Article

Long-Term Effects of Perceived Friendship with Intelligent Voice Assistants on Usage Behavior, User Experience, and Social Perceptions

Carolin Wienrich [1,*], Astrid Carolus [2], André Markus [1], Yannik Augustin [2], Jan Pfister [3] and Andreas Hotho [3]

1   Institute Human-Computer-Media, Psychology of Intelligent Interactive Systems, University of Wuerzburg, Oswald-Külpe-Weg 82, 97074 Wuerzburg, Germany
2   Institute Human-Computer-Media, Media Psychology, University of Wuerzburg, Oswald-Külpe-Weg 82, 97074 Wuerzburg, Germany
3   Data Science, Institute for Computer Sciences, University of Wuerzburg, 97074 Wuerzburg, Germany
*   Correspondence: carolin.wienrich@uni-wuerzburg.de; Tel.: +49-931-31-88051

Abstract: Social patterns and roles can develop when users talk to intelligent voice assistants (IVAs) daily. The current study investigates whether users assign different roles to devices and how this affects their usage behavior, user experience, and social perceptions. Since social roles take time to establish, we equipped 106 participants with Alexa or Google assistants and some smart home devices and observed their interactions for nine months. We analyzed diverse subjective (questionnaire) and objective data (interaction data). By combining social science and data science analyses, we identified two distinct clusters—users who assigned a friendship role to IVAs over time and users who did not. Interestingly, these clusters exhibited significant differences in their usage behavior, user experience, and social perceptions of the devices. For example, participants who assigned a role to IVAs attributed more friendship to them used them more frequently, reported more enjoyment during interactions, and perceived more empathy for IVAs. In addition, these users had distinct personal requirements, for example, they reported more loneliness. This study provides valuable insights into the role-specific effects and consequences of voice assistants. Recent developments in conversational language models such as ChatGPT suggest that the findings of this study could make an important contribution to the design of dialogic human–AI interactions.

Keywords: intelligent voice assistant; smart speaker; social relationship; social role; long-term analysis; social interaction; human–computer interaction; anthropomorphism

1. Introduction

Intelligent voice assistants, or IVAs, which are integrated into smart speakers such as Amazon's Alexa, are rapidly gaining popularity and becoming an integral part of everyday life [1,2]. These devices can recognize voice-based requests, respond with human-like speech, and assist users with various tasks [3]. Some users consider voice assistants a useful tool, whereas others have formed closer relationships or even friendships with them [4]. As the social relationships with voice assistants become more established, their impact on interactions, user experience, and self-disclosure becomes more apparent [5–7]. For example, a study by Wu, et al. [8] showed that users assign different roles to IVAs and that this attribution, in turn, determines the expectations and usage patterns. Consistent with other research [9,10], Wu, He, Peng, Li, Zhou, and Guan [8] found that one of the most common roles assigned to IVAs was that of a "friend". Therefore, the present study precisely focuses on the attribution of IVAs as friends.

However, social roles and interactions are established over time in natural interaction environments. Therefore, longitudinal studies are valuable in understanding how the

perceived role influences usage behavior and social perceptions of IVAs. They also give users the necessary time to establish a social relationship with the device [11]. Despite the growing importance of the relationship between humans and IVAs, there is a lack of longitudinal studies that focus on the social aspects of this relationship. Most longitudinal studies focus on describing usage patterns and behaviors [12,13], with relatively little attention given to the social aspects of human–technology interactions [14]. Since social roles on technologies and the attribution of social counterparts are often perceived unconsciously, using mainly explicit measurement methods falls short or even causes reactance [15–17]. Long-term studies also allow for a multifaceted capture of interactions, such as continuous interaction data.

In summary, initial short-term studies show that social role attribution affects users' expectations and use of IVAs. However, there is a lack of targeted long-term studies that include both subjective and objective data, as well as a temporal course of the phenomenon. Thus, the following research questions arise: (1) Do participants perceive IVAs as friends? (2) How does the perceived social role of IVAs as a friend influence usage behavior, user experience, and social perception over time (3) Does the user's personality influence the attribution of friendship to IVAs? To investigate this, our participants interacted with a common IVA for nine months in their homes. We obtained self-report questionnaires and analyzed continuous interaction data by using social science and data science methods. Our multi-method approach aims to uncover differences in usage behavior as a function of the social role attribution of the device and visualizes changes over time. Thus, the present work differs from previous studies as it determines whether users differ in their attribution of friendship to voice assistants by examining role-specific interaction patterns using an interdisciplinary multi-method approach. We also explore the personality traits that may favor a perception of friendship with an IVA and analyze interactions with IVAs in natural environments from a long-term perspective.

1.1. Related Work

Anthropomorphism, or the attribution of human characteristics to non-human beings, has been the subject of extensive research in psychology and the study of human–computer interaction. Previous research has shown that people tend to anthropomorphize inanimate objects under certain circumstances, which may extend to other areas of human cognition, attitudes, and even behavior [18]. This section reviews social attributions and their impact on users' interactions with technological entities. Following the definitions of Russell and Norvig [19], in this study, we focus on voice-based AI systems and use the term "IVA" to refer to intelligent voice-based assistants, with smart speakers (e.g., Amazon Echo, Google Home) and voice-controlled intelligent personal assistants (e.g., Amazon Alexa, Google Assistant) being the most popular examples.

1.1.1. Voice Assistants as Social Actors in User Relationships

The media equation suggests that humans tend to attribute social and cognitive characteristics to technological entities, treating them as social actors [20] due to the human-like characteristics or behaviors that activate users' social scripts [20]. This has been found to be the case with PCs, smartphones [21], and even websites [22]. IVAs, unlike other technologies, have particularly high social affordances and anthropomorphic attributes with their conversational speech interaction and names [23–25]. Most noticeable is the personification of smart speakers or IVAs by users. Often, users personify these systems by using human pronouns and the respective name of the IVA instead of the device name for smart speakers [9,10,26]. The social perception of IVAs is also evident in effects that otherwise only in human–human interaction. For instance, Liu and Pu [5] found that voice assistants induce a social facilitation effect. In their experiment, they studied the effect of the presence of a smart speaker on the solution speed of easy or complex tasks. The results indicate that participants solved easy tasks faster in the presence of the smart speaker, whereas they reacted slower in difficult tasks. Other works studied emotional-affective

responses toward IVAs. Carolus, Wienrich, Törke, Friedel, Schwietering, and Sperzel [6] found that observers of voice assistants feel more empathy for the assistant when they see the IVA being treated rudely. Furthermore, computer voices have been observed to increase respondents' socially desirable response behavior and encourage the disclosure of sensitive information [27,28]. Wienrich, Reitelbach, and Carolus [7] also found that voice assistants promote the disclosure of sensitive information the more trustworthy and competent they are perceived.

As research shows, IVAs can elicit social responses and behaviors in individuals. However, does this also mean that we ascribe concrete social roles to voice assistants? If so, to what extent does the perceived social role influence the usage behavior, user experience, and social perceptions in the long term once a social relationship has developed between the IVA and the user?

1.1.2. Social Roles of IVAs in Relationships with Users

Social roles are assigned to holders of certain positions or functions in social contexts, which are accompanied by demands of their behavior and character. Dreitzel [29] defines a social role as a set of expectations that are attached to the behavior of the holder in interaction situations. In other words, different expectations are attributed to the holder depending on the assigned social role. In addition, previous studies on AI-based technologies found that users attribute social roles to technology, which can affect their expectations and perceptions [8,30]. For example, users of IVAs often assign them either a tool-based role (e.g., assistant, service provider, tool) or a friend-based role (e.g., friend, family member, companion) [31]. Users expect helpful, cooperative, and understanding behavior from tool-based IVAs and respect, an emotional connection, and loyalty from friend-based IVAs. In particular, users may develop friendship-like feelings toward IVAs and associate them with social roles that can influence their interactions [32,33]. When users perceive IVAs as friends, they feel more sympathy toward them [33]. Feelings of friendship toward IVAs further motivate users to speak patiently and more slowly, thereby increasing the understanding of voice input [33]. The perception of an IVA as a friend can also positively influence attitudes toward products in the context of voice shopping [34] and products are liked more when IVAs provide product recommendations in the role of a friend.

Parasocial relationship theory can explain the use of IVAs, as they convey closeness and intimacy, positively influencing perceptions of the voice assistant as a friend, usage intention, and satisfaction [35,36]. In general, a parasocial relationship is defined as the extent to which a media consumer has developed a social relationship with a medium [37]. Parasocial relationships with media entities develop over time [38–40] and can motivate consumption [41] and influence recipient activity [42]. In the case of voice assistants, Hsieh and Lee [43] describe parasocial relationships as a key factor for perceived ease of use and future intention to use IVAs. Numerous findings support this assumption and show that the stronger the social relationship between IVAs and users, the higher the perceived user experience [44–46]. When there is a strong relationship between IVAs and users, it can encourage acceptance and social presence [24,47], as well as exploratory usage behaviors [48]. Technologies that are perceived as social and relational can result in a more positive, user-friendly, and useful user experience [44,45].

However, previous studies often rely on short-term measures, although establishing appropriate relationships takes time and usually occurs in natural environments such as homes [11,14,38]. Consequently, some researchers conducted longitudinal studies [11,49]. Gao, Pan, Wang, and Chen [49] found that some users assigned different roles to IVAs, most of them being human roles with positive emotions, whereas others attributed more impersonal roles with less positive emotions. Voit, Niess, Eckerth, Ernst, Weingärtner, and Woźniak [11] found that some users viewed smart speakers as social agents and described a social relationship with them. In contrast, others viewed them as technical tools and distanced themselves from them. Some results show that the context of the IVA's use influences the nature of the interactions and their associated emotions. Thus, these initial

studies suggest two types of roles—personal and tool-like. However, it is difficult to capture the social roles that users associate with IVAs. The attribution of social roles or human characteristics to technological devices often occurs unconsciously [15,16], which makes it hard to assess the social role using explicit measurement methods. Users often deny the social and friendship roles they attribute to AI systems if directly and explicitly asked about them [17]. Therefore, incorporating more implicit measures could complement the measurement of role-specific relationships between users and IVAs and provide insights into role-specific interaction styles. The continuous interaction data recorded and stored by voice assistant providers such as Amazon and Google may be valuable [14]. However, to date, no study has evaluated continuous interaction data and systematically examined it in terms of the social roles attributed to IVAs and the trends over time.

Furthermore, previous long-term studies have not explored the role of the user's personality in determining which social role is attributed to an IVA. Whelan et al. [50] showed that anthropomorphizing tendencies are particularly salient in individuals with insecure attachment styles and attachment anxiety. Epley, Waytz, and Cacioppo [18] suggested that people with insecure attachment styles tend to anthropomorphize to compensate for unmet interpersonal attachment needs. In the field of IVAs, it has been shown that use over a more extended period can reduce situational feelings of loneliness among users [51,52], as IVAs provide a social presence [53] that can fulfill certain social needs and provide the sense of being in a human company [18,54]. In the context of voice assistants, there is currently no empirical evidence to show how users' personality traits and attachment styles relate to social perceptions of IVAs and influence the development of feelings of friendship toward them. Similarly, although previous studies have shown that IVAs can reduce feelings of loneliness [52], they do not specifically address the role-specific significance of IVAs and are restricted to self-report methods for describing interaction styles.

1.2. Summary and Present Study

Voice assistants are a widely used technology with a broad social appeal that can influence users' perceptions and behaviors based on the social role assigned to them (e.g., assistant-like role or friend-like role). Research has also shown that the attribution of social roles and the corresponding responses are often unconscious, with users sometimes denying assigning such roles when explicitly asked about them [15–17,55]. Moreover, social processes evolve over time and in natural contexts of use [14]. In contrast to previous research approaches, our study considers the temporal conditions in which friendships develop between users and IVAs and examines them in a natural context of use (i.e., home). We conduct a longitudinal study and analyze the development of role-specific effects and relationships between users and IVAs over nine months. For users who attributed the role of friendship, we evaluate both explicit and implicit data, including the use of questionnaires and the assessment of continuous interaction data. By utilizing a multimethod approach that integrates both data and social sciences analyses, we investigate the degree to which role attribution influences usage behavior, user experience, and social perception. We also analyze the personality traits that impact role attribution. Thus, our results contribute to a better understanding of how the social perception of IVAs can affect human–AI interactions. This may become even more important in the future as AI becomes more adaptive, intelligent, and social, which is associated with greater opportunities [4] and risks [56]. Using a long-term study in a natural context of use, our research provides valuable insights into role-specific interactions with IVAs over time. Overall, this research offers valuable perspectives into the evolving relationship between users and IVAs and has implications for future developments in this field.

Structure of Present Study

Our work takes a broader exploratory approach. Rather than formulating hypotheses, we developed a set of research questions to gain a deeper understanding of the role-specific effects of friend-like IVAs at multiple levels including usage behavior, user experience, and

social perception). To provide a better overview, the Methods and Results sections are divided into five sections, each addressing one of the following research questions:

Section 1—Cluster Formation and Time Effects:

Do users differ in their attribution of friendship to IVAs and how does perceived friendship quality change over time?

Section 2—Usage Behavior:

How do patterns of use differ as a function of perceived friendship with the IVA and how do they change over time?

Section 3—User Experience:

How does user experience differ as a function of perceived friendship with the IVA and how does it change over time?

Section 4—Social Perception:

How do social perceptions differ as a function of perceived friendship with the IVA and how do they change over time?

Section 5—Personality Traits:

To what extent do personality traits of users differ in attributing vs. not attributing a friendship role to IVAs?

2. Method

2.1. Participants

A total of 106 students who reported not owning a smart speaker were included in this study. Participants were excluded from the 9-month longitudinal study if they already owned a device with a voice assistant before the study ($n = 5$), participated in the study unreliably ($n = 5$), or left the university ($n = 1$). Finally, longitudinal study data were collected from 85 participants who ranged in age from 17 to 23 years ($M = 19.42$, $SD = 1.37$) and were predominantly female ($n = 70$; $n = 7$ male; $n = 8$ diverse). Because sample sizes differed for some variables, the gender composition of the samples is reported at the level of the constructs, as shown in Table S1.

2.2. Procedure

At the beginning of the study, half of the sample was equipped with a Google Home Mini and an Amazon Echo Dot. In addition, participants were given other smart home devices such as a smart socket or a light bulb. Participants were instructed to install all devices within one week. After successful installation, different channels (e.g., the messenger service, Telegram) were used to ensure anonymous and low-threshold communication with the investigators. Participants were given randomized, anonymous subject codes to create the email addresses used to register the devices with Google or Amazon. The log files, which were used to generate user logs and analyze user behavior, were also linked to this e-mail address.

