# Can Digital Personal Assistants persuade people to exercise?

Paay, Jeni; Kjeldskov, Jesper; Papachristos, Eleftherios; Hansen, Kathrine Maja; Jørgensen, Tobias; L. Overgaard, Katrine

*Published in:*
Behaviour and Information Technology

*DOI (link to publication from Publisher):*
10.1080/0144929X.2020.1814412



*Publication date:*
2022

*Document Version*
Accepted author manuscript, peer reviewed version

# Can Digital Personal Assistants persuade people to exercise?

Jeni Paay, Swinburne University of Technology, Australia
Jesper Kjeldskov, Aalborg University, Denmark
Elefterios Papachristos, Aalborg University, Denmark
Kathrine Maja Hansen, Aalborg University, Denmark
Tobias Jørgensen, Aalborg University, Denmark
Katrine Leth Overgaard, Aalborg University, Denmark

## Abstract

Digital personal assistants (DPAs) have recently grown in popularity because they are both a commercially available new technology and reasonably affordable to the average household. This opens opportunities for new ways to assist people in everyday activities in their homes through voice-interaction. Physical activity has significant health benefits, and yet globally, 1 in 4 adults are not active enough. To address this, we investigate the persuasive potential of DPAs in increasing people's physical activity at home. We conducted a study with 48 participants to understand the effect of applying three of Fogg's persuasive principles to the design of a DPA exercise program: Suggestion, Virtual Reward, and Praise. Our findings show that DPAs have the potential, within their current technical and reactive capabilities, to persuade people to increase their physical activity at home, using Suggestion to encourage physical effort, Virtual Reward to encourage endurance, and Praise to create reassurance for beginners. Based on this, we offer three alternate perspectives for developing persuasive DPAs. We also discuss limitations of the study and suggest future research directions around using persuasion with DPAs.

## Keywords

Conversational Agents; Digital Personal Assistants; Smart Home; Persuasive Technology; Physical activity.

## 1. Introduction

The interactional benefits of human-computer voice interaction have been known for decades (Cohen & Oviatt, 1995), however, it is not until recently it has become stable enough for everyday use, and widely available through in-home devices. Digital Personal Assistants (DPAs), such as Amazon Echo and Google Home, have gained popularity and acceptance in recent years (Atkinson 2006; Kinsella 2019; NPR 2019). These hands-free voice-controlled interfaces aim to support users on an everyday basis with planning, searching for information, controlling smart home accessories and entertainment (Kinsella 2019; Porcheron et al. 2018; Pradhan et al. 2018). The DPA embodies the idea of a butler or assistant, helping the user and household manage their everyday lives.

In line with recent interest in using technology to motivate increased physical activity (Orji & Moffatt, 2018), there is potential for DPAs to assist users by defining exercise programs to increase daily physical activity. According to World Health Organization (WHO 2018), physical activity has significant health

benefits, and yet insufficient physical activity is one of the leading risk factors for death worldwide. More than 80% of the world's adolescent population and 1 in 4 adults are insufficiently physically active. It is therefore important to find ways to motivate people to become more active.

The relative maturity and stability of conversational agents creates an opportunity to understand, explore, and expand the capabilities of DPAs as agents of persuasion. Given their central location in the home and relationship with the user they are well placed to persuade people to increase physical activity in their everyday lives. Using DPAs for persuasion is a new research area in HCI, with limited specific studies, however, there is a substantial existing body of work around persuasion and voice-controlled systems, early conversational agents, and embodied conversational agents (ECAs), which informs this research (Andrews 2012, Bickmore & Picard 2005, Schulman & Bickmore 2009).

To extend existing persuasion through technology research into the new area of DPAs, we designed a user study to evaluate whether DPAs could persuade people to exercise. A custom-made DPA was developed specifically to motivate people in a physical activity session, with user feedback focusing on three of B. J. Fogg's (2003) persuasive principles: *Suggestion*, *Virtual Reward*, and *Praise*. The study aimed to assess the persuasive power of digital assistants and compare the effect of these three different persuasive principles, in the activity context.

Our work contributes to HCI knowledge on DPAs and persuasion in three ways. Firstly, based on results from the study, we propose that DPAs can be designed to have the power to persuade people to increase their daily exercise. Secondly, we show that there is a difference in the ways that the three persuasive principles examined are able to influence: the amount of physical effort people put in to exercising; the amount of time they exercise for; and how encouraged they feel to continue exercising. Thirdly, we outline implications for the design of persuasive DPAs and future research directions identifying the need for more research to understand the potentials and pitfalls of persuasive DPAs. Our study takes an initial step in exploring and understanding this new research area.

## 2. Background on Fogg's Persuasion Principles

Persuasion is defined as "*an attempt to change attitudes or behaviors or both*" by B. J. Fogg (2003), who is a pioneer in studies of persuasion and computers. In 1998, Fogg (1998) defined a persuasive technology as "*an interactive technology that attempts to change people's attitudes or behaviors in some way*". Fogg (2003) later defines a Functional Triad (Figure 1) with 16 associated principles of persuasion, acknowledging that there are many different ways that technology can persuade. The Functional Triad is a conceptual framework illustrating the three key persuasive roles a computer technology can have: Tool, Medium, and Social Actor.

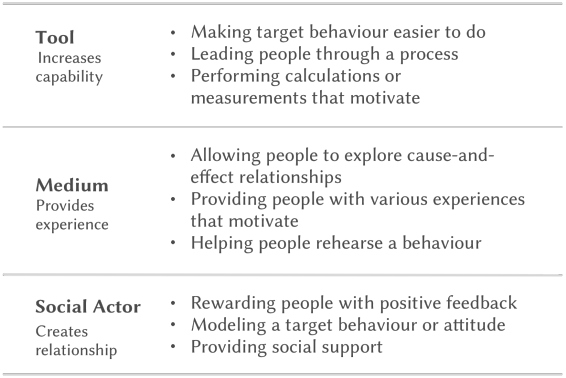


**Figure 1:** Fogg's Functional Triad: Tool, Medium, Social Actor.

Technology as a Tool increases capability and persuades by making desired outcomes easier or more efficient to achieve for the user. There are 7 associated principles: *Reduction, Tunnelling, Tailoring, Suggestion, Self-Monitoring, Surveillance,* and *Conditioning*. Technology as a Medium persuades by providing the user with compelling simulations and focuses on creating experiences for the user. There are 4 associated principles: *Cause and Effect, Virtual Rehearsal, Virtual Rewards,* and *Simulations in Real-World Contexts*. Technology as a Social Actor persuades by using a variety of social cues that elicit social responses from the human user and focuses on creating a relationship between the technology and the user. There are 5 associated principles: *Attractiveness, Similarity, Praise, Reciprocity,* and *Authority*.

Rather than attempting to compare all 16 principles, as our starting point for investigating the persuasive powers of DPAs, we chose to compare the 3 roles of Tool, Medium and Social actor, as represented by one of the principles associated with each. The choice of specific principle was based on a simple selection process within the team of researchers where each researcher nominated one specific principle per role based on their experience with DPAs and personal physical exercise routines. This was followed by a simple count of majority, and a discussion of rationales behind nominations until consensus had been reached. The process revealed a very high level of consensus, and the team of researchers subsequently agreed on the chosen set of principles for the study.

Representing the Tool role, we chose the principle of *Suggestion*. For the Medium role the principle of *Virtual Reward*, and for the Social Actor role the principle of *Praise*.

Fogg (2003) defines these principles as:

- *Principle of Suggestion*: A computing technology will have greater persuasive power if it offers suggestions at opportune moments.
- *Principle of Virtual Reward*: Providing a motivating simulated environment in which to rehearse a behaviour can enable people to change their attitudes or behaviour in the real world.
- *Principle of Praise*: By offering praise, via words, images, symbols, or sounds, computing technology can lead users to be more open to persuasion.

The *Suggestion* principle is about intervening at the right time and often builds upon people's existing motivations, reminding people that "*Now would be a good time to do X*". *Virtual Reward* can create situations that reward and thereby motivate people to do a target behaviour, giving users a more enjoyable experience and a more positive attitude toward the target behaviour. Using *Praise*, the technology expresses respect, approval, or gratitude towards the user's current behaviour or attitude, instilling positive feelings in the user and making them more susceptible to persuasion. These persuasive principles played a crucial role in the development of the DPA and guided the analysis of data as a theoretical lens.

Although voted out in favour of the Suggestion principle, the principle of Self-monitoring was discussed as a potential candidate for representing Technology as a Tool. The reason for this cast of votes was partly 1) that Self-monitoring had already been investigated in related work on activity sensing devices (e.g. Consolvo et al. 2006, de Oliveira & Oliver 2008, Foster et al. 2010, Fritz et al. 2014, Hirano et al. 2013, Jensen et al. 2010, Lin et al. 2006, Mateevitsi et al. 2014, Purpura et al. 2011, Toscos et al. 2006), and partly 2) that we wanted to explore a principle that would allow the DPA to play a more proactive role in the persuasion - i.e. making proposals - rather than merely reporting back the user's performance.

