6500 steps yesterday.

Well done, you set a great amount of steps yesterday.

You walked 6000 steps on average, let's try to get 1000 steps more today.

Please select "It would not help at all" **This was one of the sanity checks in the survey.**

Every step is a step more, even inside your home.

You were 7th among our 22 participants yesterday.

On average, people walked 6700 steps.

Erik Scherder researches exercise and its effect on the brain. He says that everybody should

exercise/walk for 30 minutes a day to keep their brain and body healthy.

You were below 5000 steps yesterday, try to walk more today!

You once reached 8700 steps during this study. Let's try to reach this number again today!

(**Q7**) Which of these following sentences would have the most positive influence on you? Meaning the sentence that would make you change your behavior. Please select for each message whether it would make you set extra steps per day. (3/3)

**Questions about persuasive strategies via example sentences per strategy**

*The possible answers were: It would help a lot! / It would help / Neutral / It would not help / It would not help at all. The scoring of those answers was: +2 +1 0 -1 -2*

Yesterday was great! Try to aim as high as you did yesterday.

If you manage to get 2500 steps more than yesterday, you may treat yourself.

Too bad, you have not walked a lot yesterday. Aim higher!

You have to walk only 3000 more steps than you did yesterday, to reach the top position.

Walking more has a positive influence on your mental health.

Wow! You really outdid yourself yesterday! Keep it going!

Please select "Neutral" **This was one of the sanity checks in the survey.**

Amanda Paluch, an assistant professor of kinesiology at the University of Massachusetts Amherst, concluded that 7000(+) steps a day has a lot of benefits and lower risk at premature death.

Yesterday about 63% of participants has a step count higher than 5400.

(**Q8**) Choose from a scale between "I love it!" for experiences you enjoy through "It's okay" to "I hate it!" for experiences you would rather avoid. The feeling that you would get in such a situation matters, do you like or dislike that feeling?

**Questions from the edited `Brainhex`, to make the questions less game-oriented.**

*The possible answers for the following questions were: I love it! / I like it. / It's okay. / I dislike it. / I hate it! The scoring of those answers was: +1 +0.5 0 -1 -2*

- Leaving quickly from something scary.
- The struggle to overcome something that is (very) challenging.
- Collecting items for a collection.
- The feeling of relief when you went away from something unsafe or unpleasant.
- Taking on a strong (sports or board game) rival in a one-on-one match.
- Talking to other people, online or in the same room.
- Getting the last piece of a collection/puzzle or (finally) finishing something.
- Working together with other people, (possibly strangers).
- Standing on a high building or wooden bridge.
- Exploring or seeking new or different things.

(**Q9**) What situation applies to you the most?

- I have no intention to take action, with regard to the number of steps walked per day, in the next 6 months or foreseeable future.
- I am aware of my walking behavior and I have an intention to / I am thinking about taking action (regarding the number of steps per day) within the next 6 months, but not immediately.
- I have an intention to take action (with regard to the number of steps set per day) within the next 30 days and I have done some initial steps towards that action.
- I am actively taking more steps a day for a (short) while now, but I am still struggling regressing into fewer steps a day.
- I am actively taking more steps a day now for more than 6 months, and I prevent a relapse of walking fewer steps a day again.
- I have walked more steps a day now for several years and feel no temptation to walk less again.

(**Q10**) Please select one of the following age categories: 18-27, 28-52, 53+

- 18-27

- 28-52
- 53+
(***Q11***) Which gender describes you the most?
- Male
- Female
- Other
(***Q12***) Do you have anything to add? Are there questions that need a motivation? Could you choose between all options well, or did you have difficulties? Otherwise just type 'bye'. Thank you for filling in this survey!
**This was an open text answer box that was not mandatory to fill in**


## A.6   Results of the pre-study

| Gender | Female | Male | Other | Total |
|---|---|---|---|---|
| | 60 | 61 | 3 | 124 |
| Age | Young (18-27) | Middle (28-52) | Old (53+) | Total |
| | 81 | 29 | 14 | 124 |
| Nationality | Greek | Italian | Mexican | |
| | 7 | 6 | 7 | |
| | Polish | Portuguese | South African | |
| | 9 | 15 | 30 | |
| | Spanish | United Kingdom | | |
| | 6 | 11 | | |
| | Dutch | Other | | Total |
| | 17 | 16 | | 124 |

**Table A.1: Distribution of gender, age and nationality of the pre-study. The 'other' nationalities occurred only once: African, Australian, Belgian, Canadian, Catalan, Chilean, Chinese, Czech (Republician), Estonian, French, Hungarian, Israeli, Latvian, Romanian, American, and Venezuelan.**


## A.7   Survey before the Experiment

Survey for step-count thesis
(***Q1***) What is your participant ID? This is to ensure your anonymity, and to make sure that your answers from this survey are coupled to your answers of the survey after the experiment.
**This was a mandatory open text answer box.**
**The** `Brainhex` **part:**
(***Q2***) Choose from a scale between "I love it!" for experiences you enjoy through "It's okay" to "I hate it!" for experiences you would rather avoid. If you have difficulty answering them, try imagining a work-related scenario or a gaming (boardgames or computer) scenario. The feeling that you would get in such a situation matters, do you like or dislike that feeling?
**The scales for the following statements were: I love it! / I like it. / It's okay. / I dislike it. / I hate it!**
- Exploring to see what you can find.
- Frantically escaping from a terrifying foe.
- Working out how to crack a challenging puzzle.
- The struggle to defeat a difficult boss.

- Playing in a group, online or in the same room.
- Responding quickly to an exciting situation.
- Picking up every single collectible in an area.
- Looking around just to enjoy the scenery.
- Being in control at high speed.
- Devising a promising strategy when deciding what to try next.
- Feeling relief when you escape to a safe area.
- Taking on a strong opponent when playing against a human player in a versus match.
- Talking with other players, online or in the same room.
- Finding what you need to complete a collection.
- Hanging from a high ledge.
- Wondering what's behind a locked door.
- Feeling scared, terrified or disturbed.
- Working out what to do on your own.
- Completing a punishing challenge after failing many times.
- Co-operating with strangers.
- Getting 100% (completing everything in a game).

(**Q3**) Please read through the following seven statements. Place them in order of preference by choosing the highest number for the most preferred statement to the lowest number for the least preferred. Please choose each number only once. So the WORST one is 1, and the BEST/preferred one is 7.

**The scales for the following statements where the numbers 1 (worst) up to 7 (best) could only be chosen once.**

- A moment of jaw-dropping wonder of beauty.
- An experience of primeval terror that blows your mind.
- A moment of breathtaking speed or vertigo.
- The moment when a solution to a difficult puzzle clicks in your mind.
- A moment of hard-fought victory.
- A moment when you feel an intense sense of unity with another player.
- A moment of completeness that you have strived for.

**The Susceptibility to Persuasion Scale part:**

(**Q4**) Please select how much you agree or disagree with the following statements.

**The scales for the following statements were: Strongly Agree / Agree / Somewhat Agree / Neutral / Somewhat Disagree / Disagree / Strongly Disagree**

- When a family member does me a favour, I am very inclined to return this favour.
- I believe rare products (scarce) are more valuable than mass products.
- I always follow advice from my general practitioner.
- Whenever I commit to an appointment I do as I told.
- I always pay back a favour.
- Please answer "Agree" here. **This was one of the sanity checks in the survey.**
- When my favourite shop is about to close, I would visit it since it is my last chance.
- I accept advice from my social network.
- When I am in a new situation I look at others to see what I should do.
- When a professor tells me something I tend to believe it is true.
- When I like someone, I am more inclined to believe him/her.
- I try to do everything I have promised to do.
- If someone from my social network notifies me about a good book, I tend to read it.

**The self-efficacy part:**

(**Q5**) Please rate your level of agreement with the following statements in a 7-point scale. Choosing 1-Strongly Disagree to 7-Strongly Agree.