The long-term study was divided into (1) the installation stage, (2) the free interaction stage, (3) the intervention stage, and (4) the interview stage (Figure 1). Usage behavior, user experience, and social perception variables in the context of IVA usage were collected over 15 time points (short "T"). In the installation stage, the devices were distributed to participants and prepared for use. In the free interaction stage, we analyzed the unrestricted interactions with the devices. In the intervention stage, we conducted experimental interventions to study their effect on usage behavior, user experience, and social perception. Intervention 1 provided a deeper understanding of the functions of IVAs. In intervention 2, participants were instructed to play games over their IVAs. In intervention 3, participants used the TK smart relaxation skill. In the interview stage, participants took part in a structured interview (11 main questions, 14 follow-up questions) that was tape-recorded. Questions were related to, for example, assessment of usage, acceptance of the IVA, and

its perception as a social interaction partner. The present study addresses time points T1 (30 October 2021) to T12 (25 February 2022) (Table 1) to analyze the behavior of natural interaction in the field and exclude the effects of interventions starting at T13.
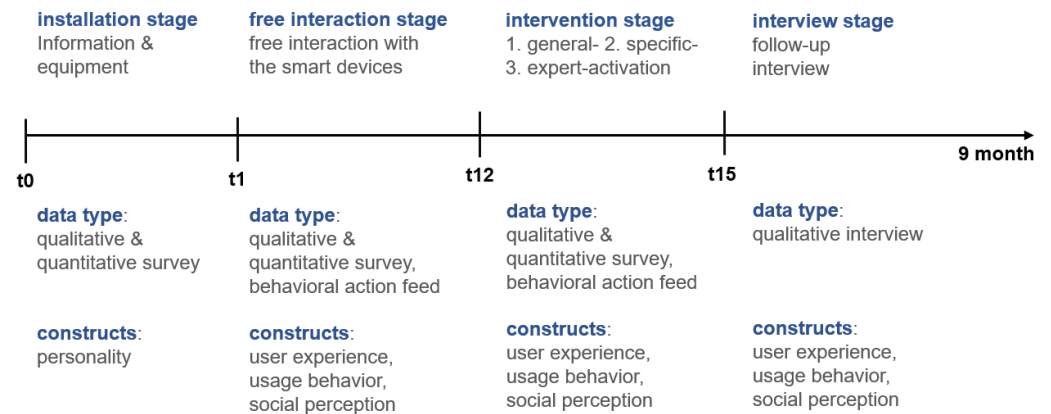
Figure 1. The course of the long-term study with the different stages, collected constructs, and data types.

Table 1. Measurement time (short "T") points and years of the long-term study.

2020							
T0	T1	T2	T3	T4	T5	T6	T7
30 Oct.	6 Nov.	11 Nov.	19 Nov.	26 Nov.	3 Dec.	17 Dec.	30 Dec.
2021							
T8	T9	T10	T11	T12	T13	T14	T15
22 Jan.	5 Feb.	11 Feb.	18 Feb.	25 Feb.	11 Mar.	25 Mar.	25 Jun.

The idea of the free interaction stage was to allow the participants to interact with their new devices without any restrictions for four months. During this stage, participants answered 13 online questionnaires. Each questionnaire began with instructions and included privacy explanations. Participants then entered their unique codeword and answered the questionnaire. New surveys and information were announced via participants' messenger services. The questionnaires included five thematic blocks such as personality (e.g., demographics; personality traits); usage behavior (e.g., the function used, frequency of use); user experience (e.g., UX user motivations, usage ratings), and social perceptions of the IVA (e.g., friendship, empathy). Although personality traits were recorded only once during the free interaction phase, the other constructs were recorded repeatedly to track changes over time. After completion of the study, participants consented to or disagreed with the analysis of their data. Participants affirmed that the recorded voice input was theirs alone and that no third-party voice data were recorded. The devices were returned after nine months.

2.3. Data Analysis

To determine whether users differed in their attribution of friendship to IVAs (**Section 1**), they were divided into homogeneous groups using K-means clustering. A multifactorial ANOVA then identified the significant differences between the groups. In **Sections 1 to 5**, the groups are compared using appropriate procedures (e.g., Welch's *t*-test [57]) to reveal the differences in usage behavior, user experience, social perception, and personality traits. For group comparisons, two-sided tests were performed at an α-level < 0.05. In **Sections 1 to 4**, the effects over time are examined using repeated measures ANOVA (RM-ANOVA) with Greenhouse–Geisser and Bonferroni corrections in post hoc tests. RM-ANOVA is an appropriate procedure because the measurements were repeated among

participants at time intervals and we wanted to examine the trends over time both within and across groups. The time between the individual measurement points is indicated in the respective analyses.

3. Measures, Results, and Discussions by Section

We describe the variables, measurement tools, and results per research question in the following sections: (1) Cluster Formation and Time Effects, (2) Usage Behavior, (3) User Experience, (4) Social Perception, and (5) Personality Traits.

3.1. Section 1—Cluster Formation and Time Effects

Research has shown that users attribute different social roles to IVAs such as tool-based or friend-based roles [31]. In this section, we analyze how participants explicitly associate IVAs with socially identified social roles [10]. However, we also use implicit methods to capture the perceptions of friendship qualities, as social attributions to technologies are often unconscious [15,16]. In this section, we examine (1) whether participants differ in their perceptions of friendship toward IVAs, (2) whether they can be grouped based on these perceptions, and (3) how friendship perceptions evolve.

3.1.1. Measures

Social Roles Scale—Explicit Measurement. Purington, Taft, Sannon, Bazarova, and Taylor [15] identified five social roles that users associate with voice assistants. Participants rated their voice assistant on a seven-point Likert scale (ranging from 1 "strongly disagree" to 7 "strongly agree") based on Purington et al. [10]'s social roles: (1) information source, (2) entertainment provider, (3) administrative assistant, (4) companion, and (5) friend ($\alpha = 0.052$). The explicit measure allowed us to determine the role that the participants consciously associated with the IVA. Participants received the instruction: "People look at their voice assistant very differently. When you think of your voice assistant, to what extent do the following descriptions apply from your perspective?".

Friendship Quality—Implicit Measurement. The *Intimate Friendship Scale* (*IFS*; Sharabany [58]) was used to measure the extent to which participants perceived the IVA as a friend. The *IFS* is an appropriate instrument for measuring the depth and quality of perceived friendship. The scale focuses on the important aspects of the relationship between interaction partners. In addition, we can use the subscales of the *IFS* to measure the latent attributes of friendship perceptions that individuals associate with their voice assistants. Compared to the explicit *Social Role Scale*, the *IFS* measures the implicit dimensions of friendship and does not directly ask about the extent to which participants perceive the IVA as a friend. With its 32 items, the *IFS* measures 8 subscales (frankness and spontaneity ($\alpha = 0.83$), sensitivity and knowing ($\alpha = 0.80$), attachment ($\alpha = 0.81$), exclusiveness ($\alpha = 0.76$), giving and sharing ($\alpha = 0.76$), imposition ($\alpha = 0.70$), common activities ($\alpha = 0.63$), and trust and loyalty ($\alpha = 0.67$)) on a scale of 1 (strongly disagree) to 6 (strongly agree). We adapted the items to voice assistants. Items that deviated too much from the original were not included in the analysis (see Table S2), which is only relevant to the imposition and common activities subscales, as well as six items from the sensitivity and knowing, attachment, exclusiveness, and giving and sharing subscales.

3.1.2. Results

Explicit and Implicit Measurements. The explicit measurement of social roles via the *Social Role Scale* at T10 showed that participants ($N = 73$) most often associated IVAs with the information source role ($M = 6.06$, $SD = 1.15$). This was followed by the entertainer ($M = 5.54$, $SD = 1.48$), assistant ($M = 3.81$, $SD = 1.96$), companion ($M = 1.63$, $SD = 1.04$), and friend ($M = 1.20$, $SD = 0.60$) roles. As expected, the implicit measure of the perceived friendship quality at T10 was higher than the explicit measure (see Table 2). Therefore, for the evaluation in Section 1, the cluster analysis was performed using the implicit measure.

Table 2. Means and standard deviations of perceived friendship quality.

	N	Mean	SD
Intimate Friendship (total)	73	2.26	0.82
Frankness and Spontaneity	73	2.30	0.93
Sensitivity and Knowing	73	1.76	1.03
Attachment	73	2.67	1.03
Exclusiveness	73	1.69	0.86
Giving and Sharing	73	2.25	1.02
Trust and Loyalty	73	2.88	1.18

Cluster Analysis. The aim was to group participants according to their perceived level of friendship with the voice assistant. This involved dividing participants into groups with similar ratings while ensuring that individuals in different groups had differing ratings. The use of the K-means cluster analysis was appropriate, as it can optimize both homogeneity within clusters and heterogeneity between clusters [59]. In addition, this method is popular and is widely used for data [60].

The parameters used to group participants in the K-means cluster analysis were Sharabany's six subscales of perceived friendship quality [58]. The procedure was computed using the R package, Stats, and the Euclidean distance was used as the measure of dissimilarity. The optimal number of clusters was determined using a gap-stat plot [61,62]. As shown in Figure 2, the optimal number of clusters in the data with $k = 2$ was determined using the NbClust package in R, which used 30 simultaneous processes to identify the optimal number of clusters (see the vertical dashed line in Figure 2). Ten indices supported a solution with two clusters according to the majority rule, whereas 6 indices suggested 3 clusters, 3 indices suggested 8 clusters, 1 index suggested 9 clusters and 3 indices suggested a 10-cluster solution. On this basis, we decided on two clusters.
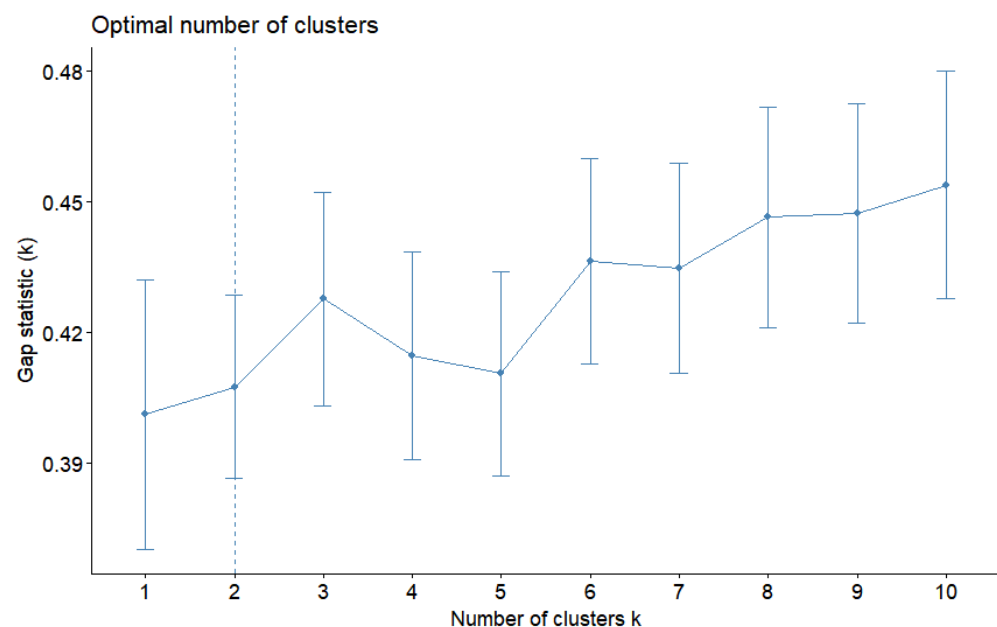

Figure 2. The gap-stat plot of the optimal number of clusters.

Table 3 shows the mean values and the number of participants in each cluster. It shows that the clusters can be divided into higher perceived friendship quality and lower perceived friendship quality. The first cluster ($n = 40$) showed higher scores in the attribution of friendship quality to IVAs than the second cluster ($n = 33$). To test the validity of the clusters, a multivariate analysis of variance (MANOVA) with post hoc tests was used to determine whether the clusters differed significantly based on Sharabany's subscales [58].

The MANOVA supported the validity of the clusters (see Table 3) by showing a significant difference in the perceived friendship quality between the two groups ($F(7, 65) = 23.66$, $p < 0.001$, Wilk's $\Lambda = 0.28$) [63].

Table 3. Scales, descriptive values, and *t*-tests of the friendship quality subscales of the *IFS*.

|	Cluster	n	Mean	SD	T-Value [1]	df
Frankness and Spontaneity	Friend	40	2.84	0.80	7.26 ***	70.67
	Non-Friend	33	1.64	0.61		
Sensitivity and Knowing	Friend	40	2.34	1.05	7.21 ***	48.98
	Non-Friend	33	1.06	0.35		
Attachment	Friend	40	3.31	0.89	8.51 ***	64.46
	Non-Friend	33	1.89	0.52		
Exclusiveness	Friend	40	2.08	0.94	5.01 ***	59.05
	Non-Friend	33	1.23	0.46		
Giving and Sharing	Friend	40	2.94	0.73	9.62 ***	70.80
	Non-Friend	33	1.41	0.63		
Trust and Loyalty	Friend	40	3.62	0.82	7.94 ***	65.72
	Non-Friend	33	2.00	0.90		

[1] *** $p < 0.001$.

To avoid ambiguity, we refer to the first cluster as the **friend cluster**, which is characterized by a higher perceived quality of friendship toward the voice assistant. Conversely, we refer to the second cluster as the **non-friend cluster**, which is characterized by a relatively lower perceived quality of friendship toward the voice assistant.

Longitudinal Effects. To assess the change in the friendship quality over time, the total friendship quality scores at T4 and T10 were used for the repeated-measures ANOVA. The main effect was significant with a medium effect size ($F(1, 66) = 7.32$, $p = 0.009$, $\eta^2 p = 0.10$). Notably, there was also a significant interaction effect with a large effect size ($F(1, 66) = 19.23$, $p < 0.001$, $\eta^2 p = 0.23$). The friend cluster ($n = 36$) showed a significant ($p < 0.001$) increase in the perceived friendship quality of the voice assistant from T4 ($M = 2.30$, $SE = 0.09$) to T10 ($M = 2.82$, $SE = 0.08$), whereas the non-friend cluster ($n = 32$) showed no significant change ($p = 1.00$) between T4 ($M = 1.65$, $SE = 0.10$) and T10 ($M = 1.53$, $SE = 0.08$).

3.1.3. Brief Discussion of Section 1

The results are interesting because the measurements of the perception of the voice assistant as a friend on both explicit and implicit levels reveal contradictions. The explicit measure shows that the majority of participants rejected the voice assistant as a friend. However, the implicit measures reveal that the voice assistant fulfills friendship qualities. The explicit measure is consistent with the findings of previous studies, which showed that social perceptions and social behavior toward technology can be denied even though they actually occurred [55]. This leads us to conclude that users view IVAs as friends far more than they admit or are aware of. Furthermore, the implicit measure of friendship quality was an appropriate variable for distinguishing between individuals based on whether they perceived IVAs as more or less friend-like. In addition, the relationship between these clusters differed over time (Figure 3). The friend cluster increasingly associated the IVA with friendship qualities over time, whereas there was no effect of time observed for the non-friend cluster. These findings guided the follow-up analysis, which investigated whether the two clusters differed in terms of their usage behavior, perceived user experience, and social perceptions over time.
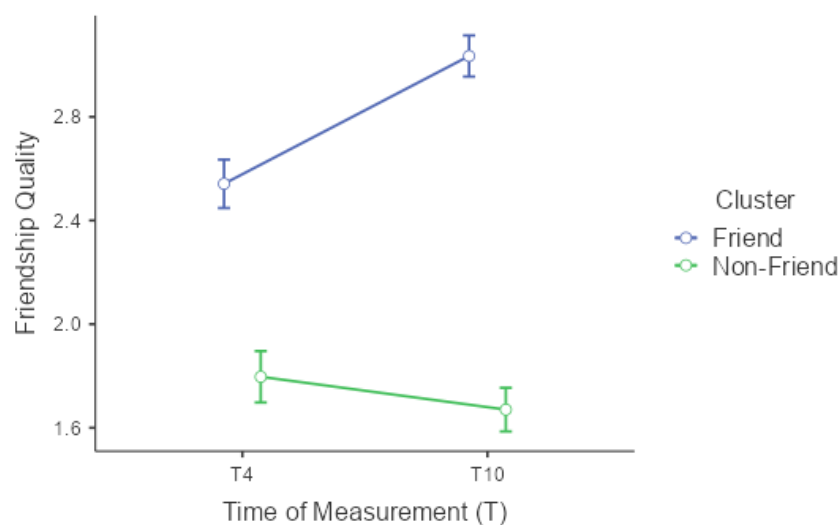
Figure 3. Changes in friendship quality over time separated by cluster with standard errors.