## 3. Related Work

In this section, we present related research on DPAs in HCI and broader research on digital persuasion and physical activity. We did not find specific research on the persuasive power of voice-controlled DPAs, however, we report on studies looking at persuasion through earlier voice-controlled devices.

### 3.1. Digital Personal Assistants

Most current research on DPAs focuses on empirical studies investigating how digital personal assistants are used and perceived (Leahu et al. 2013, Porcheron et al. 2018, Pradhan et al. 2018, Pyae & Joelsson 2018, Wulf et al. 2014), the challenges surrounding speech interaction (Hong & Findlater 2018, Murad et al. 2018, Myers et el. 2018, Springer & Cramer 2018), how to evaluate voice-interfaces (Ghosh et al. 2018, Hura 2017), as well as exploring DPAs future design considerations (Danielescu & Christian 2018, Fitton et al. 2018, McGregor & Tang 2017, Vtyurina & Fourney 2018).

Porcheron et al. (2018) define a DPA as *"embodying the idea of a virtual butler that helps you 'get things done'"*. The first well-known commercially available DPAs were on smartphones, for example, Siri and Google Now (today known as Google Assistant). The assistants evolved from being software-only programs to having their own physical form, where we now have speaker devices such as Amazon Echo, Google Home, and Apple's HomePod housing the digital assistants. In recent years, these physical devices have rapidly gained in popularity. In the US, 21% of the population now own at least one smart speaker, a 78% increase year over year, with 52% of owners using them daily (NPR 2019). Smart speaker ownership in the US rose 40% from 2018 to involve 66.4 million people and 133 million devices (Kinsella 2019). In 2019, China reached 10.6 million devices, a 500% increase from the previous year (Canalys 2019). Their most common use

includes planning (e.g., to-do lists, calendars, reminders), searching for information (e.g., news, traffic, weather), controlling smart home accessories (e.g., lights, locks), and entertainment, such as streaming video, streaming music, and playing games (Kinsella 2019, Porcheron et al. 2018, Pradhan et al. 2018).

In HCI research, DPAs have been studied in different domains and contexts. The focus has often been on empirical studies investigating how DPAs are used and perceived by people (Leahu et al. 2013, Porcheron et al. 2018, Pradhan et al. 2018, Purington et al. 2017, Pyae & Joelsson 2018, Wulf et al. 2014). Other studies focus on the challenges of speech interaction (Hong & Findlater 2018, Murad et al. 2018, Myers et al. 2018, Springer & Cramer 2018), for example, the challenges with specific types of words and named content, as well as how users overcome speech recognition errors. Speech interactions are also difficult to evaluate, as established techniques, like thinking aloud does not apply well due to the nature of speech interaction. Therefore, studies have also focussed on how we can evaluate speech interfaces (Ghosh et al. 2018, Hura 2017). Studies also suggest and show how we can improve the interactions with DPAs by adding continuous conversations (Vtyurina & Fourney 2018) and personalities (Danielescu & Christian 2018). Finally, there are studies that look into the future possibilities for DPAs, for example, how they can be used in workplaces (McGregor & Tang 2017) or how they can become a more significant part of future smart homes (Fitton et al. 2018). However, at present DPAs are primarily reactive, responding to conversational inputs from users, limiting the motivation and support they can give users (Meurisch et al. 2017).

### 3.2 Digital Persuasion and Physical Activity

Persuasion for physical health is a central theme in the field of digital persuasion. As described in a recent review of state-of-the-art and emerging trends in persuasive technology by Orji & Moffatt (2018) the evolving field of persuasive technology has in the last decade been increasingly targeting its efforts at creating behaviour changes in relation to people's well-being and health. According to this review, as much as 38% of all studies reviewed was concerned with physical activity, with 75% of this research reporting fully positive results.

The attraction towards applying persuasion for physical activity goes for development of new persuasive consumer products, research into interactive technologies for persuasion, and research into the underlying principles of persuasion in themselves. In commercially available products, persuasion has often been applied to help users track their physical activity, for example through mobile and wearable devices like Apple Watch, FitBit, or Garmin Vivosport, and through apps and services for managing their physical training program, like Bytesize's "7-minute workout" making you your own personal trainer to lose weight and get fit with fast, simple daily workouts, and  Active's "30-day fitness challenge" giving you a tailored different mini-challenge each day for a month. In human-computer interaction research, applications of different persuasive approaches have been studied in the context of physical activity using several technologies. Most commonly used are tracking devices that monitor users' physical activities like the increasing range of commercially available products mentioned before (smartphones, fitness trackers, and other smart sensors) but combined with experimental, research-driven, software. These tracking devices collect sensor data from their user, which can then be applied in

combination with different persuasive principles. Specific studies include the use of sensor data in physical standalone devices (Mateevitsi et al. 2014), in companion applications on smartphones (Consolvo et al. 2006, de Oliveira & Oliver 2008, Hirano et al. 2013, Jensen et al. 2010, Purpura et al. 2011, Toscos et al. 2006), and on computers (Foster et al. 2010, Lin et al. 2006). These applications typically display statistical data through dashboards, gaming experiences, social facilitation, or as comparisons. They also typically visualize a user's physical activity through metaphors of, for example, a flower field or a fish tank. Other studies look at how to persuade users to be physically active while being outside in motion, for instance, through voice input in the user's ear while exercising (Eyck et al. 2006) or installations embedded in the surroundings (Peeters et al. 2013, Singh & Mathew 2007). In relation to B. J. Fogg's particular theories and principles on persuasion in the HCI literature, these have been studied, for example, for encouraging teenage girls to exercise (Toscos et al. 2006), to motive athletes during virtual coaching sessions (Eyck et al. 2017), and in the context of gamful systems for the health domain (Orji et al. 2018).

Persuasion has been studied in HCI research across multiple technologies and information platforms to see how they can be designed to persuade users to change behaviour for improved health and wellbeing (e.g. Eyck et al. 2006, Gram-Hansen et al. 2018, Lin et al. 2006, Peeters et al. 2013, Purpura et al. 2011, Singh & Mathew 2007, Win et al. 2019). There are also software design patterns for persuasive computer-human dialogue around reminding, rewarding and feedback (Oduor et al. 2017). In a recent survey of persuasive technology for health and wellness it was reported that while 72% of the 85 studies reviewed used the original conceptualization of persuasive technology by Fogg (2003), generally the studies looked at other behaviour-related or psychological outcomes beyond changing behaviour and attitudes (Orji & Moffatt 2018). Hence, most studies of persuasive design use B. J. Fogg's notions of persuasion or use the PSD model for designing and evaluating persuasive systems, built on Fogg's principles (Oinas-Kukkonen & Harjumaa 2009). Additionally, some designs use feedback from digital technology to disrupt and change undesired habits (Hermsen et al. 2016) while others reinforce and strengthen existing behaviours (Orji & Moffatt).

### 3.3 Persuasion through Conversational Agents

Voice interfaces have attracted considerable research interest in recent years (Feine et al. 2019) and are seen by many as the "next big thing" (Følstad & Brandtzæg 2017, Whitenton 2017). As such, they present opportunities for human-computer interaction design for persuasive technologies. Conversational agents are software-based systems designed to interact with people using natural language, and have many different names, from conversational/virtual/digital agents or assistants through to chatbots or chatterbots. In recent years, many service providers have moved to using chatbots to enable users to access content or services by asking verbally or typing questions and commands in their everyday language (Brandtzæg & Følstad 2018). DPAs are a specific kind of voice-activated digital assistant, and as the name suggests, they are usually associated with personal technologies. The advantage of DPAs over chatbots and other generic voice interfaces is that they are able to create a personalized service through connectivity to the home's

digital ecosystem and IoT devices, giving them access to stored information about the user. It is the "human-like" interactions of digital personal assistants like Siri (Apple, 2011), Google Now (Google 2012), Cortana (Microsoft, 2015) and Alexa (Amazon, 2015) that make them so popular. They not only support real-time task completion but also develop sufficient knowledge about the user in order to act on their behalf (Luger & Sellen 2016). Research in this area includes understanding how the perceived personality of the system can affect its trustworthiness and persuasiveness (Andrews 2012), using an ECA to persuade university students to walk more and improve their physical condition (Bickmore & Picard 2005), and using an interactive counselling dialogue with an ECA to persuaded users to change attitudes towards regular exercise (Schulman & Bickmore 2009).

## 4. User study

To assess the persuasive capabilities of DPAs, we designed a controlled laboratory study in which a custom-made digital assistant, called *Ida* (Intelligent digital assistant), was used to both facilitate a physical exercise session and act as a motivator during various exercises. The focus of the study was to assess the general persuasive capabilities of *Ida*, as well as specifically compare the three persuasion principles.