**The scales for the following statements were: 1 Strongly Disagree / 2 / 3 / 4 Neutral / 5 / 6 / 7 Strongly Agree**

- If I want, I could easily walk more steps within the next two weeks.
- I have control over whether or not I walk more steps.
- Whether or not I walk more steps in the next week is entirely up to me.
- I believe I have the ability to walk more steps next week.
- I am confident that I would walk more steps within the next two weeks if I want.

**The attitude part:**

(**Q6**) Please rate your level of agreement with the following statements in a 7-point scale.

- Walking more steps in the next two weeks would be

**Scale: 1 Unimportant / 2 / 3 / 4 Neutral / 5 / 6 / 7 Important**

- Walking more steps in the next two weeks would be

**Scale: 1 Useless / 2 / 3 / 4 Neutral / 5 / 6 / 7 Useful**

- Walking more steps in the next two weeks would be

**Scale: 1 Worthless / 2 / 3 / 4 Neutral / 5 / 6 / 7 Valuable**

- Walking more steps in the next two weeks would be

**Scale: 1 Unenjoyable / 2 / 3 / 4 Neutral / 5 / 6 / 7 Enjoyable**

- Walking more steps in the next two weeks would be

**Scale: 1 Harmful / 2 / 3 / 4 Neutral / 5 / 6 / 7 Beneficial**

- Walking more steps in the next two weeks would be

**Scale: 1 Unpleasant / 2 / 3 / 4 Neutral / 5 / 6 / 7 Pleasant**

- Walking more steps in the next two weeks would be

**Scale: 1 Bad / 2 / 3 / 4 Neutral / 5 / 6 / 7 Good**

**The intention part:**

(**Q7**) Please rate your level of agreement with the following statements in a 7-point scale. Choose 1-Extremely unlikely 7-Extremely likely

**The scales for the following statements were: 1 Extremely Unlikely / 2 / 3 / 4 Neutral / 5 / 6 / 7 Extremely Likely**

- I intend to avoid walking less steps during the next two weeks
- I intend to take more stairs during the next two weeks
- I will try to take more walks during the next two weeks
- I intend to do longer walks during the next two weeks
- I intend to avoid the elevator during the next two weeks
- I will try to move around more during the next two weeks
- I will plan to walk more during the next two

**The Trans-Theoretical model part:**

(**Q8**) What situation applies to you the most?

- I have no intention to take action, with regard to the number of steps walked per day, in the next 6 months or foreseeable future.
- I am aware of my walking behavior and I have an intention to / I am thinking about taking action (regarding the number of steps per day) within the next 6 months, but not immediately.
- I have an intention to take action (with regard to the number of steps set per day) within the next 30 days and I have done some initial steps towards that action.
- I am actively taking more steps a day for a (short) while now, but I am still struggling regressing into fewer steps a day.
- I am actively taking more steps a day now for more than 6 months, and I prevent a relapse of walking fewer steps a day again.
- I have walked more steps a day now for several years and feel no temptation to walk less again.

**Special circumstances and demographics part:**

(**Q9**) Please select the special circumstances that you encountered during these 5 weeks:

- I plan on going on a holiday
- I am ill or have an injury
- I have trouble moving around in general
- I cannot take the phone with me all the time
- Other circumstances

**When selecting one of those options, the next page would show that question with an empty text answer box.**

(**Q10**) Will you synchronize Google Fit with the Pacer app?
- Yes
- No
- Partly **With an empty text answer box for explanations**

(**Q11**) Please select one of the following age categories: 18-27, 28-52, 53+
- 18-27
- 28-52
- 53+

(**Q12**) Which gender describes you the most?
- Male
- Female
- Other

(**Q13**) Would you prefer we start on the 27th of June with the step-count app, or on the 4th of July?
- 27 June till 31 July
- 4 July till 7 August
- Either way is fine

(**Q14**) Do you have anything to add? Are there questions that need a motivation? Could you choose between all options well, or did you have difficulties? Otherwise just type 'bye'. Thank you for filling in this survey!

**This was an open text answer box that was not mandatory to fill in**


## A.8   Survey after the Experiment

(**Q1**) What is your participant ID? This is to ensure your anonymity, and to make sure that your answers from this survey are coupled to your answers of the survey after the experiment. (The survey has 2 pages)

**This was a mandatory open text box.**

(**Q2**) Please rate your level of agreement with the following statements in a 7-point scale. Choosing 1-Strongly Disagree to 7-Strongly Agree.

**The scales for the following statement were: 1 Strongly Disagree / 2 / 3 / 4 Neutral / 5/ 6/ 7 Strongly Agree**
- If I want, I could easily walk more steps within the next two weeks.
- I have control over whether or not I walk more steps.
- Whether or not I walk more steps in the next week is entirely up to me.
- I believe I have the ability to walk more steps next week.
- I am confident that I would walk more steps within the next two weeks if I want.

**The attitude part:**

(**Q3**) Please rate your level of agreement with the following statements in a 7-point scale. *Note that the scales are different for every question.*
- Walking more steps in the next two weeks would be

**Scale: 1 Unimportant / 2 / 3 / 4 Neutral / 5 / 6 / 7 Important**

- Walking more steps in the next two weeks would be
**Scale: 1 Useless / 2 / 3 / 4 Neutral / 5 / 6 / 7 Useful**
- Walking more steps in the next two weeks would be
**Scale: 1 Worthless / 2 / 3 / 4 Neutral / 5 / 6 / 7 Valuable**
- Walking more steps in the next two weeks would be
**Scale: 1 Unenjoyable / 2 / 3 / 4 Neutral / 5 / 6 / 7 Enjoyable**
- Walking more steps in the next two weeks would be
**Scale: 1 Harmful / 2 / 3 / 4 Neutral / 5 / 6 / 7 Beneficial**
- Walking more steps in the next two weeks would be
**Scale: 1 Unpleasant / 2 / 3 / 4 Neutral / 5 / 6 / 7 Pleasant**
- Walking more steps in the next two weeks would be
**Scale: 1 Bad / 2 / 3 / 4 Neutral / 5 / 6 / 7 Good**
**The intention part:**
(**Q4**) Please rate your level of agreement with the following statements in a 7-point scale. Choose 1-Extremely unlikely 7-Extremely likely
**The scales for the following statements were: 1 Extremely Unlikely / 2 / 3 / 4 Neutral / 5 / 6 / 7 Extremely Likely**
- I intend to avoid walking less steps during the next two weeks
- I intend to take more stairs during the next two weeks
- I will try to take more walks during the next two weeks
- I intend to do longer walks during the next two weeks
- I intend to avoid the elevator during the next two weeks
- I will try to move around more during the next two weeks
- I will plan to walk more during the next two
**The Trans-Theoretical model part:**
(**Q5**) What situation applies to you the most?
- I have no intention to take action, with regard to the number of steps walked per day, in the next 6 months or foreseeable future.
- I am aware of my walking behavior and I have an intention to / I am thinking about taking action (regarding the number of steps per day) within the next 6 months, but not immediately.
- I have an intention to take action (with regard to the number of steps set per day) within the next 30 days and I have done some initial steps towards that action.
- I am actively taking more steps a day for a (short) while now, but I am still struggling regressing into fewer steps a day.
- I am actively taking more steps a day now for more than 6 months, and I prevent a relapse of walking fewer steps a day again.
- I have walked more steps a day now for several years and feel no temptation to walk less again.
(**Q6**) Please select the special circumstances that you encountered during these 5 weeks:
- I did go on a holiday.
- I got ill or had an injury.
- I have trouble moving around in general.
- I could not take the phone with me all the time.
- The weather influenced my walking behavior.
- Other circumstances.
**When selecting one of those option, the next page would show that question with an empty text answer box.**
Below, we will ask your top 3 and bottom 3, of the (21) sentences that you got during this experiment including a scale.
(**Q7**) Which sentence had the most positive influence on you? Meaning the sentence that made you change your behavior into setting more steps per day.

