technologies

Article

Insights on the Effect and Experience of a Diet-Tracking Application for Older Adults in a Diet Trial †

Laura M. van der Lubbe [1,*], Michel C. A. Klein [1], Marjolein Visser [2], Hanneke A. H. Wijnhoven [2] and Ilse Reinders [2]

1 Computer Science, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands; michel.klein@vu.nl
2 Department of Health Sciences, Faculty of Science, and the Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, 1081 HV Amsterdam, The Netherlands; m.visser@vu.nl (M.V.); hanneke.wijnhoven@vu.nl (H.A.H.W.); ilse.reinders@vu.nl (I.R.)
* Correspondence: l.m.vander.lubbe@vu.nl

† This paper is an extended version of our paper published in PETRA2021, Corfu, Greece, 29 June–2 July 2021.

Abstract: With an ageing population, healthy ageing becomes more important. Healthy nutrition is part of this process and can be supported in many ways. The PROMISS trial studies the effect of increasing protein intake in older adults on their physical functioning. Within this trial, a sub-study was performed, researching the added effect of using a diet-tracking app enhanced with persuasive and (optional) gamification techniques. The goal was to see how older adult participants received such technology within their diet program. There were 48 participants included in this sub-study, of which 36 completed the study period of 6 months. Our results on adherence and user evaluation show that a dedicated app used within the PROMISS trial is a feasible way to engage older adults in diet tracking. On average, participants used the app 83% of the days, during a period of on average 133 days. User-friendliness was evaluated with an average score of 4.86 (out of 7), and experienced effectiveness was evaluated with an average score of 4.57 (out of 7). However, no effect of the technology on protein intake was found. The added gamification elements did not have a different effect compared with the version without those elements. However, some participants did like the added gamification elements, and it can thus be nice to add them as additional features for participants that like them. This article also studies whether personal characteristics correlate with any of the other results. Although some significant results were found, this does not give a clear view on which types of participants like or benefit from this technology.

Keywords: persuasive technology; gamification; older adults; evaluation study; diet compliance; user experience

check for updates

Citation: van der Lubbe, L.M.; Klein, M.C.A.; Visser, M.; Wijnhoven, H.A.H.; Reinders, I. Insights on the Effect and Experience of a Diet-Tracking Application for Older Adults in a Diet Trial. *Technologies* **2022**, *10*, 31. https://doi.org/10.3390/technologies10010031

Academic Editors: Mario Munoz-Organero and Fillia Makedon

Received: 24 December 2021
Accepted: 8 February 2022
Published: 16 February 2022

1. Introduction

As populations are ageing, healthy ageing becomes more important. One of the aspects contributing to healthy ageing is healthy nutrition. Research has shown that smartphone applications (apps) used during dietary interventions have a higher adherence compared with websites or paper diaries [1]. Moreover, it is shown that smartphone apps can improve the diet compliance of participants, e.g., see [2]. Creating a specific app that meets the needs of the target group is preferred over using a general journal [3].

Gamification is used in many different domains, among which is healthy lifestyle [4]. Gamification can be described as "the intentional use of game elements for a gameful experience of non-game tasks and contexts" [5]. One of the main goals of adding gamification to an app is that it stimulates the motivation of users so that the outcomes of the application are increased. Although the ease of use of gamification declines with age [6], it can be used for older adults in different contexts; e.g., different rehabilitation or prevention games for older adults exist, for example, to make physical training to prevent falls more fun, e.g., see [7].

The PROMISS trial aims to increase protein intake in older adults with a relatively low protein intake [8,9]. Good physical functioning is important for healthy ageing and living independently at home. In the 6-month PROMISS trial, the effect of protein on physical function is studied. Part of this trial was the persuasive technology sub-study, in which participants were provided with a tablet app and a so-called foodbox, both designed specifically for diet tracking within this trial [10,11]. Persuasive communication and gamification techniques were used in the app to enhance adherence to food intake registration and diet compliance. The goal was to see how such technology was received by the older adult users and, as part of this assessment, whether it has an effect on, for example, the outcome of the dietary advice and experience with the diet.

The current article describes the outcomes of the persuasive technology sub-study of the PROMISS trial. First, in Section 2, a short overview of the background of this work is given. The research question, materials, and methods used are explained in Section 3. In Section 4, an overview of the participants is given, followed by analyses of different research questions. At the end of each subsection about a research question, a summary of the results can be found. Finally, the discussion (Section 5) summarises the results and discusses possible limitations and future possibilities. Finally, the most important lessons learned are highlighted in the conclusions in Section 6.

2. Background

In previous work, more details on the background and related work of the work presented in the current article were presented [10,11]. This section gives a brief overview of the relevant background literature.

In the field of healthy lifestyle applications, different types of apps can be found, with the majority focusing on physical activity and (healthy) nutrition [4]. In the review by Johnson (2016), seven promises of gamification for health and well-being are mentioned. Among those promises are the broad appeal and applicability and the motivation it gives to users. Apps for a healthy lifestyle often aim to change the behaviour of the users. Achieving this is often more successful when behaviour change techniques (BCTs) are incorporated in the intervention [12,13].

Previous research has shown that tailoring the design of an application to an older adult target group is good for, among other things, the effectiveness, efficiency, and user satisfaction [14]. Examples of guidelines that can be used when designing for older adults are using larger screens, fonts, and spaces between buttons [15]. In addition, the methods of interaction, navigation, and information loading are aspects that need tailoring to the target audience of older adults. In general, for all target groups, it is found that, in the domain of food intake registration, an easy and quick administration of food intake is important [2].

3. Materials and Methods

The PROMISS trial addresses whether dietary advice to increase protein intake to $\geq$1.2 g/kg adjusted body weight (aBW)/day (this measure is explained in Table 1) is beneficial for physical functioning in community-dwelling older adults. Products are valued according to a protein point system designed for this trial. More details on the design of the trial can be found in its design paper [8]. The PROMISS app is designed using the same information and instructions on products and protein as all participants of the PROMISS trial receive in an information booklet.

The PROMISS app is provided to our participants on a dedicated tablet. The app is a diet journal, in which participants can register their intake in two ways: directly entering the protein value of the meal or choosing products to create a meal. In the latter case, the app calculates the protein value of the meal. The app can also help to replace products in a meal, keeping the protein value the same. Registering intake can be performed for each instance of eating in the participants' personal dietary advice. The app gives an overview of the protein value of each meal and the whole day and compares it with the personal thresholds in the dietary advice. Moreover, the application uses persuasive communication in its notifications that remind participants to eat. Next to the app, the foodbox is a custom-made box in which specific protein products provided from the trial, such as puddings and protein bars, can be stored. Taking them out of the box triggers a notification in the app that helps users register the protein value associated with the product. More details on the design of the system can be found in a previously published paper [10]. In a second variant of the app, gamification elements (rewards, achievements, a profile page, and mini-games) were added. These elements were added to study their added value for adherence, for increasing protein knowledge (via the mini-games), and for making the app more fun to use. More details on the design of the gamified variant can be found in a previously published paper [11].

3.1. Participants

The persuasive technology sub-study was only performed in the PROMISS trial site in the Netherlands. This trial consisted of three study groups: a control group and two intervention groups, in which participants received personalised dietary advice. Recruitment and selection was part of the PROMISS trial and is described in [8,9]. The study included community-dwelling older adults (≥65 years) with a lower protein intake. Exclusion criteria, such as health concerns, are listed in Reinders et al. [8]. Participants in the intervention groups could opt-into the sub-study, as they need to obtain dietary advice that can be put in the app. Only one participant per household could be included in the sub-study because it was likely that participants living together influence their outcomes by working together. The sub-study shared the measure moments with the trial, namely baseline, 3 months, and 6 months. Participants were free to stop using the technology provided in the sub-study earlier, and the time at which instances were measured remained the same.

An overview of the participants in the trial in the Netherlands and the persuasive technology sub-study (PT) can be found in Figure 1. In total, 48 participants took part in our sub-study. Participants who did not fill in the persuasive technology evaluation questionnaire or who had ≤20 days of input on the tablet were excluded from the analysis and considered drop-outs. In total, 12 participants dropped out, of which 3 also dropped out of the trial. For this data analysis, the focus was on the 36 PT participants and the 41 participants who were not in the PT group (no PT). Five participants from the PT group had data input on the tablet for more than 200 days; those five participants were considered the COVID-19 extension group. The PT group contained 15 females and 21 males, with a mean age of 74 (SD = 4.7, minimum = 68, maximum = 85). The majority of the participants had an education level of secondary education or higher. There was no information available about the experience of participants with technology. There were 17 participants in the normal condition and 19 participants in the gamification condition. In total 26 participants received a foodbox. Participants could stop using the foodbox whenever they wanted, without becoming a drop-out.

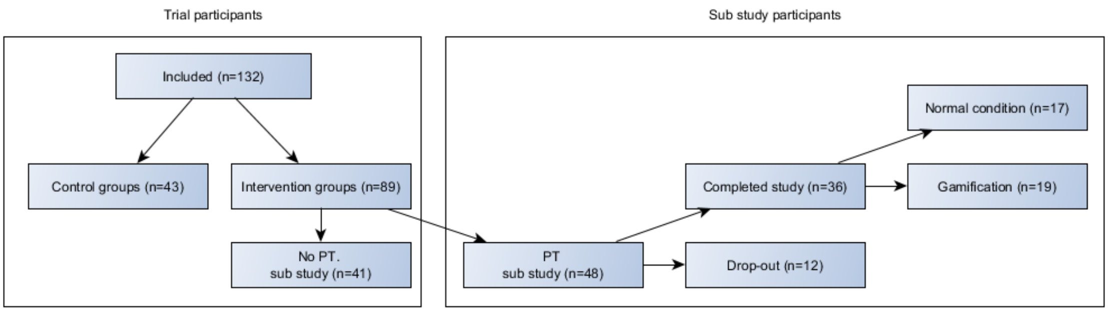

Figure 1. Overview of participants in the trial and sub-study [16].

Statistical tests were performed to determine whether participants in the PT group have comparable characteristics with the no-PT participants at baseline. No significant results are found for age, sex, education level, and mean protein intake at baseline in the grams per kilogram-adjusted body weight per day. This means that the PT and no-PT groups were not significantly different. It is compared whether there were significant differences between participants in the normal and gamification conditions. Again, no significant differences were found; thus, it can be concluded that the participants in both conditions were comparable.

Drop-Out

As mentioned, twelve participants dropped out from the sub-study. Three of these dropped out of the main trial. For the other participants, remarks they made about dropping out were noted.

One participant returned the tablet as he/she did not see the added value of the tablet over using a paper diary. This participant also told us to not use (and that they are not familiar with) technology such as a tablet or smartphone. The lack of familiarity with technology and understanding of the tablet was a reason for dropping out for at least two other participants as well.