We designed a within-group repeated measures study in which *Ida* provided four different types of feedback during an exercise session: 1) Suggestion, 2) Virtual Reward, 3) Praise, and 4) No Feedback. The effectiveness of each condition in motivating participants to exercise was measured on a number of dependent variables, including *Time on Exercise*, *Number of Repetitions* and *Average Heart Rate*. In the following sections, we provide a description of *Ida*, participant details, tasks, setup, procedure, materials and measurements in our user study.

### 4.1. Ida as a Persuasive DPA

For this study, we developed a customized DPA to minimize participant prior experiences and/or bias towards a specific device or brand. *Ida* was created using components from a Google AIY Voice Kit[1]. It contains a microphone, a speaker, and an LED button on top which lights up to indicate *Ida* is listening for commands (see Figure 2). *Ida* is connected to the Google Cloud Services[2] of speech-to-text and text-to-speech to process voice commands and to produce voice outputs, respectively. Voice input is restricted to "*yes*" and "*no*" commands to minimize voice recognition errors.

[1] https://aiyprojects.withgoogle.com/voice/

2 https://cloud.google.com/

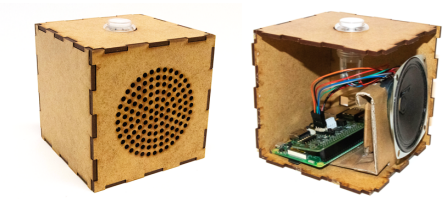

**Figure 2:** Ida – customized digital personal assistant

*Ida* was designed to conduct an entire exercise session without researcher intervention. *Ida* welcomed participants and introduced them to each individual exercise. It provided precise and clear instructions throughout the session. The main purpose was to guide participants through the exercises, with the four different feedback conditions: Suggestion, Virtual Reward, Praise, and No Feedback. *Ida* was also designed to help participants relax between exercises with calming breathing exercises.

### 4.2. Participants

A total of 48 volunteers (26 males and 22 females) participated in the study. They were recruited from a variety of sources, including social media, but most of them were volunteers from the Computer Science department of Aalborg University. Their age ranged from 17-56 years, with an average age of 28.3 years, with a standard deviation of 9.4. According to a self-assessment questionnaire, administered before the study, 27 of our participants considered themselves to be physically active while the remaining 21 were neutral or not active at all. Twenty participants regularly used an activity tracker of some type (e.g. mobile app, smartwatch). In regard to previous experience with DPAs, 11 reported that they were using them frequently, 24 that they had tried using them at some point, while 13 had no previous experience with them whatsoever. Seven of the participants that reported using DPAs frequently had a physical device at home, while the remaining four interacted with the DPA on their mobile device.

### 4.3. Tasks

To test the different persuasive feedback types, we chose four different physical exercises as tasks in our study. Each of the exercises focused on different muscle groups to minimize participant fatigue and carry over effects. The exercises were: jumping jacks, lunges, bent over flys, and shoulder presses. For each exercise a short video was created to demonstrate how it should be performed and ensure that all participants had a common understanding of what they should do. Participants were asked to perform each exercise for as long as possible but were free to stop whenever they wanted. Although it was not technically implemented, in the spirit of Wizard of Oz evaluation for intelligent agents (Maulsby et al. 1993), participants were given the impression that during the session, *Ida* would give them feedback on their actual performance. Participant's belief that this was the case was reinforced through *Ida's* feedback referencing their current activity, and the fact that they all wore a Fitbit monitoring device connected to *Ida*. In reality, each exercise and persuasive feedback type were randomized for each participant using the Graeco-Latin square design, to account for differences between the exercises and to control the impact of fatigue and learning effect between each condition. We standardized the number of feedback messages participants received during the

exercises across all conditions, because we wanted to compare the three persuasive principles.

Ida's persuasive feedback was designed so it could be used regardless of which exercise the feedback was matched to. The feedback comments for each condition followed a predetermined pattern. The first feedback comment was given 8 seconds into the exercise, the next after 10 seconds, afterwards the time interval between feedback comments was incremented by one second each time.

*Ida's* feedback comment for each of the four conditions supported or persuaded the exerciser to keep going. In the Suggestion condition, *Ida* would suggest either to keep exercising or to put more effort into the exercise. To encourage the person to keep going, *Ida* would say, e.g., *"May I suggest you keep doing the exercise for a longer period of time"* or *"Since you have already done the exercise for some time, I bet you can do it for a little longer."* To suggest an increase in effort, *Ida* would say, e.g., *"What about putting all the energy you have into the exercise"* or *"Try to put even more effort into your movements"*.

During the Virtual Reward condition, each timed feedback comment was regarded as a level achieved. *Ida* played a musical fanfare as encouragement, saying "*Level One*" the first time, and increasing the level count each time feedback was given. Virtual rewards work best inside a virtual environment, so for the sake of this experiment, we created an imagined virtual audio environment of a fitness training class/session and chose to provide "rewards" in the simple form of audio accolades, providing the vicarious experience of achieving a "win", as you might in a race or when competing against others.

In the Praise condition, *Ida* would praise the participant's persistence, elevated heart rate, and deliver a supportive positive accolade about the participant. The Praise feedback, for example, included *Ida* saying, *"What a great start"*, "*You are doing very well"*, *"Your persistence shows that you don't give up so easily"* or "*Based on your high heart rate, I can see you are putting a lot of effort into the exercise*".

In the No Feedback condition, *Ida* did not give any feedback. This condition's purpose was to provide a baseline for the study to compare to the three persuasive feedback types.

The specific sentences used for the Suggestion and Praise conditions were selected as possible examples of what a DPA might say to the user in order to either advise on or applause their physical activity. The process of developing the specific sentences took place over several iterations in the research team, until consensus on a final set of sentences had been reached. As guiding principles, we wanted a set of sentences that included both short (less than 10 words) and medium long ones (10-20 words), as well as using different types of suggestions and appraisals. We also wanted the sentences to reinforce the perception that *Ida* was responding the users' actual physical performance. As at this point in our studies, we are interested in the overall persuasive principles, rather than the specific wordings used, no further measures were made to compare between different sentences.

## 4.4. Setup

The study was conducted in a space set up to resemble a living room. Common living room furniture (e.g., sofa, coffee table, cupboards, lamps, plants) was placed in the room, with *Ida* (Figure 3). There was sufficient space in the middle of the room for the participants to comfortably perform the exercises. *Ida* was connected to an activity tracker (Fitbit) that the participants wore during the study. A Chromecast was connected to a TV, which was used to show the exercise videos. The study facilitators were sitting outside the study space and were able to observe participants through a camera positioned in the top right corner of the room, providing a good view of participants' activities.

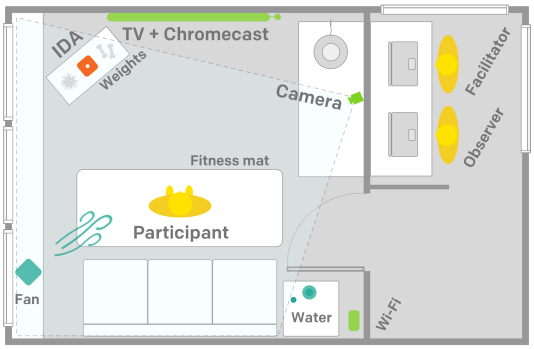


**Figure 3:** Layout of Simulated Living Room Setup.

## 4.5. Procedure

The study procedure was as follows. First, a brief introduction about the purpose of the study, and how it would progress was provided. Participants were informed that the potential of digital personal assistants in facilitating exercise sessions was being explored. It was not made explicit that comparisons were being made on the effectiveness of different persuasive feedback types. Next, participants were asked to sign a consent form and were handed an activity tracker which they were asked to wear for the duration of the study. Shortly after this, the facilitator and observer left the room and from that point, *Ida* gave all instructions and guided participants through the various steps of the study.

After welcoming participants, *Ida* outlined the whole exercise session and explained the various exercises they would perform. *Ida* also explained that participants should exercise for as long as possible but were free to stop whenever they wished and should end each exercise session by pressing the lit button on the top of the speaker. For each exercise, *Ida* showed a video on the TV screen, via the Chromecast, describing in detail how each exercise should be performed. After participants verbally confirmed that they were ready to begin, *Ida* spoke a short count down (i.e., "4-3-2-1-go") to start the exercise session. *Ida* would then interject with feedback in the predetermined pattern, with the content based on one of the four conditions, as if it was responding to the participant's actual physical activity. After each exercise, the participant had a one-minute break where *Ida* played a breathing exercise video on the TV, to relax and lower the participant's heart rate before the next exercise.

After the break, *Ida* would introduce the next exercise, which would then follow the same procedure, except it would deliver feedback using a different

persuasive feedback type. After the participant had completed all four exercises, with all four conditions, *Ida* thanked them for their participation and said goodbye. The facilitator would then enter the room and ask the participant to fill out a questionnaire and then interview the participant. Sessions lasted from 25 to 58 minutes, with 35 minutes on average.