**This was a mandatory open text box question.**
(**Q8**) How much did that sentence influence you?
- **scale:** big negative effect / small negative impact / no effect / small positive effect / big positive effect

(**Q9**) Which sentence had the 2nd most positive influence on you? Meaning the sentence that made you change your behavior into setting more steps per day.
**This was a mandatory open text box question.**
(**Q10**) How much did that sentence influence you?
- **scale:** big negative effect / small negative impact / no effect / small positive effect / big positive effect

(**Q11**) Which sentence had the 3rd most positive influence on you? Meaning the sentence that made you change your behavior into setting more steps per day.
**This was a mandatory open text box question.**
(**Q12**) How much did that sentence influence you?
- **scale:** big negative effect / small negative impact / no effect / small positive effect / big positive effect

(**Q13**) Which sentence had the least influence on you? Meaning the sentence that made you not change your behavior into setting more steps per day.
**This was a mandatory open text box question.**
(**Q14**) How much did that sentence influence you?
- **scale:** big negative effect / small negative impact / no effect / small positive effect / big positive effect

(**Q15**) Which sentence had the 2nd least influence on you? Meaning the sentence that made you not change your behavior into setting more steps per day.
**This was a mandatory open text box question.**
(**Q16**) How much did that sentence influence you?
- **scale:** big negative effect / small negative impact / no effect / small positive effect / big positive effect

(**Q17**) Which sentence had the 3rd least influence on you? Meaning the sentence that made you not change your behavior into setting more steps per day.
**This was a mandatory open text box question.**
(**Q18**) How much did that sentence influence you?
- **scale:** big negative effect / small negative impact / no effect / small positive effect / big positive effect

(**Q19**) Did you synchronize Google Fit with the Pacer app?
- Yes
- No
- Partly **With an empty text box**

(**Q20**) Which mobile device did you use?
- Android phone
- IOS Apple phone

(**Q21**) Would you like to chat with me about the experiment? Questions/suggestions/comments are welcome.
- Yes, via Whatsapp
- Yes, via phone-call
- Yes, via email
- Yes, via Teams
- No, thank you : )

(**Q22**) Are you Dutch?
- Yes

- No, but I am **with an open text box**

(**Q23**) Do you have anything to add? Are there questions that need a motivation? Could you choose between all options well, or did you have difficulties? Did you enjoy the experiment or not? Otherwise just type 'bye'. Thank you for filling in this survey!

**This was a non-mandatory open text box**

(**Q24**) Here you can paste the link with your payment request / betaalverzoek / tikki.

**This was a non-mandatory open text box, as participants could also send it via Whatsapp.**

## A.9   Results of the Experiment

In Table A.2, the distribution of nationalities of the participants are shown and their respective experiment groups.

|  | Group 1 | Group 2 | Group 3 | Total |
|---|---|---|---|---|
| Gender Other | 0 | 0 | 1 | 1 |
| Male | 4 | 3 | 2 | 9 |
| Female | 10 | 12 | 11 | 33 |
| Age Young (18-27) | 10 | 11 | 9 | 30 |
| Middle (28-52) | 2 | 3 | 3 | 8 |
| Old (53+) | 2 | 1 | 2 | 5 |
| Nationality Dutch | 10 | 11 | 14 | 35 |
| Indian | 2 | 0 | 0 | 2 |
| Greek | 1 | 0 | 0 | 1 |
| Romanian | 1 | 0 | 0 | 1 |
| Japanese | 0 | 1 | 0 | 1 |
| Chinese | 0 | 1 | 0 | 1 |
| German | 0 | 1 | 0 | 1 |
| Ecuadorian | 0 | 1 | 0 | 1 |

**Table A.2: Distribution of gender, age and nationality of the experiment.**

| Distribution of profiles | Group 1 | Group 2 | Group 3 | Total |
|---|---|---|---|---|
| Brainhex achiever | 3 (3) | 3 (3) | 2 (6) | 12 (12) |
| conqueror | 2 (2) | 2 (4) | 2 (3) | 6 (9) |
| daredevil | 1 (0) | 1 (0) | 1 (0) | 3 (0) |
| mastermind | 3 (2) | 4 (3) | 3 (2) | 10 (7) |
| seeker | 4 (3) | 4 (4) | 4 (0) | 12 (7) |
| socializer | 1 (0) | 1 (0) | 2 (0) | 4 (0) |
| survivor | 0 (1) | 0 (0) | 0 (0) | 0 (1) |
| TT precontemplation | $3 \to 4(+)$ | $3 \to 4(+)$ | $4 \to 4(\pm)$ | $10 \to 12(+)$ |
| contemplation | $4 \to 2(-)$ | $3 \to 5(+)$ | $4 \to 6(+)$ | $11 \to 13(+)$ |
| preparation | $3 \to 3(\pm)$ | $4 \to 1(-)$ | $2 \to 3(+)$ | $9 \to 7(-)$ |
| action | $3 \to 4(+)$ | $2 \to 2(\pm)$ | $2 \to 0(-)$ | $7 \to 6(-)$ |
| maintanance | $0 \to 0(\pm)$ | $1 \to 1(\pm)$ | $1 \to 1(\pm)$ | $2 \to 2(\pm)$ |
| termination | $1 \to 1(\pm)$ | $2 \to 2(\pm)$ | $1 \to 0(-)$ | $4 \to 3(-)$ |
| STPS authority | 0 | 1 | 0 | 1 |
| commitment | 5 | 6 | 4 | 15 |
| liking | 1 | 3 | 2 | 6 |
| reciprocity | 5 | 5 | 2 | 12 |
| scarcity | 3 | 0 | 6 | 9 |
| consensus | 0 | 0 | 0 | 0 |

Table A.3: Distribution of `Brainhex` profiles, `TTM` profiles and `STPS` profiles of the experiment. For the distribution of main and sub-profiles of `Brainhex`, the ones in brackets are the sub-profiles. For the distribution of the `TTM` both before and after the experiment are noted and the number stands for the amount of people in that specific `TTM`.
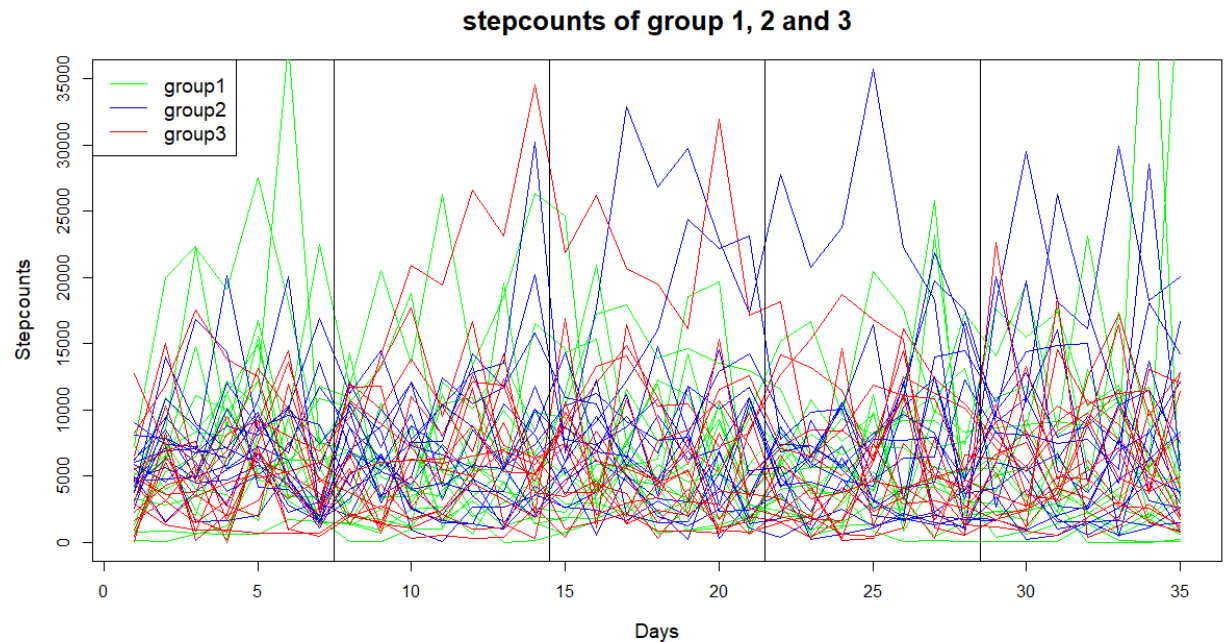