Two other participants noted that using the tablet was educational, but after some time, they no longer needed to use it or that it was time-consuming and bothersome. Moreover, finding products in the tablet was a problem for one of these participants, as the ordering was sometimes unclear and products were missing. Another participant also mentioned the tablet being time-consuming as a reason to drop out.

Another participant used the tablet once but told us to forget about it and therefore completely quit using the tablet. While another participant used the tablet on some days during the trial, but they did not use the tablet for the full trial for an unspecified reason. It needs to be noted that this participant also missed some appointments for the main study as well.

3.2. Design

The sub-study had two tablet conditions: the normal and the gamification conditions. The assignment to these conditions was semi-random because the inclusion for the gamification condition started later due to logistical reasons, and the aim was to keep the size of the groups equal. Aside from this, the foodbox was a storage box for protein products provided to some participants in the trial and can automatically register the consumption of these products. As not all participants received the products that can be registered by the foodbox and as the number of foodboxes available was limited, not all participants received one, regardless of their study condition.

The main research question of the sub-study was as follows:

> How is a gamified persuasive diet-tracking system as part of a diet program received by older adults?

To answer this question, different research tasks (RT) were defined:

1. To explore the **adherence** of participants who use the diet-tracking system.
2. To statistically test if using persuasive technology increases **protein intake**.
3. to statistically test if the **experience of the dietary advice** improves using persuasive technology.
4. To explore the **experience of the persuasive technology**.
5. To explore the effect of **gamification** on participants' **knowledge** about the amount of protein in food products.

3.3. Materials

Different data sources were used to address these research tasks. Table 1 shows the different data sources and for which RTs those sources are relevant. Within this study, the focus was to evaluate the current version of the application and its effects. Thus, mainly quantitative data in the form of questionnaire answers, protein intake, and log data were collected.

The measures for the dietary advice evaluation can be found in Appendix A.1; these were designed within the larger PROMISS trial [8]. The persuasive technology questionnaire evaluates the experienced effectiveness and user-friendliness of the tablet and the foodbox, the notifications shown in the app and the gamification elements, using seven-point Likert scales. All questions of this questionnaire can be found in Appendix A.2. This questionnaire was custom made for this study. The goal was to obtain more details on specific aspects of the system, and therefore, it was decided to custom make our questionnaires. Moreover, to ensure that older adults understand the wordings, the statements were made concrete about specific aspects. The structured questionnaires with (mainly) closed questions aimed to give insight into the evaluation of the current system. As the PROMISS trial was only performed once, more in-depth information for improvements was not gathered. In cases where statistical tests were used to complete a research task, a threshold of 0.05 was used for significance. In the tables in Section 4, significant results are marked with an asterisk (*).

Table 1. Overview of different data sources.

Data Type	Explanation	Measured	Participant Group	RT
Tablet data	Logging of participant interaction with app (e.g., which buttons clicked, which products chosen).	During tablet use	Sub-study	1, 2, 4, 5
Mean protein intake in g/kg adjusted BW/d	Mean protein intake in gram per kilogram body weight per day, adjusted for the BMI of the participant. For participants with a low/high BMI, their protein intake was calculated based on a healthy BMI.	At baseline, 3 months, and 6 months	All	2
Dietary advice evaluation	General evaluation questions about how participants experienced/value the dietary advice.	At 6 months	All	3, 5
Persuasive technology evaluation	Questionnaire about user experience and expected effectiveness of the persuasive technology.	At 6 months	Sub-study	4

The tablet collected different data, mainly to ensure correct functioning of the tablet. Every click was logged into the database (action and timestamp), with these data, we can analyse what participant clicked in the application. For each meal that was registered, the protein value and timestamp were saved, and if provided, the products and amounts were also added to this logging.

Due to the COVID-19 pandemic in 2020 (https://time.com/5791661/who-coronavirus-pandemic-declaration/, accessed on: 31 November 2021), it was not possible to conduct final measurements for participants for some time. Due to this, 14% of the participants were able to use the tablet for an extended period. As a cut-off point, 200 days (normal maximum duration) of usage was assumed. Above that, this is marked as a 'COVID-19 extension'. Whether the fully extended trial or the trial period (of maximum 200 days) was used is indicated in the analyses.

3.4. Procedure

Participants received their tablets and verbal instructions, together with written information, at the beginning of their diet trial period. After that, participants could use the tablet for six months, during which they could contact the researcher if they experienced technical difficulties. Table 1 shows which data were collected at which measure moment. All questionnaires were part of the procedure of the main trial.

4. Results

For each research task, data analysis and statistical testing (if applicable) were performed. Throughout this sections, summary boxes highlight the most important results that are found.

4.1. Adherence (RT1)

Figure 2 shows the duration of the trial for the PT participants, specifying the number of days in which they did and did not provide input. On average, participants used their tablets for 133 days (SD = 51.03, minimum = 27, maximum = 198). To measure the length of participants' tablet usage period, the first and last dates on which they registered any input were used. Figure 3 shows the duration of the trials of five participants with an extended trial period of >200 days. They used their tablet on average for 196.4 days (SD = 57.19, minimum = 101, maximum = 239). Moreover, Figure 3 shows that three of these participants were active users for most of the trial, while the others used the tablet less frequently but continued to use the tablet during the extended part of the trial.

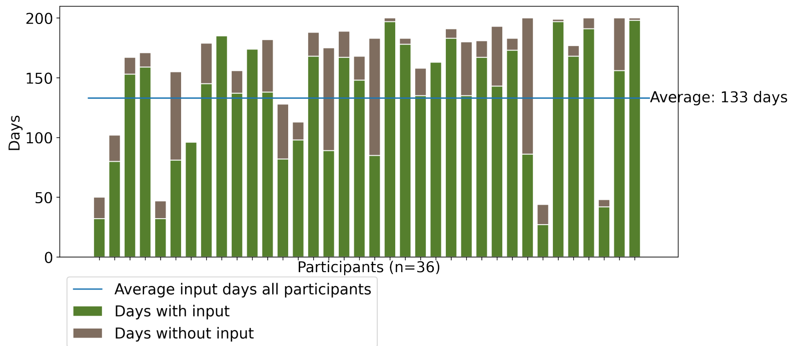

Figure 2. Number of days with and with tablet input in the trial per participant [16].

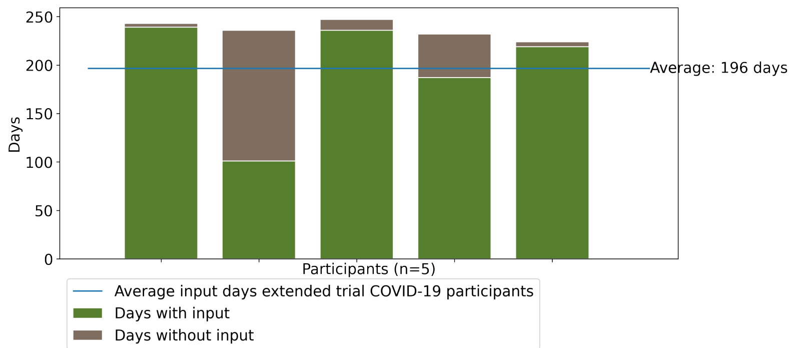

Figure 3. Number of days with and with tablet input in the trial per participant for the COVID-19 extension case [16].

Figure 4 shows the division of days with and without inputs; for this figure, the full duration of the participants in the extended trial period was used, as this is a relative graph. On average, participants used the tablet 82.7% (SD = 16.60%, minimum = 43.80%, maximum = 100%) of the total tablet use days. Based on the average number of input days and the relative adherence, the adherence of participants was considered high.

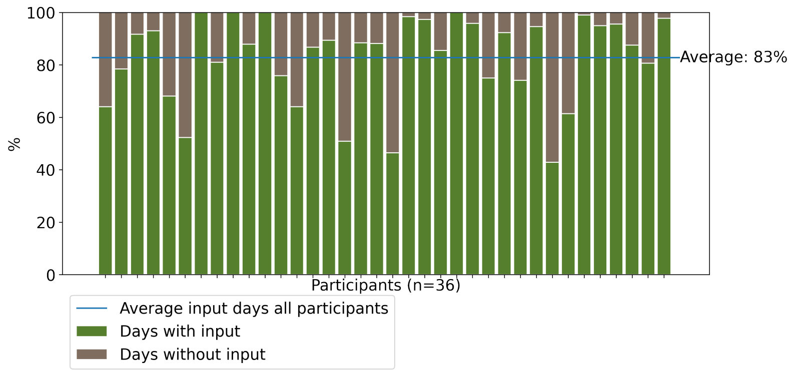

Figure 4. Percentage of (in)active days in the trial per participant [16].

The adherence of participants is considered high:

- The average number of active days of participants without COVID-19 extension is 133 (SD = 51.03, minimum = 27, maximum = 198), see Figure 2.
- The average number of active days of participants with COVID-19 extension is 196.4 (SD = 57.19, minimum = 101, maximum = 239), see Figure 3.
- The relative number of input days for participants is 82.7% (SD = 16.60%, minimum = 43.80%, maximum = 100%), see Figure 4.

4.1.1. Differences between Tablet Conditions for Adherence

When studying the differences between the normal condition and the gamification condition (excluding COVID-19 extension) for the adherence, a slight difference in the mean duration length is found; see Figures 5 and 6. However, a *t*-test shows that this difference is not significant (*p*-value = 0.378).

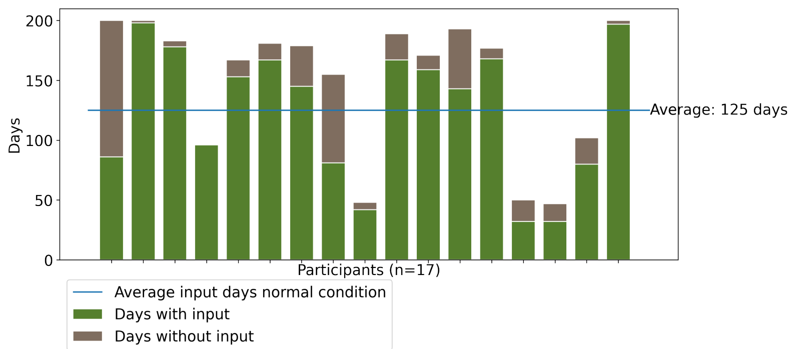

Figure 5. Number of days with and with tablet inputs in the trial per participant—normal condition.