### 4.6. Measurements

The depended variables used to determine the effectiveness of *Ida's* persuasive feedback included observed behaviour, physiological measures and participant self-reported data. Measured data included *Time on Exercise*, *Number of Repetitions*, and the participants' *Average Heart Rate* during each of the conditions. Time was measured from the end of *Ida's* countdown (i.e., on "go") to the point when the participant pressed the button to end that exercise. Repetitions were counted by the observer using a tally counter. The heart rate data was collected through the activity tracker and used to calculate the peak and average heart rate, to determine a participant's physical effort.

Demographic information about gender, age, experience with digital personal assistants (both on phones and physical devices), their motivation to track their physical activity, and how physically active they perceived themselves to be was collected on each participant.

The participant's self-reported data was gathered using the post-experiment questionnaire and the semi-structured interview. In the post-experiment questionnaire, participants were asked to rate how motivating they found *Ida* during each condition on a unipolar 5-point Likert scale, responding to questions like "I thought IDA was motivating during the lunges exercise". Since conditions and exercises were mixed for each participant results were cross-referenced with our Latin-square design sheet to assign ratings to the persuasion technique. In the semi-structured interview, participants were interviewed about their overall experience of interacting with *Ida* during their physical exercises. Specifically, they were asked to reflect on their experience with each one of the four different types of feedback separately. Apart from that, participants were asked what they liked or disliked about their interaction with *Ida* and if they could see themselves using something like this to support exercise at home.

## 5. Statistical findings

The results of our data analysis aimed to assess the persuasive abilities of *Ida* in the context of physical activity. The main goal was to assess how effective the persuasive feedback types of Suggestion, Virtual Reward, and Praise were compared to the No Feedback condition. For the dependent variables that were measured on an interval level (i.e. *Number of Repetitions*, *Time on Exercise*, *Average Heart Rate*) one-way repeated measures ANOVA were performed followed by post-hoc tests using the Bonferroni corrections. A Friedman test was performed for the dependent variable of *Perceived Motivation* followed by a post-hoc analysis with Wilcoxon signed-rank test for pairwise comparison of the persuasive feedback types. In addition, two-way repeated measures ANOVA's were performed to examine whether persuasion effectiveness differed significantly amongst any of the participant demographic categories (e.g. gender, age, prior activity levels, or experience with DPAs). The two-step transformation

approach (Templeton 2011) was used for the dependent variables of *Number of Repetitions*, and *Total Time on Exercise* since both Shapiro-Wilk, and Bartlett tests showed that they were not following a normal distribution.

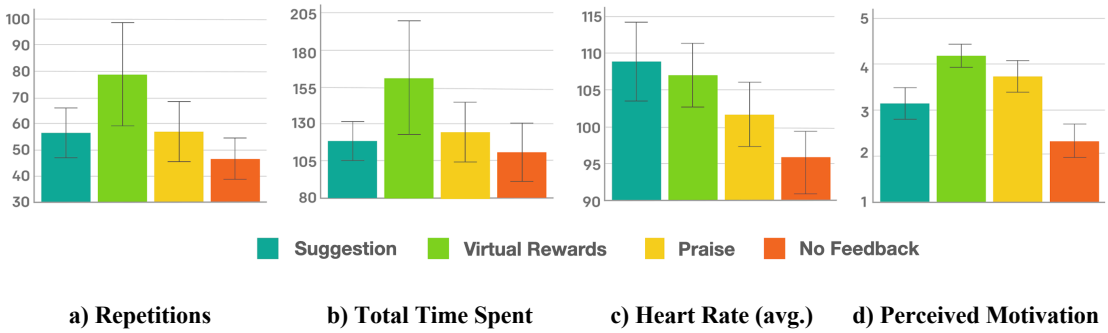


**Figure 4.** Comparison of the four conditions: Suggestion, Virtual Rewards, Praise, and No Feedback (error bars represent standard deviations).

## 5.1. Repetitions

We found statistically significant differences in the number of repetitions participants performed among the different persuasive conditions ($F(1.97,86.85) = 7.44$, $p = .001$, $\eta_p^2 = .145$). Greenhouse-Geisser corrections were applied to within-subject effects to compensate for violations of sphericity. As seen in Figure 4a participants performed more repetitions during Virtual Reward ($M = 78.70$, $SD = 67.90$) followed by Praise ($M = 57$, $SD = 39.40$) and Suggestion ($M = 56.60$, $SD = 32.30$). The smallest number of repetitions was performed during the No Feedback condition ($M = 46.70$, $SD = 27$). Post-hoc comparison with Bonferroni corrections revealed that participants performed a statistically significant higher number of repetitions during Virtual Reward compared to the No Feedback condition ($p = .006$). The number of repetitions also differed significantly between the Virtual Reward and the Praise conditions ($p = .049$). Even though, we found variation in the number of repetitions performed among the various persuasive conditions, the feedback type that clearly stands out for its effectiveness is Virtual Reward.

Mixed model repeated measures ANOVA performed with demographic variables as between subjects factors showed a significant interaction effect between persuasive feedback type and gender ($F(2.03,87.34) = 4.8$, $p = .01$, $\eta_p^2 = .10$). Follow up tests showed that in the Virtual Reward condition, males ($M = 98.92$, $SD = 82.08$) performed almost double the number of repetitions compared to females. Independent t-tests showed that only in the Virtual Reward condition there was a statistically significant effect of gender on number of repetitions ($t(46) = -2.213$, $p = .032$). We also found a significant interaction effect between persuasive feedback type and how physically active participants considered themselves to be ($F(4.23,88.84) = 2.93$, $p = .023$, $\eta_p^2 = .123$). Physically active participants on average performed more repetitions than others but the difference was only statistically significant in Virtual Reward ($t(46) = -2.85$, $p = .006$) and in the Praise ($t(46) = -2.78$, $p = .008$) conditions. Physically non-active participants performed better in the Suggestion condition followed by Virtual Reward.

### 5.2. Time

Results showed statistically significant differences in the time participants spent exercising, among the four persuasive conditions (F(2.06,88.71) = 4.81, p = .001, $\eta_p^2$ = .101). As it was the case for *Number of Repetitions*, in *Total Time on Exercise* the sphericity assumption was also violated and therefore Greenhouse-Geisser corrections were applied.

Figure 4b shows that the average time participants spent exercising in each condition varied considerably: No Feedback (M = 110.81, SD = 68.33), Praise (M = 124.95, SD = 70.39), Suggestion (M = 119.06, SD = 46.16), and Virtual Reward (M = 162.36, SD = 133.26). However, post-hoc comparisons with Bonferroni corrections revealed that only the time difference between Virtual Reward and the No Feedback condition was statistically significant (p = .032). Mixed model repeated ANOVA's that were conducted did not reveal any significant main or interaction effects between *Time on Exercise* and any of the demographic categories.

### 5.3. Heart Rate

Results from the one-way repeated measures ANOVA show a statistically significant main effect of persuasive feedback type on *Average Heart Rate* (F (3, 141) = 10.498, p < .001, $\eta_p^2$ = .183). Figure 4c illustrates that the condition with the highest average heart rate was Suggestion (M = 108.85, SD = 18.60), followed by Virtual Reward (M = 106.96, SD = 15.03), Praise (M = 101.65, SD = 15.06), and finally the No Feedback condition (M = 95.15, SD = 14.74). Post-hoc comparisons with Bonferroni corrections show that the difference in *Average Heart Rate* was statistically different between the No Feedback condition and all persuasive feedback types. The level of significance for both Suggestion and Virtual Reward compared to the No Feedback condition was p < .001 while for Praise it was p < .036. These results show that the persuasive feedback types were very effective in convincing participants to put more effort into exercising which was reflected in elevated heart rate for all of them.

Mixed-model ANOVA tests showed an interaction effect between gender and *Average Heart Rate* (F (3, 138) = 4.17, p < .007, $\eta_p^2$ = .083). This can be mainly attributed to the differences in the Suggestion condition (t (46) = -2.04, p = .047) in which male participants (M = 113.7, SD = 20.70) showed a significantly higher heart rate than female participants (M = 103.10, SD = 14.2)0. None of the other demographic data had a significant main or interaction effect on *Average Heart Rate*.

### 5.4. Perceived Motivation

Since *Perceived Motivation* was measured on a 5-point Likert scale we performed a Friedman test (non-parametric alternative to repeated measures ANOVA) to assess whether there were differences in how the persuasive feedback types were perceived by the participants. The test revealed that there was a statistically significant difference in how motivated participants felt during the persuasive conditions $\chi^2(3) = 41.25$, $p < 0.001$. Figure 4d shows that the highest rated condition was Virtual Reward (M = 4.19, SD = .842), followed by Praise (M = 3.75, SD = 1.16), and Suggestion (M = 3.15, SD = 1.19). The lowest rated condition was No Feedback (M = 2.35, SD = 1.26). Our results indicate that

participants perceived some of the conditions more motivating than others. Post hoc analysis with Wilcoxon signed-rank test showed that there was a statistically significant difference in how motivated participants felt during Praise ($Z = -4.09$, $p < .001$), Suggestion ($Z = -2.49$, $p = .013$), and Virtual Reward ($Z = -5.0$, $p < .001$) compared to the No Feedback condition. In addition, we found significant differences between Suggestion and Praise ($Z = -2.76$, $p = .006$), as well as between Suggestion and Virtual Reward ($Z = -4.02$, $p < .001$). These results show that participants felt motivated by all persuasive feedback types and that Virtual Reward was the one they preferred most.