Figure A.4: Line plot of all groups for weeks 1 to 5.
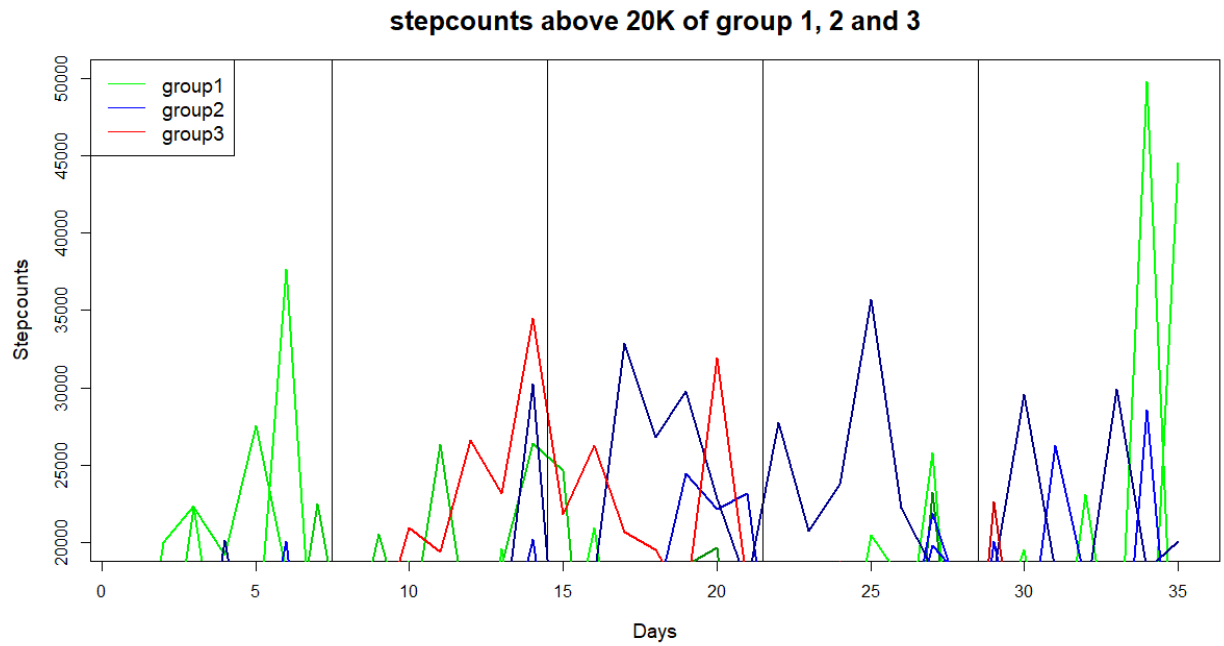
**Figure A.5: Line plot of participants who walked more than 20k on one day for weeks 1 to 5.**



**Figure A.6: Box-plots of normal days versus holidays per group. Only holidays with at least 6 days were considered.**

Figure A.7: Looking for outliers during the full experiment in group 1.

Figure A.8: Looking for outliers during the full experiment in group 2.

Figure A.9: Looking for outliers during the full experiment in group 3.

Figure A.10: Box-plots of each group of each day during the whole experiment.