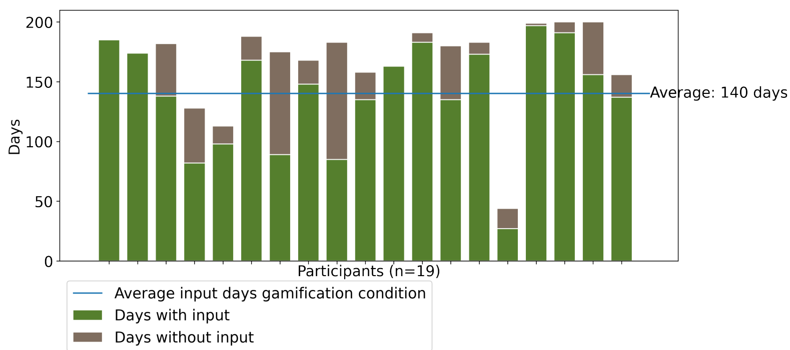

Figure 6. Number of days with and with tablet inputs in the trial per participant—gamification condition.

The relative graphs (including the COVID-19 extension), Figures 7 and 8, for the two conditions show that the average relative active usage is almost the same, namely 82% (SD = 17.03%) for the normal condition and 83% (SD = 16.67%) for the gamification condition. This difference in significance is tested with a *t*-test. The difference is not significant (*p*-value = 0.927). Based on these results, it is concluded that the version of the app did not influence the adherence of participants.

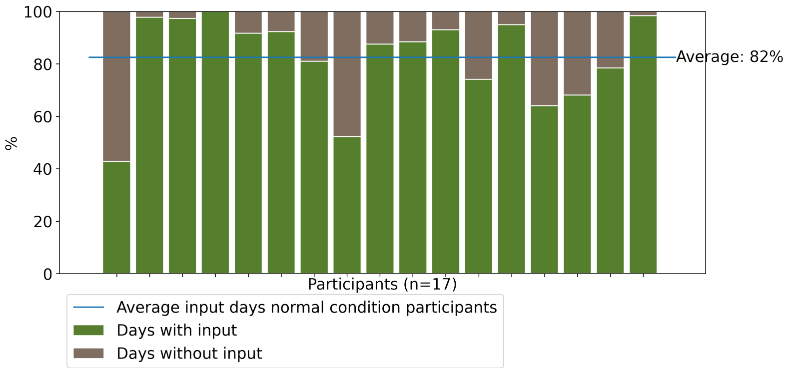

Figure 7. Percentage of (in)active days in the trial per participant—Normal condition.

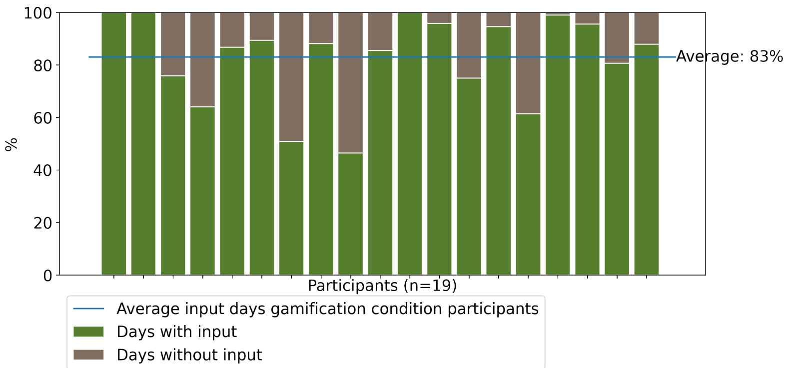

Figure 8. Percentage of (in)active days in the trial per participant—Gamification condition.

The version of the tablet application does not influence the adherence of participants:

- The average number of active days of participants (excluding COVID-19 extension) in the normal condition is 124.9 (SD = 56.94, minimum = 32, maximum = 198), see Figure 5.
- The average number of active days of participants (excluding COVID-19 extension) in the gamification condition is 140.2 (SD = 45.45, minimum = 27, maximum = 197), see Figure 6.
- The *t*-test for the difference between active days for the normal and gamification conditions is not significant (*p*-value = 0.378).
- The relative number of input days for normal-condition participants (including COVID-19 extension) is 82.4% (SD = 17.03%, minimum = 42.8%, maximum = 100%), see Figure 7.
- The relative number of input days for gamification-condition participants (including COVID-19 extension) is 83% (SD = 16.67%, minimum = 46.4%, maximum = 100%), see Figure 8.
- The *t*-test for the difference between the relative number of input days for the normal and gamification conditions is not significant (*p*-value = 0.927).

4.1.2. Differences by Personal Characteristics for Adherence

It is also statistically tested whether differences in adherence are caused by personal characteristics. It is tested whether the adherence of male and female participants is significantly different with a *t*-test. Both the mean number of days and the relative adherence are not significantly different (*p*-value = 0.990 and *p*-value = 0.725). With the help of the Pearson correlation, it is determined whether age or protein intake at baseline is correlated with the adherence of participants, both for the days of input and relative adherence. Both age (*p*-value = 0.254 and *p*-value = 0.904) and mean protein intake at baseline (*p*-value = 0.599 and *p*-value = 0.433) do not show significant differences. Finally, a one-way ANOVA is performed to test whether there are significant differences in the days of input or relative adherence between the education levels. Again, there are no significant differences found (*p*-value = 0.529 and *p*-value = 0.099). Based on all these statistical tests, it is concluded that personal characteristics do not influence the adherence of participants.

The personal characteristics of participants do not influence the adherence of participants. Different statistical tests showed no significant differences caused by personal characteristics (sex, age, protein intake at baseline, and education levels).

4.2. Persuasive Technology and Protein Intake (RT2)

Table 2 shows the *p*-values from *t*-tests comparing the protein intake of the PT group with the protein intake of the no-PT group. Moreover, it shows whether the mean of all participants included in our sub-study (PT + drop-out) is different from the no-PT group, based on *t*-tests.

When comparing the PT participants with the no-PT participants, no significant differences in their protein intake were found. However, when comparing all the participants who were included in the sub-study with the no-PT group, significant differences for the intake at 6 months and the change between baseline and 6 months were found: the no-PT group has a higher change. It is however unclear what could explain this difference. If there would be an effect of the intention to treat, it is expected that this effect is also found for the active participants.

Table 2. Results *t*-tests for protein intake (expressed in g/kg adjusted BW/d) between the PT and no-PT groups. * $p < 0.05$.

Measure	Mean PT	Mean All Sub-Study Participants	Mean No PT	*p*-Value PT	*p*-Value All Sub-Study
Mean protein intake—3 months	1.2 SD = 0.24 (n = 34)	1.2 SD = 0.27 (n = 43)	1.3 SD = 0.23 (n = 38)	0.194	0.175
Mean protein intake—6 months	1.2 SD = 0.19 (n = 36)	1.2 SD = 0.23 (n = 44)	1.3 SD = 0.28 (n = 38)	0.086	0.019 *
Change protein—between baseline and 6 months	0.4 SD = 0.25 (n = 36)	0.4 SD = 0.27 (n = 44)	0.5 SD = 0.30 (n = 38)	0.132	0.026 *
Change protein—between 3 and 6 months	0.01 SD = 0.28 (n = 34)	−0.02 SD = 0.29 (n = 42)	0.05 SD = 0.26 (n = 37)	0.549	0.230
Change protein—between baseline and 3 months	0.4 SD = 0.24 (n = 34)	0.4 SD = 0.24 (n = 43)	0.5 SD = 0.25 (n = 38)	0.351	0.236

To investigate this further, the correlation between the number of input days in the tablet and the diet outcomes is studied. Table 3 shows the results of these analyses. For these analyses, the full duration of the trial, including the extension due to COVID-19, is taken into account. Again, only a significant correlation for the intake at 6 months and the change between baseline and 6 months for all included sub-study participants is found.

Table 3. Results from the Pearson correlation test between days with tablet registration and protein intake (expressed in grams per kilogram-adjusted body weight per day). * $p < 0.05$.

Measures	All Participants Sub-Study Pearson Correlation Coefficient	*p*-Value	PT Participants Pearson Correlation Coefficient	*p*-Value
Mean protein intake—3 months	0.17 (n = 43)	0.289	0.21 (n = 34)	0.231
Mean protein intake—6 months	0.42 (n = 44)	0.005 *	0.21 (n = 36)	0.217
Change protein—between baseline and 6 months	0.32 (n = 44)	0.033 *	0.11 (n = 36)	0.529
Change protein—between 3 months and 6 months	0.16 (n = 42)	0.316	−0.09 (n = 34)	0.626
Change protein—between baseline and 3 months	0.11 (n = 43)	0.482	0.06 (n = 34)	0.722

Figure 9 shows the percentage of participants in the PT and no-PT groups that reached the goal of the main trial 1.2 protein intake in grams per kilogram-adjusted body weight per day for the 3-month measurement. The difference between how many participants reached the goal in each group is not significant (*p*-value = 0.490).

Figure 10 shows this for the 6-month measurement. Again, the differences between the two groups are not significant (*p*-value = 0.621). In both groups, the number of participants who stuck with the goal protein intake decreased at 6 months compared with at 3 months. However, the change in the group of tablet users was smaller compared with the other participants of the intervention groups.

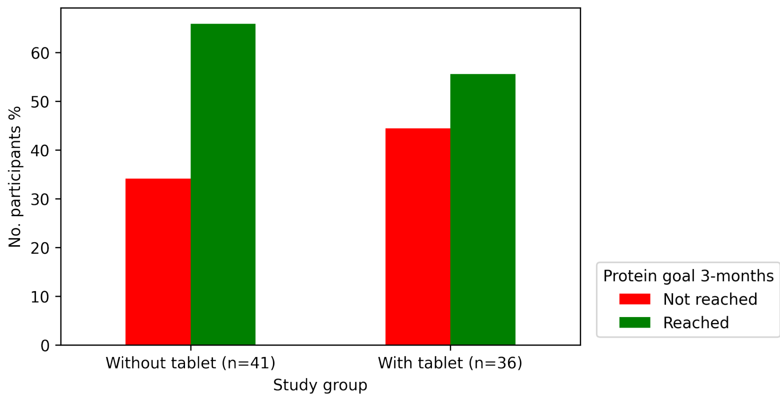

Figure 9. Participants reaching 1.2 protein intake in grams per kilogram-adjusted body weight per day at 3 months [16].

We also analysed whether there was a difference between the groups when looking at participants who reached their protein goal at both instances of measured; see Figure 11. In the group without a tablet, half of the participants reached the goal at both moments and half did not. In the group with the tablet, this was more divided and the majority did not reach the goal at both moments. However, the differences between the two groups are not significant (*p*-value = 0.521).

