



Acceptance of a Virtual Coach as Guided Intervention for Smoking Cessation
A Mixed-Methods Analysis

Jaap Dechering

Supervisors: Nele Albers, Dr. Willem-Paul Brinkman
Intelligent Systems Department, Delft University of Technology, Delft, The
Netherlands

22-6-2022

A Dissertation Submitted to EEMCS faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering

Abstract

As guided interventions are seen as more effective compared to unguided interventions and because the use of eHealth applications is rising over the years, the need to study the effects of a virtual coach as a substitute for guided interventions was created. A virtual coach does not cost as much time and money as a real coach and would thus be beneficial in combination with a virtual smoking cessation application. Therefore, the reasons for preferring to use a questionnaire or virtual coach were studied in this research. Through three questionnaires, 500 participants have indicated their experience with a virtual coach Sam, helping them to prepare to quit smoking. Based on questions about their preference for a virtual coach or a questionnaire, seven themes have been identified. Most of the themes were connected to making the activities proposed by the virtual coach more appealing and therefore the intervention more effective. All in all, the most important aspects of accepting a virtual coach for smoking cessation had to do with the proposed content by the virtual coach.

1 Introduction

In 2014, there were over 500 English based smoking cessation apps in the Apple Store and Google Play Store combined, with over 20 million downloads [1]. These eHealth applications are especially preferred by young people as almost all of them are familiar with the digital domain [2]. Therefore, over the past years, researchers have studied the effects of eHealth applications for smoking cessation in numerous countries [1]–[5].

Additionally, research has shown that guided interventions are far more effective for symptom improvement compared to unguided interventions [6]. As it is not yet fully understood how guided interventions are more effective [7] and because guided interventions need people which makes them less scalable and more expensive [6], the relevance of studying the effects of virtual agents as guided interventions was emphasized.

Although many studies have focused on virtual agents in combination with treatments [8]–[10], not many studies have looked at the effect of combining a virtual chat agent with the cognitive-behavioral therapy needed for smoking cessation [6]. Therefore, the Interactive Intelligence group of Delft University of Technology in collaboration with Leiden University has developed a conversational agent called Sam which aids people with smoking cessation [11].

In 2021, a longitude study was conducted in which over 500 participants communicated with Sam over 5 sessions. During these sessions, Sam proposed activities that should help to prepare to quit smoking. Before this study, the participants completed two questionnaires about their characteristics and their will to stop smoking. After the experience with Sam, they completed another questionnaire about their encounter with Sam. Using the data gathered during this longitude study, the following question was answered:

What are the reasons to prefer using a questionnaire or a virtual coach?

The questionnaire can be seen as an alternative to an unguided intervention while the virtual coach can be seen as the guided intervention. Thus, this question tries to analyse if the use of a virtual coach as a guided intervention can aid smoking cessation.

Using a thematic analysis, seven themes were identified about the preference for a questionnaire or a virtual coach. Four of these themes identified a positive aspect of the virtual agent: personal connection, engagement, helpful content and natural flow. The other three were connected to the positive aspects of a questionnaire and therefore the negative aspects of the virtual agent: inconsiderate, cumbersome and uneasiness.

In this report, these seven themes were studied in detail to understand why people prefer to use a questionnaire or a virtual coach and recommendations are given to improve the virtual coach.

2 Methodology

This research was solely based on data gathered during a longitude study on the virtual coach Sam. Research participants were selected using the platform Prolific, on which people can answer questionnaires in return for a monetary contribution. The collected data was primarily analysed using a thematic analysis. In order to increase the reliability of the identified themes, researcher and method triangulation were used.

2.1 Virtual coach

Sam is a virtual agent that simulates a coach helping the participant with smoking cessation (for implementation see [12]). The participants interacted with Sam in five different sessions over a period of a few weeks. During these sessions Sam proposed activities to help prepare for smoking cessation and used a persuasion type to do so [11].

2.2 Data

The dataset gathered during the longitude study consisted of 500 participants who completed a total of three questionnaires. Two questionnaires were administered before the engagement with the virtual coach, ending with one questionnaire afterwards (for dataset see [13]). The participants were identified by a random string of alphanumeric characters but for the sake of conciseness a P followed by a number between 1 and 500 was used to identify the participants in this report (for dataset used in this study see [14]).

In the questionnaire that was conducted after the interaction with Sam, research participants had to answer the following question with a score between -5 (definitely prefer questionnaire) and 5 (definitely prefer conversation):

Would you rather have followed this preparation course with a conversation with Sam or a questionnaire about your thoughts about your previous and new activity in each session?

This question and the associated ‘Why do you think so?’ were the central questions for this study. In general, most of the research respondents (n = 340, 68%) preferred to interact with a virtual agent in comparison to a questionnaire. Moreover, the rating having a median of 3 also indicated that the general consensus was remarkably positive. Nonetheless, by coding the qualitative data, somewhat negative impressions were identified as well.

Of all the participants that were selected during the pre-screening phase (n = 1221), only the ones that passed the attention checks and indicated to want to stop smoking were selected for the study. These participants (n = 500) were aged between 18 and 74 (median = 33) with 244 (48.9%) identifying as male, 247 (49.5%) identifying as female and 8 (1.6%) identifying as other, a combination of all the other answers. At the time of the study, 179 participants (35.9%) had at most a high school diploma, 239 of them (48.0%) had a community college or undergraduate degree and 80 (16.1%) had more than an undergraduate degree.

2.3 Preprocessing

One of the pre-study questionnaires scored the participants on their ability to see themselves as a quitter based on multiple questions to increase the reliability of this quitting identity. Before the analysis of the data started, these multiple items had to be combined into one measure and the Cronbach’s alpha (α) was computed to test the internal consistency of the measure [15]. As there is no standard for the Cronbach’s alpha [15] [16], this study deemed a value over 0.7 as sufficient as 0.7 indicates that most variance is explained by the measures and can thus be seen as reliable, following the article by Keith S. Taber [17].