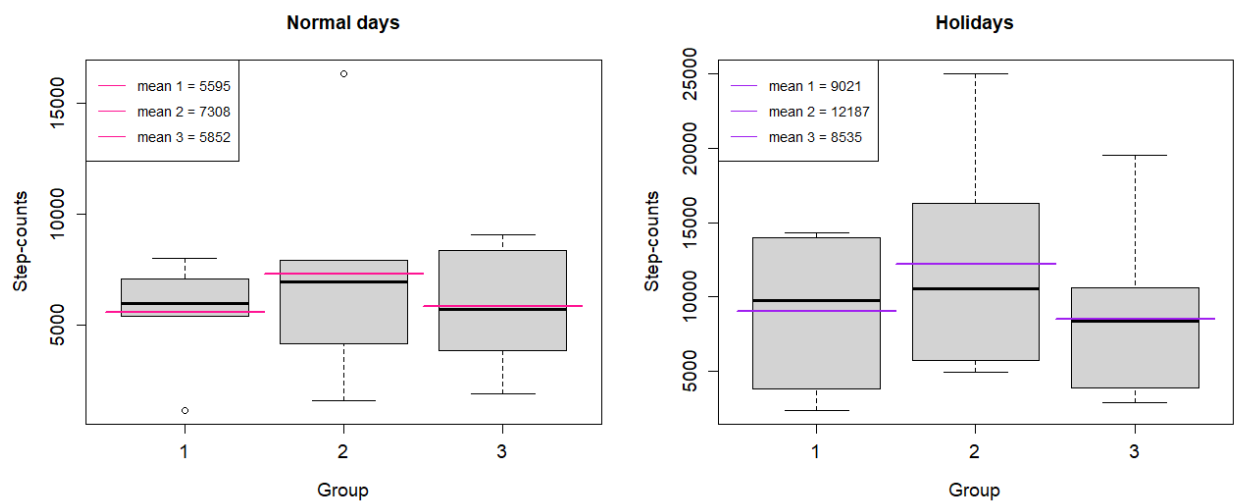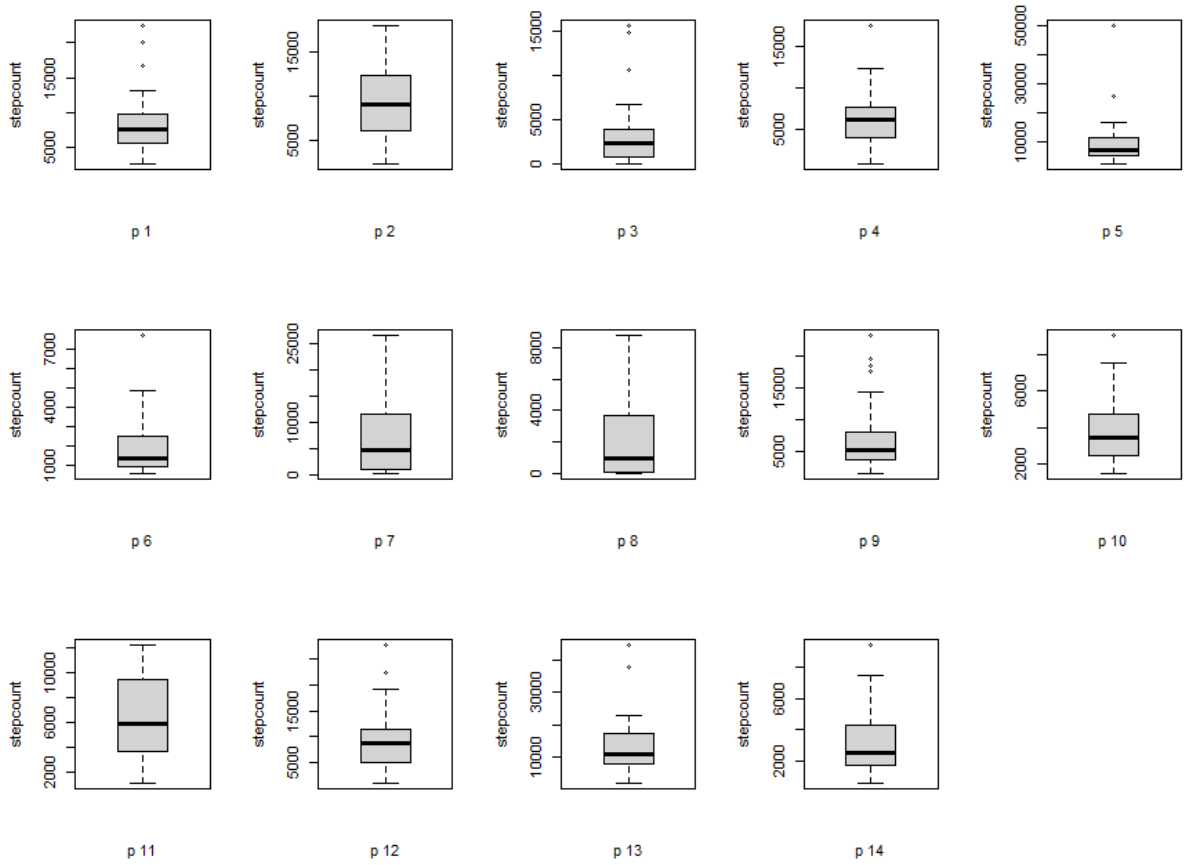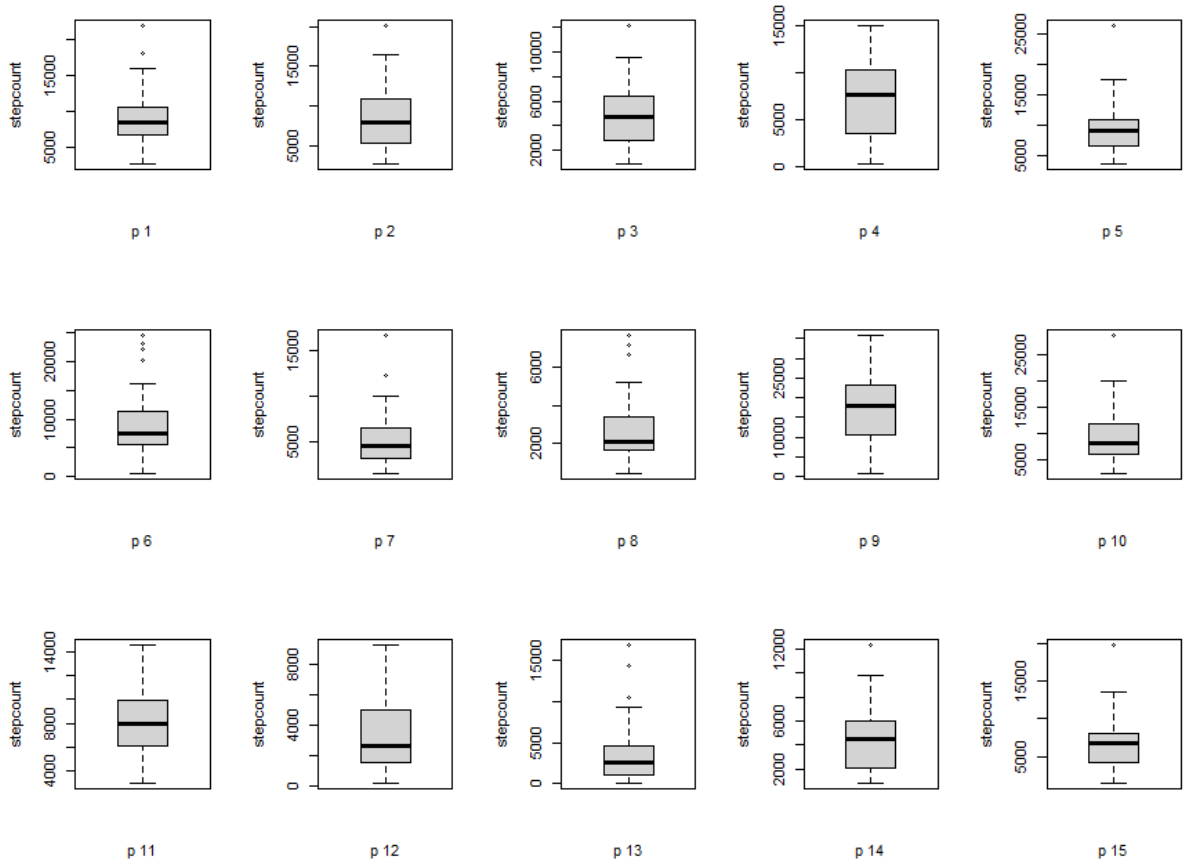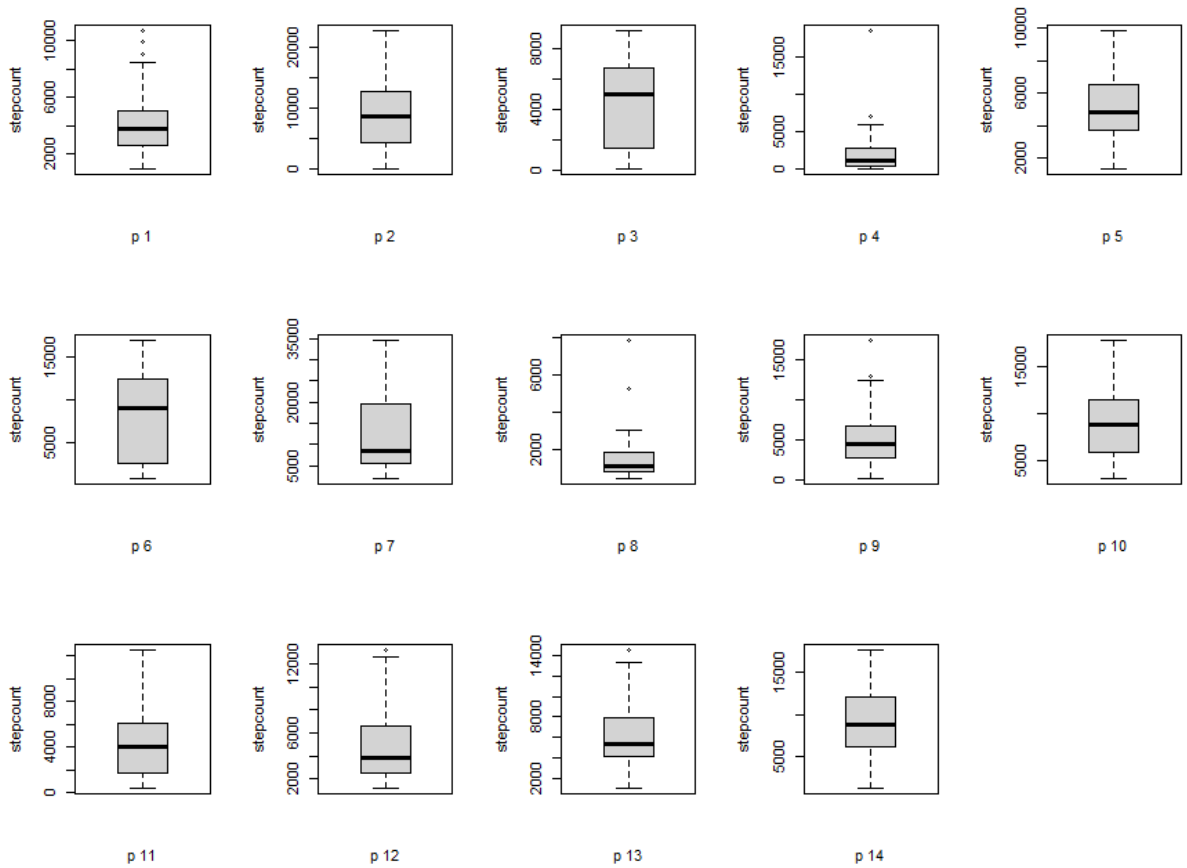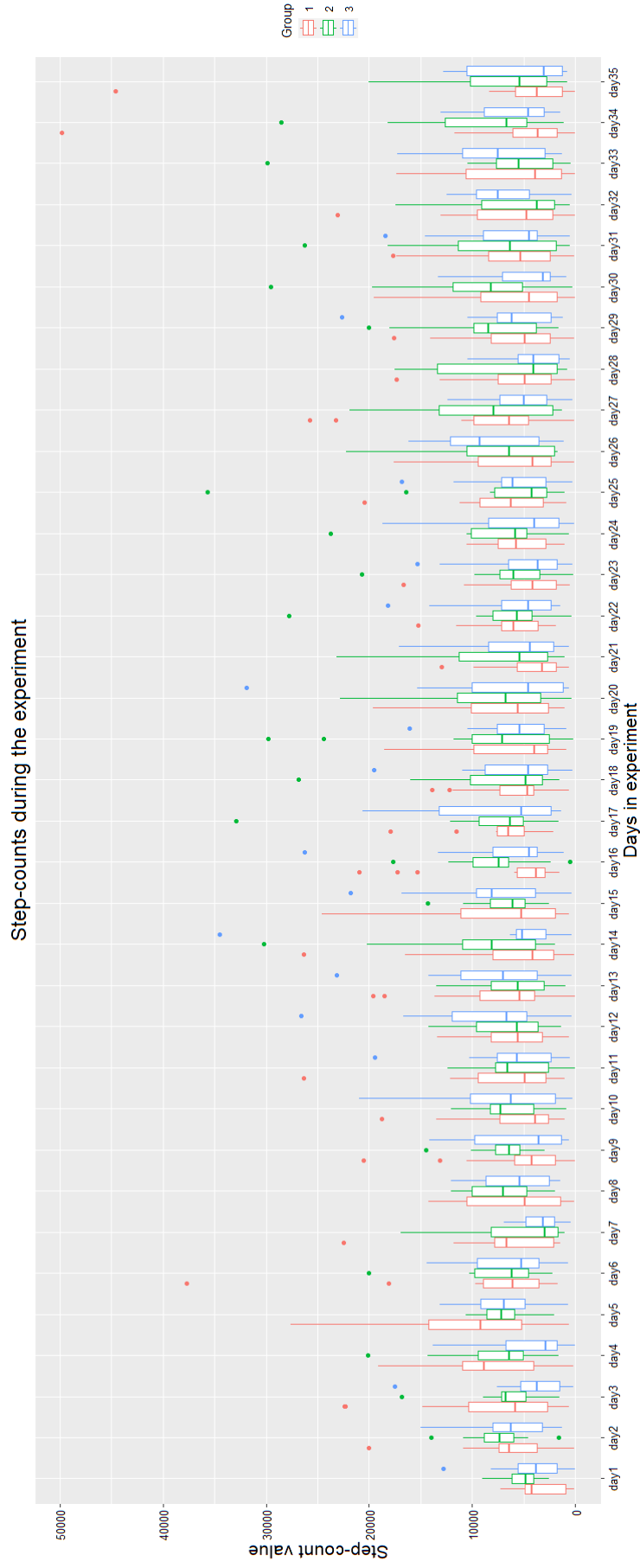| Day | Temp max | Temp avg | Temp min | Rain mm | Rain hours | Wind (m/s) |
|---|---|---|---|---|---|---|
| June 27 | 19.9 | 16.2 | 12.3 | 6.3 | 2.8 | 2 |
| June 28 | 23.0 | 17.7 | 10.3 | 0 | 0 | 2 |
| June 29 | 26.9 | 20.6 | 11.5 | 0 | 0 | 3 |
| June 30 | 23.6 | 17.8 | 13.8 | 9.5 | 2.8 | 3 |
| July 1 | 20.5 | 16.2 | 11.3 | 0 | 0 | 4 |
| July 2 | 23.9 | 18.1 | 10.2 | 0 | 0 | 3 |
| July 3 | 21.1 | 16.7 | 10.7 | 0.6 | 0.2 | 3 |
| July 4 | 22.1 | 16.7 | 9.4 | 0 | 0 | 2 |
| July 5 | 20.7 | 15.9 | 9.8 | 0 | 0 | 3 |
| July 6 | 21.7 | 16.8 | 9.0 | 0 | 0 | 3 |
| July 7 | 18.9 | 16.4 | 14.6 | 0.6 | 0.9 | 4 |
| July 8 | 23.5 | 17.9 | 9.7 | 0 | 0 | 2 |
| July 9 | 21.7 | 17.1 | 12.3 | 0 | 0 | 3 |
| July 10 | 21.1 | 17.3 | 12.9 | 0 | 0 | 3 |