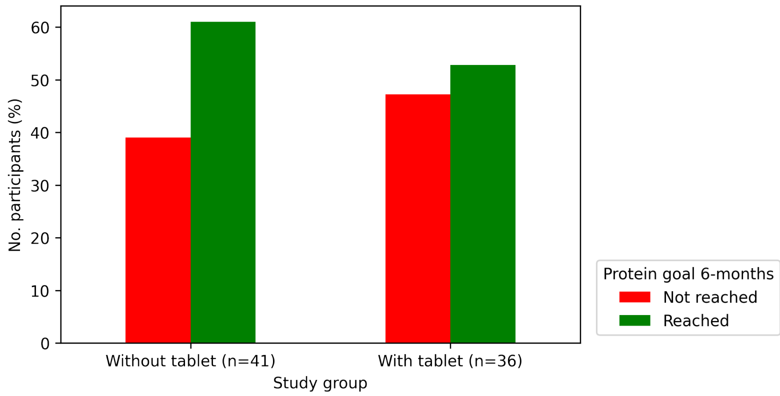

Figure 10. Participants reaching 1.2 protein intake in grams per kilogram-adjusted body weight per day at 6 months [16].

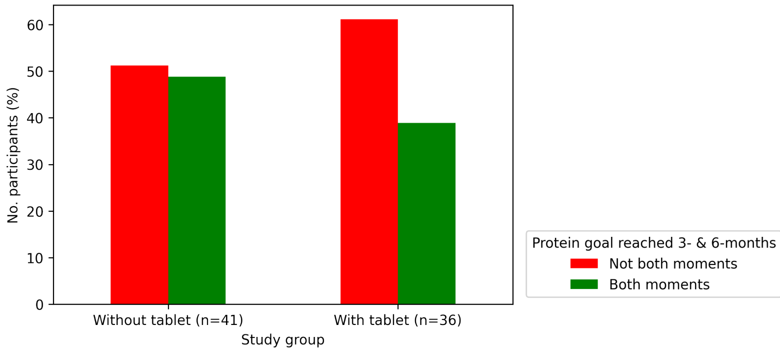

Figure 11. Participants reaching 1.2 protein intake in grams per kilogram-adjusted body weight per day at 3 and 6 months [16].

Using the persuasive technology did not increase the protein intake of participants:

- T-tests were performed to see if PT participants (n = 36) reached better diet outcomes compared with the no-PT participants (n = 41) and to compare all included sub-study participants (n = 48) with the no-PT participants; see Table 2. Significant results are found for all sub-study participants and the no-PT participants for the mean intake at 6 months (p-value = 0.019) and changed between baseline and 6 months (p-value = 0.026).
- Pearson correlations were calculated for the protein intake and the days with tablet registration for all sub-study participants and PT participants; see Table 3. Significant correlations were found for mean intake at 6 months (p-value = 0.005) and change intake between baseline and 6 months (p-value = 0.033) for all sub-study participants.
- The result of the Chi^2-test for the relation between the groups (PT and no PT) and reaching a threshold value of 1.2 g/kg adjusted BW/d at 3 months was not significant (p-value = 0.490).
- The result of the Chi^2-test for the relation between the groups (PT and no PT) and reaching a threshold value of 1.2 g/kg adjusted BW/d at 6 months was not significant (p-value = 0.621).

4.2.1. Differences between Tablet Conditions for Protein Intake

For the PT participants in the normal and gamification conditions, the diet outcomes were compared using t-tests. The results of these tests are summarised in Table 4. No significant results are found, so the tablet condition did not influence the change in protein intake of participants.

Table 4. Resulting *t*-tests for protein intake—per condition (expressed in grams per kilogram-adjusted body weight per day).

Measure	Mean Normal Condition	Mean Gamification Condition	*p*-Value
Mean protein intake in g/kg adjusted BW/d—3 months	1.2 SD = 0.20 (n = 15)	1.3 SD = 0.27 (n = 19)	0.244
Mean protein intake in g/kg adjusted BW/d—6 months	1.2 SD = 0.23 (n = 17)	1.3 SD = 0.15 (n = 19)	0.302
Change protein in g/kg adjusted BW/d—between baseline and 6 months	0.4 SD = 0.29 (n = 17)	0.4 SD = 0.22 (n = 19)	0.418
Change protein in g/kg adjusted BW/d—between 3 months and 6 months	0.04 SD = 0.31 (n = 15)	−0.01 SD = 0.27 (n = 19)	0.633
Change protein in g/kg adjusted BW/d—between baseline and 3 months	0.4 SD = 0.21 (n = 15)	0.4 SD = 0.26 (n = 19)	0.331

The version of the tablet application does not influence the protein intake of participants:

- T-tests to see if participants from normal conditions had different changes in protein intake than participants from the gamification condition (Table 4) show no significant results.

4.2.2. Differences by Personal Characteristics for Protein Intake

It is interesting to see whether the personal characteristics of participants in the PT group influenced protein intake. For the different measures of protein intake, statistical tests were performed: a *t*-test for sex, Pearson correlation coefficients for age and protein intake at baseline, and a one-way ANOVA for education level. The *p*-values resulting from these tests can be found in Table 5.

Table 5. *p*-values for protein intake (expressed in grams per kilogram-adjusted body weight per day) and personal characteristics. * $p < 0.05$.

	p-Value			
Measure	Sex	Age	Protein Baseline	Education
Mean protein intake in g/kg adjusted BW/d—3 months	0.886	0.100	0.052	0.219
Mean protein intake in g/kg adjusted BW/d—6 months	0.959	0.235	0.393	0.766
Change protein in g/kg adjusted BW/d—between baseline and 6 months	0.561	0.105	1×10^{-5} *	0.113
Change protein in g/kg adjusted BW/d—between 3 months and 6 months	0.957	0.036 *	0.036 *	0.058
Change protein in g/kg adjusted BW/d—between baseline and 3 months	0.661	0.269	0.192	0.648

The protein intake at baseline correlates significantly with the change between baseline and 6 months. To see if this is due to the tablet, the same test for the no-PT participants was conducted. This also showed a significant difference (*p*-value = 0.045), so we cannot conclude that this is an effect of the tablet. Moreover, age and protein intake at baseline correlate significantly with protein change between 3 and 6 months. However, when studying the data in more detail, it was found that one protein change is much higher compared with others. When removing this outliers, the statistical significance disappears. It, therefore, cannot be concluded that age or protein intake at baseline influences the effect of tablet use on protein intake.

The personal characteristics and tablet use of participants did not influence the protein intake of participants:

- Statistical tests for personal characteristics and protein intake (Table 5) show no significant differences for personal characteristics and tablet use.

4.3. Experience of the Diet (RT3)

After 6 months, the dietary evaluation questionnaire (see Appendix A.1) was used to assess the experience of users with the dietary advice. The PT group was compared with the no-PT group to see if there were significant differences for these questions. The results are shown in Table 6. No find significant differences were found for any of the measures. However, for both the usefulness of the dietary advice and the ease of finding products, the sub-study participants had a slightly better appreciation. The ease of finding products was one of the goals of the app.

Table 6. Results of the statistical tests for diet evaluation (range 1–5).

Measures	**Means PT**	**Means No PT**	***p*-Values**
Rating dietary advice (range 1–10)	8.4 SD = 1.03 (n = 36)	8.5 SD = 1.04 (n = 37)	0.608
Usefulness dietary advice	4.2 SD = 0.65 (n = 36)	4.1 SD = 0.70 (n = 37)	0.475
Extend to which dietary advice is followed	4.3 SD = 0.61 (n = 36)	4.4 SD = 0.49 (n = 37)	0.087
Ease to stick to dietary advice	3.9 SD = 0.86 (n = 36)	4.0 SD = 0.70 (n = 37)	0.143
Intend to continue to follow dietary advice	3.9 SD = 0.67 (n = 36)	4.0 SD = 0.71 (n = 37)	0.093
Ease to find products with similar protein amount	3.9 SD = 0.85 (n = 36)	3.8 SD = 0.71 (n = 37)	0.402

Using persuasive technology does not seem to affect the experience of the diet:

- Statistical tests comparing the PT group with the no-PT group (Table 6) show no significant differences in the experience of the diet.

4.3.1. Differences between Tablet Conditions for the Experience of the Diet

The same analyses were performed to see if there were differences in the experience of the diet between the two tablet conditions. The results can be found in Table 7. A significant difference is found for the rating of the diet, which is higher for participants in the normal condition.

Table 7. Results of the statistical tests for diet evaluation—per condition (range 1–5). * $p < 0.05$.

Measure	**Mean Normal Condition**	**Mean Gamification Condition**	***p*-Value**
Rating dietary advice (range 1–10)	8.8 SD = 0.90 (n = 17)	8.0 SD = 1.04 (n = 19)	0.036 *
Usefulness dietary advice	4.3 SD = 0.69 (n = 17)	4.2 SD = 0.63 (n = 19)	0.336
Extend to which dietary advice is followed	4.2 SD = 0.56 (n = 17)	4.3 SD = 0.67 (n = 19)	0.314
Ease to stick to dietary advice	4.0 SD = 0.71 (n = 17)	3.9 SD = 0.99 (n = 19)	0.500
Intend to continue to follow dietary advice	3.8 SD = 0.75 (n = 17)	4.1 SD = 0.57 (n = 19)	0.079
Ease to find products with similar protein amount	3.8 SD = 0.97 (n = 17)	4.0 SD = 0.74 (n = 19)	0.266

Different statistical tests compared the normal and gamification condition (Table 7). A significant difference was found for rating their diets (p-value = 0.036); the rating was higher for the normal conditions. Other aspects were not significantly different.

4.3.2. Differences by Personal Characteristics for the Experience of the Diet

It was tested whether the personal characteristics of participants in the PT group influenced their evaluation of the application. For the rating, the same statistical tests as those used for the personality characteristics tests described in Section 4.2.2 were used. For the other measures, the Mann–Whitney U test was used for the sex characteristic and the Spearman correlation was used for the other characteristics. One significant difference was found: female tablet users score the question about continuing to follow the dietary advice significantly higher compared with male participants (p-value = 0.029). For the no-PT users, this difference was not found (p-value = 0.417). It, therefore, seems that the tablet has a higher effect on female participants' intentions to continue to follow dietary advice.

Different statistical tests for the associations between personal characteristics and the evaluation of the diet were performed. A significant result for intention to continue to follow dietary advice and sex (p-value = 0.029) was found, and the score was higher for women. Other differences associated with personal characteristics were not found.

4.4. Experience Persuasive Technology (RT4)

This section analyses how participants experience persuasive technology. To do so, the results from the evaluation questionnaire as well as some data on the interaction that participants make with the system were analysed. For the evaluation, differences between the two tablet conditions and the effects caused by personal characteristics were analysed.