The quitting self identity was constructed by combining 3 different questions, having a Cronbach’s alpha of 0.76. The motivation for activities and the easiness of the activities were both combined using 2 questions, having a Cronbach’s alpha of 0.90 and 0.80 respectively. Therefore, these characteristics were deemed reliable enough for further analysis.

Additionally, ten items were used to calculate the Big-5 personality based on the TIPI questionnaire [18]. These personality traits were extraversion, agreeableness, conscientiousness, emotional stability and openness to experience. The constructed variables can all be found in the dataset connected to this study [14].

2.4 Thematic analysis

The primary method used during this research was a thematic analysis on the open answer question for using a virtual agent or a questionnaire (see 2.2). A thematic analysis is a qualitative research method that can be used to find overarching themes and patterns in open answer data [19]. The thematic analysis was done following these four steps, based on the study by Braun and Clarke [19]:

1. Getting familiar with the data
2. Coding the data
3. Identifying themes
4. Reviewing themes and method triangulation

Additionally, as the themes identified by a thematic analysis are grounded in the bias of the researcher, it is important to take extra steps to increase validity. Therefore, triangulation, more than one way of collecting or analysing data to decrease bias [20], was used at multiple occasions during the research.

2.4.1 Step 1: Getting familiar with the data

In order to start a thematic analysis, it was important to be familiar with the data [19]. During the first analysis of the free text responses, it became apparent that some answers were difficult to interpret as they referred to earlier questions not included in this research, were written in another language or simply not understandable.

Therefore, the following methods were used to deal with these unreliable answers. The answer which referred to an earlier question were identified and the ones in another language translated, both were added to the free text response between brackets to indicate this addition. The incomprehensible answers were marked as not applicable (see [14]).

2.4.2 Step 2: Coding the data

After the initial analysis had been carried out, the data could be coded. The data was coded in a mostly inductive matter, meaning that the codes were based on the free text responses instead of on theories found in literature, as well as using both semantic themes, literal codes, and latent themes, interpretive codes [19].

During the coding of the qualitative data, 25 different codes were identified of which the majority was positive (for dataset see [14]). For the purpose of training the second coder, the codes were separated into four main categories (see Figure 1): codes about the virtual agent; codes about the questionnaire; codes about personal topics and other codes. The codes had also been divided into codes about experience, codes about aim and codes about feeling.

Investigator triangulation

Finding codes in the free-text responses is a subjective process influenced by the biases of the researcher. Therefore, investigator triangulation was used to increase the reliability of the codes. Investigator triangulation is done by comparing the results and conclusions of two or more researchers to decrease bias [21].

Thus, after the codes were created by the first coder, an associated Computer Scientist was trained with this coding scheme. This second coder coded the whole dataset as well, independently from the first coder. Both coded datasets were then compared using Cohen’s Kappa, indicating the amount of agreement. A Cohen’s Kappa below 0.4 indicates a fair agreement while a Cohen’s Kappa above 0.8 shows an almost perfect agreement [22]. The coded datasets of both researchers had a Cohen’s Kappa of 0.6, a substantial agreement, which was deemed appropriate for this coding scheme [22]. Disagreement between the two coders was subsequently resolved through discussion.

2.4.3 Step 3: Identifying themes

After the double coding had been finished, the found codes were used to identify overarching themes (for dataset see [14]). These themes should capture concepts related to the acceptance of a virtual coach for smoking cessation.

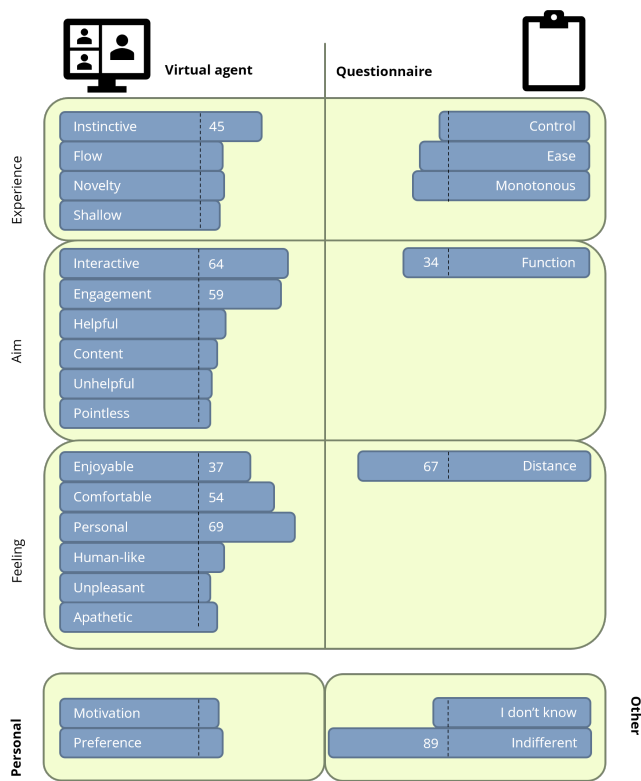


Figure 1: An overview of the coding scheme used during the thematic analysis. The width of the bars represents the frequencies of the code, with the most occurring codes being shown.

The thematic analysis was conducted within a constructionist ontology which assumes that experiences and social phenomena are constructed by people and cannot be seen as an inherited fact [20]. The context and the structural conditions of the research participants will be taken into account within this ontological position which allows the research to connect the context of the participants to their answers [19]. Therefore, the themes identified in the thematic analysis were seen as constructs which differed based on the context of the participants.

2.4.4 Step 4: Reviewing themes and method triangulation

The identification of themes was a subjective process as well, as it was influenced by the ontological position and the biases of the researcher. Therefore, to decrease the biases in themes, method triangulation was used.

Method triangulation is done by making use of multiple methods to cross-validate conclusions [21]. In this study, the themes from the thematic analysis were combined with statistical analysis of participant characteristics and relevant literature to increase the reliability of the identified themes.