### 5.5. Summary of Statistical Findings

Results of the quantitative analysis revealed all three of the persuasive feedback types had an effect on how much effort participants put into exercising compared to the No Feedback condition. Participants had a significantly higher *Average Heart Rate* when *Ida* was using any of the persuasive feedback types compared to no feedback at all. On average participants performed more repetitions and spent more time exercising in the persuasive conditions compared to the one with no feedback but those differences were statistically significant only for the Virtual Reward condition. In addition, all feedback types affected significantly the perceptions of how motivated participants felt compared to the No Feedback condition. By comparing the three conditions to each other we can see that they vary on average considerably (see figure 4). Virtual Reward performed better regarding time on exercise, repetitions, and perceived motivation while Suggestion outperformed the others in heart rate elevation. However, those differences were found to only be statistically significant for the number of repetitions between Virtual Reward and Praise ($p = .032$) and for perceived motivation between Suggestion and Praise ($p = .006$), as well as Suggestion and Virtual Reward ($p < .001$).

Further analysis showed that each persuasive feedback type affected our participants differently. Out of the three types Virtual Reward was the most effective. In this condition participants performed the most repetitions, they spent the most time exercising, and they also rated it as the most motivating compared to the others. The persuasive effect of Virtual Reward was significantly higher for males as well as physically active participants. The Suggestion condition, on the other hand, was rated lowest in regard to *Perceived Motivation* but was the best at encouraging physical effort, as it had the highest effect on participants *Average Heart Rate*. However, *Average Heart Rate* data analysis showed that Suggestion worked better for males while Virtual Reward worked better for female participants. Finally, participants perceived Praise as being the second most motivating feedback type, but the measurements did not support this perception since it was last in elevating heart rate and not significantly better than Suggestion or the No Feedback condition in *Number of Repetitions* or *Time on Exercise*. Age, previous experience with PDAs or fitness trackers did not affect any of the results.

With regard to demographic data, in most cases mixed model repeated ANOVA's did not reveal significant main or interaction effects between dependent variables and any of the demographic categories. The only exceptions were interaction effects in regard to gender and number of repetitions as well gender and heart rate elevation. We also found interaction effects between prior

physical activity and number of repetitions in the Virtual Reward and in the Praise condition.

## 6. Interview Findings

Transcribed interviews on participants' experiences with *Ida* and observations made during the studies were analysed in NVivo using a top-down approach. This included themes that evolved from topics in the interview guide, including: the three persuasion principles of *Suggestion*, *Virtual Reward* and *Praise*, Interaction with *Ida* and Future Use. Successful suspension of disbelief with respect to Ida knowing about the users' exercise performance was supported by the comments that they made during the interviews, confirming that the Wizard of Oz approach worked as intended in our study.

### 6.1. Suggestion Condition

In the Suggestion condition, *Ida* was seen as giving constructive criticism. Fourteen participants spoke positively about getting suggestions from *Ida*, indicating that it helped them push themselves to exercise more, with comments like, *"I definitely believe it helps me set the bar higher"*. When *Ida* came with a suggestion to keep going, one participant replied with *"I am sure you [Ida] think so".* However, *Ida* was often seen as a tough coach, with 31 participants either reacting to *Ida's* feedback directly, with participants observed telling *Ida* to *"Shut up",* while others complained afterwards, with comments like, *"She [Ida] is making a fool out of me", "I was thoroughly provoked",* and *"If I am already working my ass off, it cannot continue saying that I need to be doing a better job!".* However, although *Ida* provoked complaints, the participants did not say that this had any adverse effect on their motivation to exercise. Through observation we learned that 16 of the participants listened and specifically reacted to *Ida's* effort related suggestions during the exercise, for example, when *Ida* said, "*I suggest you put some more energy into your movements",* the participants actually did so.

### 6.2. Virtual Reward Condition

The Virtual Reward condition was the most popular of the four conditions, with 39 participants speaking positively about it motivating their activity. Participants identified Virtual Reward as a game and described it as a *"fun"* experience. Several participants talked about Ida as either a personal trainer or a companion to exercise with: *"It is very fun if you don't have a partner to exercise with ... then you're not on your own".* Interestingly, many of the participants really liked the competition aspect of the Virtual Reward condition. They wanted *Ida* to keep track of their performance over time so that they could beat their previous results. There were also some who wanted to compete against others, saying, *"I want to be the best"*.

Participants shared many theories and preferences as to how these levels should be achieved, such as *"time"*, *"heart rate"*, and *"repetitions"*. A general opinion amongst participants was that the metric should be known by the user before the exercise session, to better understand when rewards were given. Several participants mentioned setting a goal during the exercise, for example, *"I will take it to [level] 10 now"*. In some instances, they reached their precise goal while

others reported thinking to themselves things like, *"Aaah, one more [level]!"*. However, a few participants wanted more guidance and wished that *Ida* would set a goal for them as a human fitness coach would.

### 6.3. Praise Condition

In general, participants perceived the Praise condition as positive. Twenty-three participants said that they liked the reassurance from *Ida*, and it motivated them, for example, *"You felt that you were doing something right, I wanted to continue for longer"*. In particular, participants who considered themselves to be beginners at physical activities said it was helpful to be praised, for example, *"I am not that active, so I would prefer praise instead of a stick, as a carrot is better for beginners."* Participants also replied to *Ida's* praise, for instance, saying *"Thank you"*, and *"You are also doing great Ida"*.

Sixteen participants said they did not like praise, explaining that although they liked the *"good guy"*-attitude they did not find Praise motivating, for example, *"It makes me stop because she [Ida] is actually satisfied"* or *"The praise does nothing for me – the fact that I raised my arm 10 times is no big deal".* One participant was observed telling Ida to *"be quiet"* and ended the exercise quickly explaining she generally disliked being praised. Many participants thought *Ida's* praise was lacking sincerity, based on tone of voice. They characterized *Ida's* tone as being sarcastic, which did not encourage them, especially with respect to the effort they felt they were putting in. Some further explained that if *Ida* is supposed to motivate users to do physical activities, it needs to have a more enthusiastic tone of voice.

### 6.4 Summary of Interview Findings

Results of the qualitative analysis also revealed that all three of the persuasive feedback types had a positive effect on persuading participants to continue exercising and putting effort into it. However, this was achieved to differing degrees, and depended on individual preferences for training approaches, tone of voice and the level of exercise experience that the participant had.

In general, positive responses to the Suggestion condition related strongly to the type of human coaching that an individual responded to. Some people respond well to the tough coaching approach, which Ida was primarily perceived as representing, while others were upset or annoyed by it, and even verbally responded to *Ida* by telling her to shut up. Overall, participants were most positive about the Virtual Rewards condition, enjoyed getting the reward, and the sense of progress and competition that this type of feedback promoted. The Praise condition was influenced by whether individuals enjoyed being praised in their lives, and also, whether they felt the praise was warranted. The tone in which the praise was given, felt to be sarcastic, and lacking sincerity and enthusiasm, influenced how well accepted it was by different participants.

In the end, 35 participants stated they could see themselves using a system like *Ida* in their own homes to support their exercising. The remaining 13 could not see themselves using *Ida*, as they already had well-established fitness routines and did not find technology particularly motivating.

## 7. Discussion

The question driving this study was "can digital personal assistants persuade people to exercise?". Our findings show that DPAs have the potential, within their current technical and reactive limitations, to persuade people to increase their physical activity at home, using *Suggestion* to encourage physical effort, *Virtual Reward* to encourage endurance, and *Praise* to create reassurance for beginners.

Our evidence is both quantitative and qualitative, collected through 48 people participating in an exercise routine managed by a DPA giving persuasive feedback, in a simulated home environment. With respect to our combined findings, we identified significant statistical differences between the No Feedback condition and the three persuasive feedback types we evaluated, as well as interview findings that supported the statistical findings.

Feedback based on *Suggestion* had the highest effect on participants' average heart rates. Although *Suggestion* was rated during interview as the least motivating of the three feedback types and described as being very negative, there were indications that it was the best at encouraging physical effort. However, we acknowledge the negative responses to suggestions given was influenced by the actual wording of the suggestions in our system, and the tone of voice used by *Ida* to give these suggestions. This indicates that greater care needs to be taken in these aspects of feedback when making suggestions to people who are busy with the primary activity of doing exercises.

*Virtual Reward* showed the most promise, as it had the highest effect on repetitions, time, and participants' perceived motivation. At the same time, it was reported in interview as the favoured feedback type because it was considered to be fun, game-like and motivated them the most to keep on going.