4.4.1. Evaluation Questionnaire Persuasive Technology

Figure 12a shows the scores of the *experienced effectiveness of the tablet*. This shows that most participants find the tablet helpful (Q1) and the functionalities sufficient (Q2). However, the participants are more divided and less positive about how the tablet met their expectations (Q3). Participants are on average slightly negative about continuing to use the tablet after the study finished (Q4). One of the participants added to this that he/she was now familiar with the protein diet. Another participant added to their answer to Q4 that it would be nice to use the tablet from time to time, instead of every day. Overall, concerning the experience effectiveness (Q1–Q4), participants were slightly positive (mean = 4.6, SD = 1.94).

Figure 12b shows the scores for the *user-experience of the tablet*. The only question that scored below neutral was the question about irregularities (Q6). During the trial, different irregularities were found, some of which could not be solved during the trial. Overall, the user-friendliness was evaluated slightly positive, with an average score of 4.9 (SD = 1.99) for Q5–Q10.

Figure 12c shows the scores for the *foodbox questions* (Q11–Q13), which were on average evaluated negatively (mean = 2.8, SD = 1.79). Only the questions about ease to use (Q12) were evaluated a bit more positive.

There was some confusion among participants about what was meant by *notifications*. Moreover, for Q18, one participant answered both 4 and 5; in this case, a score of 4.5 was used as an exception. One participant added to the questions that he/she did not use the reminders but filled in his/her consumption whenever it suited him/her. Most statements were evaluated neutral and with diverse scores; see Figure 12d. For the friendly tone (Q14), the score was a bit higher. In general, the notifications (Q14–Q21) were evaluated with an average score of 4.4 (SD = 1.72).

The messages that participants received could be of one of three styles, depending on their personality. It turned out that only two styles were used: COM (30 participants) and AUT (6 participants). The COM-style has no personality requirements, which means that most participants did not have a personality that suited a specific communication style.

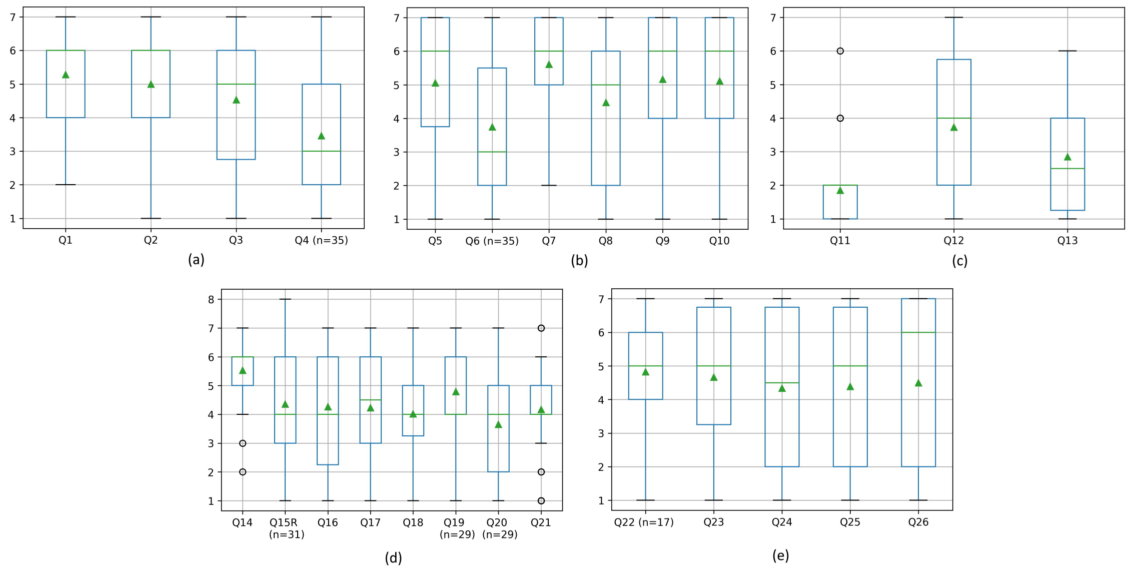

Figure 12. Boxplots of the PT evaluation questionnaire per topic for all active participants—scale 1-7 [16], Triangles: mean scores; Circles: outliers; 1 = strongly disagree (negative evaluation), 7 = strongly agree (positive evaluation); the questions can be found in Appendix A.2. (**a**) Experienced effectiveness tablet (n = 36). (**b**) User-friendliness of the tablet (n = 36). (**c**) Foodbox (n = 26). (**d**) Notifications (n = 30). (**e**) Gamification (n = 18).

Table 8 shows the results from the statistical tests for the means of the two styles. Participants with the COM style were slightly more positive about the messages they received, mainly about the friendliness of the messages. Although for individual questions no significant differences were found, a significant difference was found when combining all of the scores. The COM style was appreciated significantly better compared with the AUT style.

The evaluation of the gamification elements was only relevant for participants in the gamification tablet condition. Figure 12e shows that the scores about *gamification* (Q22–Q26) were very divided and, on average, slightly above neutral (mean = 4.5, SD = 2.25). For the mini-games, participants were the most positive about the learning aspect of the games.

Table 8. Average scores of the notifications for different communication styles—scale 1–7, 1 = strongly disagree, 7 = strongly agree. * $p < 0.05$.

	Both Styles	**AUT Style**	**COM Style**	***p*-Value**
Friendly tone	5.5 SD = 1.31 (n = 30)	5.0 SD = 2.00 (n = 5)	5.6 SD = 1.15 (n = 25)	0.389
Compelling (reversed)	4.3 SD = 1.84 (n = 31)	4.6 SD = 1.14 (n = 5)	4.3 SD = 1.9 (n = 26)	0.311
Relevant personal situation	4.3 SD = 1.91 (n = 30)	3.6 SD = 2.19 (n = 5)	4.4 SD = 1.87 (n = 25)	0.341
Motivational	4.2 SD = 1.81 (n = 30)	3.2 SD = 1.64 (n = 5)	4.4 SD = 1.8 (n = 25)	0.184
Interesting	4.0 SD = 1.62 (n = 30)	3.0 SD = 1.73 (n = 5)	4.2 SD = 1.55 (n = 25)	0.134
Believable and trustworthy	4.8 SD = 1.40 (n = 29)	4.6 SD = 1.95 (n = 5)	4.8 SD = 1.31 (n = 24)	0.439
About obstacles encountered	3.6 SD = 1.70 (n = 29)	2.8 SD = 1.64 (n = 5)	3.8 SD = 1.68 (n = 24)	0.218
Suitable for age and perceptions	4.17 SD = 1.62 (n = 30)	4.2 SD = 2.28 (n = 5)	4.2 SD = 1.52 (n = 25)	0.456
Average score topic	*4.4 SD = 1.72*	*3.9 SD = 1.86*	*4.5 SD = 1.68*	*0.000 **

The analysis of the evaluation of the persuasive technology questionnaire shows that experienced effectiveness, user-friendliness of the tablet app, notifications, and gamification were evaluated positively, but the foodbox was evaluated negatively. See Figure 12. The differences in evaluation for communication styles was analysed (Table 8). The COM style was used for 30 participants, and the AUT style was used for 6. The COM style was evaluated significantly better compared with the AUT style (p-value = 0.000).

4.4.2. Differences between Tablet Conditions for the Experience with the Technology

For the experienced effectiveness and the user-friendliness of the tablet, the differences between the tablet conditions were analysed. As the usage of the foodbox was the same in both tablet conditions, the differences for the foodbox questions were not studied. The same holds for the notification messages.

Table 9 shows the average scores for the questions about experienced effectiveness and user-friendliness of the tablet for both tablet conditions. It also shows the combined scores for both topics. Only for the question about the motivation to stick to the dietary advice was a significant difference found. The gamification condition gave a higher score to this question.

Table 9. Average scores for the follow-up questions about the tablet for different conditions—scale 1–7, 1 = strongly disagree, 7 = strongly agree. * $p < 0.05$.

	Both Conditions	**Normal Condition**	**Gamifi- Cation Condition**	***p*-Value**
Motivation stick dietary advice (Q1)	5.3 SD = 1.60 (n = 36)	4.8 SD = 1.63 (n = 17)	5.7 SD = 1.49 (n = 19)	0.037 *
Functionalities needed present (Q2)	5.0 SD = 1.71 (n = 36)	4.65 SD = 1.69 (n = 17)	5.3 SD = 1.70 (n = 19)	0.105
Meets expectations (Q3)	4.5 SD = 2.08 (n = 36)	4.2 SD = 1.89 (n = 17)	4.8 SD = 2.25 (n = 19)	0.175
Continue to use (Q4)	3.5 SD = 1.93 (n = 35)	3.2 SD = 1.65 (n = 16)	3.6 SD = 2.17 (n = 19)	0.468
Average score experienced effectiveness	*4.6 SD = 1.94*	*4.2 SD = 1.79*	*4.8 SD = 2.04*	*0.242*
Easy to use (Q5)	5.06 SD = 1.90 (n = 36)	4.8 SD = 1.60 (n = 17)	5.3 SD = 2.14 (n = 19)	0.081
No irregularities encountered (Q6)	3.7 SD = 2.09 (n = 35)	3.5 SD = 2.10 (n = 16)	3.9 SD = 2.12 (n = 19)	0.368
Quickly way around (Q7)	5.6 SD = 1.50 (n = 36)	5.3 SD = 1.61 (n = 17)	5.9 SD = 1.37 (n = 19)	0.103
Recommend to others (Q8)	4.5 SD = 2.13 (n = 36)	4.2 SD = 2.19 (n = 17)	4.7 SD = 2.11 (n = 19)	0.269
Comprehensible (Q9)	5.2 SD = 1.76 (n = 36)	5.0 SD = 1.62 (n = 17)	5.3 SD = 1.92 (n = 19)	0.195
Fun to use (Q10)	5.1 SD = 2.07 (n = 36)	4.6 SD = 2.15 (n = 17)	5.6 SD = 1.92 (n = 19)	0.066
Average score user-friendliness	*4.9 SD = 1.99*	*4.6 SD = 1.94*	*5.1 SD = 2.01*	*0.241*

Statistical tests were performed for the difference between conditions for experienced effectiveness and user-friendliness of the tablet app (Table 9). The results show that participants from the gamification condition rated experiencing motivation significantly better (p-value = 0.037).

4.4.3. Differences by Personal Characteristics for the Experience with the Technology

It is interesting to see if the evaluation of participants is influenced by personal characteristics. For this, several statistical tests were performed: the Mann–Whitney U test for the difference between men and women, and Spearman correlations for the other characteristics. For the questions about experienced effectiveness, a significant difference was found: men evaluated the experienced effectiveness higher compared with women (p-value = 0.016). Moreover, there was a significant positive correlation (p-value = 0.030) between the age of participants and the motivation to stick to the dietary advice provided by the tablet (Q1).

For the questions about the user-friendliness as well as for the foodbox and for the averages for each topic, no significant differences or correlations between personal characteristics and evaluation were found.