Quantitative analysis

The combined data from the three questionnaires was used for the statistical analysis. Correlation tests between the preference rating of a questionnaire compared to a virtual agent and other user characteristics were used to identify interre-

relationships between these variables as well as motivations of the participants to use the virtual agent. However, by conducting numerous correlation tests, the chance of getting a random significant correlation gets higher. Thus, in order to prevent this p-hacking (significance hacking), the statistical relationships of interest were defined before the correlations were conducted.

As Wilson and Peterson showed that digital access is not only based on having the tools but also on having the knowledge [23], it is interesting to look at participant characteristics that might influence this knowledge. Studies have shown age is an important indication for a decrease in digital skills [24], [25] and therefore this might be an interesting characteristic. Educational level might also be interesting as studies have shown that the education level is associated with a lower digital knowledge as well [24], [25]. Gender was another interesting characteristic as research has shown that men prefer to be autonomous and in control, which is something that might not be feasible when interacting with a virtual agent [2].

Additionally, the extend to which the participants perceived the activities to be motivational and easy was also interesting, as studies indicated that performance expectancy, the perceived benefit of the virtual agent, and effort expectancy, the easiness, are an important aspect of the acceptance of technology [3]. As the behaviour changing activities proposed by Sam were mostly based on becoming physical active, the physical activity identity as well as the transtheoretical model (TTM)-stage for becoming physically active were interesting as well. Both these measures might indicate a reluctance for the proposed activities and therefore for using the virtual agent.

Furthermore, the relation between some psychometric tests and the rating might be interesting to identify personal characteristics making the acceptance of a virtual agent more appealing. Studies have shown that an internal positive attitude towards digital tools increases the extend to which digital tools are used [26]. Therefore, the personality measures openness to experience and extraversion were used, as these might indicate how easy a participant converses with the virtual agent and how susceptible to technology the participant is. The quitting self efficacy and the nonsmoker self identity indicate the extend to which the participant is able to see themselves as a quitter and as a non-smoker. Thus, these measures might also influence the acceptance of the virtual coach.

As the rating could only be analysed on an interval level of measurement, the correlations had to be computed using a Spearman's rho analysis [20]. The effect sizes of these Spearman's correlations were interpreted based on the guidelines by Schober and Schwarte [27]. The variable gender identity could only be analysed on a nominal level of measurement meaning that the correlation between gender and the rating could not be measured using Spearman's rho. Thus, this correlations was measured using a Chi-square test [20].

Additionally, a Categorical Principal Component analysis (CATPCA) [28] was used to visually represent the statistical relationships between the identified themes and other characteristics of the participants. These relationships gave an overview of possible interesting connections with other variables as well as interesting correlations between the themes.

As the analysis level of all analysed variables were nominal, the CATPCA was constructed by making use of reciprocal averaging. This calculates the principal component, the average, of each answer category and places them close to the answer categories that were scored by the participants who answered similarly on other questions [28]. It will also locate the more frequent categories closer to the origin [28]. In this way, clusters of similar scored answer categories could be identified as well as general directions of relationships between variables.

However, it should be noted that a CATPCA is sensitive to low frequency categories as these would be located far away from the origin, skewing the plot. Therefore, a bottom threshold of 6 responses per answer category was used during this research. This meant that the data had to be re-coded, constructing new categories and dealing with low frequencies (see dataset [14]). The variables were re-coded in the following ways:

The variable age was re-coded into age groups of 6 years and the outlier 74 was excluded from the CATPCA. The motivation for activities and the easiness of the activities had to be rounded to the closest integer to use in the CATPCA. The quitting self identity had to be rounded as well and the outlier category 1 had to be excluded as only 6 participants scored this category. This is almost identical to the re-coding process of openness to experience as this measure was rounded as well and the category 2 was left out of the CATPCA as it was only scored 4 times.

The variables of the themes were constructed by calculating the percentage of codes associated with the theme that a participant scored on. These were distributed from 'a lot', almost all codes were mentioned associated with the theme, to 'none', no codes were mentioned associated with the theme. Thus, the amount of categories per variable was depended on the amount of codes per theme.

By constructing the themes in this way, they could be analysed on an ordinal level of measurement. This allowed for a Spearman's rho analysis between the themes and participant characteristics, which combined with a contingency table could expose statistical relationships. The tables used in the sections on themes are mostly constructed using this method.

Literature study

The themes from the thematic analysis as well as the findings from the quantitative analysis were combined with relevant literature to cross-validate the findings. The topics of the literature covered the acceptability of and access to technology as well as the preparedness for behaviour change. This resulted in reliable themes as they were connected with established theories.

3 Results

Seven themes were identified using a thematic analysis (see Table 1 for the mentioning frequency). Three of them were more positive toward the usage of a questionnaire and thus negative towards the virtual agent: inconsiderate, cumbersome and uneasiness, having a moderate, weak and weak negative correlation with the preference rating respectively. The

other four were positive about the usage of a virtual agent: personal connection, engagement, helpful content and natural flow, all having a weak positive correlation (Table 1).

Theme	Frequency	%	ρ
Personal connection	125	25.0	0.36**
Engagement	111	22.2	0.37**
Helpful content	86	17.2	0.23**
Natural flow	74	14.8	0.31**
Inconsiderate	47	9.4	-0.42**
Cumbersome	30	6.0	-0.33**
Uneasiness	19	3.8	-0.24**

% of all the participants mentioned this theme

** Spearman's rho correlation between theme and rating is significant at the 0.01 level (2-tailed), all having a significance value of 0.00

Table 1: The frequency, percentages and correlation effect sizes of the themes mentioned

Additionally, as mentioned in 2.4.4, correlations between the preference rating and characteristics of the participants were tested. Table 2 shows that there is a weak significant correlation between age, motivated activities, easy activities, openness to experiences and the rating. Quitting self identity was significant at the 0.05 level but the effect size indicates that there is almost no correlation.