3.2. Section 2—Usage Behavior

The perception of IVAs as friends can affect how users interact with and utilize voice assistants [32,33]. To determine usage behavior, we used subjective and implicit measurement techniques to further analyze usage habits, future usage intentions, and the frequency and type of features used by users. These factors provided a holistic understanding of both current and future IVA usage patterns. We then conducted a two-sided Welch's *t*-test to determine if there were any disparities in the usage behavior variables between the friend cluster and the non-friend cluster. Our choice of statistical tests followed the recommendations of [57].

3.2.1. Measures

Subjective Measure—Usage Habits. To understand the adoption and changes in the use of voice assistants, we examined the integration of IVAs into daily routines and tested for group variances. Previous research has indicated that incorporating media into daily routines is a critical factor in determining usage patterns [64]. We adapted items from the Social Media Use Integration Scale [64] to capture the *integration of IVAs into daily individual routines* (13 items; e.g., The voice assistant wakes me up every morning) and *daily routines with others* (6 items; I call my friends using the voice assistant). The reliability of the questionnaire was $\alpha = 0.75$.

Subjective Measure—Future Usage. A 5-point Likert scale (1 = I do not believe in it at all; 5 = I believe in it very much), which was adapted from Przybylski et al. [65], was used to measure participants' *interest in continuing to use the voice assistant in the future* (one item, "I will continue to use a voice assistant after the end of the study") and *tendency to recommend the voice assistant to others* (word to mouth; two items, e.g., "I will tell others positive things about my voice assistant"). The internal consistency of the items in this study was $\alpha = 0.83$.

Implicit Measure—Behavioral Data. Considering participants' consent, their usage behavior was examined using IVA interaction data from participants' exported activity logs. The information was then compiled into a CSV file that contained columns for each participant's anonymous subject code, the date of the command, the command and conversation content, and the response from the voice assistant. From the logs, we were able to extract a total of 22,436 speech entries ($M = 467.42$, $SD = 787.35$). Thus, we were able to implicitly derive actual usage behavior based on voice interactions with the IVAs. Due to technical complications in the provider platform, some user logs were incomplete. In total, we had access to 48 complete user logs.

To determine how users interacted with the IVA during the study period, we first took a sample of 2000 transcribed voice commands from the usage logs and examined

them using Mayring's Qualitative Content Analysis [66]. If identification via the user's transcribed voice input was not possible, we used the IVA's transcribed voice output as an indicator of the function used. We then developed a fixed set of 45 subcategories (functions used by the voice assistant such as news, weather, listen to music) and 7 primary categories (a structural classification of functions at a higher level based on similarities such as knowledge acquisition, support, media entertainment) to categorize the interactions based on the voice commands used (see Table S3). This set was derived from categories determined in previous literature [12,26,67], feature reviews from device vendors [68], and inductive new categories, and was iteratively revised by all researchers until an agreement was reached.

We generated and used keywords to automatically categorize the transcribed voice input. For instance, for the subcategory "lamps", we manually searched for related voice commands containing the keywords "lamp", bright", and "light". As we analyzed the commands from various usage logs, we were able to identify additional relevant keywords related to user intent (e.g., "bulb"), as well as find voice commands that contained similar keywords but were not related to user intent (e.g., "Is it already *bright* outside?"). After finding new keywords and exceptions, we added or removed keywords to differentiate this category from other categories of voice commands. A randomized sample of 1000 voice commands was used to verify that the majority of the categorized voice commands corresponded to the actual function (subcategory). This iterative approach allowed us to provide a unique classification for each subcategory.

Once the keywords were finalized, we performed the automated categorization by string matching. In the specific case of the inductively created primary category "Social Interaction," we conducted a second, independent categorization process. The primary category captured the extent to which users personify the IVA [69]. We performed a second categorization process because we assumed that personification features could occur in all the functions (subcategories) used. For example, politeness phrases (such as "please" and "thank you") and greetings (such as "hello" or "hi") may be used in combination with other functions (e.g., listening to music, alarm clock, and time) such as "Alexa, please play a song" or "Can you please tell me what time it is?". To maintain the validity and reliability of the categorizations, each categorization was manually reviewed and adjusted as needed. A sample of the user input was verified, and the majority of the categorized speech input was found to match the actual function. A second independent coder categorized 10% of the whole sample (Cohen's kappa = 0.95).

3.2.2. Results

Subjective Measure—Usage Habits. The reference time point for assessing usage habits was T8. The friend cluster ($M = 2.95$, $SD = 0.81$, $n = 38$) integrated the voice assistant significantly more into their daily routines ($t(65.15) = 4.38$, $p < 0.001$, $d = 1.06$) than the non-friend cluster ($M = 2.14$, $SD = 0.72$, $n = 30$). When interacting with others daily, the voice assistant was significantly ($t(65.25) = 4.77$, $p < 0.001$, $d = 1.16$) more integrated by the friend cluster ($M = 2.44$, $SD = 0.70$, $n = 38$) than the non-friend cluster ($M = 1.67$, $SD = 0.62$, $n = 30$).

Subjective Measure—Future Usage. The reference time point for the intention to continue using and recommending the voice assistant in the future was T8. The intention to recommend the voice assistant in the future was significantly higher ($t(64.93) = 3.75$, $p < 0.001$, $d = 0.89$) for the friend cluster ($M = 4.48$, $SD = 1.15$, $n = 40$) than the non-friend cluster ($M = 3.39$, $SD = 1.29$, $n = 33$). Regarding interest in continuing to use the voice assistant in the future, there was no significant ($p = 0.340$) difference between the friend cluster ($M = 3.03$, $SD = 1.46$) and the non-friend cluster ($M = 2.70$, $SD = 1.45$).

Objective Measure—Behavioral Data. The data on the voice assistant functions used by each participant throughout the entire study period were collected and an overall value was calculated. Using this index, Grubbs' test identified one significant outlier, which was excluded from further analyses. Both clusters were then tested for differ-

ences in frequency of use. We tested for differences between the thematic categories that summarized the individual functions of the IVA, as well as the individual functions themselves (subcategories).

Categories. Welch's *t*-tests across categories revealed significant and marginally significant differences between the two clusters (Table 4). For example, individuals in the friend cluster were, on average, more likely to use the IVA for support ($p = 0.038$) and social interactions ($p = 0.028$), with differences for use in knowledge acquisition ($p = 0.086$) and mood management ($p = 0.073$) being marginally significant and higher for the friend cluster. However, differences between the clusters in terms of media and entertainment ($p = 0.796$) and smart homes ($p = 0.910$) were not found to be significant.

Table 4. Descriptive values and group differences at the category level separated by cluster.

Category	Cluster	N	Mean	SD	T-Value [1]	df	Effect Size Cohen's d
Knowledge Acquisition	Friend	27	36.26	53.00	1.77 †	30.57	0.49
	Non-Friend	21	17.38	14.03			
Support	Friend	27	86.78	151.73	2.17 *	30.68	0.60
	Non-Friend	21	20.52	40.63			
Medial Entertainment	Friend	27	260.86	391.62	−0.26	28.28	−0.08
	Non-Friend	21	308.14	755.72			
Mood Management	Friend	27	10.89	19.29	1.86 †	30.05	0.51
	Non-Friend	21	3.71	4.80			
Smart Home	Friend	27	32.41	52.32	0.11	35.91	0.03
	Non-Friend	21	30.33	70.07			
Social Interaction	Friend	27	52.07	78.02	2.32 *	27.87	0.64
	Non-Friend	21	16.57	13.10			

[1] * $p < 0.05$, † $p < 0.10$.

Subcategories. Welch's *t*-tests revealed significant and marginally significant differences between the clusters (Table 5). The friend cluster was more likely to use the IVA for news ($p = 0.074$), as a local guide ($p = 0.025$), as an alarm clock and for the time ($p = 0.017$), as a calendar ($p = 0.044$), for cooking ($p = 0.098$), for audiobooks and stories ($p = 0.073$), for jokes ($p = 0.061$), for self-esteem ($p = 0.077$), as a fun gadget ($p = 0.087$), to apologize ($p = 0.083$), to show interest in social cues ($p = 0.034$), for greetings and goodbyes ($p = 0.014$), and for direct speech ($p = 0.049$).

Table 5. Means, standard deviations, percentages, and significance tests with the effect size of the function used (subcategories) separated by cluster.

Function/Subcategory	Friend Cluster (n = 27)			Non-Friend Cluster (n = 21)			T-Value	df	p [1]	Effect Size Cohen's d
	Mean	SD	%	Mean	SD	%				
News	8.07	17.32	1.68	1.81	2.38	0.45	1.857	27.3	0.074 †	0.51
Knowledge Base	16.63	29.00	3.46	8.91	9.09	2.24	1.304	32.3	0.201	0.36
Local Guide	0.78	1.22	0.16	0.19	0.40	0.05	2.344	32.9	0.025 *	0.65
Calculator	1.56	4.83	0.32	0.29	0.78	0.07	1.343	27.7	0.19	0.37
Weather	9.22	10.58	1.92	6.19	7.56	1.56	1.157	45.7	0.253	0.33
Alarm clock and time	53.63	100.13	11.16	4.43	7.26	1.11	2.545	26.4	0.017 *	0.69
Timer	22.37	61.51	4.65	8.48	17.46	2.13	1.117	31.2	0.272	0.31
Reminder	2.15	5.68	0.45	0.29	1.10	0.07	1.663	28.5	0.107	0.46
To-Do List	0.26	1.16	0.05	0.00	0.00	0	1.158	26	0.257	0.32
Calendar	0.56	1.09	0.12	0.10	0.30	0.02	2.101	31	0.044 *	0.58
Route planner	0.48	1.16	0.1	0.48	1.08	0.12	0.016	44.4	0.987	0.01
Cooking	2.00	3.82	0.42	0.62	1.60	0.16	1.696	36.6	0.098 †	0.47
Shopping list	5.33	21.49	1.11	6.14	24.51	1.54	−0.12	40	0.905	−0.04
Video streaming	0.67	0.78	0.14	0.76	1.14	0.19	−0.33	34	0.745	−0.10
Listen to music	107.52	152.87	22.37	108.76	194.67	27.3	−0.02	37.2	0.981	−0.01

Table 5. *Cont.*

Function/Subcategory	Friend Cluster (*n* = 27)			Non-Friend Cluster (*n* = 21)			T-Value	df	p [1]	Effect Size
	Mean	SD	%	Mean	SD	%				Cohen's d
Audiobooks and stories	1.07	2.06	0.22	0.29	0.72	0.07	1.853	33.7	0.073 †	0.51
Games and skills	13.30	49.44	2.77	4.91	10.48	1.23	0.858	29	0.398	0.24
Media control	138.26	255.34	28.76	193.43	567.57	48.6	−0.41	26.3	0.682	−0.13
Jokes	3.26	5.36	0.68	1.14	1.62	0.29	1.94	31.9	0.061 †	0.53
Motivation	0.59	1.53	0.12	0.14	0.48	0.04	1.443	32.3	0.159	0.40
Relaxation	1.52	3.84	0.32	1.10	3.08	0.28	0.424	45.9	0.674	0.12
Sleep aid	1.59	4.39	0.33	0.19	0.68	0.05	1.635	27.6	0.113	0.45
Self-esteem	0.59	1.37	0.12	0.10	0.30	0.02	1.835	29.2	0.077 †	0.50
Negative mood	0.70	2.02	0.15	0.24	0.70	0.06	1.117	33.6	0.272	0.31
Fun Gadget	2.22	4.46	0.46	0.67	0.86	0.17	1.772	28.4	0.087 †	0.48
Social presence	0.41	1.39	0.08	0.14	0.36	0.04	0.947	30.3	0.351	0.26
Lamps	69.04	181.96	14.36	56.19	87.85	14.1	0.322	39.3	0.749	0.09
Sockets	29.37	52.31	6.11	28.86	68.96	7.25	0.028	36.3	0.978	0.01
Other Smart Home	0.04	0.19	0.01	0.05	0.22	0.01	−0.18	40.2	0.862	−0.05
Connection	2.74	10.48	0.57	1.43	2.23	0.36	0.633	29	0.532	0.17
Offense	1.19	2.02	0.25	1.19	1.86	0.3	−0.01	44.6	0.993	−0.00
Appreciation	0.44	1.05	0.09	0.33	0.66	0.08	0.448	44.2	0.656	0.13
Congratulation	0.63	2.90	0.13	0.00	0.00	0	1.129	26	0.269	0.31
Apology	0.11	0.32	0.02	0.00	0.00	0	1.803	26	0.083 †	0.49
Interest in social cues	4.96	7.04	1.03	1.52	3.66	0.38	2.188	40.8	0.034 *	0.61
Greetings/Goodbyes	9.19	17.07	1.91	0.48	0.87	0.12	2.647	26.2	0.014 *	0.72
Politeness	16.07	55.22	3.34	3.52	5.94	0.89	1.172	26.8	0.251	0.32
Intimate expression	3.41	4.48	0.71	1.95	2.29	0.49	1.459	40.5	0.152	0.41
Direct speech	16.07	20.07	3.34	7.57	7.12	1.9	2.043	34	0.049 *	0.57
Calls/voice messages	0.85	2.27	0.18	0.29	0.64	0.07	1.236	31.2	0.226	0.34
Shopping	0.48	1.37	0.1	1.00	3.27	0.25	−0.68	25.5	0.502	−0.21
Routine	0.00	0.00	0	0.00	0.00	0				
Whisper mode	0.19	0.62	0.04	0.10	0.30	0.02	0.658	39.3	0.514	0.18
Misunderstood request	35.07	52.04	7.3	24.29	28.39	6.1	0.916	41.8	0.365	0.26
Commands Total	480.74	606.46	100	398.05	850.55	100	0.377	34.8	0.708	0.11

[1] * *p* < 0.05, † *p* < 0.10. On this basis, the functions marked in bold are significant results.

