Acceptance of a virtual coach for quitting smoking and becoming more physically active: A thematic analysis

Traits for a virtual coach to be a "friend"

Mahira Ali
Supervisor: Nele Albers
Responsible Professor: Willem-Paul Brinkman
EEMCS, Delft University of Technology, The Netherlands

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

Seeing the virtual coach as a friend is beneficial as it increases the motivation, confidence, and perseverance of the accompanying. Here, an investigation has taken place to what characteristics the virtual coach must possess to establish this friendly relationship. Thus, the main research question is: **What are the reasons for seeing the virtual coach as a stranger or friend?** This research made use of pre-gathered data. Here, five hundred participants interacted with the text-based virtual coach Sam in five separate sessions. Afterwards, each participant rated the relationship with the virtual coach, followed by an explanatory free-text response to which thematic analysis was applied. This resulted in five main themes: **Relation**, **Positive Characteristics**, **Perception**, **Impersonal**, and **Chat Opinions**. These were used to determine the characteristics the virtual coach must have to be considered a friend. Furthermore, these themes were used to obtain correlations. With these, it has been concluded that the user is more likely to develop a closer relationship with the virtual coach when: perceiving it as a human, pleasantly conversing, and being positive about its character. Whereas the user's age has a negligible influence on the relationship.

1 Introduction

Annually, 8 million people lose their lives due to smoking [1]. Even though 90% of smokers are aware of the life-threatening diseases that come with smoking, this is not the main reason to quit smoking [2]. For many, the greatest motivation to stop smoking is disapproval of their surroundings [3]. Since the drive to stop smoking usually comes from a personal perspective, and smoking cessation is a lifestyle change, quitting smoking requires consistent social support [4].

eHealth refers to health services provided through the Internet and other technologies [5]. Often, virtual coaches are deployed in eHealth applications to provide the desired user interaction [6]. Research showed that eHealth leads to increased motivation, confidence, and perseverance [7]. These properties are enhanced by a friendly relationship between the coach and the accompanying [8]. This supports the fact that it is beneficial to see the virtual coach as a friend - one that always keeps your "secrets." However, the characteristics the virtual coach must possess to be seen as a friend are still undefined.

During this research, the virtual coach Sam was used. Sam is a text-based virtual coach developed for the Perfect Fit project to convince smokers to quit by performing small activities [9] [10]. Five hundred participants interacted with the virtual coach for up to five sessions. Afterwards, each participant gave a rating on how they would describe their relationship with the conversational agent Sam followed by a free-text response explaining the rating.

Thematic analysis was applied to five hundred free-text responses. These have been gathered in the conducted experiment by Albers and Brinkman [11]. Subsequently, the results were used to determine which characteristics a virtual coach must possess such that it is regarded as a friend.

Triangulation was used to ensure the quality of the performed thematic analysis, as described by Peersman [12]. For this purpose, another researcher was involved during the thematic analysis. This enhanced the reliability of the findings by reducing the bias in the analysis [13]. Additionally, using multiple data sources supported conclusions about results and decreased the likelihood of false interpretations [14]. These data sources are, in addition to the free-text responses and ratings, literature, and quantitative data. The quantitative data were collected in the experiment by Albers and Brinkman [11]. The quantitative data of each participant consisted of demographic characteristics such as age and gender. But also smoke-related characteristics, for instance, smoking frequency.

This research aimed to identify the characteristics the virtual coach must have to develop a closer relationship with the user. Hence, the main research question is: **What are the reasons for seeing the virtual coach as a stranger or friend?**

2 Methodology

In the study of Albers and Brinkman, five hundred participants have interacted with a virtual coach named Sam in five separate sessions [11]. The participants were persuaded in the first two sessions based on a random persuasion type. Then, the participants were persuaded by the persuasion type that fitted them the best for the remaining sessions.

Prior to these five sessions, demographic and smoke-related characteristics were gathered. Figure 1 provides an overview of the main characteristics of the participants.
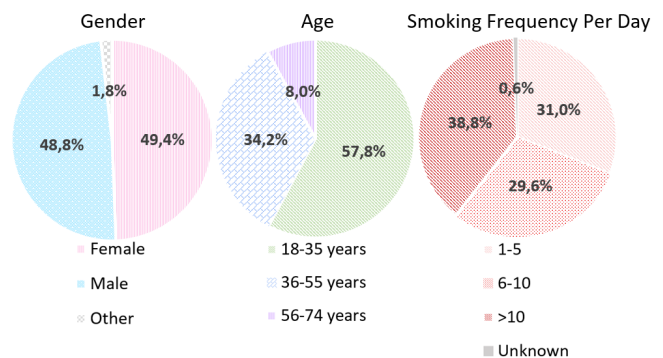

Figure 1: Main characteristics of the participants ($n = 500$).

After these five sessions, the participants took part in a questionnaire. During this questionnaire, they rated how they would characterise the relation they experienced with the virtual coach. The rating scaled from a stranger (-5) to a friend (5), where 0 indicated neutral. After the rating, a free-text response followed where each participant explained their rating.

Next, the five hundred free-text responses were used during the thematic analysis. Braun and Clark described the phases of thematic analysis [15]. These phases were used as a guideline to perform the thematic analysis. Additionally, the quality of the executed thematic analysis was ensured by applying two types of triangulation: researcher triangulation and method triangulation. Researcher triangulation, where multiple researchers were involved, determined the reliability of the coding scheme generated during thematic analysis. Method triangulation, which supported the findings, was executed with the free-text responses, literature, and quantitative data, containing demographic and smoke-related characteristics. Hence, the analysis setup was structured as described in Figure 2.
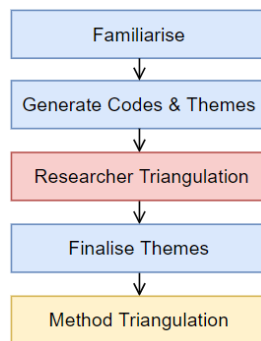

Figure 2: The setup of the analysis where the blue boxes are the three phases of the thematic analysis. Researcher triangulation and method triangulation were executed to ensure the quality of the thematic analysis.

2.1 Thematic Analysis Phase 1 : Familiarisation

The thematic analysis started with the familiarisation with the free-text responses. This phase has been completed by reading twice through all the free-text responses. Meanwhile, initial ideas for potential codes were written down.

2.2 Thematic Analysis Phase 2 : Generating codes and themes

During the second phase of thematic analysis, codes were generated in two rounds. The codes were produced inductively in the first round of coding, and in the second round, the coding was done deductively [16]. Inductive coding was executed by starting with no codes and developing the codes while analysing the data [17]. Subsequently, deductive coding was carried out by coding the responses with the codes determined in the first coding round [18]. Lastly, these codes were used to search for themes. This was done before double coding to prevent inconveniences.

2.3 Researcher Triangulation

The use of researcher triangulation determined the reliability of the coding scheme. This has been expressed by the average Cohen's Kappa, calculated by summing up the Cohen's Kappa of each code divided by the total number of codes. The Cohen's Kappa has been interpreted according to the guidelines provided by Landis and Koch [19].