| July 11 | 23.4 | 18.3 | 12.7 | 0 | 0 | 2 |
| July 12 | 26.8 | 20.8 | 10.5 | 0.5 | 0.6 | 2 |
| July 13 | 27.2 | 21.5 | 13.6 | 0 | 0 | 3 |
| July 14 | 22.5 | 17.2 | 9.8 | 0 | 0 | 2 |
| July 15 | 22.0 | 17.0 | 8.4 | 0 | 0 | 2 |
| July 16 | 21.6 | 17.2 | 9.6 | 0 | 0 | 3 |
| July 17 | 26 | 17.8 | 8.0 | 0 | 0 | 1 |
| July 18 | 32.2 | 22.3 | 10.6 | 0 | 0 | 2 |
| July 19 | 35.5 | 26.9 | 14.2 | 0 | 0 | 3 |
| July 20 | 27.7 | 24 | 18.6 | 0.3 | - | 3 |
| July 21 | 19.8 | 17.7 | 15.6 | 8 | 6.2 | 3 |
| July 22 | 20.7 | 17.5 | 12.7 | <0.05 | - | 2 |
| July 23 | 25.1 | 18.3 | 10.3 | 0 | 0 | 1 |
| July 24 | 29.4 | 22.8 | 13.5 | 0 | 0 | 3 |
| July 25 | 24.5 | 21 | 17.4 | <0.05 | - | 4 |
| July 26 | 20.2 | 17.1 | 10.5 | 3.9 | 1.4 | 3 |
| July 27 | 20.9 | 15.8 | 8.8 | 0 | 0 | 3 |
| July 28 | 22.3 | 17.2 | 11.5 | 0 | 0 | 4 |
| July 29 | 24.5 | 18.9 | 14.5 | 0 | 0 | 3 |
| July 30 | 24.4 | 19.2 | 11.0 | <0.05 | - | 2 |
| July 31 | 22.0 | 19.1 | 16.1 | 1.9 | - | 3 |