Preference rating	ρ
Age	-0.12**
Education level	0.06
Motivated activities	0.31**
Easy activities	0.23**
Physical activity identity	0.07
TTM-stage	-0.07
Openness to experiences	0.15**
Extraversion	0.05
Quitting self identity	0.09*
Non-smoker self identity	0.03

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Table 2: Correlations between the preference of the virtual agent or the questionnaire and other user characteristics.

As gender cannot be analysed by a Spearman's rho analysis as the variable is not of ordinal level of measurement, a Chi-square test was used. Table 3 shows that there is no apparent relationship between gender identity and the rating. This is confirmed by the Chi-square test between gender and the preference rating split into three categories (prefer virtual agent, prefer questionnaire or neutral), $\chi^2(4, N = 499) = 1.13$, $p = 0.89$, meaning that there is no significant relation.

Furthermore, as mentioned in 2.4.4, to identify statistical relationships between the themes and the participant characteristics a CATPCA analysis was used. Figure 2 shows the relation between the principal components of the themes and the significant correlations shown in Table 2.

In this figure a negative cluster can be identified in the top left corner, including the themes uneasiness and inconsiderate. At the bottom the themes about engagement and helpful content are located, while at the top right the themes about

Gender identity	Preference rating					
	Questionnaire		No preference		Virtual agent	
	No.	%	No.	%	No.	%
Male	35	14.3	42	17.2	167	68.4
Female	41	16.6	39	15.8	167	67.6
Other	2	25.0	1	12.5	5	62.5

% within gender identity

Table 3: Contingency table showing the frequencies and percentages of gender identity falling in each category

natural flow and personal connection are found. Therefore, the graph can be seen as a v shape transforming from a disconnect from the virtual agent, via engagement, to a personal connection with the virtual agent. This shape will be used as guide in the explanation of the seven themes.

3.1 Personal connection

“[E]ven if it was obvious that I wasn’t talking to a real person, it still felt like I had someone checking on me and my progress, making me feel more motivated” (P219) .

Like P219, many participants (n = 125, see Table 1) mentioned that they felt a personal connection to the virtual agent, resulting in this theme being the most frequently mentioned one. The codes that make up this theme are ‘the app being comfortable or personal’ and ‘the questionnaire feeling distant’. The relationships between this theme and participant characteristics can be identified by looking at the correlations (see Table 2) and the CATPCA (see Figure 2).

The CATPCA plot shows that participants indicating to feel a strong personal connection to the virtual agent are located close to the participants that indicate that the activities were motivational and easy (see right top Figure 2). This corresponds with earlier studies in which the performance expectancy, the perceived benefits of the system (motivation), and the effort expectancy, the ease of using the system, were correlated with the acceptance of the technology [3].

Theme Personal Connection	ρ
Motivated activities	0.17**
Easy activities	0.10**
Quitting self identity	0.09
Openness to experiences	0.06

** Correlation is significant at the 0.01 level (2-tailed)

Table 4: Spearman’s rho test between participant characteristics and theme personal connection.

The frequencies between participants that feel motivated by the activities and view the activities as easy and the feeling of a personal connection show that participants who are positive about the motivation and easiness of activities are more prone to feel a personal connection with the virtual agent (see contingency table in appendix A). This can be identified in the free-text response as well: “I feel more comfortable with Sam’s format and it makes me more attractive” (P406) .

A Spearman’s rho analysis was used to study the relation between the feeling of a personal connection and the easiness and motivation of activities. Both motivated activities

($\rho = 0.17$) and easy activities ($\rho = 0.10$) showed a significant weak correlation with a feeling of a personal connection, indicating a relationship between these characteristics and this theme.

Additionally, in the CATPCA (see Figure 2) a feeling of a strong personal connection is also located close to being really open to experience and a high ability to see oneself as a quitter. However, as can be seen in Table 4, there is no significant correlation between these characteristics.

Therefore, as a recommendation it would be good to focus on the personal connection to make the activities more attractive to participants, as there is a relationship between how motivated and easy the activities are and a personal connection with the virtual agent. However, as we do not know the direction of this relationship, it might also be worth creating more tailored activities to increase the feeling of a personal connection.

3.2 Engagement

“A questionnaire feels more distant which I think would make me less motivated to do the tasks. Where as knowing you are going to have the conversation style makes you want to do the tasks and be successful. I think as it is more informal and supportive than just filling in a form” (P272) .

Engagement was another important aspect mentioned by many respondents (n = 111, see Table 1). This theme was constructed by the codes ‘engaging’, ‘motivation’, ‘enjoyable’ and ‘the questionnaire being monotonous’.

The importance of engagement can be identified in the CATPCA analysis (see Figure 2) as all the answer categories of engagement are positioned closely to the origin, meaning that many participants indicated the feeling of engagement. The figure also shows that the extend to which participants indicated to feel engaged slightly corresponds with the different age groups as ‘a lot of engagement’ is located at the younger age groups and ‘not engaged at all’ is located towards the older ages. This relation can also be indicated in studies showing that older people are less likely to have access to digital tools [24], [25].

The frequencies show that 26.9% of all participants younger than the median indicated to feel engaged while only 17.5% of people older than or equal to the median said to feel engaged (see appendix A). This relationship was also confirmed by a Spearman’s rho of -0.13 ($p < 0.01$), showing a weak negative correlation.

Therefore, it is recommended to study this age knowledge gap [24], [25] with regards to a virtual agent. As bridging this gap can increase the feeling of engagement with the virtual coach and in turn increase the effectiveness of the intervention.

3.3 Helpful content

“Sam brought a unique opportunity to talk things through rather than force myself to be introspective” (P344) .

As P344 indicated, the perceived helpfulness of the virtual agent was an important aspect for accepting the technology as this theme was mentioned 86 times. Moreover, many people that preferred the questionnaire or were neutral, indicated that the virtual agent was just a questionnaire and thus not helpful.