The average Cohen's Kappa of the coding scheme is 0.63. This means that a *substantial agreement* has been reached. Additionally, the Cohen's Kappa of each of the themes and subthemes can be found in Figure 3 (Section 3). This was done to indicate the reliability of the themes.

During the discussion of the final coding, the coding scheme was also revised. Here, codes were eliminated, renamed, and merged in cooperation with the double coder. For example, the code *Lend an ear* was implied by the code *Caring*. Therefore, *Lend an ear* was removed.

While amending the coding scheme, Cohen's Kappa was considered. As a result, the codes with a Cohen's Kappa beneath 0.60 were reconsidered to eliminate as it is the upper bound of *moderate agreement*. While reviewing, the following was taken into account: the code's frequency, the code's descriptiveness, and the insight the code provided. Additionally, codes with a Cohen's Kappa of 1.00 were removed when used scarcely. The changes to the coding scheme and their associated Cohen's Kappa can be found in Appendix A, Figure 1. Above that, the frequency of the codes can be seen in Appendix B, Figure 2.

In addition, responses were encountered during the discussion, of which the coding could not be determined. This was because both codings could be argued to be valid, where one was interpreted positively while the other negatively. Therefore, based solely on the response, which interpretation should be considered 'correct' could not be decided. Thus, the rating was taken into account to make a decision. Above that, there were responses where no agreement on the assignment of codes was reached. Therefore, a third researcher was involved in making the final decision.

Furthermore, responses in other languages were translated into English. Next, codes were assigned to the response. For example, participant P0 responded in Spanish.

2.4 Thematic Analysis Phase 3 : Finalising themes

The themes were reviewed, defined, and named in the final phase of the thematic analysis. The revision initially involved two levels. In the first level, candidate themes were determined to form a coherent pattern. During this level, two decisions were made. Firstly, the theme of general remarks on the application was eliminated. This was done because this theme was implied by the theme containing positive characteristics of the virtual coach. Secondly, the theme of remarks on the experiment's setup was removed. This decision was made as there was inadequate data to support the theme.

Here, several codes were removed due to insufficient supportive data and awareness that the codes did not suit the responses. For instance, it was decided to eliminate the code *personal* as the code did not cover the content of the data response despite having a reasonable Cohen's Kappa.

The relation between the data set and the individual themes was checked at the second level. Here, a thematic map was derived where a data response can have one or more themes.

Lastly, the themes were named and defined. The final themes have been presented in the results (Section 3).

2.5 Method Triangulation

The performance of method triangulation supported the themes. The execution of the thematic analysis on the five hundred free-text responses resulted in these themes. Additionally, quantitative analysis showed correlations between these themes and the relation with the virtual coach. Lastly, literature contributed to assisting these results.

Before performing method triangulation, the quantitative data were pre-processed. For this purpose, Cronbach's α was calculated. Since the reliability for each data item was sufficiently high based on the guidelines provided by Habadin *et al.*, the means of the items were used as index measures [20]. In Table 1, the corresponding Cronbach's α can be seen with the respective internal consistency.

Table 1: Cronbach's Alpha for each pre-questionnaire data with the corresponding internal consistency according to Habadin *et al.* [20].

Data Item	Cronbach's α	Internal Consistency
Physical Activity Identity	0.89	Good
Quitter Self-Identity	0.76	Acceptable
Smoker Group-Identity	0.79	Acceptable
Smoker Self-Identity	0.74	Acceptable
NonSmoker Self-Identity	0.67	Questionable
Quitting Self-Efficacy	0.83	Good
NonSmoker Group-Identity	0.79	Acceptable
Need for Cognition	0.77	Acceptable

During the thematic analysis, presumptions were made based on the responses. For instance, the response *"It's still a chat bot, not a close friend"* (P7f) raised the conjecture that the awareness of being a machine had an impact on the relationship with the virtual coach. Subsequently, these presumptions were supported by consulting literature before applying quantitative analysis.

Based on these presumptions, quantitative analysis was performed by determining correlations. Here, correlation analysis was carried out with the ratings, themes, and quantitative data. Various correlation coefficients were utilized. This is because dichotomous data (theme) and continuous data (rating and characteristics) were used. An overview of these correlation coefficients can be seen in Table 2. The corresponding interpretation of each correlation coefficient can be found in Appendix C.

Table 2: Correlation coefficients used during the quantitative analysis. The data specifies in what case which coefficient was used.

Correlation Coefficient	Guideline	Data 1	Data 2
r_ϕ	Akoglu [21]	Theme	Theme
r_{pb}	Dichoso and Cabauatan [22]	Theme	Rating
r_s	Dancey and Reidy [23]	Rating	Characteristic

r_ϕ, Phi Coefficient; r_{pb}, Point Biserial Correlation Coefficient; r_s, Spearman's Rank Correlation Coefficient

For each correlation, the p-value has been reported. The p-values have been interpreted by the use of the guidelines provided by Raiola and Tore [24].

3 Results

In this section, the results of the executed analysis are presented. Here, participant identifiers were used to refer to participants. More details on each participant can be found in Appendix D, Table 4.

3.1 Themes

During the execution of the thematic analysis, five main themes and subthemes were derived. The five main themes are: **Relation**, **Positive Characteristics**, **Perception**, **Impersonal** and **Chat Opinions**.

The theme **Relation** contained responses of participants which gave a remark on the relation with the virtual coach. Thereafter, these were divided into three subthemes, namely **Positive Relation**, **Neutral Relation**, and **Negative Relation**. Firstly, the subtheme **Positive Relation** consisted of responses where the participant experienced a positive relationship with the virtual coach. This was indicated by either calling it *"a good relation"* (P1a), referring to the virtual coach as *"a friend"* (P1b) or stating that the virtual coach was *"more than just a stranger"* (P1c). Secondly, in the subtheme **Neutral Relation** the virtual coach was considered to be *"neither a stranger nor a close friend"* (P2a) or *"simply a professional"* (P2b). Lastly, in the third subtheme **Negative Relation** a negative relation was being experienced with the virtual coach. Here, there was either *"no relationship"* (P3a), the virtual coach was said to be *"not a close friend"* (P3b) or it felt *"not possible to create a relation with a robot"* (P3c).

Participants, who noticed positive characteristics of the virtual coach, were included in the theme **Positive Characteristics**. For instance, the virtual coach was referred to as *"friendly"* (P4a), *"polite"* (P4b) and *"supportive"* (P4c).

The theme **Perception** consisted of responses where participants indicated how they perceived the virtual coach. Subsequently, these were split up into the following three subthemes: **Human-Like**, **Artificial but Human-Like** and **Clearly Artificial**. The first subtheme, **Human-Like**, includes responses where the virtual coach was addressed as *"he"* (P5a, P5b) or *"him"* (P5c, P5d). In the second subtheme, **Artificial but Human-Like**, the virtual coach *"felt more like a real person than an AI"* (P6a). Finally, in the third subtheme, **Clearly Artificial**, the virtual coach was *"just a bot"* (P7a) and *"not human"* (P7b).