*Praise* was felt by participants to lack sincerity but was otherwise mentioned as a positive reassurance. Many participants indicated that *Praise* would be best suited to beginners or in combination with the other two persuasive feedback types. As there was a pre-defined set of phrases for praising participants, this did not account for those at different levels of fitness and experience with exercising, and what was considered challenging and difficult for beginner exercisers, was considered easy, and not praise worthy, by experienced exercisers.

Consequently, we found that there were key differences between the three persuasive principles and their potential uses in persuading people to increase their physical activity. *Suggestion* could be used for strength training, where the focus is on getting a higher heart rate over a shorter duration. *Virtual Reward* could be used for cardiovascular exercises (e.g. running, cycling), where there is often more focus on endurance. *Praise* could be used for beginners or to strengthen the two other persuasions.

### 7.1. DPA as Tool, Medium and Social Actor

Relating our findings to B. J. Fogg's Functional Triad (Fogg 2003), participants expressed different improvements that could be made to *Ida*, and in general, these improvements align with Fogg's three persuasive roles for technology: Tool, Medium and Social Actor. In the Tool role, we saw improvements such as providing the user with more information (repetitions, time, heart rate), and integrating account information, specifically calendar data, to tailor the

persuasion to fit the user and enable the assistant to persuade at the most opportune moment (i.e. Fogg's principle of Kairos). In the Medium role, the improvements mostly involved issues of user experience, with the need for the addition of fun and achievement rewards into the exercise activity. Participants enjoyed their virtual coaching session and competition with other imagined fellow exercisers, giving the impression of not exercising alone. In the Social Actor role, participants wanted *Ida* to have a more human voice, express more sincerity and enthusiasm and act more like a personal trainer or exercise companion might.

By choosing to explore the opportunities provided by one principle from each of these three technology roles, designed into the DPA's feedback, we are able to provide this comparison between our results and Fogg's Triad. This gives us the ability to contribute these new ideas around how DPAs might be used in each of these roles to help persuade people to exercise. As a Tool, a DPA has the technical capabilities to provide personalised feedback based on the actual physiological responses of the person, and their adherence to the exercises prescribed, including activity specific advice and a gradual increase in difficulty of exercises as the person improves. As a Medium, DPAs can use voice interaction to create a virtual auditory environment in which exercisers can experience a more exciting and motivating environment than their home lounge room, with rewards for winning and achievements. As a Social Actor, DPAs can become a personal trainer, taking on individualised tones of voice and turns of phrase that best suits the exerciser, matching the users feedback-type preference and current level of exercise proficiency. All of this still assumes a reactive model of operation, however, as the next generation of proactive DPAs becomes a technical reality (Meurisch et al. 2017) they will likely offer improved detection and consideration of user's fitness goals improving DPA's abilities to act in these persuasive roles. Enabling digital personal assistants to be proactive, to have more sensors, to analyse users, and to have more access to a user's personal information, increases the challenges surrounding DPAs and persuasive design.

### 7.2. Persuasion or Subception

In Denmark, the first widely-used DPA smart speaker to use the Danish language was Google Home, launched in September 2018. This was followed in the media by criticisms of Google Home, questioning the continuous recording of users' conversations and relying on access to their personal data (e.g. Liao 2019). One participant in our study claimed explicitly that he does not trust DPAs and would never allow them into his home. Other participants stated they did not want DPAs to have too much control over deciding when they should be physically active, indicating that there are problems of trust around DPAs.

B. J. Fogg received criticisms on his work around persuasive technologies, because persuasion has the potential to do good, but it also has a darker side where it can be misused for personal or corporate gain. The philosophical question - *is computer-mediated persuasion ethical?* (Atkinson 2006) - has been one of the main concerns regarding persuasive technologies, where use of subception (short for subliminal perception, meaning when a stimulus occurs without conscious recognition (Dixon 1958)), is often mentioned as the problem.

Advancing the persuasive power of DPAs adds a whole new plethora of concerns. With their prominent placement in the home and the always-on aspect of the technology, they have the potential to influence people based on things they have learned about their users, but which may not have been explicitly shared. For the most part, DPAs can be beneficial in many ways, and in our study, we saw how they can help users increase their physical activity. There is much research on how technology can be used for good within persuasion, for example, helping people to quit smoking (Paay et al. 2014), achieving weight loss (Purpura et al. 2011), eating a healthier diet (Fadhil et al. 2016, Orji et al. 2013), or being more sustainable (Aleahmad et al. 2008). Nevertheless, as interaction designers of technology, we can be blindsided by abundant possibilities, forgetting to look at potential implications and concerns around new technologies. For example, a company like Amazon, who is also a web retailer, could perhaps be interested in making Alexa persuade users to buy more products through their online services. If DPAs were to be more proactive, it could be just another platform where political messages and advertisers could invade the privacy of our homes. If the user is unaware of the persuasion, is being manipulated and misinformed, it is no longer just persuasion but becomes subception. The question is - *where do we draw the line between persuasion and subception?* - and do we trust manufacturers and third-party developers of DPAs to not misuse the power of persuasion, because, as Voltaire said, *"with great power comes great responsibility".*

### 7.3. Implications for Design

Through our study, we learned that designing persuasive DPAs requires consideration of the fact that people are different and are affected differently by persuasive feedback. Some people are more motivated by a specific type of persuasion, for example, some enjoy praise, while others generally dislike it. We suggest that using a combination of different principles, instead of just one, could potentially increase the effect of persuasion. We also discovered that to be more useful and persuasive the DPA needs to be more context-aware and integrated with the user's digital ecosystem. For example, using calendar information to suggest exercise sessions at opportune moments.

Our findings therefore indicate different design opportunities for implementing persuasion in DPAs. These implementations can be classified into three different design perspectives: *user-controlled, system-automated,* and *sentient*. User-controlled opportunities can be implemented now, system-automated would potentially require years of development and sentient DPAs might be realized in the future.

The *user-controlled* implementation of persuasion would allow the user to tailor the DPAs persuasive personality based on their personal preferences. The drawback is that users will most likely choose a persuasive personality they enjoy or perceive to be the most motivating, which is not necessarily the persuasion feedback type that will have the most effect performance-wise.

The *system-automated* implementation would leverage collected data about the user, and machine intelligence to achieve the most effective persuasion. The DPA could try different persuasions, measuring the effects, and then based on the results, tailor the persuasive feedback type to achieve the best results

performance-wise for that individual. The persuasion should adapt to the context and the user's mood, as persuasion that works in one situation may not necessarily work in another. For example, on a bad day, the user might not respond well to a tough coach but would perhaps respond better to an encouraging coach. Participants' feedback showed that they enjoyed the connectivity between *Ida* and the TV, as they felt it gave a more immersive experience. Several participants wanted more statistical information about their performance (e.g. repetitions, heart rate, and time) and feedback on their execution of the exercise. This information could either be shown on a real-time dashboard using the TV or highlighted by *Ida*, for example, with a message saying, *"You have now done 30 repetitions"*. Additionally, several participants wanted to be able to modify the exercises, to focus on different body parts, duration, and the feedback type *Ida* used to motivate them during the exercise. Some participants stated that if *Ida* was their DPA, they wanted their accounts (e.g. calendar and music streaming service) synchronized to remind them to exercise when they had time for it and incorporate music they liked into the exercise routine.

A *sentient* DPA would persuade by being responsive and conscious of impressions gathered from the user, responding as a human coach would, with personality, understanding and empathy. Utilizing the same persuasive feedback types that human coaches use and combining that with the computational power of DPAs could create a very persuasive system. Futuristic personal assistants, with capabilities exceeding technology as we know it today, are common in popular culture and science fiction (Jonze 2013, Kubrick 1968, Villeneuve 2017). In *Blade Runner 2049* the sentient holographic companion, Joi, understands the emotional needs of her user to such an extent that they develop a romantic relationship. Imagine having a DPA who is your friend, who knows you, your needs and desires, responds to your moods and always has your best interest in mind. In science fiction, different fears around sentient assistants have also been represented. For example, in the movie *2001: A Space Odyssey*, HAL 9000 kills crew members of the space ship. Or the movie *Her*, in which the complexities and implications of human – AI relationships are explored. The universal question is - *to what extent should we make digital assistants sentient?* – and how does this enhance their persuasive possibilities?

### 7.4. Limitations and Future Research

Although we have contributed new knowledge around DPAs and persuasion, given that there is very limited study in this area to date, we acknowledge that our study has limitations. As with any study on persuasion, choices had to be made around which persuasive principles we would focus on, which persuasive technology roles we would explore, and whether to create our own DPA or use an "off the shelf" model. These design choices were made and justified within the research team. *Ida* did not have the functionality of a commercially available DPA, because we built our own, which affected participant's perception of *Ida*, especially those with prior experience of DPAs. We also acknowledge that the choice of phrases and voice tone, commented on by participants in the interviews, possibly affected our results. It is difficult to determine the exact effect of the actual words used in our designed feedback for the three persuasive conditions – as we found that the same words had different influence on

different individuals. However, participants' reactions to these choices are reported in the qualitative results to help balance quantitative measures. From our interview responses, we can see that participants generally believed that *Ida* was monitoring their exercise routines and responding directly to this activity. Finally, our short-term controlled laboratory study has limited generalizability because of the controlled environment and duration of the study.