Figure 2: A Categorical Principal Components Analysis between the seven identified themes and the correlating characteristics. The analysis places answer categories closer together that are scored by participants that have similar responses on other answer categories, thus cluster between similar answer categories can be identified.

Thus, although the CATPCA does not clearly show relations between this theme and other characteristics, the relation with the variable motivated activities might be interesting as three positive answer categories (2,3,4) are located in the same area as the theme helpful content (see right bottom Figure 2). Additionally, studies have shown that the performance expectancy, the perceived benefit and motivation, has a positive impact on the acceptance of technology [3]. Therefore, the helpfulness of the content might be related to how motivated the participants were by the activities.

The frequency distribution shows that the theme helpful content is indeed related to the code indifferent as only 4.5% indicates that the content was helpful from the participants that indicated to feel indifferent, compared to 20.0% from the participants that were not indifferent (see appendix A). Additionally, a correlation between helpful content and motivated activities has been found as well, as the relation had a Spearman's rho of 0.17 ($p < 0.00$) indicating a significant weak relation.

Therefore, it would be recommended to continue developing activities that are seen as helpful and that motivate the participants by tailoring these activities, as these have a positive influence on the acceptance of the virtual coach.

3.4 Natural flow

"More dynamic, in the sense of being challenged and able to give answers and get a reaction from the other side" (P68).

74 participants indicated the natural flow as a reason for preferring the virtual agent. This theme was constructed using the codes 'flow', 'instinctive' and 'human-like' and is thus focused on the naturalness of using the virtual agent. As a virtual coach can feel more natural it might be easier to use compared to a questionnaire, which can increase the acceptance of the virtual coach.

When looking at the CATPCA (Figure 2) it seems that the participants who scored on many codes associated with natural flow are located closely to participants that indicated that the activities were easy and motivating. For the same reasons

as mentioned in 3.1, this relation is interesting as the effort expectancy and thus the ease of use has an influence on use behaviour [3].

Theme Natural Flow	ρ
Motivated activities	0.12**
Easy activities	0.12**

** Correlation is significant at the 0.01 level (2-tailed)

Table 5: Spearman’s rho test between participant characteristics and theme natural flow.

A Spearman’s rho shows that motivated and easy activities both have a positive correlation with the theme natural flow, having both $\rho = 0.12$ ($p < 0.01$). Although the correlation is weak, only 8.8% of the participants that found the activities difficult indicated that they found the virtual agent natural compared to 16.3% that found the activities easy (see appendix A). Thus, there is a small relation in this dataset between this theme and the easiness of activities.

Therefore, it would be recommended to study the flow of the virtual agent and to try to make this more natural as this might increase how motivated people are to do the activities and the perceived easiness of the activities. This will in turn increase the effectiveness of the intervention [3].

3.5 Inconsiderate

“Id rather have a free format to express my thoughts rather than pre-filled out responses” (P128) .

A small amount of participants ($n = 47$, see Table 1) indicated that the virtual agent was inconsiderate, just as P128. This theme was constructed by combining the codes for ‘unhelpful’, ‘pointless’, ‘shallow’ and ‘questionnaire has a better functionality’. Therefore, this theme encompassed many participants who found the conversation with the virtual agent too limited or pointless.

The CATPCA shows that the participants who mentioned multiple codes corresponding with this theme were closely related to participants that indicated that the activities were not motivating as well as the activities being difficult to complete (see left top Figure 2). As already mentioned in 3.1, this is an interesting relation as studies show that the performance expectancy and the effort expectancy is important for the acceptance of technology [3].

Theme Inconsiderate	ρ
Motivated activities	-0.15**
Easy activities	-0.09*

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Table 6: Spearman’s rho test between participant characteristics and theme inconsiderate.

A Spearman’s rho analysis shows that there is a weak negative correlation ($\rho = -0.15$, $p < 0.01$) between the extend to which participants found the activities motivating and the inconsiderate theme (see Table 6). The correlation between the

easiness of activities and the theme is significant at the 0.05 level but would be seen as a negligible correlation.

As this theme combines participants who found the virtual agent limited and because there is a small relationship between this theme and motivation, it might be good to give the participants more options to freely respond to the conversational agent. Additionally, it might also be beneficial to further develop the virtual agent to be able to have more natural conversations which makes it more considerate and therefore less frustrating to use.

3.6 Cumbersome

“There is no gimmicky conversation. Clear questions and answer” (P249 about a questionnaire) .

A few participants ($n = 30$, see Table 1) indicated that they found the virtual agent cumbersome. Out of the 30 participants that mentioned this theme, 11 mentioned that it would be easier using a questionnaire (for instance P31, and P359). This is also caused by the codes combined in this theme, namely the codes ‘the easiness of a questionnaire’ and ‘being in control with a questionnaire’.

The CATPCA shows that the participants that found the virtual coach cumbersome are located in the same area as the participants that were not motivated by the activities (see left bottom Figure 2). Additionally, the categories of the cumbersome theme are slightly in the same clusters as the categories of openness to experience. This could be interesting as people who are not open to new experiences might indicate that they find an questionnaire easier. This tendency is also found in studies in which people with a more positive attitude towards digital tools incorporate it more extensively [26].

Theme Cumbersome	ρ
Motivated activities	-0.11*
Openness to experiences	-0.09

* Correlation is significant at the 0.05 level (2-tailed)

Table 7: Spearman’s rho test between participant characteristics and theme inconsiderate.

When looking at the frequencies, 8.1% of participants that found the activities not motivating indicated that the virtual agent was cumbersome compared to only 4.5% of participants that found the activities motivated (see appendix A). However, only a weak correlation could be found that is significant at the 0.05 level ($\rho = -0.11$, $p = 0.02$). Additionally, there is also no clear relation between participants that find the virtual agent cumbersome and their openness to experience as $\rho = -0.09$ ($p = 0.05$) (see Table 7).