Responses, which indicated that the participant had the feeling that the virtual coach was not personal enough or lacked personality, were assigned to the theme **Impersonal**. Here, the participant had *"no personal connection"* (P8a) or *"don't know Sam"* (P8b, P8c) since the virtual coach *"shows no emotion nor can show empathy"* (P8d).

The theme **Chat Opinions** comprised responses where remarks were made on the chat. Hereafter, these were divided into the subtheme **Good Chat** and **Objections to the**

Chat. In the first subtheme, **Good Chat**, the participants *"enjoyed talking to Sam"* (P9a) since they had the feeling as if *"talking to a real person"* (P9b) and *"Sam generally responded well"* (P9c). The second subtheme, **Objections to the Chat**, included dissatisfaction about the chat. Here, the *"conversations were just a series of questions and present responses"* (P10a) which consisted of *"mainly clicking and typing"* (P10b) and was *"just a simple script"* (P10c).

The themes, with their respective average Cohen's Kappa and usage frequency, can be found in Figure 3.
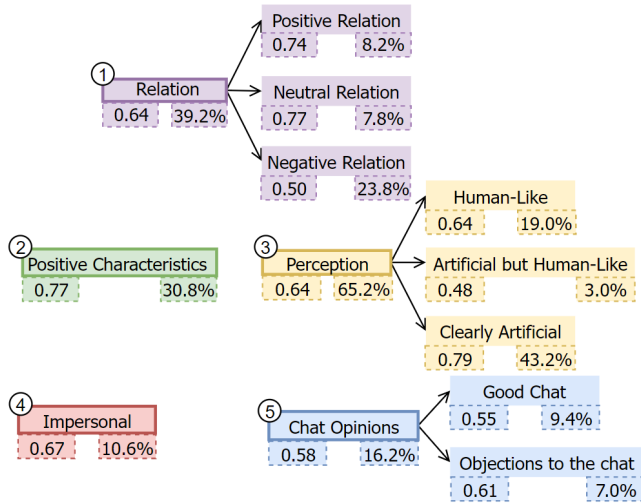

Figure 3: Final five themes with their subthemes. On the left side of each entry is the average Cohen's Kappa and on the right side the frequency percentage ($n = 500$).

3.2 Impact of positive characteristics on the relation

"He felt friendly" (P4a)

Wagner has investigated the influence of characteristics on the development of a relationship [25]. These results showed that positive characteristics are essential in a friend. Hence, a correlation was obtained between the positive characteristics of the virtual coach and the relation with the virtual coach. These phi coefficients (r_ϕ) can be found in Figure 4, which were interpreted according to the guidelines by Akoglu [21].
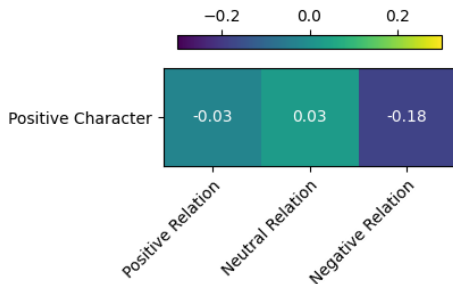

Figure 4: An heatmap of the correlations - expressed in r_ϕ - between the theme positive chacteristics and the subthemes of relation.

Here, the observance of positive characteristics caused a *strong* decrease in experiencing a negative relation with the virtual coach. On the contrary, a positive or neutral relation was negligibly affected by being pleased with the character of the virtual coach.

This finding was supported by performing a correlation analysis between the virtual coach's positive characteristics and the participants' ratings. This resulted in a significant *very good* relation ($r_{pb} = 0.45$, $p < 0.001$). By this means, a higher rating was induced by finding the character of the virtual coach positive. Here, the virtual coach was referred to as *"a kind person"* (P4d), *"a friendly guy"* (P4e), *"a good coach"* (P4f), and *"a helpful person"* (P4g). Hence, characteristics the virtual coach must possess to be considered positive were kindness, friendliness, guidance, and helpfulness.

Therefore, experiencing the character of the virtual coach as positive leads to a closer relationship with the virtual coach. Hoyos-Valdés endorsed this by stating that *"a friendship is grounded in the good qualities of the friends"* [26]. Here, honesty, kindness, loyalty, and authenticity define goodness.

3.3 Impact of the perception on the relation

"It's a bot, there was no relationship" (P7c)

Becker *et al.* have researched the benefits of incorporating emotions into a conversational agent [27]. They showed that integrating emotions into a conversational agent leads to the perception of being human by the user. This causes an increment in the likeability of the conversational agent. Hence, the correlation was determined between the perception of the virtual coach and the relation with the virtual coach. In Figure 5 the findings (r_ϕ) can be viewed.
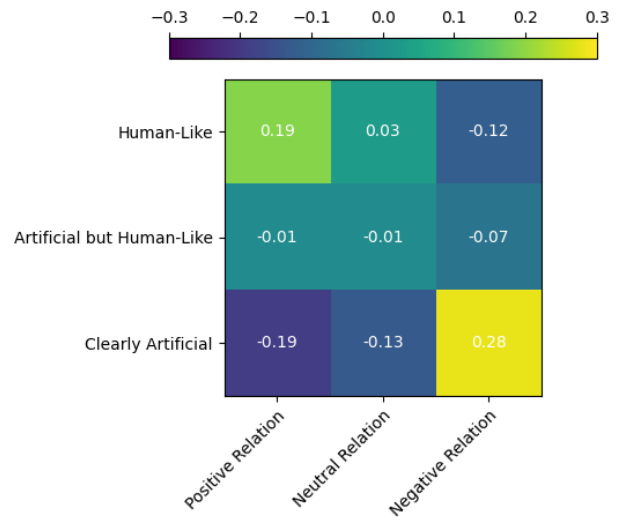

Figure 5: An heatmap of the correlations - expressed in r_ϕ between the subthemes of perception and the subthemes of relation.

As shown in Figure 5, the awareness of the virtual coach being artificial but still feeling that it is human negligibly

impacted the relation with the virtual coach. Above that, experiencing a neutral relation was minorly influenced by the perception of the virtual coach.

However, perceiving the virtual coach as human provoked a significant ($p < 0.001$) *strong* increase in having a positive relation with the virtual coach. This coincided with the findings of Becker *et al.* [27]. Additionally, the contrapositive of this research was supported. This is because regarding the virtual coach as clearly artificial resulted in a significant ($p < 0.001$) *very strong* increase in experiencing a negative relation.

To support these findings, correlation analysis was performed between the perception and the ratings given by the participants. Figure 6 shows these results (r_{pb}).
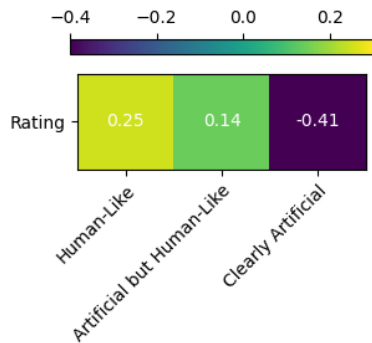

Figure 6: An heatmap of the correlations - expressed in r_{pb} - between the perception and the rating given by the participants.