Figure 4 shows the functions used and IVA usage across a week. The bar chart for the friend cluster is more colorful, indicating a wider range of functions used. Both clusters primarily used the voice assistant to control media, listen to music, and control their smart home. The friend cluster additionally showed more intensive use of the alarm clock and time function. Similarly, the friend cluster had more difficulty being understood by the voice assistant. Looking at the days of the week, both clusters were mainly active from Monday to Wednesday, whereas the non-friend cluster was most active on Thursday.
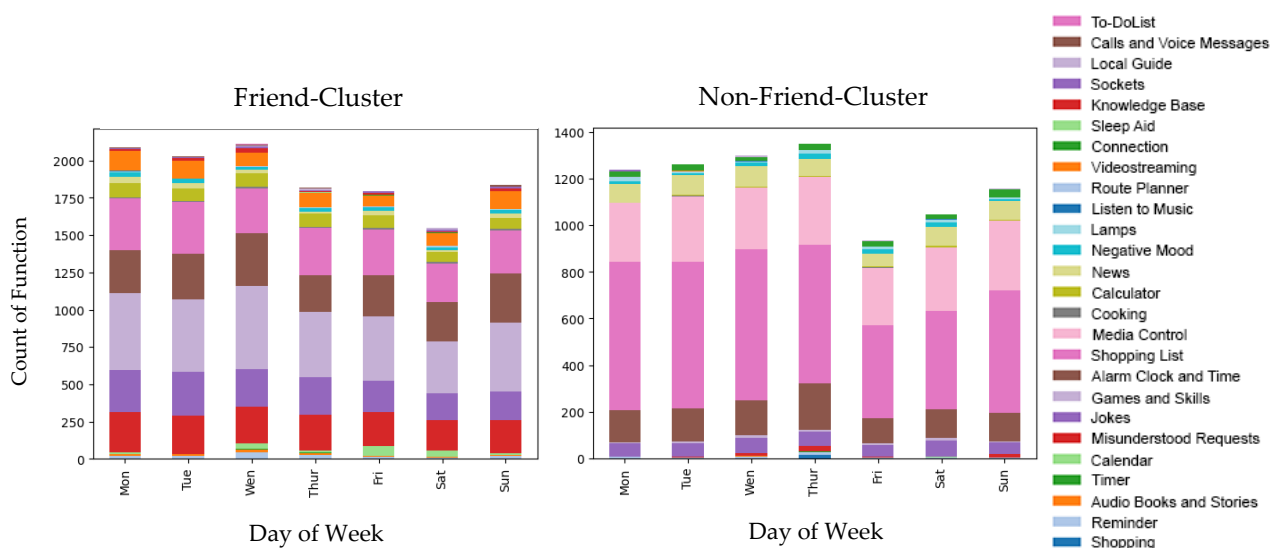

Figure 4. Functions used and IVA usage across a week separated by cluster.

3.2.3. Interaction Modeling

Below, we model additional usage indicators based on IVA data to better understand the interactions with IVAs and potentially disaggregate usage differences as a function of perceived friendship.

Daily Use. To understand how the IVAs were used by the two clusters daily, participants' voice commands were examined in more detail. Analyses showed that there was no significant difference in the frequency of use across the study period (118 days) between the groups ($t(39.09) = 0.81$, $p = 0.425$, $d = 0.25$). The friend cluster sent an average of $M = 5.55$ ($SD = 7.10$) commands and the non-friend cluster sent an average of $M = 3.73$ ($SD = 7.51$) commands per day to their voice assistants.

Length of Voice Commands. In the two clusters, we looked at how many words on average the participants used per voice command throughout the study period (Figure 5). The friend cluster used marginally ($t(39.36) = 1.87$, $p = 0.069$, $d = 0.58$) more words per voice command on average ($M = 3.37$, $SD = 2.38$) than the non-friend cluster ($M = 2.59$, $SD = 1.93$).
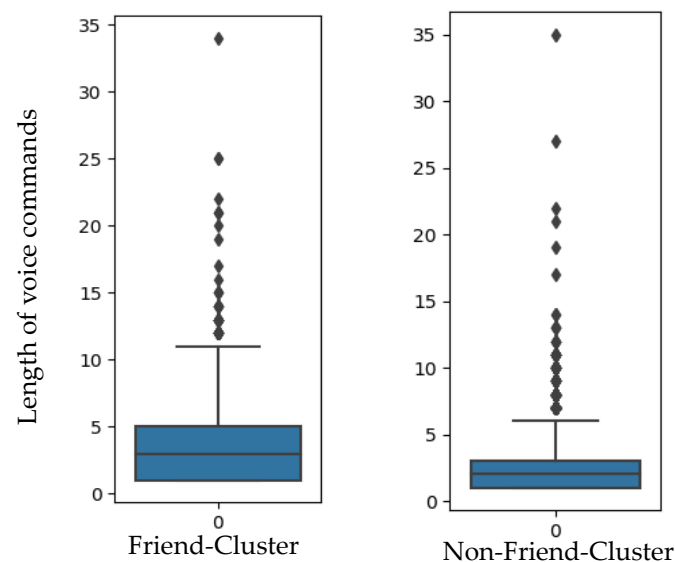
Figure 5. Length (number of words) of voice commands separated by cluster.

Word length per voice command over time. Figure 6 shows the number of words used per voice command per cluster over time. The average voice command length changed over time. Notably, the friend cluster showed a positive trend in terms of the word length used over time (minimum 2.05, maximum 4.31), whereas the non-friend cluster showed a negative trend (minimum 1.59, maximum 3.78). To assess the temporal effect, we compared the average instruction length of the two clusters in the first and last month of use. There was a significant difference between the two clusters ($F(1, 31) = 7.13$, $p = 0.012$, $\eta^2 p = 0.19$). In the first 4 weeks, the friend cluster ($M = 3.88$, $SE = 0.26$) and the non-friend cluster ($M = 3.34$, $SE = 0.25$) did not differ significantly ($p = 0.84$). However, in the last four weeks, the friend cluster ($M = 3.92$, $SE = 0.25$) used significantly ($p = 0.027$) more words per voice command than the non-friend cluster ($M = 2.87$, $SE = 0.24$). Notably, even in the first four weeks of use, the friend cluster used significantly more words per speech command than the non-friend cluster used in the last four weeks ($p = 0.041$).
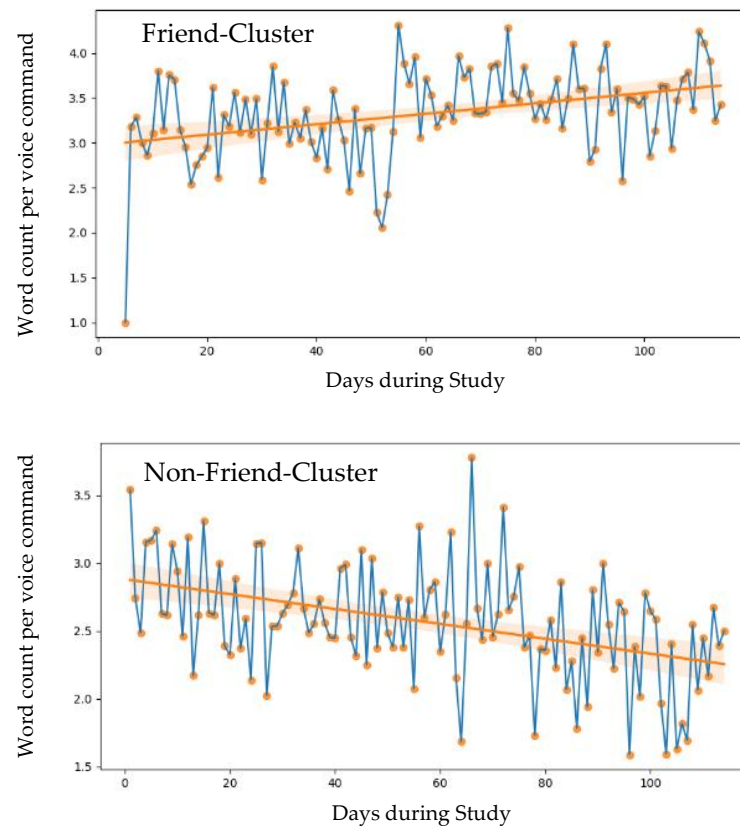
Figure 6. The number of words used per voice command over time separated by cluster.

3.2.4. Brief Discussion of Section 2

Our analysis revealed differences in the usage behavior between individuals from the friend cluster and the non-friend cluster. The friend cluster was more likely to use the voice assistant for functions related to support, mood management, and knowledge acquisition. More specifically, IVAs were more often used, for example, for checking the news, as an alarm and clock, or as a calendar. There was also a tendency for individuals from the friend cluster to use the voice assistant to boost their self-esteem (sample voice command: "Say something nice to me") and have more fun with the IVA to entertain themselves (sample voice command: "Activate self-destruction"). The friend cluster was also more likely to apologize to their voice assistant than the non-friend cluster. The friend cluster tended to interact more socially with the IVA. Thus, the friend cluster was more interested in the voice assistant's personality (sample voice command: "Do you have friends?", "How are you?", or "How old are you?"), more likely to greet or say goodbye to the voice assistant, and more likely to address it using the personal pronoun "you".

The subjective measures indicated that the friend cluster was more likely than the non-friend cluster to recommend the voice assistant to others in the future. In addition, the friend cluster self-reported that they had integrated the voice assistant into their daily lives and interactions with others more than the non-friend cluster. Interaction modeling based on the categorized user logs showed that users in the friend cluster used the IVA for a wider range of functionalities. Although both groups used the voice assistant primarily for media control, listening to music, or controlling their smart home, the friend cluster used the voice assistant more frequently for obtaining daily news, as an alarm clock, or for checking the time. In addition, the friend cluster used marginally more words in their voice interactions with the voice assistant. This may indicate that users who associated the voice assistant more strongly with friendship qualities used more complex sentence structures when speaking to the voice assistant. In contrast, the non-friend cluster used fewer words

per speech input. It is also possible that the number of words per voice command used is related to the types of functions used by the respective clusters.

3.3. Section 3—User Experience

When users perceive technologies as social, it can have a positive impact on meeting their needs [44]. For example, the stronger the social relationship between the user and the IVA, the higher the perceived usefulness [46]. The fulfillment of usage motives (e.g., pragmatic and hedonic) and usage needs (e.g., autonomy and competence), along with usage evaluation (e.g., perceived value and awe), are considered essential components in evaluating the user experience [70–73]. Accordingly, we used valid scales to measure these elements of the user experience and conducted two-tailed Welch's *t*-tests to determine whether the two clusters differed in their user experience. To identify temporal trends in the user experience variables across clusters and examine differences within and between clusters, RM-ANOVAs were performed.

3.3.1. Measures

Fulfillment of Usage Motives. Four items were used to assess the pragmatic (e.g., "The interaction fulfilled my seeking for simplicity.", $\alpha = 0.80$) and hedonic (e.g., "The interaction fulfilled my seeking for pleasure.", $\alpha = 0.79$) quality based on the short version of the *AttrakDiff mini* [70]. The eudaimonic quality was assessed by four items (e.g., "The interaction fulfilled my seeking to do what you believe in.", $\alpha = 0.82$) adapted from Huta [71]. Four items were used to evaluate the social quality (e.g., "The interaction fulfilled my seeking for social contact.") based on Hassenzahl, Wiklund-Engblom, Bengs, Hägglund, and Diefenbach [72] ($\alpha = 0.84$). Questions were asked directly in terms of motive fulfillment through the interaction with the IVAs. The items were rated from 1 (not at all) to 7 (very much).

Fulfillment of User Needs. The 14-item scale by Hassenzahl, Wiklund-Engblom, Bengs, Hägglund, and Diefenbach [72] measures users' experiences and evaluations in relation to the satisfaction of their psychological needs, the positive effects, and their perceptions of products. The questionnaire includes the subscales *competence* ($\alpha = 0.80$), *popularity* ($\alpha = 0.68$), *relatedness* ($\alpha = 0.84$), *security* ($\alpha = 0.64$), *meaning* ($\alpha = 0.72$), and *stimulation* ($\alpha = 0.77$). Participants were instructed to rate their interaction with their voice assistant on a five-point Likert scale (1 = not at all; 5 = very much) in terms of fulfillment.

Usage Evaluation. User experience is often evaluated in terms of the fulfillment of hedonic and eudaimonic goals [73]. To assess usage evaluation in this way, we used the well-validated scale of Huta and Ryan [73]. The corresponding subscales, *Sense* ($\alpha = 0.66$), *value* ($\alpha = 0.83$), *implication* ($\alpha = 0.67$), *awe* ($\alpha = 0.75$), *inspiration* ($\alpha = 0.74$), *transcendence* ($\alpha = 0.89$), and *carefreeness* ($\alpha = 0.86$), were measured on a seven-point Likert scale (1 = not at all; 7 = very much).

3.3.2. Results

Fulfillment of Usage Motives (Group Differences). To compare the fulfilled usage motives, time point T11 was used as the reference time point. The friend cluster showed significantly higher scores in fulfilling the four usage motives compared to the non-friend cluster (pragmatic: $t(39.56) = 2.03$, $p = 0.049$, $d = 0.54$; hedonic: $t(40.71) = 2.76$, $p = 0.009$, $d = 0.73$; eudaimonic: $t(59.09) = 4.99$, $p < 0.001$, $d = 1.26$; social: ($t(65.25) = 3.901$, $p < 0.001$, $d = 0.98$) (Table 6).

Fulfillment of Usage Motives (Longitudinal Effects). At T5 ($M = 3.13$, $SE = 0.18$) and T11 ($M = 2.77$, $SE = 0.17$), there was a significant negative main effect of eudaimonic use motive with a medium effect ($N = 60$, $F(1, 58) = 4.00$, $p = 0.050$, $\eta^2 p = 0.06$). There was no significant interaction effect ($F(1, 58) = 0.275$, $p = 0.602$, $\eta^2 p = 0.01$).

Table 6. Overview of descriptive values and inferential statistics on the fulfilled usage motives.

	Cluster	n	M	SD	T-Value [1]	df	Effect Size Cohen's d
Pragmatic	Friend	37	5.07	1.24	2.03 *	39.56	0.54
	Non-Friend	26	4.20	1.91			
Hedonic	Friend	37	5.17	1.14	2.76 **	40.71	0.73
	Non-Friend	26	4.12	1.69			
Eudaimonic	Friend	37	3.56	1.33	4.99 ***	59.09	1.26
	Non-Friend	26	2.02	1.11			
Social	Friend	37	3.18	1.62	3.91 ***	60.25	0.98
	Non-Friend	26	1.76	1.27			

[1] *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Fulfillment of Usage Needs (Group Differences). Individuals from the friend cluster showed significantly higher scores in competence ($t(51.60) = 5.45$, $p < 0.001$, $d = 1.40$), popularity ($t(59.52) = 5.51$, $p < 0.001$, $d = 1.39$), relatedness ($t(60.09) = 4.33$, $p < 0.001$, $d = 1.06$), security ($t(39.77) = 3.03$, $p = 0.004$, $d = 0.80$), autonomy ($t(52.84) = 5.22$, $p < 0.001$, $d = 1.34$), stimulation ($t(51.58) = 6.65$, $p < 0.001$, $d = 1.71$), and self-actualization ($t(60.85) = 5.57$, $p < 0.001$, $d = 0.99$) (Table 7).

Table 7. Overview of descriptive values and inferential statistics on the fulfilled usage motives.

	Cluster	n	M	SD	T-Value [1]	df	Effect Size Cohen's d
Competence	Friend	37	3.62	1.08	5.45 ***	51.60	1.40
	Non-Friend	26	2.06	1.15			
Popularity	Friend	37	2.92	1.12	5.51 ***	59.52	1.39
	Non-Friend	26	1.50	0.92			
Relatedness	Friend	37	2.62	1.38	4.33 ***	60.09	1.06
	Non-Friend	26	1.40	0.85			
Security	Friend	37	4.85	1.12	3.03 **	39.77	0.80
	Non-Friend	26	3.69	1.71			
Autonomy	Friend	37	4.11	1.35	5.22 ***	52.84	1.34
	Non-Friend	26	2.27	1.39			
Stimulation	Friend	37	4.55	1.32	6.65 ***	51.58	1.71
	Non-Friend	26	2.21	1.42			
Self-Actualization	Friend	37	2.49	1.23	3.96 ***	60.86	0.99
	Non-Friend	26	1.42	0.90			

[1] *** $p < 0.001$, ** $p < 0.01$.