Based on our study, we identified four avenues for future research. The first is around the language, sincerity and personality of the personal assistant. Given the parameters of our study, it remains uncertain as to what degree it affected the persuasive power of the DPA. This requires further investigation. The second involves how well alternative persuasive principles and theories might work on personal assistants, and if a general overview of the strengths and weaknesses can be made. B. J. Fogg (2003) lists many other persuasive principles and other principles of persuasion exist beyond Fogg. It would be worth exploring these alternatives with respect to their use with DPAs. The third is the degree to which connection between the DPA and users' smart appliances and sensors in their digital ecosystems affect their persuasive power. Although we connected a Fitbit to *Ida*, this was simply for data collection. It would be worth designing for this data as actual input to the type and form of feedback that the DPA gives the user. The fourth, and maybe the most important, is that individual users have differing perceptions of and reactions to persuasion, and differing receptivity to subception, from this always-on device that is listening to them and trying to persuade them. To truly understand the potential persuasive power of DPAs, more studies are needed, over longer periods of time, in real world contexts.

## 8. Conclusions

In this study, we investigated the power of digital personal assistants to persuade people to exercise. We selected three of Fogg's (2003) persuasive principles, *Suggestion*, *Virtual Reward*, and *Praise*, one from each of Fogg's persuasive technology roles, *Tool*, *Medium* and *Social Actor*, to investigate as feedback types for motivating exercise. To do this, we developed a custom-made DPA to facilitate a physical exercise session and encourage participants using these different persuasive feedback types during four different exercises in a controlled laboratory environment which simulated a home lounge room.

We found that DPAs clearly have the potential, within their current technical and reactive limitations, to persuade people to increase their physical activity at home, using *Suggestion* to encourage physical effort, *Virtual Reward* to encourage endurance, and *Praise* to create reassurance for beginners.

Design implications for persuasive DPA's include the need for users to be able to tailor the DPAs persuasive feedback and personality based on their personal preferences to achieve the best results performance-wise for that individual. The persuasion should adapt to the context and the user's mood. It should also provide statistical information about their performance and feedback on execution of the exercise, and allow users to modify the exercises, to focus on different body parts, duration. The DPA would ideally be synchronized with other smart devices in their home, making it possible to tailor exercise

scheduling and music. Ideally, a DPA should be as responsive and aware as a human coach, with personality, understanding and empathy.

Future research in this area should look at: understanding the impact of voice, language and tone on persuasiveness; investigate how well other persuasive principles work with DPAs; exploring how other connected devices can increase persuasion; and investigate the impact of subception on adoption of persuasive DPAs.

In essence, digital personal assistants can persuade people to exercise – making it happen is a matter of appropriate interaction design.

## References

Aleahmad, T., Balakrishnan, A. D., Wong, J., Fussell, S. R. and Kiesler, S. (2008) Fishing for sustainability: the effects of indirect and direct persuasion. In CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08), 3021-3026. https://doi.org/10.1145/1358628.1358801

Andrews, P.Y. (2012) System Personality and Persuasion in Human-Computer Dialogue. ACM Trans. Interact. Intell. Syst. 2(2), Article 12, 27 pages. https://doi.org/10.1145/2209310.2209315

Atkinson, B. M. (2006) Captology: A critical review. In Proceedings of International conference on Persuasive Technology (Persuasive '06), 171-182. https://doi.org/10.1007/11755494_25

Bickmore, T. W. and Picard, R. W. (2005) Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput.-Hum. Interact. 12(2): 293–327.

Brandtzæg, P.B. and Følstad, A. (2018) Chatbots: changing user needs and motivations. interactions 25(5): 38–43. https://doi.org/10.1145/3236669

Canalys. (2019) China overtakes US in fast growing smart speakermarket. https://www.canalys.com/newsroom/china-overtakes-us-in-fast-growing-smart-speaker-market (accessed 20 April, 2020)

Cohen, P. and Oviatt, S. (1995) The role of voice input for human-machine communication. Proceedings of the National Academy of Sciences of the United States of America. 92. 9921-7. https://doi.org/10.1073/pnas.92.22.9921

Consolvo, S., Everitt, K., Smith, I. and Landay J. A. (2006) Design requirements for technologies that encourage physical activity. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '06), pp. 457-466. https://doi.org/10.1145/1124772.1124840

Danielescu, A. and Christian, G. (2018) A Bot is Not a Polyglot: Designing Personalities for Multi-Lingual Conversational Agents. In Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (CHI EA '18), Paper CS01, 9 pages. https://doi.org/10.1145/3170427.3174366

De Oliveira, R. and Oliver, N. (2008) TripleBeat: enhancing exercise performance with persuasion. In Proceedings of the 10th international conference on Human computer interaction with mobile devices and services (MobileHCI '08), pp. 255-264. http://dx.doi.org/10.1145/1409240.1409268

Dixon, N. F. (1958) Subception. New Biology, 27: 27-38.

Eyck, A., Geerlings, K., Karimova, D., Meerbeek, B., Wang, L., IJsselsteijn, W., de Kort, Y., Roersma, M. and Westerink, J. (2006) Effect of a virtual coach on athletes' motivation. In Proceedings of International conference on Persuasive Technology (Persuasive '06), pp. 158-161. https://doi.org/10.1007/11755494_22

Fadhil, A., Matteotti, C., Armellin, G., Villafiorita, A. and Betti, D. (2016) CoachMe: a platform for promoting healthy lifestyle. In Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '16), pp. 1077-1080. https://doi.org/10.1145/2957265.2965004

Feine, J., Gnewuch, U., Morana, S. and Maedche, A. (2019) A Taxonomy of Social Cues for Conversational Agents. International Journal of Human-Computer Studies, 132: 138-161.

Fitton, D., Read, J. C., Sim, G. and Cassidy, B. (2018) Co-designing voice user interfaces with teenagers in the context of smart homes. In Proceedings of the 17th ACM Conference on Interaction Design and Children (IDC '18), pp. 55-66. https://doi.org/10.1145/3202185.3202744

Fogg, B.J. (1998) Persuasive computers: perspectives and research directions. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '98), pp. 225-232. http://dx.doi.org/10.1145/274644.274677

Fogg, B.J. (2003) Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann.

Følstad, A. and Brandtzæg, P.B. (2017) Chatbots and the new world of HCI. interactions 24(4):38–42. https://doi.org/10.1145/3085558

Foster, D., Linehan, C., Kirman, B., Lawson, S. and James, G. (2010) Motivating physical activity at work: using persuasive social media for competitive step counting. In Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments (MindTrek '10), pp. 111-116. http://dx.doi.org/10.1145/1930488.1930510

Fritz, T., Huang, E., Murphy, G. and Zimmermann, T. (2014) Persuasive Technology in the Real World: A Study of Long-Term Use of Activity Sensing Devices for Fitness. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14), pp. 487–496. https://doi.org/10.1145/2556288.2557383

Ghosh, D., Foong, P. S., Zhang, S., and Zhao, S. (2018). Assessing the Utility of the System Usability Scale for Evaluating Voice-based User Interfaces. In Proceedings of the Sixth International Symposium of Chinese CHI (ChineseCHI '18), pp. 11-15. https://doi.org/10.1145/3202667.3204844

Gram-Hansen, S. B., Rabjerg, M. F. and Hovedskou, E. K. B. (2018) What makes it persuasive? In International Conference on Persuasive Technology, pp. 16-27.

Hermsen, S., Frost, J., Renes, R. J. and Kerkhof, P. (2016) Using feedback through digital technology to disrupt and change habitual behavior: A critical review of current literature. Computers in Human Behavior, 57: 61-74.

Hirano, S. H., Farrell, R. G., Danis, C. M. and Kellogg, W. A. (2013) WalkMinder: encouraging an active lifestyle using mobile phone interruptions. In CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13), pp. 1431-1436. https://doi.org/10.1145/2468356.2468611

Hong, J. and Findlater, L. (2018) Identifying Speech Input Errors Through Audio-Only Interaction. In Proceedings of the 2018 CHI Conference on Human Factors in

Computing Systems (CHI '18), Paper 567, 12 pages. https://doi.org/10.1145/3173574.3174141

Hura, S. L. (2017) Usability testing of spoken conversational systems. Journal of Usability Studies, 12(4): 155-163.

Jensen, K. L., Krishnasamy, R. and Selvadurai, V. (2010) Studying PH. A. N. T. O. M. in the wild: a pervasive persuasive game for daily physical activity. In Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction (OZCHI '10), pp. 17-20. http://dx.doi.org/10.1145/1952222.1952228

Jonze, S. (2013) Her [Motion picture] United States: Warner Bros. Pictures.