The evaluation of notifications seems to be more dependent on personal characteristics. Two questions, asking whether any obstacles were encountered (Q20) and whether the messages were suited to the participant's age and perception (Q21), were evaluated significantly higher by women (p-value = 0.028 and p-value = 0.008). Moreover, the average score for the messages was significantly higher for women (p-value = 0.001). Moreover, a significant negative correlation between the protein intake at baseline and the question about the friendly tone of messages (Q14) was found (p-value = 0.023). Finally, there were different significant negative correlations between education level and the evaluation of messages: whether messages were relevant (Q16—p-value = 0.000), were motivational (Q17—p-value = 0.007), were interesting to read (Q18—p-value = 0.002), asked about any obstacles faced (Q20—p-value = 0.001), or were suitable for age and perceptions (Q21—p-value = 0.037). Moreover, the combined evaluation was also significantly negatively correlated (p-value = 0.005).

Lastly, only one significant correlation was found for the gamification questions: a negative correlation between education level and motivation of the profile page (Q22—p-value = 0.047).

Several significant differences and correlations are found between personal characteristics and the evaluation of the experience with the persuasive technology:

- Men evaluated the experienced effectiveness significantly higher compared with women (p-value = 0.016)
- There was a significant positive correlation between the age of participants and the motivation to stick to the dietary advice provided by the tablet (p-value = 0.030).
- Different significant negative correlations were found for education level and questions about the messages: messages were relevant, were interesting to read, asked about any obstacles faced, or were suitable for age and perceptions. The overall evaluation of the messages was also significantly negatively correlated with education level.
- The motivation provided on the profile page has a significant negative correlation with education level (p-value = 0.047).

4.4.4. Interaction with the Tablet

It was studied how participants used the technology, which can give insights into which elements are used more often. Figure 13 shows how often, relative for each user, the three different input methods (fast input screen, meal composer, or via the foodbox) are used. As this is a relative graph, the extended trial due to COVID-19 was also taken into account. This shows that foodbox input is only used by a few participants. Moreover, some participants use fast input for most inputs, while other participants use the meal composer most of the time. Only in a few cases, this is a bit more balanced. It needs to be noted that this graph does not correct for repeated inputs. Repeated inputs can be due to two reasons: a participant changes an input later or a participant repeats its input (for example to check it or to redo it).

In general, the meal composer is used 52% of the inputs, while fast input is used for 47% of the inputs and the foodbox is only used for 1% of the inputs.

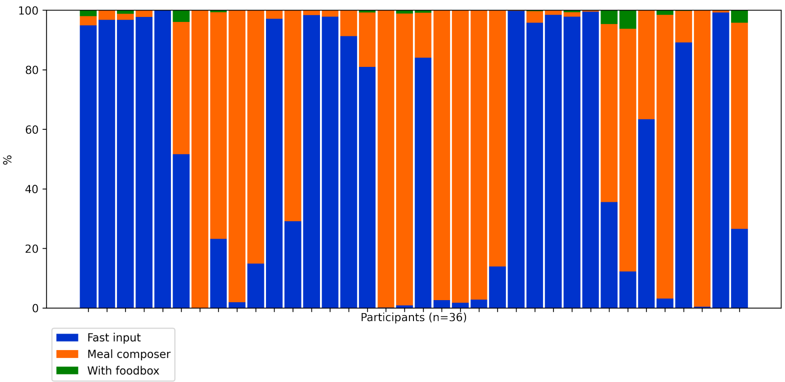

Figure 13. Relative number of registrations on the tablet with different methods of input [16].

Participants can react in two ways to a notification: they can either confirm, meaning they start to enter their intake, or postpone the notification, which enables a reminder. Of the total number of notifications sent, 87% was ignored, 12% was confirmed, and 1% was postponed.

To study how the gamification components were used, different types of interaction were studied. The data were not cleaned to take into account the extension of the trial due to COVID-19. Only two participants in the gamification condition fell into this category. It was however checked whether their data were different from the other participants.

It was studied how often the participants visited their profile page. Figure 14 shows that there were quite some participants that only visited it less than 20 times. There were also two extremes of participants who visited it more than 100 times, which was not caused by the extension of the trial due to COVID-19. On average, participants visited their profile page 41.5 times (SD = 41.61). It can be tested whether there was a Pearson correlation between the days with inputs and the number of profile visits for participants. This correlation (0.032) was not significant (p-value = 0.179). This means that the number of profile visits does not correlate with the number of days that the participants use their tablets.

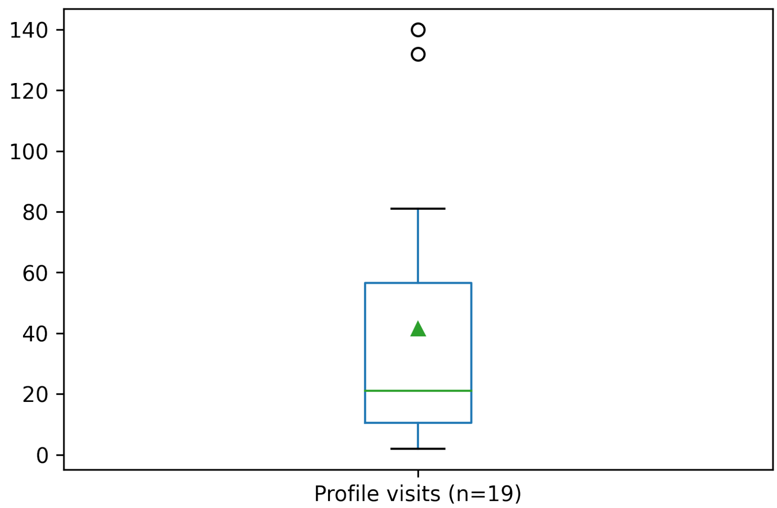

Figure 14. Boxplot of profile page visits for the gamification-condition participants [16].

When a participant's intake met the protein criteria for that eating moment, a mini-game is triggered. When asked to play a game, it is also possible to skip the game. Moreover, it was found that not all games that are started are also finished. This can be due to a bug, which closes the screens when the mini-game was offered to the participant, but it can also be because the participant put the tablet aside during the game. Figure 15 shows that only six participants (32%) of this condition played more than 50% of the mini-games that they were allowed to play. Some participants did not finish a single game.

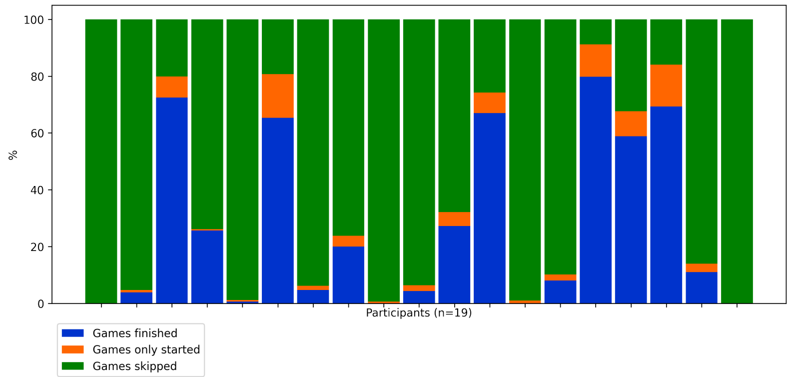

Figure 15. Relative number of mini-games played and skipped per gamification condition participant [16].

In Figure 16, it is analysed whether participants have a preference for a specific mini-game. To analyse this, the number of finished games was compared with the total number of proposed games (the games finished plus the games that were skipped). The games that were started but not finished were ignored because whether this is a bug or intentional cannot be detected. For some participants, this figure shows that they have a preference for some games over other games, but this pattern is different for each participant.

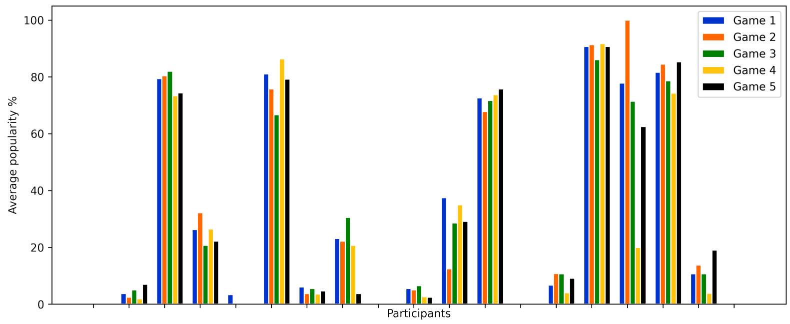

Figure 16. Popularity difference for mini-games.

Figure 17 shows that the popularity of all games, on average, is comparable. For this graph, all popularity scores are summed and divided by the number of gamification participants. To compare the averages a one-way ANOVA test was performed. This showed that no significant difference between the different averages could be found (p-value = 0.994).

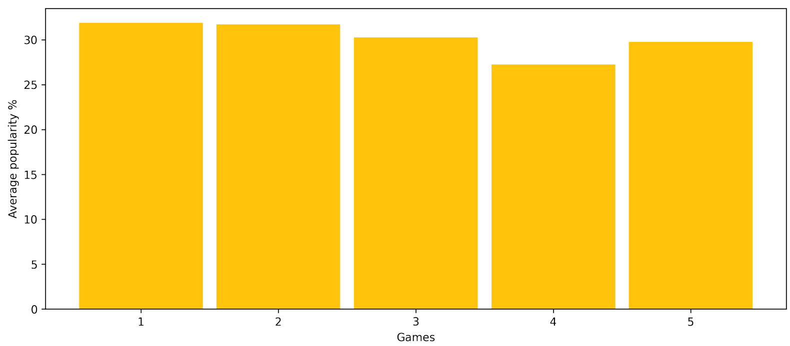

Figure 17. Average popularity of each mini-game.

As the studied aspects of interaction are either the same in both conditions or only present for the gamification condition, the differences between the two study groups for these aspects were not studied.

For the different methods of input (Figure 13), again, it was shown that the foodbox was not often used and that the other methods were almost equally often used. Overall, participants had one preferred way of input (fast input or meal composer) and only some used them both in a balanced manner. Most participants did not use notification but instead choose when to input their consumption when it suited them: 87% of the notifications were ignored, 12% were confirmed, and 1% were postponed. The number of profile page visits is presented in Figure 14. The average number of visits is 41.5 (SD = 41.61, minimum = 2, maximum = 140). The Pearson correlation between days of use and profile visits is not significant (p-value = 0.179). For the gamification, some participants liked and played the mini-games, while others did not or played only some games (Figure 15). For the popularity of the games, only slight differences can be seen for some participants (Figure 16), but overall (Figure 17), no significant differences were found (p-value = 0.994).