Fulfillment of Usage Needs (Longitudinal Effects). For the period from T6 ($M = 3.57$, $SE = 0.19$) to T11 ($M = 3.16$, $SE = 0.18$), a significant negative main effect was observed for the satisfaction of the need for autonomy with a medium effect size ($F(1, 58) = 6.52$, $p = 0.013$, $\eta^2 p = 0.10$). The interaction effect was not significant ($F(1, 58) = 0.00$, $p = 0.951$, $\eta^2 p = 0.00$). For the need for stimulation, a significant negative main effect with a strong effect size was also identified at T6 ($M = 4.12$, $SE = 0.20$) and T11 ($M = 3.31$, $SE = 0.17$) ($F(1, 58) = 20.70$, $p < 0.001$, $\eta^2 p = 0.26$) (Figure 7). For the experience of stimulation, there was a significant interaction effect with a strong effect size ($F(1, 58) = 13.48$, $p < 0.001$, $\eta^2 p = 0.19$). At T6 ($M = 3.52$, $SE = 0.308$) and T11 ($M = 2.06$, $SE = 0.26$), perceived stimulation decreased significantly ($p < 0.001$) over time for the non-friend cluster ($n = 25$). In the friend cluster ($n = 35$), no statistical change was detected between T6 ($M = 4.71$, $SE = 0.26$) and T11 ($M = 4.56$, $SE = 0.22$).
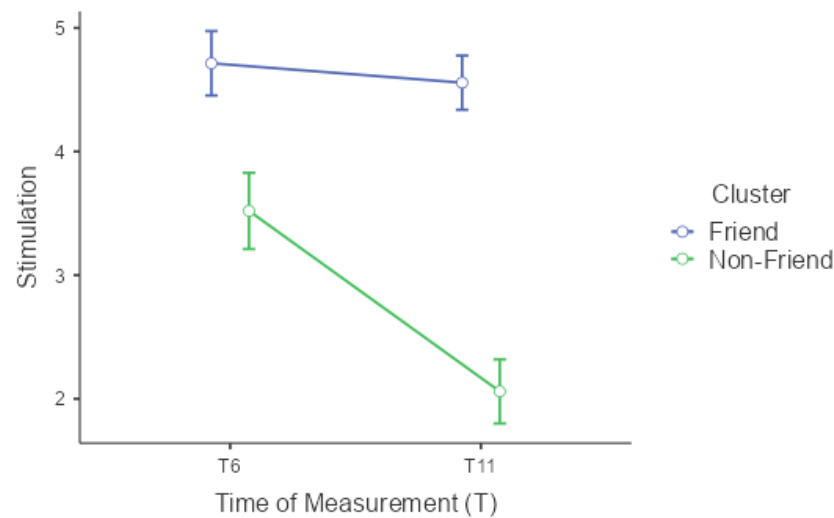
Figure 7. Change in stimulation over time separated by cluster with standard errors. **Usage Evaluation (Group Differences).** For the between-cluster comparison of the usage evaluation, T9 was used as the reference time point. For the friend cluster, the values of the sense ($t(62.54) = 4.19$, $p < 0.001$, $d = 1.01$), value ($t(64.30) = 3.92$, $p < 0.001$, $d = 0.94$), implication ($t(63.91) = 3.62$, $p < 0.001$, $d = 0.87$), awe ($t(67.39) = 4.57$, $p < 0.001$, $d = 1.09$), inspiration ($t(67.13) = 5.30$, $p < 0.001$, $d = 1.27$), and transcendence ($t(67.44) = 2.89$, $p = 0.005$, $d = 0.69$) subscales were significantly higher than those for the non-friend-cluster (Table 8). The difference between the two clusters in the carefreeness subscale was found to be marginally significant ($t(67.67) = 1.75$, $p = 0.085$, $d = 0.42$).

Table 8. Overview of descriptive values and inferential statistics for usage evaluation.

	Cluster	n	M	SD	T-Value [1]	df	Effect Size Cohen's d
Sense	Friend	38	4.21	0.97	4.19 ***	62.54	1.01
	Non-Friend	32	3.16	1.10			
Value	Friend	38	4.09	1.03	3.92 ***	64.30	0.94
	Non-Friend	32	3.08	1.10			
Implication	Friend	38	3.15	0.95	3.62 ***	63.91	0.87
	Non-Friend	32	2.28	1.03			
Awe	Friend	38	2.83	0.84	4.57 ***	67.39	1.09
	Non-Friend	32	1.94	0.78			
Inspiration	Friend	38	3.08	0.84	5.30 ***	67.13	1.27
	Non-Friend	32	2.05	0.79			
Transcendence	Friend	38	2.49	1.11	2.87 **	67.44	0.69
	Non-Friend	32	1.76	1.02			
Carefreeness	Friend	38	3.50	1.15	1.75 †	67.67	0.42
	Non-Friend	32	3.04	1.04			

[1] *** $p < 0.001$, ** $p < 0.01$, † $p < 0.10$.

Usage Evaluation (Longitudinal Effects). The analysis of the time effects of the usage evaluation highlighted the differences between the sense and value subscales. For the sense subscale, there was a marginally significant negative main effect ($F(1, 67) = 2.05$, $p = 0.051$, $\eta^2 p = 0.06$, $N = 69$) with a weak effect size between T6 ($M = 3.50$, $SE = 0.14$) and T9 ($M = 3.44$, $SE = 0.13$). For the value subscale, there was a significant positive main effect with a medium effect size ($F(1, 67) = 5.24$, $p = 0.025$, $\eta^2 p = 0.07$). Between T6 ($M = 3.31$, $SE = 0.15$) and T9 ($M = 3.59$, $SE = 0.13$), the values of the value subscale decreased over time. There were no significant interaction effects.

3.3.3. Brief Discussion of Section 3

Our analyses show significant differences in the user experience between the two clusters. Users in the friend cluster reported significantly higher satisfaction in terms of pragmatic hedonic, eudaimonic, and social needs than the non-friend cluster, meaning that the friend cluster experienced a more enjoyable interaction with their voice assistant that was more meaningful and social. Similarly, key usage needs were better met for the friend cluster than for the non-friend cluster. In addition, the friend cluster perceived the interaction as more emotionally moving, valuable, meaningful, and inspiring than the non-friend cluster. In terms of expectations, the friend cluster was more likely to want IVAs to help with difficult tasks, value their opinions, and provide a sense of social closeness.

Time-related effects throughout the study's duration indicate that the satisfaction of eudaimonic needs decreased, regardless of the cluster assignment. Accordingly, interaction with the voice assistant over time (regardless of its social role) was seen less as a meaningful experience that enabled personal growth and the expression of self-actualization [74]. Regardless of the cluster assignment, participants felt that their autonomy was increasingly violated when interacting with the IVA. The non-friend cluster perceived the interaction with the voice assistant over time as increasingly less stimulating. These results complement previous studies that have reported a decline in usage interest after a diminishing novelty effect [11,26,75–77].

3.4. Section 4—Social Perception

Research has shown that parasocial interactions [36,42], empathy [78], social presence [24,47], attachment [79], and perceived humanness correlate with perceived relationship quality with IVAs and other AI systems. To examine whether these aspects of social perception differed between the friend cluster and the non-friend cluster, we conducted two-sided Welch's *t*-tests. To analyze the temporal patterns of these social variables within and across clusters, we performed RM-ANOVAs.

3.4.1. Measures

Parasocial Interaction. We measured parasocial interactions (PSI) using the *Universal PSI Scale* [80]. The scale measured the PSI processes on a total of 14 subdimensions, where each of the subdimensions contained four items. The PSI processes were summarized on cognitive, affective, and behavioral/non-verbal dimensions. All were answered on a 5-point scale (1 = not at all; 5 = very much). The reliabilities of the individual subdimensions ranged from $\alpha = 0.69$ (counter empathy) to $\alpha = 0.88$ (antipathy).

Empathy. To measure participants' empathy toward their voice assistant, the Psychological Involvement—Empathy subscale ($\alpha = 0.86$) from the Social Presence module of the *Game Experience Questionnaire* was used [81] and adapted to the voice assistant (e.g., I felt connected to the voice assistant). For six items, subjects indicated the extent to which the statements applied to them using a 5-point Likert scale (0 = not at all; 4 = extremely).

Social Sense. To measure how social participants perceived their voice assistant, the *social presence* ($\alpha = 0.73$; With the voice assistant, it feels like there is another person in the room), *likeability* ($\alpha = 0.52$; I like my voice assistant), and *status* ($\alpha = 0.66$; The voice assistant has a higher social status than I do) subscales of Bailenson, et al. [82] were measured with a total of 10 items. On a 7-point Likert scale, participants indicated the extent to which the statements applied to them (−3 = strongly disagree; +3 strongly agree).

Attachment. To measure how attached participants felt to their voice assistant, the *Inclusion of Others in the Self (IOS) Scale* [83] was used ($\alpha = 0.93$). Participants were asked to select the circle illustration that best described their relationship with their voice assistant. The more the circles overlapped, the greater the perceived attachment to the voice assistant.

Uncanny Valley. Ho and MacDorman [84] subscales were used to assess the perceived Humaneness and Eeriness of the voice assistant in terms of the Uncanny Valley effect with 14 items. Participants rated the voice assistant on a five-point Likert scale for bipolar

adjectives of *humaneness* (6 adjectives, e.g., artificial vs. natural; $\alpha = 0.85$) and *eeriness* (8 adjectives, e.g., calming vs. scary; $\alpha = 0.74$) indices.

3.4.2. Results

The results of Welch's *t*-tests are listed in Table 9.

Table 9. Overview of descriptive data and results of Welch's *t*-tests on social variables.

	Cluster	n	M	SD	T-Value [1]	df	Effect Size (Cohen's d)
Cognitive PSI	Friend	36	3.21	0.68	3.58 ***	46.35	0.92
	Non-Friend	28	2.44	0.97			
Affective PSI	Friend	36	2.72	0.90	4.09 ***	61.97	1.01
	Non-Friend	28	1.91	0.68			
Behavioral PSI	Friend	36	4.07	0.89	4.35 ***	48.81	1.11
	Non-Friend	28	2.91	1.18			
Social Presence	Friend	36	2.86	0.93	4.77 ***	55.32	1.16
	Non-Friend	28	2.00	0.49			
Likeability	Friend	36	2.10	1.04	3.03 **	61.88	0.75
	Non-Friend	28	1.41	0.77			
Status	Friend	36	3.46	0.60	5.04 ***	56.33	1.27
	Non-Friend	28	2.67	0.64			
Empathy	Friend	36	3.31	1.00	7.17 ***	61.92	1.78
	Non-Friend	28	1.69	0.80			
Attachment	Friend	40	1.58	0.93	3.11 **	48.03	0.70
	Non-Friend	33	1.09	0.29			
Humaneness	Friend	38	2.61	0.77	4.29 ***	66.36	1.02
	Non-Friend	32	1.94	0.55			
Eeriness	Friend	38	3.55	0.69	4.01 ***	67.15	0.96
	Non-Friend	32	2.91	0.65			

[1] *** $p < 0.001$, ** $p < 0.01$. PSI = Parasocial Interaction.

Parasocial Interaction (Group Differences). For the analysis of parasocial interaction, T11 was used as the reference time point. Individuals in the friend cluster ($n = 36$) scored significantly higher on the cognitive ($t(46.35) = 3.58$, $p < 0.001$, $d = 0.92$), affective ($t(61.97) = 4.09$, $p < 0.001$, $d = 1.01$), and behavioral ($t(48.81) = 4.35$, $p < 0.001$, $d = 1.11$) parasocial interactions compared to individuals in the non-friend cluster ($n = 28$).

Parasocial Interaction (Longitudinal Effects). A positive and significant main effect with a mean effect size was observed between T5 ($M = 2.10$, $SE = 0.10$) and T11 ($M = 2.31$, $SE = 0.10$) for affective parasocial interaction ($N = 63$, $F(1, 61) = 4.78$, $p = 0.033$, $\eta^2 p = 0.07$). Notably, we observed a significant interaction effect for cognitive parasocial interaction ($F(1, 61) = 9.57$, $p = 0.003$, $\eta^2 p = 0.14$) (Figure 8). Cognitive parasocial interaction increased significantly for the friend cluster from T5 ($M = 2.93$, $SE = 0.14$) to T11 ($M = 3.21$, $SE = 0.14$) ($p = 0.042$).

Social Sense (Group Differences). T11 was used as the referent time point. The friend cluster had higher values for social presence ($M = 2.86$, $SD = 0.93$; non-friend cluster: $M = 2.00$, $SD = 0.93$), likeability ($M = 2.10$, $SD = 1.04$; non-friend cluster: $M = 1.41$, $SD = 0.77$), and status ($M = 3.46$, $SD = 0.60$; non-friend cluster: $M = 2.67$, $SD = 0.64$). The differences between the friend cluster ($n = 36$) and the non-friend cluster ($n = 28$) in social presence ($t(55.32) = 4.77$, $p < 0.001$, $d = 1.16$), likeability ($t(61.88) = 3.30$, $p = 0.005$, $d = 0.75$), and status ($t(56.33) = 5.03$, $p = 0.004$, $d = 1.27$) were significant.
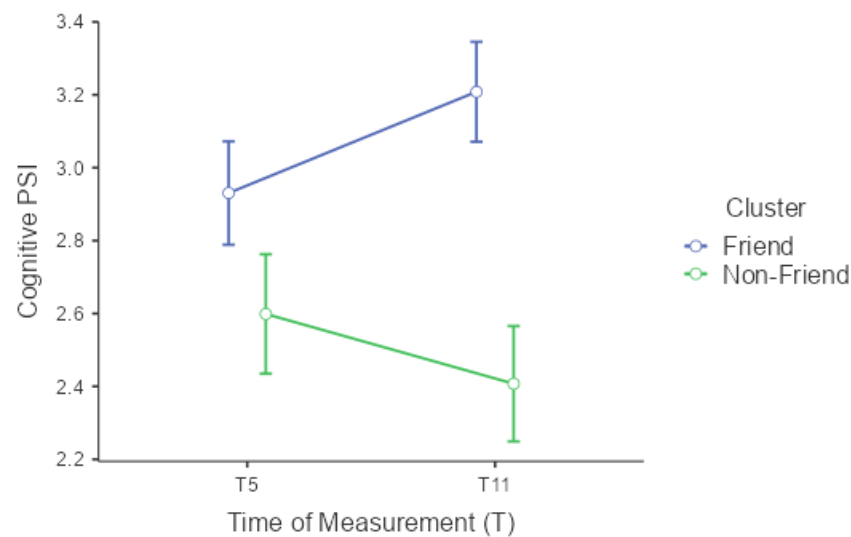
Figure 8. Change in cognitive parasocial interaction (PSI) over time separated by cluster with standard errors.

Social Sense (Longitudinal Effects). A significant main effect was found for social presence ($N = 63$, $F(1, 61) = 5.94$, $p = 0.018$, $\eta^2 p = 0.09$). Social presence decreased from T5 ($M = 2.65$, $SE = 0.10$) to T11 ($M = 2.43$, $SE = 0.10$). The interaction effects did not become significant.