Here, the perception of being clearly artificial induced a significant ($p < 0.001$) *very good* decrease in the rating. This was expressed by participants, for instance, stating that *"you can't have a relationship with a machine"* (P7d), and they *"prefer to talk to a real person"* (P7e). Furthermore, the rating was significantly ($p < 0.001$) *marginally* increased by seeing the virtual coach as a human. Here, the virtual coach *"felt like a friend"* (P5e) and *"a good connection with him"* (P5f) was being experienced.

Therefore, perceiving the virtual coach as a human leads to a closer relationship with the virtual coach. The user must not be aware that the virtual coach is built of lines of code. Human-like attributes affect the perception of the virtual coach [28]. Such attributes were human-like language or name, which showed an increase in the perception of the virtual coach as human-like. However, calling the virtual coach by his name, Sam, caused a *poor* increase in the rating ($r_{pb} = 0.12$, $p = 0.008$).

3.4 Impact of the chat on the relation

"I enjoyed our conversations" (P9d)

Research has been performed by Mesch and Talmud on how communication influences online friendship formations [29]. This study showed that virtual communication aids in the creation of new friendships. Based on this finding, the correlation was calculated between how the user experienced the chat with the virtual coach and the relation with the virtual coach. In Figure 7 the values (r_ϕ) can be found.

Figure 7: An heatmap of the correlations - expressed in r_ϕ - between the subthemes of chat and the subthemes of relation.

As shown in Figure 7, pleasantly conversing with the virtual coach resulted in a significant ($p = 0.02$) *moderate* increase when classifying the relation with the virtual coach as positive. Additionally, experiencing a negative relation significantly ($p = 0.003$) *moderately* decreased when having a good chat with the virtual coach. Finally, objections to the chat had a negligible influence on the relation.

Correlation analysis was performed between the subthemes of opinions about the chat and the participants' ratings to strengthen this presumption. The values (r_{pb}) can be found in Figure 8.
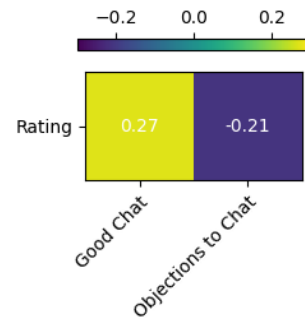
Figure 8: An heatmap of the correlations - expressed in r_{pb} - between the rating and the subthemes of chat.

Here, a good chat with the virtual coach caused a significant ($p < 0.001$) *acceptable* increase in the rating. This was because the virtual coach *"sounded like a really good friend"* (P9e) and had a *"friendly and caring tone"* (P9f). Above that, he was *"giving good advices"* (P9g); he *"had polite answers"* (P9h) and *"responded well to the answers"* (P9i). Thus, a good chat can be held with the virtual coach when his tone is pleasant; he provides relevant information and replies well.

Moreover, objections to the chat induced a significant ($p < 0.001$) *marginal* decrease in the rating. Such

objections were about the response mechanism. For instance, there was not *"much choice about answers"* (P10d), and it were *"standard answers not really personal"* (P10e), causing that one *"couldn't really respond as liked"* (P10f). Additionally, the conversation was being experienced as being monotonous. In particular, *"it was manly clicking and typing"* (P10b); the virtual coach gave *"the same answers to each response"* (P10g), and *"the conversations were just a series of questions and present responses"* (P10a). Above that, the conversation was called *"a simple script"* (P10c) with *"automated replies"* (P10h).

Therefore, experiencing a pleasant conversation with the virtual coach stimulates the development of a closer relationship with the virtual coach. Van Pinxteren *et al.* investigated which communicative behaviours positively affect the relation [30]. This was classified into two dimensions: one that considers the verbal and appearance, and the other the responsiveness. Firstly, appearances based on human similarity and supportive praise positively influenced the relationship. Secondly, responsiveness implied that the user felt understood. This was accomplished by listening signals such as expressing concern and emotions and asking questions.

3.5 Impact of impersonal on the relation

"He doesn't know me, only little parts" (P8e)

The "gifted conversationalist" Gabor has written a book on how to make friends [31]. In his book, Gabor advises to "write a short list of events going on in your life that you are willing to share with others." This will help get a conversation flowing which can potentially result in a friendship. Therefore, the correlation was determined between experiencing the virtual coach impersonal and the relation with the virtual coach. In Figure 9 the values (r_ϕ) can be viewed.
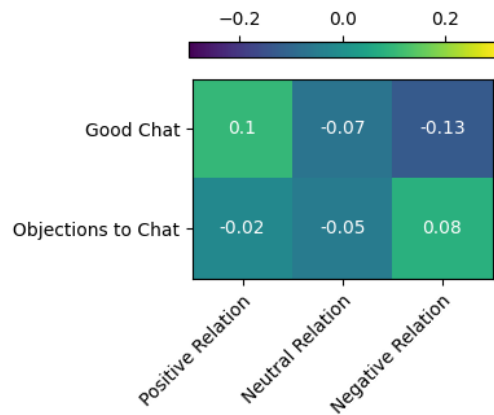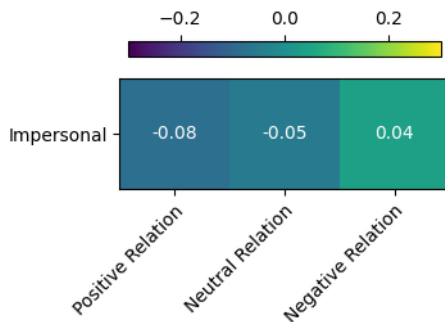

Figure 9: An heatmap of the correlations - expressed in r_ϕ - between the theme impersonal and the subthemes of relation.

Encountering the virtual coach impersonal resulted in an insignificant ($p > 0.1$) correlation with the relations experienced with the virtual coach. Correlation analysis was performed between impersonal and the participants' ratings to support this presumption. Here, considering the virtual coach impersonal caused a significant *acceptable* decrease in the rating ($r_{pb} = -0.26$, $p < 0.001$). The virtual coach said

to be *"a one size fits all"* (P8f), who *"didn't know anything about me other that I smoke"* (P8g) and *"I can't see, don't know"* (P8h). Hence, it did not *"feel personal enough to call it a friend"* (P8i).

Thus, there is evidence that not experiencing the virtual coach impersonal is beneficial in developing a closer relationship with the virtual coach. Mattar and Wachsmuth concluded that a personalised chat led to greater satisfaction with the conversation and a more positive attitude towards the agent [32]. However, this requires further investigation as, in this research, not all findings provided support.