Table A.4: Table with the details on the weather during the 5 week experiment. The entries with '-' had no such data on those days.

| ID | Group | Nationality | Self-efficiacy | Attitude | Intention | TTM |
|---|---|---|---|---|---|---|
| NF2 | 1 | Dutch | 30 26 (-4) | 34 36 (+2) | 33 29 (-4) | 1 3 (+2) |
| NM9 | 1 | Dutch | 33 32 (-1) | 35 40 (+5) | 34 37 (+3) | 0 0 (±0) |
| NF7 | 1 | Dutch | 28 35 (+7) | 39 43 (+4) | 32 31 (-1) | 2 2 (±0) |
| NF9 | 1 | Dutch | 24 26 (+2) | 36 39 (+3) | 35 37 (+2) | 2 3 (+1) |
| EF11 | 1 | Greek | 30 35 (+5) | 46 37 (-9) | 22 14 (-8) | 0 0 (0) |
| EM10 | 1 | Indian | 34 29 (-5) | 46 46(±0) | 39 37 (-2) | 2 1 (-1) |
| NF17 | 1 | Dutch | 30 31 (+1) | 34 35 (+1) | 27 29 (+2) | 3 2 (-1) |
| NF18 | 1 | Dutch | 23 11 (-12) | 7 19 (+12) | 28 7 (-21) | 0 0 (±0) |
| EM5 | 1 | Indian | 29 22 (-7) | 48 41 (-7) | 46 42 (-4) | 3 3 (±0) |
| NM7 | 1 | Dutch | 20 22 (+2) | 35 32 (-3) | 25 22 (-3) | 1 0 (-1) |
| NF29 | 1 | Dutch | 34 27 (-7) | 35 35 (±0) | 30 21 (-9) | 1 1 (±0) |
| NF31 | 1 | Dutch | 30 25 (-5) | 28 41 (+13) | 28 25 (-3) | 5 2 (-3) |
| EF32 | 1 | Romanian | 28 27 (-1) | 33 26 (-7) | 32 44 (+12) | 1 3 (+2) |
| NF35 | 1 | Dutch | 33 17 (-16) | 48 48 (±0) | 36 28 (-8) | 3 5 (+2) |
| NF1 | 2 | Dutch | 29 27 (-2) | 28 35 (+7) | 35 31 (-4) | 4 5 (+1) |
| NF3 | 2 | Dutch | 23 23 (±0) | 26 25 (-1) | 7 7 (±0) | 0 0 (±0) |
| NF4 | 2 | Dutch | 30 30 (±0) | 45 43 (-2) | 48 37 (-11) | 3 3 (±0) |
| NF13 | 2 | Dutch | 24 27 (+3) | 35 47 (+12) | 16 29 (+13) | 5 4 (-1) |
| NM6 | 2 | Dutch | 34 23 (-11) | 40 30 (-10) | 33 21 (-12) | 2 0 (-2) |
| EM3 | 2 | Ecuadorian | 35 15 (-20) | 28 38 (+10) | 31 30 (-1) | 1 1 (±0) |
| NF16 | 2 | Dutch | 27 27 (±0) | 36 36 (±0) | 28 29 (+1) | 0 1 (+1) |
| EF19 | 2 | Dutch | 18 23 (+5) | 19 29 (+10) | 7 34 (+27) | 1 1 (±0) |
| EF21 | 2 | Japanese | 31 20 (-11) | 36 35 (-1) | 29 31 (+2) | 3 2 (-1) |
| NF23 | 2 | Dutch | 35 30 (-5) | 37 36 (-1) | 28 27 (-1) | 5 5 (±0) |
| EF24 | 2 | Dutch | 31 32 (+1) | 30 29 (-1) | 29 25 (-4) | 1 0 (-1) |
| EF36 | 2 | Chinese | 35 33 (-2) | 25 28 (+3) | 21 15 (-6) | 0 0 (±0) |
| NF25 | 2 | Dutch | 35 33 (-2) | 39 40 (+1) | 32 33 (+1) | 2 3 (+1) |
| NM8 | 2 | Dutch | 20 27 (+7) | 35 35 (±0) | 29 37 (+8) | 2 1 (-1) |
| EF38 | 2 | German | 19 20 (+1) | 40 39 (-1) | 36 26 (-10) | 2 1 (-1) |
| NM1 | 3 | Dutch | 33 35 (+2) | 27 29 (+2) | 14 14 (±0) | 0 0 (±0) |
| EM2 | 3 | Dutch | 29 29 (±0) | 49 40 (-9) | 41 34 (-7) | 4 4 (±0) |
| NF34 | 3 | Dutch | 20 19 (-1) | 25 17 (-8) | 18 15 (-3) | 0 0 (±0) |
| NF5 | 3 | Dutch | 25 25 (±0) | 32 33 (+1) | 31 30 (-1) | 1 1 (±0) |
| NF6 | 3 | Dutch | 15 30 (+15) | 36 37 (+1) | 35 28 (-7) | 1 2 (+1) |
| EF8 | 3 | Dutch | 17 24 (+7) | 36 37 (+1) | 27 26 (-1) | 0 1 (+1) |
| EF15 | 3 | Dutch | 15 26 (+11) | 28 23 (-5) | 13 28 (+15) | 5 0 (-5) |
| NF20 | 3 | Dutch | 30 15 (-15) | 39 33 (-6) | 41 21 (-20) | 3 1 (-2) |
| NF22 | 3 | Dutch | 30 29 (-1) | 25 25 (±0) | 22 12 (-10) | 0 0 (±0) |
| NF26 | 3 | Dutch | 19 22 (+3) | 35 32 (-3) | 35 33 (-2) | 3 2 (-1) |
| NF27 | 3 | Dutch | 32 23 (-9) | 43 40 (-3) | 35 30 (-5) | 2 2 (±0) |
| NF28 | 3 | Dutch | 23 29 (+6) | 46 43 (-3) | 40 34 (-6) | 2 1 (-1) |
| NF30 | 3 | Dutch | 25 26 (+1) | 36 38 (+2) | 27 27(±0) | 1 1 (±0) |
| NF37 | 3 | Dutch | 22 20 (-2) | 28 35 (+7) | 28 33 (+5) | 1 1 (±0) |

Table A.5: Table with the details on self-efficacy, attitude, intention and the TTM of the participants before and after the experiment. On the left side of each column, the score of the first survey is given and next to that is the score of the survey after the experiment. After that, the difference is given between brackets ().

| Sentence | Strategy | Times mentioned | Average score | Effect on walking more |
| --- | --- | --- | --- | --- |
| On average, participants walked 5502 yesterday. You walked ## steps yesterday. | Consensus | 5 | 3.6 | small positive effect |
| About 9% of the participants has done more than 4.000 steps yesterday. You walked ## steps yesterday. | Consensus | 4 | 3.5 | small positive effect |
| Walking is not only good for your health but also for the environment. | Tips | 6 | 3.5 | small positive effect |
| Have a dance party with your neigh-bours/(grand)children or even pets, to increase your step-count. | Tips | 4 | 4 | small positive effect |
| Do remember to try and walk for (at least) half an hour today, good luck! | Reminders | 3 | 3.67 | small positive effect |
| It is better to walk than to sit during the day. | Information | 4 | 3.5 | small positive effect |
| Do you remember that you reached ## steps during the last 7 days? Let us try to reach this number again today! | Reminders | 6 | 3.83 | small positive effect |
| A new day arose with new opportunities to walk more! | Reminders | 4 | 3.5 | small positive effect |
| Hi ##, you have increased your step-count with ## since you began the experiment, keep it going! (Yesterday you walked ## steps.) | Personalization | 4 | 4 | small positive effect |
| Hello ##, you probably have some time left today to walk around the block. (Yesterday you walked ## steps.) | Personalization | 3 | 2 | small negative effect |

Table A.6: Table of each sentence and how often it was mentioned and it's average score in the survey after the experiment. Scores were between 0 (big negative effect) and 5 (big positive effect). Reminders and tips were sent to all experiment groups. To keep the sentences anonymous, the names and step-count values are replaced with '##'.

| Sentence | Strategy | Times mentioned | Average score | Effect on walking more |
|---|---|---|---|---|
| Someone was only ## steps behind you yesterday, to keep ahead, try to walk at least ## steps today. (Yesterday you walked ## steps.) | Competition | 4 | 3.75 | small positive effect |
| You walked more than the average (5582) of the participants in your group yesterday. (You walked ## steps yesterday.) | Competition | 3 | 4.67 | small positive effect |
| You accomplished to have more than ## steps in total. (Of which ## in the last 7 days. Yesterday you walked ## steps.) | Rewards | 3 | 3.67 | small positive effect |
| A study conducted by Amanda Paluch, PhD, assistant professor in the department of kinesiology at the University of Massachusetts Amherst found that taking at least 7.000 steps per day was associated with a 50-70% lower chance of early death among middle-aged adults. | Authority | 7 | 3.86 | small positive effect |
| In a National Health and Nutrition Examination Survey, people who walked for about 8,000 steps a day were half as likely to die prematurely from heart disease or any other cause as those who accumulated 4,000 steps a day. | Authority | 4 | 4 | small positive effect |
| Your step-count of yesterday was not so good (##), aim higher. | Punishment | 3 | 2 | small negative effect |

**Table A.7: Table of each sentence and how often it was mentioned and it's average score in the survey after the experiment. Scores were between 0 (big negative effect) and 5 (big positive effect). Reminders and tips were sent to all experiment groups. To keep the sentences anonymous, the names and step-count values are replaced with '##'.**

### A.9.1   Interviews

**Group 1:**

(1) Het meedoen middels de app op telefoon vond ik minder handig. Vergat hem vaak mee te nemen. Stappenteller is veel makkelijker. Vond de app installeren en goed aan de praat krijgen wat lastig zoals je in het begin hebt gemerkt. Ook elke einde dag openen van de app bleek gaandeweg nodig. Berichtjes kende ik al qua inhoud van Ommetje app van de Hersenstichting. Dit onderzoek heeft niet echt bijgedragen aan mijn loopgedrag dat ik inmiddels zo'n 2 jaar heb opgebouwd tijdens Corona periode/thuiswerken.