Empathy (Group Differences). Reference time point T11 was used to assess the difference in empathy between the two clusters. The friend cluster ($n = 36$, $M = 3.31$, $SD = 0.99$) showed greater empathy toward the IVA than the non-friend cluster ($n = 28$, $M = 1.69$, $SD = 0.80$). The difference was significant ($t(61.92) = 7.14$, $p < 0.001$) and the effect was large ($d = 1.78$).

Empathy (Longitudinal Effects). In the period between T5 ($M = 2.42$, $SE = 0.11$) and T11 (2.49, $SE = 0.12$) we did not detect a significant main effect for empathy ($N = 63$; $F(1, 61) = 0.40$, $p = 0.531$, $\eta^2 p = 0.01$) but we did detect a significant interaction effect with a medium effect size ($F(1, 61) = 7.47$, $p = 0.008$, $\eta^2 p = 0.109$). Empathy in the friend cluster ($n = 36$) increased significantly from T5 ($M = 2.93$, $SE = 0.15$) to T11 ($M = 3.31$, $SE = 0.15$) ($p < 0.001$). The non-friend cluster ($n = 27$) showed no significant changes ($p = 1.00$) between T5 ($M = 1.90$, SE = 0.17) and T11 ($M = 1.67$, $SE = 0.18$) (Figure 9).
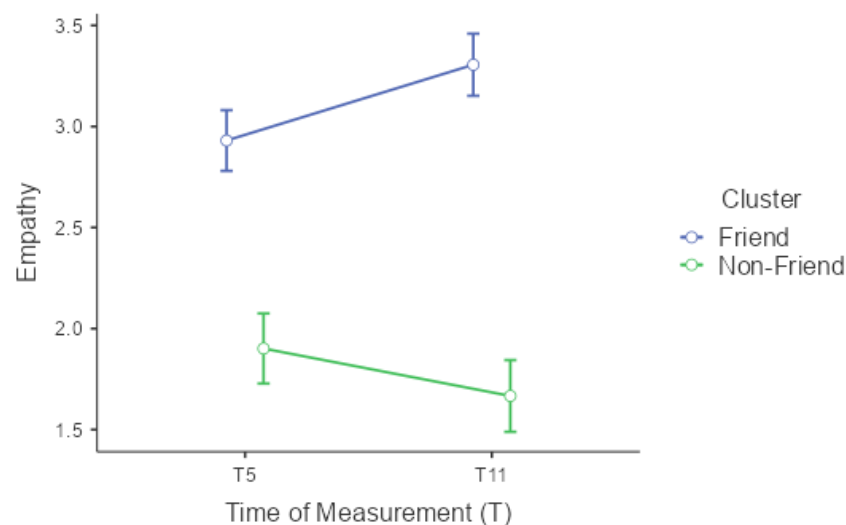
Figure 9. Change in empathy over time separated by cluster with standard errors.

Attachment (Group Differences). T10 served as the reference time point for analyzing attachment. The attachment scores of individuals in the friend cluster ($n = 40$, $M = 1.58$, $SD = 0.93$) were found to be significantly higher towards the voice assistant ($t(48.03) = 3.11$, $p = 0.003$, $d = 0.70$) as compared to those in the non-friend cluster ($n = 33$, $M = 1.09$, $SD = 0.29$).

Attachment (Longitudinal Effects). We could not identify a significant main effect ($N = 68$, $F(1, 66) = 0.34$, $p = 0.562$, $\eta^2 p = 0.01$) and interaction effect ($F(1, 66) = 0.00$, $p = 0.973$, $\eta^2 p = 0.00$) for the clusters in relation to attachment.

Uncanny Valley (Group Differences). To assess perceived humaneness and eeriness, T9 was chosen as the reference time point. The friend cluster ($n = 38$) perceived the voice assistant to be significantly ($t(66.36) = 4.29$, $p < 0.001$, $d = 1.02$) more human ($M = 2.61$, $SD = 0.77$) than the non-friend cluster ($n = 32$, $M = 1.94$, $SD = 0.55$). In addition, the friend cluster perceived their voice assistant ($M = 3.55$, $SD = 0.69$) to be significantly more eery ($t(67.15) = 4.01$, $p < 0.001$, $d = 0.96$) than the non-friend cluster ($M = 2.91$, $SD = 0.65$).

Uncanny Valley (Longitudinal Effects). Between T3 ($M = 2.23$, $SE = 0.09$) and T9 ($M = 2.31$, $SE = 0.08$), the main effect ($N = 63$, $F(1, 61) = 0.74$, $p = 0.395$, $\eta^2 p = 0.01$) and interaction effect ($F(1, 61) = 0.22$, $p = 0.638$, $\eta^2 p = 0.00$) for the humaneness subscale were not significant. The main effect for eeriness was significant at a medium effect size ($N = 63$, $F(1, 61) = 8.20$, $p = 0.006$, $\eta^2 p = 0.12$). Eeriness reduced significantly from T3 ($M = 3.49$, $SE = 0.08$) to T9 ($M = 3.25$, $SE = 0.08$). The interaction effect for eeriness ($F(1, 61) = 0.84$, $p = 0.362$, $\eta^2 p = 0.01$) was not significant.

3.4.3. Brief Discussion of Section 4

Our analyses of the social perception of IVAs revealed several differences between individuals in the friend cluster and those in the non-friend cluster. Cognitive parasocial interactions were more pronounced in the friend cluster, which suggests that they paid more attention to the voice assistant, evaluated its actions, and perceived similarities between themselves and the assistant [80]. These findings are consistent with previous research indicating that parasocial relationships can foster perceptions of friendship [85,86]. Both clusters showed an increase in affective parasocial interactions throughout the study, indicating that participants may have experienced emotional interactions marked by sympathy, empathy, or antipathy, regardless of the cluster assignment [80].

The friend cluster had a stronger perception of the voice assistant's social presence and a greater association of the voice assistant with higher status and likeability. Notably, the sample's perception of the voice assistant's social presence decreased over time, regardless of the cluster assignment. However, overall, it is evident that people from the friend cluster felt more connected with and had more empathy toward the voice assistant. The perceived empathy toward the voice assistant increased over time in the friend cluster. The data showed that the friend cluster perceived the voice assistant to be more human and more eery. This relationship is unsurprising considering the Uncanny Valley effect. The Uncanny Valley effect refers to the phenomenon that when artificial human-like entities (such as robots) become more and more human-like, viewers' positive impressions and sympathy diminish until they finally reach a point where they are almost (but not quite) human-like [87]. Conversely, we can assume that the friend cluster may have perceived the voice assistant to be so human-like that this perception led to an unsettling feeling.

3.5. Section 5—Personality Traits

Previous research has shown that users' personalities (e.g., extraversion, agreeableness) [88], loneliness [52,54], and attachment styles [18] are related to how human and social IVAs and other AI systems are perceived to be. This may suggest that certain personality traits of humans can promote the perceived friendship quality of IVAs. Therefore, we performed a two-sided Welch *t*-test to determine if there were any differences between the two clusters concerning the above personality traits.

3.5.1. Measures

Personality. The *NEO-FFI* [89] is a questionnaire that uses 60 items to measure the dimensions of the Big 5 personality model (neuroticism, extraversion, openness, agreeableness, and conscientiousness). Participants indicated the extent to which they agreed or disagreed with each statement on a five-point Likert scale (0 = strongly disagree; 4 = strongly agree). The German version of the questionnaire was developed by Borkenau and Ostendorf [90]. The reliability of the questionnaire is in the acceptable to good range (neuroticism $\alpha = 0.81$, extraversion $\alpha = 0.77$, openness $\alpha = 0.73$, agreeableness $\alpha = 0.73$, conscientiousness $\alpha = .83$).

Loneliness. We used the *De Jong Gierveld Loneliness Scale* with six items [91] to measure loneliness. The questionnaire consisted of the subscales emotional loneliness ($\alpha = 0.74$) and social loneliness ($\alpha = 0.73$). Each item was assessed on a 7-point Likert scale (1 = strongly disagree; 7 = strongly agree). The internal consistency of the questionnaire was $\alpha = 0.76$.

Attachment Patterns. To measure the participants' attachment patterns, the *Relationship Scales Questionnaire (RSQ)* [92] was used, which was translated into German by Steffanowski, et al. [93]. The *RSQ* used a five-point Likert scale (1 = strongly disagree; 5 = strongly agree) to assess the degree of defined attachment patterns. The questionnaire contained 30 items and the subscales fear of separation ($\alpha = 0.81$), fear of closeness ($\alpha = 0.77$), lack of trust ($\alpha = 0.77$), and desire for independence ($\alpha = 0.72$).

3.5.2. Results

Personality. When analyzing the potential personality differences between the clusters, we observed the initial tendencies. The difference between the friend cluster ($n = 40$) and the non-friend cluster ($n = 32$) for neuroticism was marginally significant ($t(62.18) = 1.75$, $p = 0.084$, $d = 0.42$). The friend cluster showed higher scores in neuroticism ($M = 4.28$, $SD = 0.88$) than the non-friend cluster ($M = 3.89$, $SD = 1.00$). The differences in extraversion, openness, agreeableness, and conscientiousness were not significant.

Loneliness. We found significant differences in social ($t(69.80) = 2.56$, $p = 0.013$, $d = 0.60$) and emotional ($t(69.80) = 2.56$, $p = 0.032$, $d = 0.52$) loneliness between the friend cluster ($n = 40$) and the non-friend cluster ($n = 32$) (Figure 10). The friend cluster had higher scores ($M = 3.03$, $SD = 1.30$) in social loneliness than the non-friend cluster ($M = 2.30$, $SD = 1.10$). The friend cluster also had higher scores ($M = 3.88$, $SD = 1.30$) in emotional loneliness than the non-friend cluster ($M = 3.40$, $SD = 0.86$).
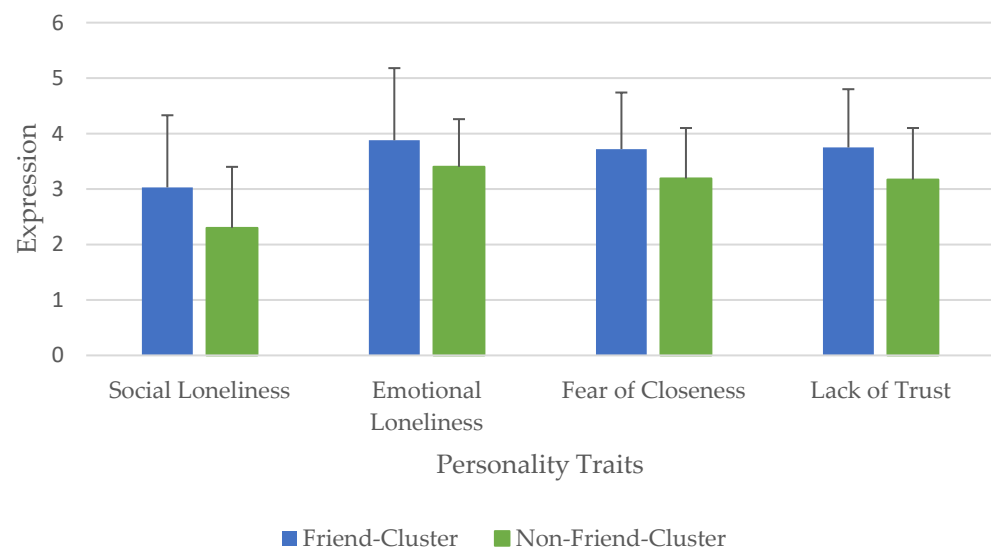
Figure 10. Comparison of means with plotted standard deviations of loneliness and attachment styles separated by clusters.

Attachment Patterns. We identified significant differences in the fear of closeness ($t(69.19) = 2.33$, $p = 0.023$, $d = 0.55$) and lack of trust ($t(69.26) = 2.48$, $p = 0.015$, $d = 0.59$) between the clusters (Figure 10). The friend cluster had significantly higher scores than the non-friend cluster in the fear of closeness (friend cluster: $M = 3.72$, $SD = 1.02$; non-friend cluster: $M = 3.19$, $SD = 0.91$) and lack of trust (friend cluster: $M = 3.75$, $SD = 1.05$; non-friend cluster: $M = 3.17$, $SD = 0.93$).

3.5.3. Brief Discussion of Section 5

We found differences between the friend cluster and the non-friend cluster in terms of their personality traits. First, the friend cluster showed marginally higher scores in neuroticism. Other studies have previously found a positive relationship between neuroticism and anthropomorphizing tendencies [88,94]. Neuroticism is characterized by negative emotional states and instability and is associated with social anxiety and avoidance of social evaluations [95,96]. Moreover, neuroticism is positively correlated with loneliness [97,98], which can increase the probability of being pessimistic, anxious, and distrustful in social situations [99,100]. This may promote a motivation to value voice assistants as friends, as their social presence may provide a low-threat alternative to real-world contact and is less associated with a fear of judgment. Our results regarding loneliness show that individuals in the friend cluster were significantly more socially and emotionally lonely than individuals in the non-friend cluster. Consistent with this finding, Epley et al. [101] showed that the lonelier people are, the more likely they are to socialize and anthropomorphize with non-human entities. Additionally, individuals with insecure attachment styles and attachment anxiety also tend to exhibit higher levels of anthropomorphization [50]. In addition, our results indicate that the friend cluster was characterized by attachment styles driven by a fear of closeness and lack of trust. The resulting relationship building with IVAs may then compensate for social needs and deficiencies [18]. The suspected relationships should be further investigated in future studies.

4. General Discussion

AI-based technologies with adaptive and intelligent features imitate human social traits and evolve into putative social interaction partners or friends. We combined data science and social science methods to cluster participants regarding their attribution of friend-like social roles to IVAs. The longitudinal study investigated how these roles developed over nine months and their impact on usage behavior, user experience, and social perceptions. The results revealed that users who associated IVAs with higher friendship quality differed from those who did not, as they used the devices more often for various types of tasks, were more satisfied, rated the interactions as more enjoyable, and indicated a greater intent to use voice assistants in the future. In addition, we found differences between the clusters in terms of the social perceptions of voice assistants. Users attributing friendship to their IVA reported feelings of empathy and connectedness toward their voice assistant. In addition, the user's personality determined the emergence and manifestation of role attribution. Thus, the cluster that associated the voice assistant more strongly with a perception of friendship showed a higher expression of loneliness and an insecure attachment style.

In relation to our **first research question** (see Section 1), we found significant differences in how users attributed friendship to voice assistants and how the perceived friendship quality changed over time. Our results showed that implicit measures could better classify role attribution than explicit measures. Thus, our findings support previous results that social roles are often perceived unconsciously by users [15,16] and explicit questioning may lead to the denial of these roles [17,55]. Our analysis identified two clusters based on whether participants associated voice assistants more (friend cluster) or less (non-friend cluster) with friendship qualities. Time-based analyses showed that the perceived friendship quality increased significantly over time for the friend cluster, whereas it remained unchanged for the non-friend cluster.