Therefore, there are no strong indications that a characteristic of this dataset has a relation with participants that find the virtual agent cumbersome. For further insights into why people perceive a questionnaire to be easier it would be recommended to conduct more research on this topic.

3.7 Uneasiness

“I didn’t relate to Sam and the bot-speak was irritating. A questionnaire would feel more natural” (P451) .

Just as P451, a small group of participants ($n = 19$, see Table 1) stated that they felt uncomfortable or frustrated using the virtual agent. This theme was created by combining the codes of ‘apathetic’ and ‘unpleasant’.

Although there was only a small group that mentioned codes connected to this theme, the CATPCA shows that the participants who mentioned this theme were closely related to the participants who mentioned the inconsiderate theme (see left top Figure 2). Thus, just as the inconsiderate theme, the relations with the easiness of activities and the motivation to do the activities is important.

Theme Uneasiness	ρ
Motivated activities	-0.11**
Easy activities	-0.08
Theme inconsiderate	0.26**

** Correlation is significant at the 0.01 level (2-tailed)

Table 8: Spearman’s rho test between participant characteristics and theme inconsiderate.

A weak negative correlation ($\rho = -0.11$, $p < 0.01$) was found between the theme uneasiness and how motivated participants were to do the activities (Table 8). There is no significant correlation between this theme and the easiness of the activities.

Additionally, as the categories of the themes inconsiderate and uneasiness in the CATPCA are located closely together, both themes might be correlated. This could be explained by the fact that participants who perceive the virtual agent as shallow and not considerate enough might also perceive it as frustrating and difficult. A Spearman’s rho analysis confirmed this weak positive correlation, having a ρ of 0.26 ($p < 0.00$) (Table 8).

Therefore, it would be recommended to further develop the virtual agent to decrease the feeling of it being inconsiderate, as mentioned in 3.5. As people who perceive it to be inconsiderate are also more prone to be frustrated with the virtual agent. Thus, making the virtual coach more considerate might in turn decrease the feeling of frustration and uneasiness as well.

4 Discussion and conclusion

Through the open answer question, seven themes have been identified which were connected to why participants preferred to use a virtual coach or a questionnaire for smoking cessation. As the general consensus was fairly positive towards the use of a virtual coach, four themes were positively correlating with a preference for the virtual coach: personal connection, engagement, helpful content and natural flow. The other three themes were negatively correlating with the preference for the virtual coach: inconsiderate, cumbersome and uneasiness. These last three appeared less frequent.

Two important factors relating to most themes were how easy the participants found the activities and how motivated the participants were to execute the activities. These two aspects were important as performance expectancy, the expected benefit and thus motivation, and effort expectancy,

connected to easiness, were important predictors for use behaviour, the acceptance of technology [3].

The themes personal connection, helpful content and natural flow all had a small positive correlation with motivated activities. Personal connection and natural flow both had a small positive correlation with easy activities as well. The themes inconsiderate, cumbersome and uneasiness all had a small negative correlation with at least one of the two. This indicates the important relation between the themes and the easiness of activities and the motivation to do the activities.

Nonetheless, based on the data and the statistical model a directional relation cannot be assumed. Thus, it is not known if the easiness of and motivation for activities is predicting the themes or the other way around. Therefore, the acceptance of the virtual coach might be increased by either focusing on tailoring the activities to the participants or by studying the themes in more detail to increase the performance and effort expectancy [3].

A personal connection could increase the attractiveness of the activities proposed by the virtual coach just as the perceived helpfulness of the content might be a big factor increasing the motivation to do the activities [3]. The natural flow of the virtual coach could also increase the perceived easiness of the activities as participants that perceived the virtual agent to be shallow and inconsiderate found the activities to be less motivational and were more prone to be frustrated with the virtual agent. Therefore, all these topics could greatly improve the effectiveness of a virtual coach.

Additionally, the theme engagement had a small negative correlation with age, which could be caused by a digital knowledge gap the older people are [24], [25]. Therefore, it might also be important to further develop the application to make it accessible for people that do not have a lot of digital know-how.

Nonetheless, this study has limitations as well, as the correlations between the characteristics and the themes are identified within the dataset. This dataset, however, does not have to represent a realistic depiction of the real world as it can be skewed by multiple factors.

One of this factors might be that the participants were selected using the Prolific platform. This causes many participants to be already familiar with the online realm and they might therefore be more acceptable of a virtual coach. To study this effect, future research could conduct a follow-up study in which participants are selected using offline methods. This could reveal undiscovered relations between participant characteristics and the acceptance of the virtual coach.

Another factor that might have skewed the outcome of this study, would be that all participants engaged with Sam, there was no group of participants that engaged with a real coach. Future research could study the difference between a virtual coach and a real coach by adding this group of participants and measuring the effectiveness of the virtual coach compared to an in-person intervention.

All in all, the remarkably positive acceptance of Sam combined with the mentioned benefits by participants indicates a new promising innovative treatment for virtual interventions, as “[i]t was nice having someone to talk to” (P395).

5 Responsible Research

During this research process, many subjective choices had to be made. Therefore, it is important to be aware of the bias this introduces and especially to be transparent about the choices being made to increase the reliability of the findings in this paper.

In order to increase the reliability of the codes, researcher triangulation was used in which two researchers, independently from each other, coded the same dataset and discussed their disagreements [21]. Likewise, method triangulation was used to increase the reliability of the themes by cross-validating these themes with statistical analysis and literature [21].

The steps taken in the statistical analysis are explained in detail in 2.4.4, so follow-up research could confirm these statistical tests or understand how the outcomes in this report are calculated. By specifying the variables of interest before the correlations were tested, p-hacking was also countered. The literature used in the report is all cited and could be easily found in the bibliography for future research.

Additionally, for the Categorical Principal Component Analysis, multiple variables had to be re-coded to counter outliers and to construct nominal level of measurement variables. This process is also based on subjective choices and is described in detail in 2.4.4.