3.6 Impact of the user's age on the relation

"Helpful tool but I don't life talking to computers" (P11a)

Research has been performed by Vaportzis *et al.* to investigate the relationship between older adults and technology [33]. Here, it has been concluded that older adults are often reluctant to use technology due to barriers, concerns, and scepticism. Therefore, correlations were calculated between three different age groups and the relation experienced with the virtual coach. The three age groups that were used are: young adults (18 to 35 years; $n = 289$), middle-aged adults (36 to 55 years; $n = 171$) and older adults (56 years and older; $n = 40$) [34]. In Figure 10 the values (r_ϕ) can be found.
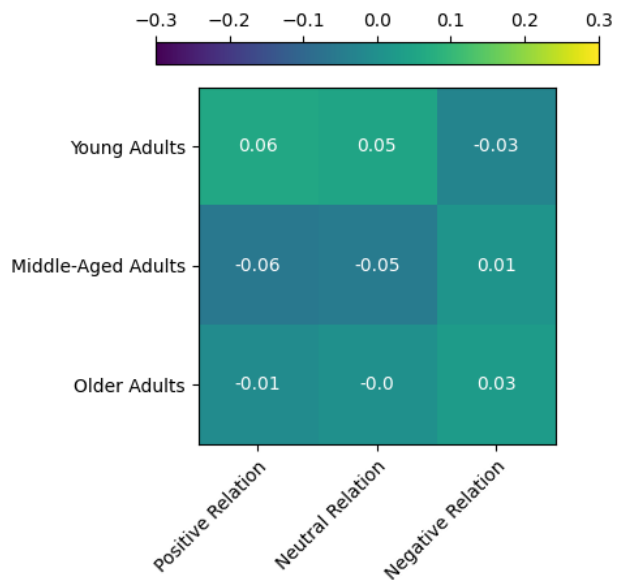
Figure 10: An heatmap of the correlations - expressed in r_ϕ - between the age groups and the subthemes of relation.

As shown in Figure 10, insignificant ($p > 0.1$) correlations were found. However, the 63-year-old participant P11a mentioned that he does not *"life talking to computers."* Whereas the 25-year-old participant P11b said to *"like talking to AIs."* Therefore, correlation analysis was performed between the rating and the age groups. This was done to investigate further the influence of the participant's age on the relation experienced with the virtual coach. Here, the age

group had a negligible impact on the rating. In Figure 11 the values (r_{pb}) can be viewed.
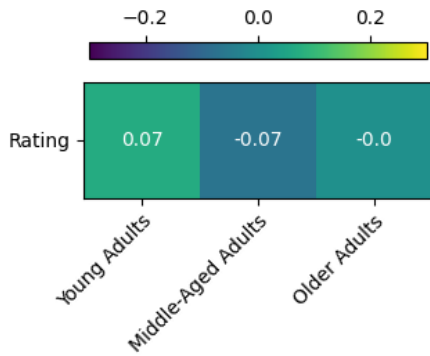

Figure 11: An heatmap of the correlations - expressed in r_{pb} - between the rating and the age groups.

Above that, the user's age caused *no* or *negligible* decrement in the rating given by the participants ($r_s = -0.03$). Hence, the user's age minorly influenced the development of a closer relationship with the virtual coach.

4 Responsible Research

This section reviews the research data (Section 9.1). Subsequently, the ethical aspects and scientific integrity of this research are reflected (Section 9.2 & 9.3). Lastly, the reproducibility of the analysis is discussed (Section 9.4).

4.1 Research Data

This research was based on data collected by Albers *et al.*, which received ethical approval from the TU Delft University Human Research Ethics Committee [35]. Subsequently, data trimming and falsification have been avoided by not making modifications to the dataset. Additionally, the inclusion and exclusion criteria for the data responses have been described in the methodology (Section 2). Above that, data fabrication has been prevented by reporting negative results. In addition, p-hacking averted by justifying the correlation with a presumption expressed as a quote and searching for supportive literature beforehand.

4.2 Ethical Aspects

Personal data was used during this research. Therefore, it was of great importance that this data was not leaked. Hence, the data was provided through a shared folder on Microsoft Teams, solely accessible to the research members. Above that, data was not shared outside the research team. Also, leakage risks were minimised by the supervisor by providing the data with random identifiers instead of the participant names.

4.3 Scientific Integrity

During this research, collaboration occurred with the research members N. Aretz, J. Dechering, A. Ekinci, and O. Sheasha. Each research has contributed to eliciting the reasons for accepting the virtual coach, which helps during smoking cessation. N. Aretz and A. Ekinci assisted in the performance of double coding. Additionally, the pre-processing results were shared to ensure that the correlation analysis happened with the same values.

Attempting to ensure transparency, the initial and final coding schemes have been provided in Appendix A, Figure 1. Here, the changes to the coding scheme have been indicated. Additionally, the frequency of the codes can be found in Appendix B, Figure 2. Above that, the initial coding, the coding performed by the second researcher, and the final coding have been archived in the 4TU.ResearchData repository [36]. Furthermore, the removal of the codes and the finalisation of the themes have been described in the methodology (Section 2). Moreover, plagiarism has been avoided by citing all referred literature.

4.4 Reproducibility

It was strived to guarantee reproducibility by uploading source code to the 4TU.ResearchData repository [36]. These have been used to calculate the Cohen's Kappa and correlation coefficients and generate the heatmaps. Additionally, instructions have been uploaded on how functions in excel have been used to perform calculations and obtain data. Above that, the steps taken during the analysis have been described in the methodology (Section 2).

5 Discussion and Limitation

This research was built upon data. While executing quantitative analysis, data responses were encountered where the participants did not fill in all the questionnaire entries. In retrospect, these responses should not have been included in the data. Then, they would not be taken into account during the thematic analysis. However, as the data was pre-collected, it was possible to conduct this research in a limited time.

While coding the responses, the rating was not considered. Considering the rating would improve the theme's reliability on how the relationship with the virtual coach was experienced. Yet, after performing the correlation analysis between the subthemes of relation and another data source, a correlation analysis was done between the rating and the data source. Hence, this provided a check to the priorly obtained correlation.

During the final coding scheme discussion, the coding scheme was reduced. Codes were eliminated based on low agreement, low frequency, or improper description. The removal of infrequently used codes caused a loss of information. But, it was checked whether another code implied these codes while removing them. The alternative code was assigned to avoid losing the response's inclusion under the theme in such cases. Besides that, these scarce codes often caused outliers, negatively impacting the correlation analysis.

An alternative thematic map was used at the second level of the third phase of thematic analysis. Here, according to Braun and Clark, the final themes should be such that each data response corresponds to one theme [15]. However, it was decided to use an alternative thematic map where a data

response can belong to one or more themes. This decision was made because the data responses happen to be sparse. Therefore, classifying them into only one theme would cause the loss of valuable information.

Spearman's rank correlation coefficient indicated the correlation between the rating and age of the participant. It was decided to utilise Spearman instead of Pearson as it was unclear whether a monotonic or linear relationship existed between the variables. This was because there would be no information loss when applying Spearman on a relationship that happened to be perfectly linear, whereas using Pearson on a non-linear relation, information would be lost [37].