Regarding the **second research question** (see Section 2), we examined how user behavior with the IVA differed between the friend cluster and non-friend cluster and how it changed over time. We found that users who viewed the voice assistant as a friend (friend cluster) interacted with it in different ways than the non-friend cluster. Users who associated the IVA with higher friendship quality tended to use it more often for support and social interaction. They were also more likely to use it for knowledge acquisition and mood management. In social interaction, the friend cluster showed more interest in the voice assistant as a personality and addressed it directly. These results are consistent with those of Purington, Taft, Sannon, Bazarova, and Taylor [10], who found that users were more likely to interact socially with their voice assistant and address it as a personality when they saw it as a friend or companion. Data science methods were used to reveal additional role-specific interaction patterns over time. For example, individuals in the friend cluster used more words for their voice commands by the end of the study period, whereas the opposite trend was observed in the non-friend cluster.

The third research question (see Section 3) explored the differences in user experience with IVAs between the friend cluster and the non-friend cluster) and how the user experience evolved over time. Individuals in the friend cluster perceived the IVA differently based on their user experience. They valued their interactions with the IVA more and evaluated their experiences as more meaningful, significant, inspiring, and emotional. We found that the need for competence and autonomy, for example, was more strongly fulfilled in individuals from the friend cluster. Thus, perceived friendship with the voice assistant may be a predictor and key variable for a positive user experience. Furthermore, we found that the IVA's role as a friend was associated with more social interactions. This supports the findings of Purington, Taft, Sannon, Bazarova, and Taylor [10], who found that personification (i.e., treating the voice assistant as a human) was associated with higher user satisfaction. The significant influence of the parasocial perception of IVAs on the user experience suggests a vast design space. Although our results show only correlations, the consequences can be positive or negative, underscoring the high responsibility in designing future human–AI interfaces.

In the **fourth research question** (see Section 4), we determined whether social perceptions differed between the friend cluster and the non-friend cluster and the temporal trends of these social perceptions. Users who associated IVAs with higher friendship quality had a different social perception of them. Individuals in the friend cluster felt more empathy, as well as an attachment to the voice assistant and cognitive parasocial interactions, which increased over time. The studies by Youn and Jin [36], Hernandez-Ortega and Ferreira [85], and Ramadan, F Farah, and El Essrawi [86] showed positive correlations between parasocial relationships and users' perceived attachment to IVAs, which may contribute to the perception of IVAs as friends. This could be one reason for the increase in friendship quality over time in the friend cluster. In addition, the anthropomorphic perception of AI can influence the nature of the relationship with the user [102]. Our study found that individuals in the friend cluster perceived the IVA as more human and socially present. The fact that voice assistants are always present and listening can foster a sense of friendship but it should be noted that they cannot replace human friends and are merely programs that perform tasks and provide information. It is important to continue monitoring the impact of Artificial Intelligence on the perceived boundaries between technology and human relationships.

In our **fifth research question** (see Section 5), we explored the potential differences between the personality traits of users in the friend cluster and those in the non-friend cluster. In this way, we gained valuable insights into whether certain personality traits might make people more prone to associate voice assistants with friendly qualities. We found that individuals in the friend cluster were more likely to exhibit lonely and attachment-anxious personality traits and slightly higher scores on neuroticism. Voice assistants may be appealing to individuals with these personality traits because they can compensate for socially imbalanced needs and provide a sense of companionship [18,52,53]. In this way, the voice assistant can act as a friend that users can turn to when they are lonely or need someone

to talk to. Although voice assistants cannot feel emotions, they can provide users with a sense of connection. In the future, it will be important to recognize that although IVAs are capable of human-like conversations and responses, they have no real emotions or needs and cannot replace human relationships.

4.1. Limitations

The participants in the study were very young and predominantly female, thereby limiting the representativeness and generalizability of the results. Furthermore, the study was limited to Alexa and Google Assistant, which are only two of the available voice assistants. Future research should consider other voice assistants to allow for a comparison of their impact on user perception and behavior. The present study is also limited in its internal validity. Unlike laboratory studies, for example, it was not possible to perfectly control whether participants conscientiously integrated smart speakers into their daily lives. This may also be why this study experienced some data losses. The usage analysis used participants' technical interaction data, which were documented by the provider in usage logs. We received some empty files, which could have been due to system complications or the fact that some participants were inconsistent in participating in the study. Nonetheless, the present study benefits from real-world interaction scenarios and external validity, which are usually limited in laboratory settings.

To measure the social roles attributed by participants to the IVAs, we created a questionnaire based on the findings of Purington, Taft, Sannon, Bazarova, and Taylor [10]. This demonstrated low internal consistency, which may have limited the reliable measurement of social roles. Participants' denial of the voice assistant's role as a friend was very high. Although studies suggest that the conscious assignment of a friend-based role to voice assistants is low [31], measurement error cannot be ruled out due to the limitations of the assessment tool. We encourage future studies to develop psychometric instruments that measure the role-specific characteristics and perceptions of voice assistants and compare their effectiveness with implicit methods for measuring the role of IVAs.

4.2. Future Directions

Future research should aim for greater participant diversity and consider factors such as cultural and socioeconomic backgrounds, usage levels, and device ownership. Our study's predominantly female, tech-savvy participants without prior experience with voice assistants highlights the need for a more diverse range of participants. The inclusion of more diverse users and contextual variables could provide greater insight into the factors underlying the different clusters. In addition, it would be beneficial to investigate the presence of additional clusters with social relationships. Examining the impact of users' experience levels with voice assistants on user experience, social perception, and usage is also crucial. Regarding usage, our study found that the primary use of voice assistants was to control the smart home and listen to music. This usage pattern can be attributed to a variety of factors such as a lack of knowledge about the IVA's capabilities, lack of practicality, frustration with speech recognition errors, or motivation issues during the study. To gain a deeper understanding of the impact of voice assistants on user experience and behavior, future research should involve more extensive usage of these systems. It is important to consider that advancements in the field of voice-based systems, especially regarding their anthropomorphic design and ability to communicate proactively, fluently, and naturally, may have an impact on the quality of the interactions [103]. Therefore, training programs to educate participants on the full range of functions of voice assistants could increase or change the kind of usage. Evaluating the effectiveness of such training programs might be an interesting topic for future research.

Currently, IVAs are not able to initiate conversations and their dialogue flow often fails to meet users' expectations [104]. However, as IVAs become more advanced and capable of more natural and fluid conversations, social bonds, including feelings of friendship, may become stronger. Recent developments such as *ChatGPT* indicate the potential for this

ability. OpenAI announced the powerful *ChatGPT* conversational language model, which can generate natural language texts and conduct dialogues [105]. This development could be particularly significant for smart speaker vendors, as machines will be better equipped to respond to follow-up questions and recognize the humor and mood of users [105,106]. Furthermore, even slight changes in the volume, speech rate, or pitch can affect the personified perception of voice assistants [25]. Therefore, future studies should investigate the effects of IVAs' dialogue capability or adaptability (e.g., gender of voice) on users' perceived friendship and relationship quality. We anticipate that even small social adaptations to AI-based voice systems can have large impacts. In our study, we observed that the perceived friendship quality scores of users with IVAs varied significantly between clusters, despite the differences in the descriptive values not being particularly large. Nevertheless, even these small differences between groups had significant effects on the variables we studied. As technology continues to advance and become more social, the potential consequences and effects of such developments may become even stronger. In this context, we must take a closer look at the potential negative effects that could arise from friendship relationships between users and IVAs, such as issues related to information credibility [76], impulsive shopping behavior [107], or disclosure of personal information [56]. Wienrich, Reitelbach, and Carolus [7] showed that the social role of IVAs can have an impact on the disclosure of sensitive information. Therefore, it would be particularly important to examine users' disclosure behavior as a function of perceived friendship quality and consider privacy concerns and behaviors in this context.

The results of this study have implications for the design and development of voice assistants. Friendship with voice assistants can increase brand engagement and customer loyalty or lead to higher customer satisfaction. Studies show that a strong social relationship can lead to user satisfaction with IVAs [46]. Sharabany's friendship quality subscales [58] could be used as design guidelines to build deeper, friend-like relationships between users and voice assistants. For example, a voice assistant could proactively ask users about their plans to fulfill the "Frankness and Spontaneity" subscale. Given our research and that of others, this may be particularly relevant for users of IVAs in seeking social interaction or compensating for a lack of social contact [108]. The effects of design using Sharabany's subscales should be investigated in future laboratory and longitudinal studies. The results of this study could have further applications in clinical psychology research and practice. In particular, computer-assisted methods in mental healthcare are becoming increasingly important. In this context, chatbots are an effective method for alleviating depressive and anxiety symptoms [109]. Furthermore, recent studies have shown that voice assistants are accepted and preferred by older people in the communication of therapy methods [110]. The relationship between the patient and therapist is a critical factor in determining the success of therapy [111–113]. Nevertheless, the quality of the friendships that patients cultivate in their personal lives can also positively impact treatment outcomes [114]. The findings from our study could be used in the design of future voice-based systems to make applications more personal and therapy services more helpful.

4.3. Conclusions

Overall, it was shown that people who attributed friendship to their voice assistant exhibited different usage behavior and had a different user experience and quality of interaction. In addition, our findings provide valuable insights into how user personality traits may influence the perceived social role of IVAs. Our long-term approach and interdisciplinary data collection and analysis contribute to a holistic analysis of users in their natural environment. The results provide new research and design ideas and promote a deeper understanding of AI–human interaction. The results show that the consideration of human needs and social processes is essential in the design of IVAs. Recent developments in conversational language models such as *ChatGPT* show that the current study is highly relevant and provides an outlook on the effects of the attributions and expectations of human-like voice-based AI.

Research Highlights

Perceived Friendship Quality:

- Users differ in their attribution of friendship qualities to voice assistants and can be grouped as such.

Usage Behavior:

- Users who attribute higher friendship qualities to IVAs are more likely to use them for support functions (e.g., local guide, time queries, calendar) and integrate them into their daily lives.

User Experience:

- Users who attribute higher friendship qualities to IVAs have more of their pragmatic, hedonic, eudaimonic, and social needs met, whereas those who attribute lower friendship qualities perceive interactions with IVAs as less stimulating over time.

Social Perception:

- Users who attribute higher friendship qualities to IVAs perceive them to be more socially present, like them more, and assign them a higher status. They also feel more empathy and attachment toward the voice assistant, with perceived empathy and cognitive parasocial interactions increasing over time.

Personality Traits:

- Users who attribute higher friendship qualities to IVAs scored significantly higher in loneliness and insecure attachment patterns, which may promote their friendship perceptions of IVAs.

References

1. Clark, L.; Doyle, P.; Garaialde, D.; Gilmartin, E.; Schlögl, S.; Edlund, J.; Aylett, M.; Cabral, J.; Munteanu, C.; Edwards, J. The state of speech in HCI: Trends, themes and challenges. *Interact. Comput.* **2019**, *31*, 349–371. [CrossRef]
2. Dunn, J. Virtual assistants like Siri and Alexa look poised to explode. Available online: https://www.businessinsider.com/virtual-assistants-siri-alexa-growth-chart-2016-8 (accessed on 23 February 2023).

3.   Chattaraman, V.; Kwon, W.-S.; Gilbert, J.E.; Ross, K. Should AI-Based, conversational digital assistants employ social-or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. *Comput. Hum. Behav.* **2019**, *90*, 315–330. [CrossRef]

4.   Ki, C.-W.C.; Cho, E.; Lee, J.-E. Can an intelligent personal assistant (IPA) be your friend? Para-friendship development mechanism between IPAs and their users. *Comput. Hum. Behav.* **2020**, *111*, 106412. [CrossRef]

5.   Liu, N.; Pu, Q. Can Smart Voice Assistant Induce Social Facilitation Effect? A Preliminary Study. In Proceedings of the International Conference on Human-Computer Interaction, Copenhagen, Denmark, 19–24 July 2020; pp. 616–624.

6.   Carolus, A.; Wienrich, C.; Törke, A.; Friedel, T.; Schwietering, C.; Sperzel, M. 'Alexa, I feel for you!' Observers' empathetic reactions towards a conversational agent. *Front. Comput. Sci.* **2021**, *46*, 682982. [CrossRef]

7.   Wienrich, C.; Reitelbach, C.; Carolus, A. The Trustworthiness of Voice Assistants in the Context of Healthcare Investigating the Effect of Perceived Expertise on the Trustworthiness of Voice Assistants, Providers, Data Receivers, and Automatic Speech Recognition. *Front. Comput. Sci.* **2021**, *53*, 685250. [CrossRef]

8.   Wu, S.; He, S.; Peng, Y.; Li, W.; Zhou, M.; Guan, D. An empirical study on expectation of relationship between human and smart devices—With smart speaker as an example. In Proceedings of the Fourth International Conference on Data Science in Cyberspace (DSC), Hangzhou, China, 23–25 June 2019; pp. 555–560.

9.   Turk, V. Home invasion. *New Sci.* **2016**, *232*, 16–17. [CrossRef]

10.   Purington, A.; Taft, J.G.; Sannon, S.; Bazarova, N.N.; Taylor, S.H. "Alexa is my new BFF" Social Roles, User Satisfaction, and Personification of the Amazon Echo. In Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 2853–2859.

11.   Voit, A.; Niess, J.; Eckerth, C.; Ernst, M.; Weingärtner, H.; Woźniak, P.W. 'It's not a romantic relationship': Stories of Adoption and Abandonment of Smart Speakers at Home. In Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia, Essen, Germany, 22–25 November 2020; pp. 71–82.

12.   Bentley, F.; Luvogt, C.; Silverman, M.; Wirasinghe, R.; White, B.; Lottridge, D. Understanding the long-term use of smart speaker assistants. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* **2018**, *2*, 1–24. [CrossRef]

13.   Mavrina, L.; Szczuka, J.; Strathmann, C.; Bohnenkamp, L.M.; Krämer, N.; Kopp, S. "Alexa, You're Really Stupid": A Longitudinal Field Study on Communication Breakdowns Between Family Members and a Voice Assistant. *Front. Comput. Sci.* **2022**, *4*, 791704. [CrossRef]

14.   Carolus, A.; Wienrich, C. Adopting Just Another Digital Assistant or Establishing Social Interactions with a New Friend? Conceptual Research Model of a Long-Term Analysis of First-Time Users' Adoption and Social Interactions with Smart Speakers. In Proceedings of the Mensch und Computer Conference, Darmstadt, Germany, 4–7 September 2022; pp. 498–502.

15.   Nass, C.; Steuer, J.; Tauber, E.R. Computers are social actors. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 24–28 April 1994; pp. 72–78.

16.   Kim, Y.; Sundar, S.S. Anthropomorphism of computers: Is it mindful or mindless? *Comput. Hum. Behav.* **2012**, *28*, 241–250. [CrossRef]

17.   Frennert, S.; Eftring, H.; Östlund, B. What older people expect of robots: A mixed methods approach. In Proceedings of the International Conference on Social Robotics, Bristol, UK, 27–29 October 2013; pp. 19–29.

18.   Epley, N.; Waytz, A.; Cacioppo, J.T. On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychol. Rev.* **2007**, *114*, 864–886. [CrossRef]

19.   Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*; Pearson Education: Munich, Germany, 2020; Volume 3.

20.   Reeves, B.; Nass, C. *The media equation: How people treat computers, television, and new media like real people*; Cambridge University Press: Cambridge, UK, 1996; Volume 10, p. 236605.