4.4.5. Differences by Personal Characteristics for Interaction

To see if the way that participants interact with the tablet is dependent on personal characteristics, several statistical tests were performed. First, differences between males and females were studied using Mann–Whitney U tests. For the percentage of inputs carried out using the fast input method, a significant difference was found (p-value = 0.031). Male participants used a fast input on average a bit more (mean = 55.8%, SD = 42.83%) compared with female participants (mean = 55.3%, SD = 43.52%). Sex differences were not found for the number of profile visits or the percentage of mini-games finished.

Next, tests for significant Pearson correlations were performed for age, protein intake at baseline for the percentage of fast input, number of profile visits, and percentage of mini-games finished. A one-way ANOVA was used to compare the different education levels for these measures. No significant results were found.

Different statistical tests for the influence of personal characteristics on the interaction with the tablet were performed. The only significant result is that fast input is used significantly more by male users (p-value = 0.031).

4.5. Gamification and Protein Knowledge (RT5)

In the dietary advice evaluation questionnaire, one question addressed whether the participants feel that they have a good understanding of the amount of protein in different types of products (see Appendix A.1). This question gives insight into whether participants in the PT group gained more insight into protein products compared with the no-PT group. In contrast to our expectations, the mean score of the PT group (n = 36, mean = 3.9, SD = 0.93) was lower compared with the no-PT group (n = 37, mean = 4.1, SD = 0.74); this difference is significant (p-value = 0.030). The normal condition (n = 17, mean = 3.9, SD = 0.83) was compared with the gamification condition (n = 19, mean = 3.8, SD = 1.03), but these results turned out not significant (p-value = 0.391).

Another way to study whether there was a learning effect for the mini-games is to look at the scores for the mini-games. To do so, the mean per five mini-games of one type was calculated. For participants who played at least one mini-game more than five times, a graph was created, which is shown in Figure 18. Overall, there was no trend towards higher scores when playing more mini-games. However, when looking at participants who play the mini-games quite often, it seems that their scores often stabilised.

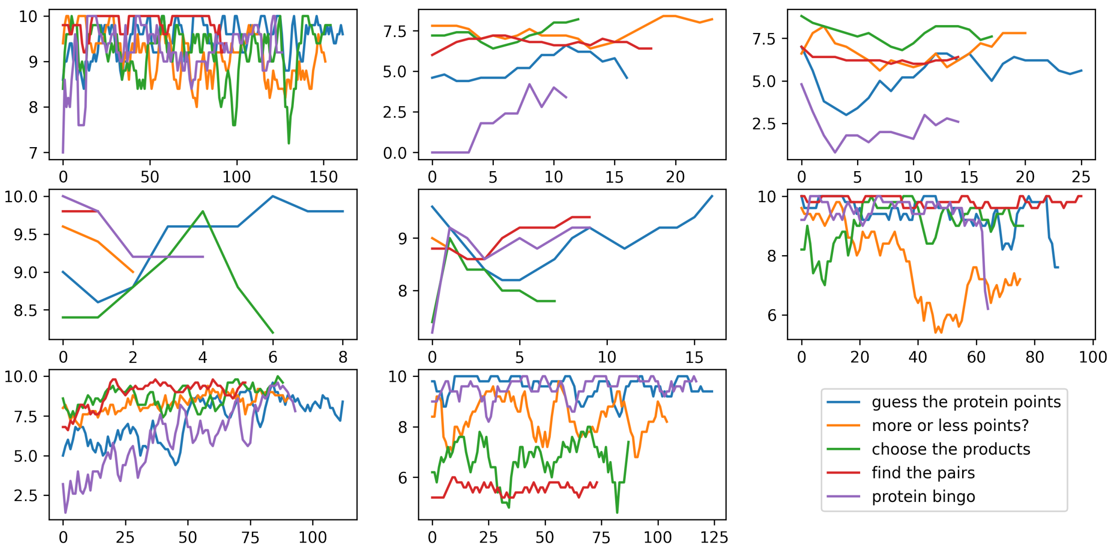

Figure 18. Average scores per five games played for gamification participants with more than five games played—each graph represents the scores of one player. The X-axis shows the number of games played, while the y-axis shows the scores (0–10).

Next, the scores in the first half of the games and the scores in the last half of the games were analysed. For this, all games were combined per participant. With a *t*-test, it was tested whether the differences between those means were significant. All participants from the gamification condition who played at least two games were included (14 participants). From these tests, it becomes clear that five participants made a significant change in their average points. For four of these participants, this is an improvement; for one, it is not, but this participant played a very low number of games compared with the other participants. It needs to be noted that the scores are not a complete representation of the knowledge of participants. It is unclear whether participants used their protein information booklet when playing the games or whether this change is because they understood the games better.

Whether there is a Pearson correlation between the number of games played and the average score of participants was tested. The coefficient was 0.689, with a significant p-value of 0.006. Figure 19 shows the scatterplot of this analysis. Again, although the correlation was significant, it cannot be concluded based on this that the games contributed to more protein knowledge, as other explanations for this finding are also possible, for the reasons explained above.

Participants in the tablet condition rated their protein knowledge significantly lower compared with participants without a tablet (p-value 0.030). No significant result was found between the two sub-study conditions concerning their self-reported protein knowledge (p-value = 0.391). When looking at the scores that participants earn in the mini-games, it is hard to observe any trends, but when a high number of games were played, the scores seemed to stabilise (Figure 18). Comparing the mean scores of the first and second halves of the games played per participant, significant differences can be found for 5 out of 13 participants. Four of those participants improved their scores in the second half. A significant Pearson correlation was found between games played and average game score (Figure 19, coefficient 0.674, p-value = 0.012).

Differences by Personal Characteristics for Protein Knowledge

Differences in the rating of the question about understanding protein points caused by personal characteristics were tested. With the Mann–Whitney U test, it was tested if there were differences caused by sex, but this was not significant (p-value = 0.472). Next, the Spearman correlation was calculated between age, protein intake at baseline, and education level, and rating for their understanding of the questions. Again, these correlations were not significant.

Moreover, to see if personal characteristics influence the average mini-game score, a Mann–Whitney U test was used for differences between male and female participants, a Pearson correlation was used for the age and protein intake at baseline, and a one-way ANOVA was used for education level. All tests were not significant, so personal characteristics do not have a significant effect on the average game scores of participants.

No significant results were found in the statistical tests for the association between personal characteristics, and reported protein knowledge and average game scores.

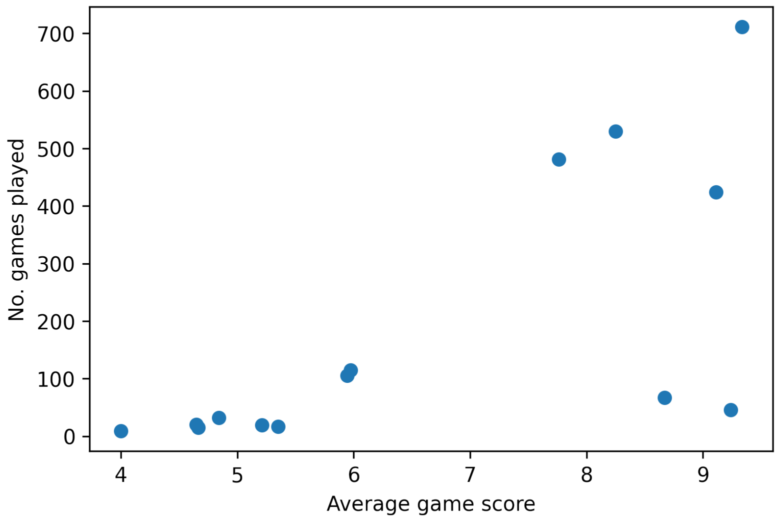

Figure 19. Scatterplot average game score and number of games played (n = 14).

5. Discussion

This article gives an overview of the results of the persuasive technology sub-study of the PROMISS trial. The aim was to learn more about how older adults receive a diet-tracking system as part of a diet program, in our specific case, the PROMISS trial. Therefore, the adherence of participants, the change in protein intake, the experience of the diet and the persuasive technology, and the impact of gamification on protein knowledge were studied in this article. If relevant, it was studied if differences between the normal and gamified conditions exist, and the possible effects of personal characteristics were studied as well.

The participants of the study were 83% adherent to using the technology during the trial period, with an average of 133 days. No differences in the adherence of participants were found between the two conditions in the study or based on personal characteristics.

Although the adherence was high, using the technology did not significantly change their protein intake compared with participants in the intervention groups without a tablet. No effect was found for the number of days that the tablet was used on the protein intake. During the study, it was noted that the tablet sometimes gave higher estimates of protein intake due to round differences with the dietitians. However, the results show that this did not have a significant negative effect on the protein intake of participants. No differences between the two conditions or caused by personal characteristics for the protein intake were found.

Participants with or without a tablet did not significantly evaluate the differently diet on a whole. The normal condition in the PT group gave the diet a significantly higher rating compared with the gamification condition. Moreover, it was found that women in the PT group have a significantly higher intention to continue to follow the dietary advice, while this difference was not found in the no-PT group.

Participants appreciated the tablet and the gamification elements, but the foodbox was negatively evaluated. For the experienced effectiveness and the user-friendliness of the tablet, differences between the two conditions were studied. Only one significant difference was found: the gamified condition reported a higher motivation to stick to the dietary advice by the tablet, compared with the normal condition. Furthermore, for all questions, it was tested whether personal characteristics influence the evaluation of the persuasive technology. Some effects of sex and age were found, but the most significant negative correlations were found for education level and the questions about the notification messages.

When studying the interaction with the app, the data showed that participants often ignore the notification message, and instead interacted with the app on their own initiative. Most participants either used the fast input or the meal composer most of the time, some participants used these methods of input in a more balanced manner. For the mini-games, a similar pattern was seen: some participants skipped all mini-games, other participants played almost all mini-games, and some participants played some mini-games and skipped some others. A significant difference was found in the percentage of inputs using the fast input between men and women (higher for men).

Finally, the learning effect of the mini-games was studied. Based on a question about the understanding of protein, this cannot be derived. From the scores of the games, it seemed that there is a learning effect for some participants. Moreover, the scores were correlated with the number of games played. However, higher scores can also be due to other aspects, such as using the information booklet. Altogether, the app was considered a feasible way to teach about protein intake. No effects of personal characteristics on the rating of the understanding or the scores in the mini-games were found.

There were some limitations to the study. The sample size of the sub-study was 36, which was limited but turned out to be sufficient for the presented analysis. When the different conditions were studied, the sample size was small (17–19). Therefore, it is unclear whether the analysis would hold with a larger sample. Moreover, it might be that the participants had some different characteristic(s) that made them volunteer to participate in the study, which could affect the results. This could not be derived from the available data.