This analysis has limitations. Namely, the obtained Cohen's Kappa values do not depict accurate reliability. This is because the coding scheme was changed during the discussion of the final coding. Here, codes were removed, merged, and renamed. Hence, an extra round of double coding should be performed with a third researcher to improve this research. This third researcher must not be involved during the first round of double coding. The reason for this is that it can lead to biased values. In short, an uninvolved third researcher should perform the double coding with the updated coding scheme to obtain more accurate Cohen's Kappa values.

6 Conclusions and Future Work

This research aimed to identify what characteristics the virtual coach must possess to be considered a friend. For this purpose, thematic analysis was executed with researcher and method triangulation. Here, it has been concluded that the user is more likely to see the virtual coach as a friend when it is perceived as human by the user. Additionally, the user can develop a closer relationship with the virtual coach when a good conversation can be held with the virtual coach, and the positive characteristics of the virtual coach stand out. On the contrary, the user's age has a negligible influence on the relationship with the virtual coach.

This research has shown that having the feeling that the virtual coach is human, pleasantly conversing with the virtual coach, and finding the character of the virtual coach positive, leads to a closer relationship between the virtual coach and the accompanying. However, further research is needed on what characteristics the virtual coach must possess to be perceived as human. Besides, an investigation must follow what makes the conversation with the virtual coach and his character pleasant.

Above that, more research is required on whether impersonality affects the relationship with the virtual coach. This research found results coinciding with literature on personalising chats to develop a closer relationship. However, as not all findings provided support, further investigation is required.

References

[1] W. H. Orginazation, "Tobacco." https://www.who.int/news-room/fact-sheets/detail/tobacco#:~:text=Tobacco%20kills%20more%20than%208, exposed%20to%20second%2Dhand%20smoke., July 2021. Visited on : 17 April 2022.

[2] M. Milcarz, K. Polanska, L. Bak-Romaniszyn, and D. Kaleta, "Tobacco health risk awareness among socially disadvantaged people - a crucial tool for smoking cessation," *International Journal of Environmental Research and Public Health*, 2018.

[3] K. Buczkowski, L. Marcinowicz, S. Czachowski, and E. Piszczek, "Motivations toward smoking cessation, reasons for relapse, and modes of quitting: results from a qualitative study among former and current smokers," *Patient Preference and Adherence*, 2014.

[4] J. Balmford, R. Borland, and P. Benda, "Patterns of use of an automated interactive personalized coaching program for smoking cessation," *Web-Assisted Tobacco Interventions*, 2008.

[5] G. Eysenbach, "What is e-health?," *Journal of Medical Internet Research*, 2001.

[6] M. J. Hashim, "User interactivity in ehealth applications: A novel taxonomy," *12th International Conference on Innovations in Information Technology (IIT)*, 2016.

[7] J. Grolleman, B. van Dijk, A. Nijholt, and A. van Emst, "Break the habit! designing an e-therapy intervention using a virtual coach in aid of smoking cessation," *Human Media Interaction*, 2006.

[8] M. E. Kamali, L. Angelini, M. Caon, D. Lalanne, O. A. Khaled, and E. Mugellini, "An embodied and ubiquitous e-coach for accompanying older adults towards a better lifestyle," *International Conference on Human-Computer Interaction*, 2020.

[9] N. Albers, M. A. Neerincx, and W.-P. Brinkman, "Reinforcement learning-based persuasion by a conversational agent for behavior change," *33rd Benelux Conference on Artificial Intelligence and 30th Belgian-Dutch Conference on Machine Learning*, 2021.

[10] "Pefect fit." https://perfectfit-onderzoek.nl/informatie, 2022. Visited on : 24 May 2022.

[11] N. Albers and W.-P. Brinkman, "Perfect fit - experiment to gather data for and test a reinforcement learning-approach for motivating people," 2021.

[12] G. Peersman, *An Introduction to Triangulation*.

[13] V. A. Thurmond, "The point of triangulation," *Journal of Nursing Scholarship*, 2001.

[14] T. D. Jick, "Mixing qualitative and quantitative methods: Triangulation in action," *Administrative Science Quarterly*, 1979.

[15] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Research designs: Quantitative, qualitative, neuropsychological, and biological*, 2012.

[16] M. S. Linneberg and S. Korsgaard, "Coding qualitative data: a synthesis guiding the novice," *Qualitative Research Journal*, 2019.

[17] R. E. Boyatzis, *Transforming Qualitative Information: Thematic Analysis and Code Development*.

[18] B. F. Crabtree and W. F. Miller, "A template approach to text analysis: Developing and using codebooks," *Doing qualitative research*, 1999.

[19] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, 1977.

[20] N. F. Habidin, A. F. M. Zubir, N. M. Fuzi, N. A. M. Latip, and M. N. A. Azman, "Sustainable performance measures for malaysian automotive industry," *World Applied Sciences Journal*, 2015.

[21] H. Akoglu, "User's guide to correlation coefficients," *Turkish Journal of Emergency Medicine*, 2018.

[22] A. A. Dichoso and R. J. M. Cabauatan, "Test item analyzer using point-biserial correlation and p-values," *International Journal of Scientific & Technology Research*, 2020.

[23] C. P. Dancey and J. Reidy, "Statistics without maths for psychology," *Pearson Education*, 2007.

[24] G. Raiola and P. A. di Tore, "Statistical study on bodily communication skills in volleyball to improve teaching methods," *Journal of Human Sport and Exercise*, 2022.

[25] L. Wagner, "Good character is what we look for in a friend: Character strengths are positively related to peer acceptance and friendship quality in early adolescents," 2019.

[26] D. Hoyos-Valdés, "The notion of character friendship and the cultivation of virtue," *Journal for the Theory Social Behaviour*, 2018.

[27] C. Becker, S. Kopp, and I. Wachsmuth, "Why emotions should be integrated into conversational agents," *Conversational Informatics: an Engineering Approach*, 2007.

[28] T. Araujo, "Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions," *Computers in Human Behavior*, 2018.

[29] G. S. Mesch and I. Talmud, "Online friendship formation, communication channels, and social closeness," *International Journal of Internet Science*, 2006.

[30] M. M. V. Pinxteren, M. Pluymaekers, and J. G. Lemmink., "Human-like communication in conversational agents: a literature review and research agenda," *Journal of Service Management*, 2020.

[31] D. Gabor, *How To Start A Conversation And Make Friends*.

[32] N. Mattar and I. Wachsmuth, "Let's get personal assessing the impact of personal information in human-agent conversations," *International Conference on Human-Computer Interaction*, 2014.

[33] E. Vaportzis, M. G. Clausen, and A. J. Gow, "Older adults perceptions of technology and barriers to interacting with tablet computers: A focus group study," *Frontiers in Psychology*, 2017.

[34] N. M. Petry, "A comparison of young, middle-aged, and older adult treatment-seeking pathological gamblers," *The Gerontologist*, 2002.

[35] N. Albers, M. A. Neerincx, and W.-P. Brinkman, "Acceptance of a virtual coach for quitting smoking and becoming physical active: Dataset," 2022. DOI: 10.4121/19934783.v1.