Besides, both the original dataset as well as the dataset used in this paper, with the re-coded variables, the identified codes and themes, are published on 4TU [13], [14]. In this dataset, the answers that were difficult to interpret, described in detail in section 2.4.1, can be found as well. This allows future researchers to reproduce and validate the findings of this paper.

Thus, in general, most ethical aspects of this research had to do with how the data is constructed, like codes and themes, and how the data is reconstructed, like re-coding for statistical tests. By being as transparent as possible and by publishing the final dataset, this research tried to counter invalid findings and ethical issues.

Acknowledgement

I would want to thank Omar Sheasha for his role as second coder and Nele Albers and Dr. Willem-Paul Brinkman for their support and advise.

References

- [1] J. B. Bricker, K. E. Mull, J. A. Kientz, *et al.*, “Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy,” *Drug and Alcohol Dependence*, vol. 143, pp. 87–94, 1 2014, ISSN: 18790046. DOI: 10.1016/j.drugalcdep.2014.07.006.
- [2] L. L. Struik, J. L. Bottorff, N. B. Baskerville, and J. L. Oliffe, “The crush the crave quit smoking app and young adult smokers: Qualitative case study of affordances,” *JMIR mHealth and uHealth*, vol. 6, 6 Jun. 2018, ISSN: 22915222. DOI: 10.2196/mhealth.9489.
- [3] K. Ghorai and P. Ray, “Factors affecting the acceptance of mobile based multi-feature service for smoking cessation using utaut,” University of New South Wales, 2019. [Online]. Available: <https://www.researchgate.net/publication/332656570>.
- [4] A. Kulhánek, R. Gabrhelík, D. Novák, and V. Burda, “Ehealth intervention for smoking cessation for czech tobacco smokers: Pilot study of user acceptance adiktologie,” 2018, pp. 81–85.
- [5] C. Free, R. Whittaker, R. Knight, T. Abramsky, A. Rodgers, and I. G. Roberts, “Txt2stop: A pilot randomised controlled trial of mobile phone-based smoking cessation support,” *Tobacco Control*, vol. 18, pp. 88–91, 2 Apr. 2009, ISSN: 09644563. DOI: 10.1136/tc.2008.026146.
- [6] S. Provoost, A. Kleiboer, J. Ornelas, *et al.*, “Improving adherence to an online intervention for low mood with a virtual coach: Study protocol of a pilot randomized controlled trial,” *Trials*, vol. 21, 1 Dec. 2020, ISSN: 17456215. DOI: 10.1186/s13063-020-04777-2.
- [7] S. Provoost, H. M. Lau, J. Ruwaard, and H. Riper, *Embodied conversational agents in clinical psychology: A scoping review*, May 2017. DOI: 10.2196/jmir.6553.
- [8] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, “Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (woebot): A randomized controlled trial,” *JMIR Mental Health*, vol. 4, 2 Apr. 2017, ISSN: 23687959. DOI: 10.2196/mental.7785.
- [9] T. Bickmore and A. Gruber, *Relational agents in clinical psychiatry*, Mar. 2010. DOI: 10.3109/10673221003707538.
- [10] A. Bresó, J. Martínez-Miranda, C. Botella, R. M. Baños, and J. M. García-Gómez, “Usability and acceptability assessment of an empathic virtual agent to prevent major depression,” *Expert Systems*, vol. 33, pp. 297–312, 4 Aug. 2016, ISSN: 14680394. DOI: 10.1111/exsy.12151.
- [11] N. Albers, M. A. Neerincx, and W. P. Brinkman, “Reinforcement learning-based persuasion by a conversational agent for behavior change,” 2021.
- [12] N. Albers, *Perfectfit-project/virtual_coach_rl_persuasion_algorithm: Virtual coach for longitudinal study*, version 1.0, Mar. 2022. DOI: 10.5281/zenodo.6319356. [Online]. Available: <https://doi.org/10.5281/zenodo.6319356>.
- [13] N. Albers, M. A. Neerincx, and W.-P. Brinkman, *Acceptance of a virtual coach for quitting smoking and becoming physically active: Dataset*, May 2022. DOI: 10.4121/19934783.v1. [Online]. Available: https://data.4tu.nl/articles/dataset/Acceptance_of_a_Virtual_Coach_for_Quitting_Smoking_and_Becoming_Physically_Active_Dataset/19934783.
- [14] J. Dechering, *Acceptance of a virtual coach as guided intervention for smoking cessation: Dataset*, Jun. 2022. DOI: 10.4121/20066186.

[15] M. Tavakol and R. Dennick, *Making sense of cronbach's alpha*, Jun. 2011. DOI: 10.5116/ijme.4dfb.8dfd.

[16] J. Barbera, N. Naibert, R. Komperda, and T. C. Pentecost, *Clarity on cronbach's alpha use*, Feb. 2021. DOI: 10.1021/acs.jchemed.0c00183.

[17] K. S. Taber, "The use of cronbach's alpha when developing and reporting research instruments in science education," *Research in Science Education*, vol. 48, pp. 1273–1296, 6 Dec. 2018, ISSN: 15731898. DOI: 10.1007/s11165-016-9602-2.

[18] S. D. Gosling, P. J. Rentfrow, and W. B. Swann, "A very brief measure of the big-five personality domains," *Journal of Research in Personality*, vol. 37, pp. 504–528, 6 2003, ISSN: 00926566. DOI: 10.1016/S0092-6566(03)00046-1.

[19] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, pp. 77–101, 2 2006, ISSN: 14780887. DOI: 10.1191/1478088706qp063oa.

[20] A. Bryman, *Social research methods*. Oxford university press, 2016, ISBN: 978-0-19-968945-3.

[21] N. Carter, D. Bryant-Lukosius, A. Dicenso, J. Blythe, and A. J. Neville, *The use of triangulation in qualitative research*, Sep. 2014. DOI: 10.1188/14.ONF.545-547.

[22] M. L. McHugh, "Interrater reliability: The kappa statistic," *Biochemia Medica*, vol. 22, pp. 276–282, 3 2012. [Online]. Available: <https://hrca.k.srce.hr/en/89395>.