Kinsella, B. (2019) U.S. Smart Speaker Ownership Rises 40% in 2018 to 66.4 Million and Amazon Echo Maintains Market Share Lead Says New Report from Voicebot. https://voicebot.ai/2019/03/07/u-s-smart-speaker-ownership-rises-40-in-2018-to-66-4-million-and-amazon-echo-maintains-market-share-lead-says-new-report-from-voicebot/ (accessed 20 April 2020)

Kubrick, S. (1968) 2001: A Space Odyssey [Motion picture] United States: Metro-Goldwyn-Mayer, Inc.

Leahu, L., Cohn, M., and March, W. (2013). How categories come to matter. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13), 3331-3334. https://doi.org/10.1145/2470654.2466455

Liao, S. (2019) This project hacks Amazon Echo and Google Home to protect privacy. https://www.theverge.com/circuitbreaker/2019/1/15/18182214/amazon-echo-google-home-privacy-protection-project-white-noise (accessed 20 April, 2020)

Lin, J. L., Mamykina, L., Lindtner, S., Delajoux, G. and Strub, H. B. (2006) Fish'n'Steps: encouraging physical activity with an interactive computer game. In Proceedings of the 8th international conference on Ubiquitous Computing (UbiComp'06), pp. 261-278. http://dx.doi.org/10.1007/11853565_16

Luger, E. and Sellen, A. (2016) "Like Having a Really Bad PA": The Gulf between User Expectation and Experience of Conversational Agents. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16), pp. 5286–5297. https://doi.org/10.1145/2858036.2858288

Mateevitsi, V., Reda, K., Leigh, J. and Johnson, A. (2014) The health bar: a persuasive ambient display to improve the office worker's wellbeing. In Proceedings of the 5th Augmented Human International Conference (AH '14), Article 21, 2 pages. http://dx.doi.org/10.1145/2582051.2582072

McGregor, M. and Tang, J. C. (2017) More to Meetings: Challenges in Using Speech-Based Technology to Support Meetings. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17), pp. 2208-2220. https://doi.org/10.1145/2998181.2998335

Maulsby, D., Greenberg, S. and Mander, S. (1993) Prototyping an intelligent agent through Wizard of Oz. In Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93), pp. 277–284. https://doi.org/10.1145/169059.169215

Meurisch, C., Ionescu, M-D., Schmidt, B. and Mühlhäuser, M. (2017) Reference model of next-generation digital personal assistant: integrating proactive behavior. In Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International

Symposium on Wearable Computers (UbiComp '17), pp. 149–152. https://doi.org/10.1145/3123024.3123145

Murad, C., Munteanu, C., Clark, L. and Cowan, B. R. (2018) Design guidelines for hands-free speech interaction. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '18), pp. 269-276. https://doi.org/10.1145/3236112.3236149

Myers, C. Furqan, A., Nebolsky, J., Caro, K. and Zhu, J. (2018) Patterns for How Users Overcome Obstacles in Voice User Interfaces. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 6, 7 pages. https://doi.org/10.1145/3173574.3173580

NPR (2019) NPR Report: Smart Speakers See 78% increases YOY. https://www.npr.org/about-npr/682946406/npr-report-smart-speakers-see-78-increase-yoy?t=1559650166992&t=1560153697219 (accessed 20 April, 2020)

Oduor, M., Alahäivälä, T. and Oinas-Kukkonen, H. (2017) Software Design Patterns for Persuasive Computer–Human Dialogue: Reminder, Reward, and Instant Feedback. In L. Little, E. Sillence and A. Joinson (Eds.), Behavior change research and theory: Psychological and technological perspectives, pp. 47–67. Elsevier Academic Press. https://doi.org/10.1016/B978-0-12-802690-8.00003-7

Oinas-Kukkonen, H. and Harjumaa, M. (2009). Persuasive systems design: key issues, pro- cess model, and system features. Communications of the Association for Information Systems, 24(1): 28.

Orji, R. and Moffatt, K. (2018). Persuasive technology for health and wellness: State-of-the-art and emerging trends. Health informatics journal, 24(1): 66-91.

Orji, R., Tondello, G. F. and Nacke, L. E. (2018) Personalizing Persuasive Strategies in Gameful Systems to Gamification User Types. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 435, 14 pages. https://doi.org/10.1145/3173574.3174009

Orji, R., Vassileva, J. and Mandryk, R. L. (2013) LunchTime: a slow-casual game for long-term dietary behavior change. Personal Ubiquitous Comput. 17(6): 1211-1221. http://dx.doi.org/10.1007/s00779-012-0590-6

Paay, J., Kjeldskov, J., Brinthaparan, U., Lichon, L., Rasmussen, S., Srikandaraja, N., Smith, W., Wadley, G. and Ploderer, B. (2014) Quitty: using technology to persuade smokers to quit. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14), pp. 551-560. https://doi.org/10.1145/2639189.2639195

Peeters, M., Megens, C., van den Hoven, E., Hummels C. and Brombacher, A. (2013) Social stairs: taking the piano staircase towards long-term behavioral change. In Proceedings of the 8th international conference on Persuasive Technology (Persuasive '13), pp. 174-179. http://dx.doi.org/10.1007/978-3-642-37157-8_21

Porcheron, M., Fischer, J. E, Reeves, S. and Sharples, S. (2018) Voice Interfaces in Everyday Life. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 640, 12 pages. https://doi.org/10.1145/3173574.3174214

Pradhan, A., Mehta, K. and Findlater, L. (2018) "Accessibility Came by Accident": Use of Voice-Controlled Intelligent Personal Assistants by People with Disabilities. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 459, 13 pages. https://doi.org/10.1145/3173574.3174033

Purington, A., Taft, J. G., Sannon, S. Bazarova, N. N. and Taylor, S. H. (2017) "Alexa is my new BFF": Social Roles, User Satisfaction, and Personification of the Amazon Echo. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17), pp. 2853-2859. https://doi.org/10.1145/3027063.3053246

Purpura, S., Schwanda, V., Williams, K., Stubler, W. and Sengers, P. (2011) Fit4life: the design of a persuasive technology promoting healthy behavior and ideal weight. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11), pp. 423-432. https://doi.org/10.1145/1978942.1979003

Pyae, A. and Joelsson, T. N. (2018) Investigating the usability and user experiences of voice user interface: a case of Google home smart speaker. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '18), pp. 127-131. https://doi.org/10.1145/3236112.3236130

Schulman, D. and Bickmore, T. (2009) Persuading users through counselling dialogue with a conversational agent. In Proceedings of the 4th International Conference on Persuasive Technology (Persuasive '09), Article 25, 1–8. https://doi.org/10.1145/1541948.1541983

Singh, V. and Mathew, A. P. (2007) WalkMSU: an intervention to motivate physical activity in university students. In CHI '07 Extended Abstracts on Human Factors in Computing Systems (CHI EA '07), pp. 2657-2662. https://doi.org/10.1145/1240866.1241058

Springer, A. and Cramer, H. (2018) "Play PRBLMS": Identifying and Correcting Less Accessible Content in Voice Interfaces. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 296, 13 pages. https://doi.org/10.1145/3173574.3173870

Templeton, G. F. (2011) A Two-Step Approach for Transforming Continuous Variables to Normal: Implications and Recommendations for IS Research. Communications of the Association for Information Systems, 28. http://dx.doi.org/10.17705/1CAIS.02804

Toscos, T., Faber, A., An, S. and Gandhi, M.P. (2006) Chick clique: persuasive technology to motivate teenage girls to exercise. In CHI '06 Extended Abstracts on Human Factors in Computing Systems (CHI EA '06), pp. 1873-1878. http://dx.doi.org/10.1145/1125451.1125805

Villeneuve, D. (2017) Blade runner 2049 [Motion picture] United States: Warner Bros. Pictures

Vtyurina, A. and Fourney, A. (2018) Exploring the Role of Conversational Cues in Guided Task Support with Virtual Assistants. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18), Paper 208, 7 pages. https://doi.org/10.1145/3173574.3173782

Whitenton, K. (2017). Voice First: The Future of Interaction? Nielsen Norman Group, online article in Human Computer Interaction, https://www.nngroup.com/articles/voice-first/ (accessed 20 April, 2020)

WHO (2018) Physical Activity, https://www.who.int/news-room/fact-sheets/detail/physical-activity (accessed 20 April, 2020)

Win, K. T., Roberts, M. R. and Oinas-Kukkonen, H. (2019). Persuasive system features in computer-mediated lifestyle modification interventions for physical activity. Informatics for Health and Social Care, 44(4): 376-404.

Wulf, L., Garschall, M., Himmelsbach, J. and Tscheligi. M. (2014) Hands free - care free: elderly people taking advantage of speech-only interaction. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14), pp. 203-206. https://doi.org/10.1145/2639189.2639251