[36] M. Ali, "Traits for a virtual coach to be a friend," 2022. DOI: 10.4121/20099102.

[37] R. Artusi, P. Verderio, and E. Marubini, "Bravais-pearson and spearman correlation coefficients: meaning, test of hypothesis and confidence interval," *The International Journal of Biological Markers*, 2002.

APPENDIX

A Coding Schemes

Initial Coding Scheme

The Relation With The Virtual Coach

Code	Kappa
Good Relation	1.00
A Friend	0.83
Regular Relation	0.50
Acquaintances	0.40
More Open to See Sam as a Friend	0.20
More than a Stranger	0.39
Strictly-Business Relation	0.86
Not a friend nor a stranger	0.67
Not a friend	0.60
Not a close relation	0.26
No relation	0.51
No desire to be friends	0.33
Unimaginable to form a relation	0.63

The Perception of the Virtual Coach

Code	Kappa
Human-Like	0.64
Artificial but Human-Like	0.48
Clearly Artificial	0.79

Objections about the setup of the Experiment

Code	Kappa
Anonymity	1.00
Frequent User of Setup	0.00
Not enough interaction	0.34
More time required	0.69

Opinions about the Chat

Code	Kappa
Good Chat	0.58
Chat with "A Friend"	0.59
Responds Good	0.49
Limited Answer Choice	0.89
Scripted	0.73
Monotonous	0.22

The Virtual Coach is being experienced as impersonal

Code	Kappa
Personal	0.49
Do not know Sam	0.43
Does not know "me"	0.55
Sam has no "face"	0.80
Sam has no emotions	0.89
Sam has no mind	0.66

General Opninions

Code	Kappa
Enjoy	0.73
Positive Feeling	0.89
Good Job	0.61
Easy	0.86
Relaxing	0.80
Interesting	1.00
Decent	0.61
Not Unique	0.00
Unnatural	0.00

Characteristics of the Virtual Coach

Code	Kappa
Friendly	0.96
Guidance	0.58
Trustworthy	0.66
Warm	0.92
Caring	0.84
Welcoming	0.45
Unprejudiced	0.40
Sympathetic	1.00
Informal	0.91
Lend an Ear	0.67
Comforting	0.00
Understanding	0.71
Polite	1.00
Patient	1.00
Thoughtful	0.50
Helpful	0.94
Motivating	1.00
Supportive	0.72
Persuasive	0.67
Informative	0.44
Straightforward	1.00
Professional	0.63
Personality Clash	0.66
Sounds Irritated	0.00
Unhelpful	1.00

Implied by Strictly-Business Relation

Final Coding Scheme

The Relation With The Virtual Coach

Positive Relation
- Good Relation
- A Friend
- More than a Stranger

Neutral Relation
- Strictly-Business Relation
- Not a friend nor a stranger

Negative Relation
- Not a friend
- No relation
- Unimaginable to form a relation

The Perception of the Virtual Coach

Human-Like

Artificial but Human-Like

Clearly Artificial

Opinions about the Chat

Good Chat
- Good Chat
- Chat with "A Friend"
- Responds Good

Objections to the Chat
- Limited Answer Choice
- Scripted
- Monotonous

The Virtual Coach is being experienced as impersonal

- Do not know Sam
- Does not know "me"
- Sam has no "face"
- Sam has no emotions
- Sam has no mind

Positive Characteristics of the Virtual Coach

- Friendly
- Guidance
- Trustworthy
- Warm
- Caring
- Welcoming
- Unprejudiced
- Sympathetic
- Informal
- Understanding
- Polite
- Patient
- Thoughtful
- Helpful
- Motivating
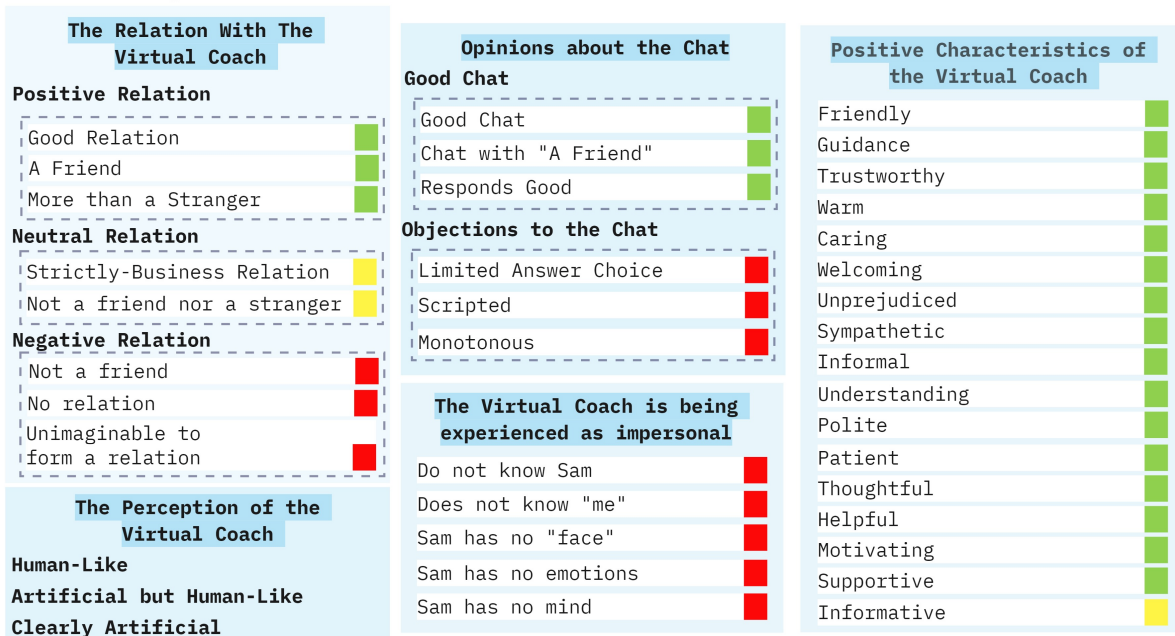- Supportive
- Informative

Figure 1: Above is the initial coding scheme. Three colour boxes have been used where green should be interpreted as positive, yellow should be interpreted as neutral, and red should be interpreted negatively. Cohen's Kappa for each code can be found in the circle on the left. The codes with a red Cohen's Kappa have been eliminated. The arrows show which codes have been merged. Below the final coding scheme can be found.