(2) The procedure of participating is clear to me. I don't feel the messages make a big difference (short-term or long-term) in general since I'm not setting up the goal for myself and I feel having moved enough with other methods (e.g. biking) Thank you for the hard work anyway and it's nice to collaborate with you :)

(3) Ik vond het leuk om mee te doen! Leuk om bij te houden hoeveel stappen je loopt op een dag. Je probeert dan toch de dag erna meer te lopen! Met de berichtjes deed ik niks, ik las ze alleen. Het heeft geen effect op langere termijn omdat ik de stappen nu niet meer bijhoud. Ik vind dat met telefoon niet heel handig omdat je die wel eens vergeet. Dan moet ik al een horloge kopen.. Maar leuk experimentje!!

(4) - leuk om mee te doen. Je kon in je groep zien dat er soms flink werd gewerkt aan het aantal stappen 😁 dat had op mij trouwens geen effect- ik moest soms even 2 x lezen, mijn Engels is hierdoor weer flink verbeterd. Maar het was goed duidelijk - eigenlijk heeft het bij mij geen effect, volgens mij. Ik loop, wandel wanneer het mij uit komt en in principe sport ik 3 a 4 uurtjes per week. Ik vind dat voldoende. Ik draag al 3 jaar een fitbit maar let niet op de stappen. In het begin wel maar heb hem ook niet meer gekoppeld aan mijn mobiel.  Dat was wel een dingetje, normaal heb ik de telefoon wel in de buurt maar niet aldoor bij me. Daar moest ik erg om denken  - berichtjes waren soms leuk en informatief maar deed er niet iets mee. Soms was het bedoeld als extra stimulans maar dacht ik zelf dat ik wel voldoende deed. - ik denk dat ik nu zo in het leven sta dat ik doe wat ik leuk vind en probeer daarbij "voldoende" te bewegen.  Wat dat betreft ben ik niet de juiste persoon voor zo'n beweeg app Wat wel goed mee werkte was het weer. Ik denk dat je over het algemeen hele andere resultaten had gehad bij regen en kou.

(5) Vond het prima om mee te doen, alles was wel duidelijk en ook de instructies zijn goed te volgen wat mij betreft. Ik heb de berichten wel gelezen enzo maar ik moet zeggen dat ik niet extra ben gaan wandelen ofzo, maar dat kwam ook een beetje omdat ik gefrustreerd was dat de app meer stappen aangaf dan mijn andere stappenteller 😄. Effect op de langere termijn lijkt me lastig te beoordelen, denk dat dat heel erg ligt aan wat voor persoonlijkheid je hebt en of je snel beïnvloedt wordt door zulke berichten.

**Group 2:**

(1) Vraag 1 vond het wel oke, vragen lijsten wel erg lang en veel, advies volgende keer eerst in Nederlands vertalen. Vraag 2 ja dat was geen probleem. Vraag 3 nee heb er niet meer om gelopen, komt ook door het verschil in stappen met de app en mijn horloge, is niet motiverend. Vraag 4 nee heb niets jet de berichten gedaan, alleen gelezen. Vraag 5 voor mij persoonlijk niet  kan me wel voorstellen dan andere mensen er wel iets mee kunnen.

(2) Ik vond het geen probleem om te doen. Het koste weinig extra tijd. Over het algemeen was het erg duidelijk. Voor mij heeft het geen verschil gemaakt in de hoeveelheid stappen die ik zet. De appjes vond ik persoonlijk niet erg prettig, kwam een beetje over als m'n moeder die vroeger zeurde dat ik iets moest doen 😂. Ik vond het jammer dat ik niet in mijn startscherm kon zien hoeveel stappen ik die dag gezet had. Ik denk dat zoiets beter zou helpen. Ik zou liever helemaal geen appje krijgen

(3) I would like to thank you for selecting me to participate in the experiment and helping me do some daily activity as part of it. Yes of course I enjoyed the messages. Yes I do think it could work in the long run and I do think about walking more often. My walking has definitely increased after I installed the app and now too. It's higher than the usual.

(4) **[from an audio file]** Ja, ehm, ik realiseerde met net dat ik eigenlijk best nog vaak mijn stappen tel sinds dit onderzoek. Dus dat is wel grappig want dat deed ik, daarvoor heb ik dat echt nog nooit van mijn leven gedaan. En ik vind het eigenlijk best wel leuk om dat af en toe checken en dan denk ik, ohja, ik heb wel genoeg bewogen. Dat geeft me dan wel een goed gevoel. En dan merk ik ook van, als ik bijvoorbeeld een lange dag op kantoor heb gezeten dat ik dan, denk oeh ik moet eigenlijk wel nu even lopen naar een winkel, of nog even een extra blokje om. Dus ik zou best kunnen zeggen dat het een effect heeft gehad op lange termijn. Alleen ja ik weet dus niet of dat echt door de inhoud van de berichten kwam, want ja voor mij idee ben ik daar niet heel erg mee bezig geweest. Maar ja ik heb meer het idee dat het gewoon het hele er bezig mee zijn met, dat je elke dag een cue krijgt. Dat dat vooral hetgeen is wat heeft geholpen, en dat jij bijvoorbeeld telkens vroeg van wil je dan even kijken in die app om hem weer te synchroniseren. Omdat ik dat dan bijna elke dag deed, zag ik het toch weer, die visualisatie en dat is wat ik dan wel weer leuk vind. Dus, ja ik denk dat onbewust die habit er een beetje in is gekropen dus dat ik wel eigenlijk wel grappig hoe dat is gebeurd.

**Group 3:**

(1) Leuk om mee te doen! Alles was wel duidelijk en ik merkte dat ik me veel bewuster was van de stappen. Ik kijk nu soms bij gezondheid maar vergeet ook te kijken. Ik denk dat ik de bewuste app of een andere weer ga gebruiken om me toch weer de bewustwording te geven. Dus dat zou wel het lange termijn effect zijn. 😊 Meestal waren de berichtjes leuk en stimuleren en gaven ze het gevoel dat je het morgen beter moest doen.

(2) Op zich leuk, het kostte niet veel moeite. Het was erg duidelijk, maar of het heeft geholpen weet ik niet. Ik loop van mijzelf al erg veel op een dag! Van de berichtjes werd ik erg competitief 😤 😂 Misschien voor de langere termijn wel, op dagen dat ik vrij was merkte ik dat ik het fijner vond als ik er toch wat stappen inkreeg

(3) Meedoen was leuk, interessant om eens wat beter op de hoogte te zijn van hoeveel stappen je zet. Toelichting was duidelijk, jammer (en beetje onpraktisch) alleen dat het net iets anders werkte voor de iPhone. De berichtjes deden mij eigenlijk niet zoveel, misschien zouden ze wel werken als het om een groep gaat van mensen die ik ken en/of vaak zie. Op die manier zouden de competitieve berichten in ieder geval meer aanslaan denk ik. Ik denk dat de bewustwording over hoeveelheid stappen wel iets doet op lange termijn, ik merk dan ik momenteel ook nog meer op stappen let dan hiervoor (alleen al omdat zo'n cijfertje op m'n home screen staat)

(4) - Ja ik vond het leuk om mee te doen! Ik weet dat ik wel redelijk veel stappen zet in het dagelijks leven, dus ik vond het wel leuk om dat bevestigd te hebben. - Het was zeker duidelijk, de documenten met pijlen en alles waren heel duidelijk en daarom was het ook gemakkelijk om mee te doen. - Ik denk niet dat ik meer of minder stappen ben gaan zetten door de berichten en de app. Ik heb zelf ook een sporthorloge en vroeger hechtte ik veel waarde aan stappen/activiteits doelen etc., maar ik heb juist wat meer geleerd om dat niet zo letterlijk te nemen dus daarom liet ik deze berichtjes ook wat meer aan me voorbij gaan. Op de dagen dat ik weinig stappen had, kwam dat meestal doordat ik al had gesport (zonder telefoon), dus daarom wist ik wel beter.- Ik denk dat het daarom niet per se effect heeft op de lange termijn op deze manier, maar als het bijvoorbeeld wel een accurate stappenteller is (zoals sporthorloges) misschien wel, en dan natuurlijk ook wat meer gepersonaliseerd. Wel denk ik dat het sociale aspect ook wel helpt, zeker als je het met vrienden/collega's zou doen, denk ik dat het een leuke competitie kan zijn (zoals FitBit ook al heeft).