[23] S. M. Wilson and L. C. Peterson, "The anthropology of online communities," *Annual Review of Anthropology*, vol. 31, pp. 449–467, 2002, ISSN: 00846570. DOI: 10.1146/annurev.anthro.31.040402.085436.

[24] T. Drabowicz, *Chapter 12: Digital skills inequality in the context of an aging society: The case of poland*, E. Hargittai, Ed., Nov. 2021.

[25] T. Elliott and J. Earl, "Online protest participation and the digital divide: Modeling the effect of the digital divide on online petition-signing," *New Media and Society*, vol. 20, pp. 698–719, 2 Feb. 2018, ISSN: 14617315. DOI: 10.1177/1461444816669159.

[26] O. Wohlfart, T. Trumler, and I. Wagner, "The unique effects of covid-19 – a qualitative study of the factors that influence teachers' acceptance and usage of digital tools," *Education and Information Technologies*, vol. 26, pp. 7359–7379, 6 Nov. 2021, ISSN: 15737608. DOI: 10.1007/s10639-021-10574-4.

[27] P. Schober and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesthesia and Analgesia*, vol. 126, pp. 1763–1768, 5 May 2018, ISSN: 15267598. DOI: 10.1213/ANE.0000000000002864.

[28] M. Linting and A. V. D. Kooij, "Nonlinear principal components analysis with catpca: A tutorial," *Journal of Personality Assessment*, vol. 94, pp. 12–25, 1 Jan. 2012, ISSN: 00223891. DOI: 10.1080/00223891.2011.627965.

A Theme contingency tables

		Theme Personal Connection		
			No	Yes
Motivated activities	Negative	Count %	127 85.2	22 14.8
	Neutral	Count %	34 82.9	7 17.1
	Positive	Count %	214 69.0	96 31.0
Easy activities	Negative	Count %	49 86.0	8 14.0
	Neutral	Count %	26 83.9	5 16.1
	Positive	Count %	300 72.8	112 27.2
Quitting self identity	< median	Count %	148 78.7	40 21.3
	≥ median	Count %	227 72.8	85 27.2
Openness to experiences	< median	Count %	190 78.5	52 21.5
	≥ median	Count %	185 71.7	73 28.3

% across rows

** Correlation is significant at the 0.01 level (2-tailed)

Table 9: Contingency table between characteristics and the feeling of a personal connection. Additionally, the result of a Spearman's rho test can be found in the left column of the table.

		Theme Engagement		
			No	Yes
Age	< median	Count %	182 73.1	67 26.9
	≥ median	Count %	207 82.5	44 17.5

% across rows

** Correlation is significant at the 0.01 level (2-tailed)

Table 10: Contingency table between characteristics and the feeling of engagement. Additionally, the result of a Spearman's rho test can be found in the left column of the table.

		Theme Helpful Content		
			No	Yes
Code Indifferent	No	Count	329	82
		%	80.0	20.0
	Yes	Count	85	4
		%	95.5	4.5
Motivated activities	Negative	Count	133	16
		%	89.3	10.7
	Neutral	Count	37	4
		%	90.2	9.8
	Positive	Count	244	66
		%	78.7	21.3

% across rows

** Correlation is significant at the 0.01 level (2-tailed)

Table 11: Contingency table between characteristics and the helpfulness of the content. Additionally, the result of a Spearman's rho test can be found in the left column of the table.

		Theme Inconsiderate		
			No	Yes
Motivated activities	Negative	Count	129	20
		%	86.6	13.4
	Neutral	Count	33	8
		%	80.5	19.5
	Positive	Count	291	19
		%	93.9	6.1
Easy activities	Negative	Count	47	10
		%	82.5	17.5
	Neutral	Count	28	3
		%	90.3	9.7
	Positive	Count	378	34
		%	90.6	9.4

% across rows

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Table 13: Contingency table between characteristics and the feeling that the virtual agent is inconsiderate. Additionally, the result of a Spearman's rho test can be found in the left column of the table.

		Theme natural		
			No	Yes
Motivated activities	Negative	Count	132	17
		%	88.6	11.4
	Neutral	Count	38	3
		%	92.7	7.3
	Positive	Count	256	54
		%	82.6	17.4
Easy activities	Negative	Count	52	5
		%	91.2	8.8
	Neutral	Count	29	2
		%	93.5	6.5
	Positive	Count	345	67
		%	83.7	16.3

% across rows

** Correlation is significant at the 0.01 level (2-tailed)

Table 12: Contingency table between characteristics and the perceived natural flow. Additionally, the result of a Spearman's rho test can be found in the left column of the table.

		Theme Cumbersome		
			No	Yes
Motivated activities	Negative	Count	137	12
		%	91.9	8.1
	Neutral	Count	37	4
		%	90.2	9.8
	Positive	Count	296	14
		%	95.5	4.5
Openness to experiences	< median	Count	220	22
		%	90.9	9.1
	≥ median	Count	250	8
		%	96.9	3.1

% across rows

* Correlation is significant at the 0.05 level (2-tailed)

Table 14: Contingency table between characteristics and the feeling that the virtual agent is cumbersome. Additionally, the result of a Spearman's rho test can be found in the left column of the table.

		Theme uneasiness		
			No	Yes
Motivated activities	Negative	Count	138	11
		%	92.6	7.4
	Neutral	Count	38	3
		%	92.7	7.3
	Positive	Count	305	5
		%	98.4	1.6
Easy activities	Negative	Count	51	6
		%	89.5	10.5
	Neutral	Count	29	2
		%	93.5	6.5
	Positive	Count	401	11
		%	97.3	2.7

% across rows

** Correlation is significant at the 0.01 level (2-tailed)

Table 15: Contingency table between characteristics and the feeling of uneasiness with the virtual agent. Additionally, the result of a Spearman's rho test can be found in the left column of the table.