B   Frequecies of the codes
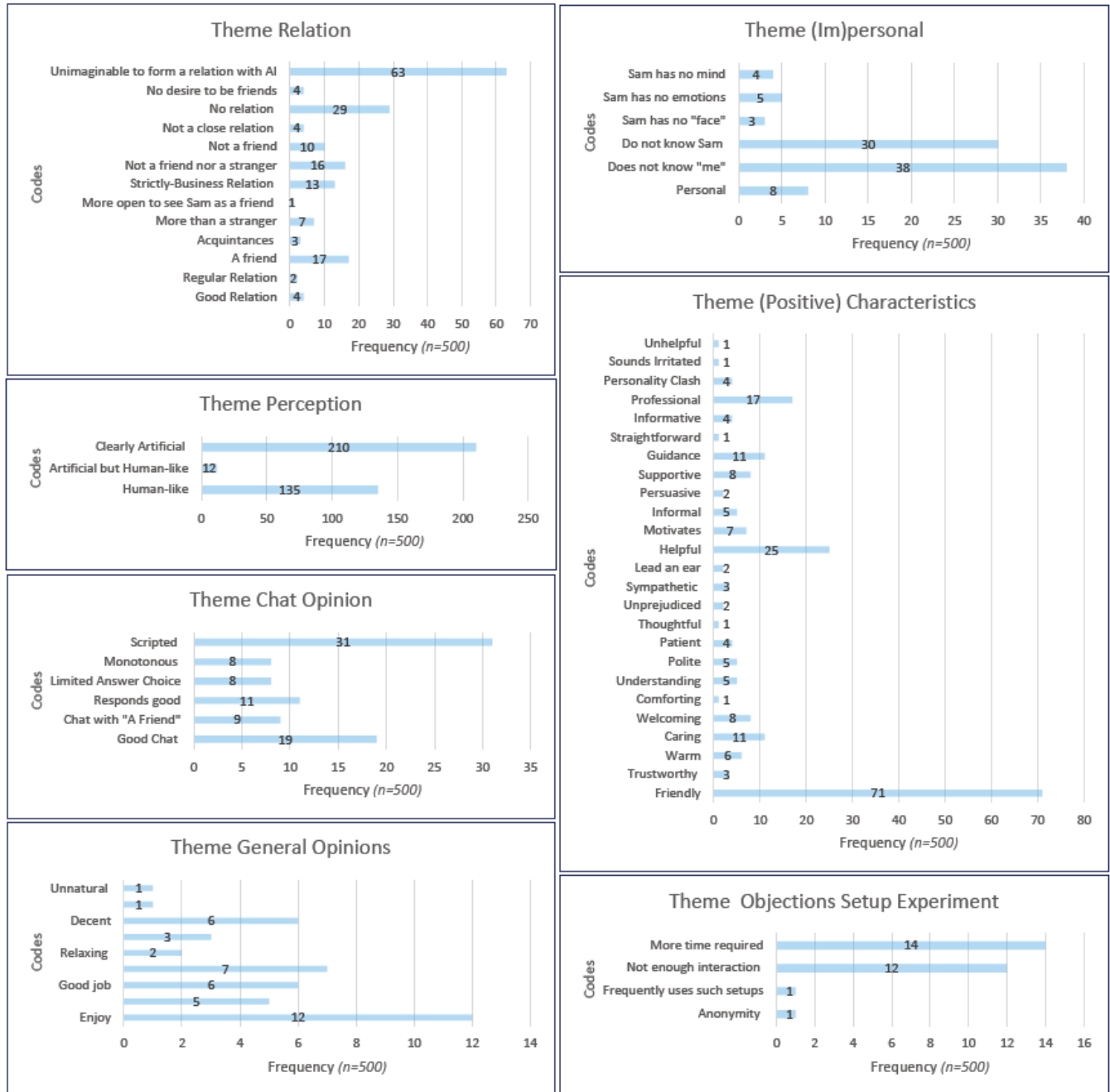

Figure 2: The number of occurrences of the codes ($n = 500$).

C   Interpretation Tables of the Correlation Coefficients

Table 1: Interpretation of the Phi Coefficient (r_ϕ) according to the guidelines by Akoglu. [21]

Phi (r_ϕ)	Interpretation
> 0.25	Very Strong
> 0.15	Strong
> 0.10	Moderate
> 0.05	Weak
> 0	No or Very Weak

Table 2: Interpretation of the Point Biserial Correlation Coefficient (r_{pb}) according to the guidelines by Dichoso and Cabauatan. [22]

Point Biserial Correlation Coefficient (r_{pb})	Discrimination
$0.40 - 1.00$	Very Good
$0.30 - 0.39$	Reasonably Good
$0.20 - 0.29$	Marginal or Acceptable
$0.00 - 0.19$	Poor Discrimination

Table 3: Interpretation of the Spearsman's Rank Correlation Coefficient (r_s) according to the guidelines by Dancey and Reidy. [23]

Spearsman's Rank Correlation Coefficient (r_s)	Correlation
> 0.69	Very Strong Relationship
$0.40 - 0.69$	Strong Relationship
$0.30 - 0.39$	Moderate Relationship
$0.20 - 0.29$	Weak Relationship
$0.01 - 0.19$	No or Negligible Relationship

D   List of Quoted Participants

Table 4: Participant identifiers with their corresponding Random ID.

Participant ID	Random ID	Participant ID	Random ID
P0	53E85C52175JD	**Theme Impersonal**	
Subtheme Positive Relation		P8a	93J49D23272UM
P1a	67J61Z33454PR	P8b	69G67A38952TK
P1b	21F55Z66983BW	P8c	98S96T85477LJ
P1c	78E87Q13350QR	P8d	36U57P84141JV
Subtheme Neutral Relation		P8e	67L20J48800FY
P2a	69R83M66494PM	P8f	14T14I79235FK
P2b	11N20D16896EB	P8g	12G95W21197WJ
Subtheme Negative Relation		P8h	36U57P84141JV
P3a	54J99Y98708XC	P8i	49D37C18043VE
P3b	16M21V26693RY	**Subtheme Good Chat**	
P3c	56V34O67163HJ	P9a	70M80Y38484YX
Theme Positive Characteristics		P9b	63O32F71435DQ
P4a	16Z15Z20553NF	P9c	88B14S69830KL
P4b	36Y61A25101BA	P9d	19G98C51472ET
P4c	57I46I21490RT	P9e	42W96N11133WQ
P4d	12O34D29417JU	P9f	63O32F71435DQ
P4e	46F56A31181VL	P9g	68T14J46648HM
P4f	17E69T57249RA	P9h	94C29S54433WZ
P4g	25B86K37242FN	P9i	88B14S69830KL
Theme Human-Like		**Subtheme Objections to the Chat**	
P5a	26C68F88241QZ	P10a	36Z50Q84163AL
P5b	52Z39P53026PV	P10b	99S56F71330CK
P5c	21F31X44258SP	P10c	52S32B30914FK
P5d	26L82L20194AD	P10d	16L28Y79565MD
P5e	49D34H64237US	P10e	63K23V64183WA
P5f	81F14B16974DZ	P10f	21C26P83665LR
Theme Artificial but Human-Like		P10g	91X75W89915CV
P6a	12D25K46543FF	P10h	71S66Z90667KY
Theme Clearly Artificial		**Age**	
P7a	11B82Y86792KT	P11a	53E85C52175JD
P7b	21K51B22265AZ	P11b	63W84N62639MJ
P7c	94C29S54433WZ		
P7d	65V33F28647PQ		
P7e	25R66I80185HQ		
P7f	16M21V26693RY		