During the study, it was sometimes necessary to fix some problems with the application. The inclusion for the gamification condition started later in the study after some bugs were fixed. The evaluation and participation of participants may have been influenced by the number of bugs they encountered during their trial period. However, no major differences were found in the average scores of the evaluation questionnaire.

For the analysis of the adherence, the 12 drop-out participants were not included. Including all of those excluded participants changed the overall adherence to 104 days and 76%. However, four participants did not have any interaction with the system. When excluding those, the average adherence was 114 days and 83%. Although the average adherence expressed in active days would have lowered, it was still more than half of the trial period (6 months) and the ratio of active days was the same as the adherence ratio of the PT participants. The results on adherence can also be influenced by the way the duration of the tablet trial was counted for each participant. The first and last days of use were used to calculate the trial duration. It could be that participants put the tablet aside for some time and started using it again right before an appointment. However, this was not found in our data set. Some participants did not use the tablet for a longer period during their trial, sometimes due to issues or holidays and sometimes without a given reason. However, they started using it again for a considerable time before ending their use. For only one participant, it was found that there were around two weeks of inactivity, then five days of activity, and then the participant stopped using the tablet.

This research does not look at the quality of the diet tracking data that participants entered into the app. It was not studied whether the data from the tablet could be used by dietitians within the trial. For example, dietitians perform phone recalls during the trial to discuss the protein intake of participants. It would be interesting to study if the data from the tablet could replace such recalls. However, another research setup would be needed to address this.

6. Conclusions

Based on the results, we conclude that using a tablet with a diet-tracking application, with persuasive communication, dedicated to the trial is a feasible way to engage participants. The target users, older adults, support the use of such persuasive technology, as shown by their evaluation of the system as well as their adherence to using it. The experienced effectiveness and the user-friendliness of the tablet were evaluated positively. The adherence was 83%, with an average of 133 days. No major differences in the evaluation of the normal condition and the gamification condition were found for the adherence or evaluation of the system or diet. Moreover, it is shown that the gamification elements were not liked/used by all participants and seems to be a personal preference. As no personal characteristics were found that might be correlated with this, it cannot be predicted which participants will like it. However, it does not seem to be a drawback for participants who do not like the gamification. For all different measures, associations with personal characteristics were studied. Although a few significant correlations were found, no patterns were found for specific groups who experience or evaluate the technology differently. Multiple significant correlations between the evaluation of the notification messages and education level were found, so this can be something to investigate in further research. Overall, it can be concluded that using diet-tracking applications in trials is well received by older adult participants and is thus a feasible way to track a diet. In future research, it would be interesting to study how such persuasive technology can be further integrated into a diet study, for example, to support the work of dietitians.

Author Contributions: Conceptualisation, L.M.v.d.L., M.C.A.K., M.V., H.A.H.W. and I.R.; methodology (overall study), M.V., H.A.H.W. and I.R.; methodology (sub-study), L.M.v.d.L. and M.C.A.K.; software, L.M.v.d.L.; formal analysis, L.M.v.d.L.; data curation, M.V., H.A.H.W. and I.R.; writing—original draft preparation, L.M.v.d.L.; writing—review and editing, M.C.A.K., M.V., H.A.H.W. and I.R.; visualisation, L.M.v.d.L.; supervision, M.C.A.K.; funding acquisition, M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by European Union Horizon2020 PROMISS Project 'Prevention of Malnutrition in Senior Subjects' grant number 678732.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of VU medisch centrum (protocol code 2018.399 and 5 June 2018).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data of this study is not publically available.

Acknowledgments: We thank the PROMISS trial dietitians: Merel Vrijmoeth, Mariska Bout, and Marjon Veeke.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Questionnaires

Appendix A.1. Dietary Advice Evaluation

This questionnaire consists of seven questions about the following measures:

- Rating dietary advice (scale 1–10);
- Usefulness dietary advice;
- Extent to which dietary advice was followed;
- Ease of sticking to the dietary advice;
- Intent to continue following dietary advice;
- Understanding of protein in products;
- Ease of finding products with similar protein amounts.

If not indicated otherwise, the measure was evaluated on a five-point Likert scale. To analyse the Likert-scale questions, the Mann–Whitney U test was used. For the rating, a *t*-test was used.

Appendix A.2. Persuasive Technology Questionnaire

Different topics were discussed in this questionnaire. The following topics and questions were asked (translated from Dutch):

- Effectiveness of the tablet application
  1. The tablet helps/motivates me to stick to my dietary advice.
  2. The tablet has all of the functionalities that I need.
  3. The tablet meets my expectation.
  4. If I continue to follow dietary advice, I would like to keep on using the tablet.
- User-friendliness of the tablet
  5. The tablet was easy to use.
  6. I did not encounter any irregularities when using the tablet.
  7. I quickly knew my way around the tablet.
  8. I would recommend the tablet to someone else.
  9. The tablet was comprehensible.
  10. It was fun to use the tablet to follow my dietary advice.
- Foodbox
  11. If I stick to my dietary advice, I would like to continue using the foodbox.
  12. The foodbox is easy to use.
  13. It was fun to use the foodbox when following my dietary advice.
- Notifications
  14. The messages have a friendly tone.
  15. The messages are compelling.
  16. The messages are relevant for my personal situation.
  17. The messages are motivational.
  18. The messages are interesting to read.
  19. The messages are believable and trustworthy.
  20. The messages are about obstacles that I encounter.
  21. The messages suit my age and perceptions.
- Gamification
  22. I was motivated by my profile page.
  23. I was motivated by achieving an achievement.
  24. I found playing games motivating.
  25. I liked playing the games.
  26. I found the games informative.

All questions for this questionnaire could be rated on a 7-point Likert scale (1 = strongly disagree, 4 = neutral, 7 = strongly agree). Question 15 is a negatively framed question, and the scores are therefore reversed in the analyses. A Mann–Whitney U test was used to compare the mean scores of the different groups, if applicable.

References

1. Carter, M.C.; Burley, V.J.; Nykjaer, C.; Cade, J.E. Adherence to a smartphone application for weight loss compared to website and paper diary: Pilot randomized controlled trial. *J. Med. Internet Res.* **2013**, *15*, e32. [CrossRef] [PubMed]
2. Coughlin, S.S.; Whitehead, M.; Sheats, J.Q.; Mastromonico, J.; Hardy, D.; Smith, S.A. Smartphone applications for promoting healthy diet and nutrition: A literature review. *Jacobs J. Food Nutr.* **2015**, *2*, 021. [PubMed]
3. Cordeiro, F.; Epstein, D.A.; Thomaz, E.; Bales, E.; Jagannathan, A.K.; Abowd, G.D.; Fogarty, J. Barriers and negative nudges: Exploring challenges in food journaling. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 1159–1162.
4. Johnson, D.; Deterding, S.; Kuhn, K.A.; Staneva, A.; Stoyanov, S.; Hides, L. Gamification for health and wellbeing: A systematic review of the literature. *Internet Interv.* **2016**, *6*, 89–106. [CrossRef] [PubMed]
5. Seaborn, K.; Fels, D.I. Gamification in theory and action: A survey. *Int. J.-Hum.-Comput. Stud.* **2015**, *74*, 14–31. [CrossRef]

6. Koivisto, J.; Hamari, J. Demographic differences in perceived benefits from gamification. *Comput. Hum. Behav.* **2014**, *35*, 179–188. [CrossRef]
7. Stanmore, E.K.; Mavroeidi, A.; de Jong, L.D.; Skelton, D.A.; Sutton, C.J.; Benedetto, V.; Munford, L.A.; Meekes, W.; Bell, V.; Todd, C. The effectiveness and cost-effectiveness of strength and balance Exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: A multi-centre, cluster randomised controlled trial. *BMC Med.* **2019**, *17*, 49. [CrossRef] [PubMed]
8. Reinders, I.; Wijnhoven, H.A.H.; Jyväkorpi, S.K.; Suominen, M.H.; Niskanen, R.; Bosmans, J.E.; Brouwer, I.A.; Fluitman, K.S.; Klein, M.C.A.; Kuijper, L.D.; et al. Effectiveness and cost-effectiveness of personalised dietary advice aiming at increasing protein intake on physical functioning in community-dwelling older adults with lower habitual protein intake. *BMJ Open* **2020**, *10*. Available online: https://bmjopen.bmj.com/content/10/11/e040637 (accessed on 31 November 2021). [CrossRef] [PubMed]
9. Reinders, I.; Visser, M.; Jyväkorpi, S.K.; Niskanen, R.T.; Bosmans, J.E.; Jornada Ben, Â.; Brouwer, I.A.; Kuijper, L.D.; Olthof, M.R.; Pitkälä, K.H.; et al. The cost effectiveness of personalized dietary advice to increase protein intake in older adults with lower habitual protein intake: a randomized controlled trial. *Eur. J. Nutr.* **2021**, *61*, 1–16.
10. van der Lubbe, L.M.; Klein, M.C.A. Designing a System with Persuasive Communication to Improve Diet Compliance for Elderly Users. In Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth'19, Trento, Italy, 20–23 May 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 234–241.
11. van der Lubbe, L.M.; Klein, M.C.A. Integrating Gamification into a System to Improve Diet Compliance for Elderly Users. In Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good, GoodTechs '20, Antwerp, Belgium, 14–16 September 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 150–155.
12. Van Genugten, L.; Dusseldorp, E.; Webb, T.L.; Van Empelen, P. Which combinations of techniques and modes of delivery in internet-based interventions effectively change health behavior? A meta-analysis. *J. Med. Internet Res.* **2016**, *18*, e155. [CrossRef] [PubMed]
13. Webb, T.; Joseph, J.; Yardley, L.; Michie, S. Using the internet to promote health behavior change: A systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J. Med. Internet Res.* **2010**, *12*, e4. [CrossRef] [PubMed]
14. Johnson, R.; Kent, S. Designing universal access: web-applications for the elderly and disabled. *Cogn. Technol. Work.* **2007**, *9*, 209–218. [CrossRef]
15. Al-Razgan, M.S.; Al-Khalifa, H.S.; Al-Shahrani, M.D.; AlAjmi, H.H. Touch-based mobile phone interface guidelines and design recommendations for elderly people: A survey of the literature. In *International Conference on Neural Information Processing*; Springer: Berlin/Heidelberg, Germany, 2012.
16. Lubbe, L.M.; Klein, M.C.A.; Visser, M.; Wijnhoven, H.A.H.; Reinders, I. Experiences with Using Persuasive Technology in a Diet Trial for Older Adults. In Proceedings of the 14th PErvasive Technologies Related To Assistive Environments Conference, Corfu, Greece, 29 June–2 July 2021; pp. 244–251.