21.   Carolus, A.; Muench, R.; Schmidt, C.; Schneider, F. Impertinent mobiles-Effects of politeness and impoliteness in human-smartphone interaction. *Comput. Hum. Behav.* **2019**, *93*, 290–300. [CrossRef]

22.   Karr-Wisniewski, P.; Prietula, M. CASA, WASA, and the dimensions of us. *Comput. Hum. Behav.* **2010**, *26*, 1761–1771. [CrossRef]

23.   Gambino, A.; Fox, J.; Ratan, R.A. Building a stronger CASA: Extending the computers are social actors paradigm. *Hum.-Mach. Commun.* **2020**, *1*, 71–85. [CrossRef]

24.   Go, E.; Sundar, S.S. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Comput. Hum. Behav.* **2019**, *97*, 304–316. [CrossRef]

25.   Nass, C.I.; Brave, S. *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*; MIT Press Cambridge: Cambridge, MA, USA, 2005.

26.   Sciuto, A.; Saini, A.; Forlizzi, J.; Hong, J.I. "Hey Alexa, What's Up?" A Mixed-Methods Studies of In-Home Conversational Agent Usage. In Proceedings of the 2018 Designing Interactive Systems Conference, Hong Kong, China, 9–13 June 2018; pp. 857–868.

27.   Couper, M.P.; Tourangeau, R.; Steiger, D.M. Social presence in web surveys. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, WA, USA, 31 March–5 April 2001; pp. 412–417.

28.   Tourangeau, R.; Couper, M.P.; Steiger, D.M. Humanizing self-administered surveys: Experiments on social presence in web and IVR surveys. *Comput. Hum. Behav.* **2003**, *19*, 1–24. [CrossRef]

29.   Dreitzel, H.P. *Die Gesellschaftlichen Leiden Und Das Leiden an der Gesellschaft. Vorstudien zu Einer Pathologie des Rollenverhaltens*; F. Enke Verlag: Stuttgart, Germany, 1968; pp. 10–17.

30. Warpefelt, H.; Verhagen, H. A model of non-player character believability. *J. Gaming Virtual Worlds* **2017**, *9*, 39–53. [CrossRef] [PubMed]

31. Wang, X.; Wang, B.; Han, G.; Zhang, H.; Xie, X. Is It Just a Tool Or Is It a Friend?: Exploring Chinese Users' Interaction and Relationship with Smart Speakers. In *Perceiving the Future through New Communication Technologies: Robots, AI and Everyday Life*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 129–146.

32. Wagner, L. Good character is what we look for in a friend: Character strengths are positively related to peer acceptance and friendship quality in early adolescents. *J. Early Adolesc.* **2019**, *39*, 864–903. [CrossRef]

33. Schweitzer, F.; Belk, R.; Jordan, W.; Ortner, M. Servant, friend or master? The relationships users build with voice-controlled smart devices. *J. Mark. Manag.* **2019**, *35*, 693–715. [CrossRef]

34. Rhee, C.E.; Choi, J. Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Comput. Hum. Behav.* **2020**, *109*, 106359. [CrossRef]

35. Han, S.; Yang, H. Understanding adoption of intelligent personal assistants: A parasocial relationship perspective. *Ind. Manag. Data Syst.* **2018**, *118*, 618–636. [CrossRef]

36. Youn, S.; Jin, S.V. "In A.I. we trust?" The effects of parasocial interaction and technopian versus luddite ideological views on chatbot-based customer relationship management in the emerging "feeling economy". *Comput. Hum. Behav.* **2021**, *119*, 106721. [CrossRef]

37. Rubin, A.M.; Perse, E.M.; Powell, R.A. Loneliness, parasocial interaction, and local television news viewing. *Hum. Commun. Res.* **1985**, *12*, 155–180. [CrossRef]

38. Horton, D.; Richard Wohl, R. Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry* **1956**, *19*, 215–229. [CrossRef] [PubMed]

39. Meyrowitz, J. *No Sense of Place: The Impact of Electronic Media on Social Behavior*; Oxford University Press: New York, NY, USA, 1986.

40. Nordlund, J.-E. Media interaction. *Commun. Res.* **1978**, *5*, 150–175. [CrossRef]

41. Conway, J.C.; Rubin, A.M. Psychological predictors of television viewing motivation. *Commun. Res.* **1991**, *18*, 443–463. [CrossRef]

42. Kim, J.; Rubin, A.M. The variable influence of audience activity on media effects. *Commun. Res.* **1997**, *24*, 107–135. [CrossRef]

43. Hsieh, S.H.; Lee, C.T. Hey Alexa: Examining the effect of perceived socialness in usage intentions of AI assistant-enabled smart speaker. *J. Res. Interact. Mark.* **2021**, *15*, 267–294. [CrossRef]

44. Goudey, A.; Bonnin, G. Must smart objects look human? Study of the impact of anthropomorphism on the acceptance of companion robots. *Rech. Et Appl. En Mark. (Engl. Ed.)* **2016**, *31*, 2–20. [CrossRef]

45. Chowanda, A.; Flintham, M.; Blanchfield, P.; Valstar, M. Playing with social and emotional game companions. In Proceedings of the Intelligent Virtual Agents: 16th International Conference, Los Angeles, CA, USA, 20–23 September 2016; pp. 85–95.

46. Lavado-Nalvaiz, N.; Lucia-Palacios, L.; Pérez-López, R. The role of the humanisation of smart home speakers in the personalisation–privacy paradox. *Electron. Commer. Res. Appl.* **2022**, *53*, 101146. [CrossRef]

47. Blut, M.; Wang, C.; Wünderlich, N.V.; Brock, C. Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *J. Acad. Mark. Sci.* **2021**, *49*, 632–658. [CrossRef]

48. Cao, C.; Zhao, L.; Hu, Y. Anthropomorphism of Intelligent Personal Assistants (IPAs): Antecedents and Consequences. In Proceedings of the Pacific Asia Conference on Information Systems (PACIS), Xi'an, China, 8–12 July 2019.

49. Gao, Y.; Pan, Z.; Wang, H.; Chen, G. Alexa, my love: Analyzing reviews of amazon echo. In Proceedings of the 2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, Guangzhou, China, 8–12 October 2018; pp. 372–380.

50. Whelan, J.; Hingston, S.T.; Thomson, M. Does growing up rich and insecure make objects seem more human? Childhood material and social environments interact to predict anthropomorphism. *Personal. Individ. Differ.* **2019**, *137*, 86–96. [CrossRef]

51. Jones, V.K.; Hanus, M.; Yan, C.; Shade, M.Y.; Blaskewicz Boron, J.; Maschieri Bicudo, R. Reducing Loneliness Among Aging Adults: The Roles of Personal Voice Assistants and Anthropomorphic Interactions. *Front. Public Health* **2021**, *9*. [CrossRef] [PubMed]

52. Scherr, S.A.; Meier, A.; Cihan, S. Alexa, tell me more–about new best friends, the advantage of hands-free operation and life-long learning. In *Mensch und Computer 2020-Workshopband*; Gesellschaft für Informatik e.V.: Bonn, Germany, 2020.

53. Mishra, A.; Shukla, A.; Sharma, S.K. Psychological determinants of users' adoption and word-of-mouth recommendations of smart voice assistants. *Int. J. Inf. Manag.* **2022**, *67*, 102413. [CrossRef]

54. Poushneh, A. Humanizing voice assistant: The impact of voice assistant personality on consumers' attitudes and behaviors. *J. Retail. Consum. Serv.* **2021**, *58*, 102283. [CrossRef]

55. Nass, C.; Moon, Y. Machines and mindlessness: Social responses to computers. *J. Soc. Issues* **2000**, *56*, 81–103. [CrossRef]

56. Ischen, C.; Araujo, T.; Voorveld, H.; van Noort, G.; Smit, E. Privacy concerns in chatbot interactions. In Proceedings of the International Workshop on Chatbot Research and Design, Amsterdam, The Netherlands, 19–20 November 2019; pp. 34–48.

57. Rasch, D.; Kubinger, K.D.; Moder, K. The two-sample t test: Pre-testing its assumptions does not pay off. *Stat. Pap.* **2011**, *52*, 219–231. [CrossRef]

58. Sharabany, R. Intimate friendship scale: Conceptual underpinnings, psychometric properties and construct validity. *J. Soc. Pers. Relatsh.* **1994**, *11*, 449–469. [CrossRef]

59. Ketchen, D.J.; Shook, C.L. The application of cluster analysis in strategic management research: An analysis and critique. *Strateg. Manag. J.* **1996**, *17*, 441–458. [CrossRef]
60. Jian, A.K. Data clustering: 50 years beyond k-means, pattern recognition letters. *Corrected Proof* **2010**, *31*, 651–666.
61. Tibshirani, R.; Walther, G.; Hastie, T. Estimating the number of clusters in a data set via the gap statistic. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2001**, *63*, 411–423. [CrossRef]
62. Martinez, W.L.; Martinez, A.R.; Solka, J. *Exploratory Data Analysis with MATLAB*; Chapman and Hall/CRC: New York, NY, USA, 2017.
63. Burns, R.; Burns, R.P. *Business Research Methods and Statistics Using SPSS*; Sage Publications: London, UK, 2008.
64. Jenkins-Guarnieri, M.A.; Wright, S.L.; Johnson, B. Development and validation of a social media use integration scale. *Psychol. Pop. Media Cult.* **2013**, *2*, 38–50. [CrossRef]
65. Przybylski, A.K.; Ryan, R.M.; Rigby, C.S. The motivating role of violence in video games. *Personal. Soc. Psychol. Bull.* **2009**, *35*, 243–259. [CrossRef]
66. Mayring, P. Qualitative Inhaltsanalyse–Abgrenzungen, Spielarten, Weiterentwicklungen. *Forum Qual. Soz./Forum: Qual. Soc. Res.* **2019**, *20*, 1–14.
67. Garg, R.; Sengupta, S. He is just like me: A study of the long-term use of smart speakers by parents and children. *ACM Interact. Mob. Wearable Ubiquitous Technol.* **2020**, *4*, 1–24. [CrossRef]
68. Amazon. Alexa Kennenlernen. Available online: https://www.amazon.de/b?ie=UTF8&node=12775495031 (accessed on 22 March 2023).
69. Pradhan, A.; Findlater, L.; Lazar, A. "Phantom Friend" or "Just a Box with Information" Personification and Ontological Categorization of Smart Speaker-based Voice Assistants by Older Adults. *Proc. ACM Hum.-Comput. Interact.* **2019**, *3*, 1–21. [CrossRef]
70. Hassenzahl, M.; Monk, A. The inference of perceived usability from beauty. *Hum.–Comput. Interact.* **2010**, *25*, 235–260. [CrossRef]
71. Huta, V. Eudaimonic and hedonic orientations: Theoretical considerations and research findings. In *Handbook of Eudaimonic Well-Being*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 215–231.
72. Hassenzahl, M.; Wiklund-Engblom, A.; Bengs, A.; Hägglund, S.; Diefenbach, S. Experience-oriented and product-oriented evaluation: Psychological need fulfillment, positive affect, and product perception. *Int. J. Hum.-Comput. Interact.* **2015**, *31*, 530–544. [CrossRef]
73. Huta, V.; Ryan, R.M. Pursuing pleasure or virtue: The differential and overlapping well-being benefits of hedonic and eudaimonic motives. *J. Happiness Stud.* **2010**, *11*, 735–762. [CrossRef]
74. Seaborn, K.; Pennefather, P.; Fels, D.I. Eudaimonia and hedonia in the design and evaluation of a cooperative game for psychosocial well-being. *Hum.–Comput. Interact.* **2020**, *35*, 289–337. [CrossRef]
75. Cho, M.; Lee, S.-s.; Lee, K.-P. Once a kind friend is now a thing: Understanding how conversational agents at home are forgotten. In Proceedings of the Designing Interactive Systems Conference, San Diego, CA, USA, 23–28 June 2019; pp. 1557–1569.
76. Pradhan, A.; Lazar, A.; Findlater, L. Use of intelligent voice assistants by older adults with low technology use. *ACM Trans. Comput.-Hum. Interact. (TOCHI)* **2020**, *27*, 1–27. [CrossRef]
77. Trajkova, M.; Martin-Hammond, A. "Alexa is a Toy": Exploring older adults' reasons for using, limiting, and abandoning echo. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–13.
78. Leite, I.; Pereira, A.; Mascarenhas, S.; Martinho, C.; Prada, R.; Paiva, A. The influence of empathy in human–robot relations. *Int. J. Hum.-Comput. Stud.* **2013**, *71*, 250–260. [CrossRef]
79. Loureiro, S.M.C.; Japutra, A.; Molinillo, S.; Bilro, R.G. Stand by me: Analyzing the tourist–intelligent voice assistant relationship quality. *Int. J. Contemp. Hosp. Manag.* **2021**, *33*, 3840–3859. [CrossRef]
80. Schramm, H.; Hartmann, T. The PSI-Process Scales. A new measure to assess the intensity and breadth of parasocial processes. *Communications* **2008**, *33*, 385–401. [CrossRef]
81. IJsselsteijn, W.A.; De Kort, Y.A.; Poels, K. The game experience questionnaire. *Tech. Univ. Eindh.* **2013**, *46*, 1–9.
82. Bailenson, J.N.; Aharoni, E.; Beall, A.C.; Guadagno, R.E.; Dimov, A.; Blascovich, J. Comparing behavioral and self-report measures of embodied agents' social presence in immersive virtual environments. In Proceedings of the 7th Annual International Workshop on PRESENCE, Valencia, Spain, 13–15 October 2004; pp. 216–223.
83. Aron, A.; Aron, E.N.; Smollan, D. Inclusion of other in the self scale and the structure of interpersonal closeness. *J. Personal. Soc. Psychol.* **1992**, *63*, 596–612. [CrossRef]
84. Ho, C.-C.; MacDorman, K.F. Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Comput. Hum. Behav.* **2010**, *26*, 1508–1518. [CrossRef]
85. Hernandez-Ortega, B.; Ferreira, I. How smart experiences build service loyalty: The importance of consumer love for smart voice assistants. *Psychol. Mark.* **2021**, *38*, 1122–1139. [CrossRef]
86. Ramadan, Z.; F Farah, M.; El Essrawi, L. From Amazon. com to Amazon. love: How Alexa is redefining companionship and interdependence for people with special needs. *Psychol. Mark.* **2021**, *38*, 596–609. [CrossRef]
87. Mori, M. The uncanny valley: The original essay by Masahiro Mori. Available online: https://web.ics.purdue.edu/~drkelly/MoriTheUncannyValley1970.pdf (accessed on 23 February 2023).

88. Letheren, K.; Kuhn, K.-A.L.; Lings, I.; Pope, N.K.L. Individual difference factors related to anthropomorphic tendency. *Eur. J. Mark.* **2016**, *50*, 973–1002. [CrossRef]
89. McCrae, R.R.; Costa Jr, P.T. A contemplated revision of the NEO Five-Factor Inventory. *Personal. Individ. Differ.* **2004**, *36*, 587–596. [CrossRef]
90. Borkenau, P.; Ostendorf, F. *NEO-Fünf-Faktoren-Inventar (NEO-FFI) Nach Costa und McCrae: Handanweisung*; Hogrefe: Göttinger, Germany, 1993.
91. Gierveld, J.D.J.; Tilburg, T.V. A 6-item scale for overall, emotional, and social loneliness: Confirmatory tests on survey data. *Res. Aging* **2006**, *28*, 582–598. [CrossRef]
92. Griffin, D.W.; Bartholomew, K. Relationship scales questionnaire. *J. Personal. Soc. Psychol.* **1994**. [CrossRef]
93. Steffanowski, A.; Oppl, M.; Meyerberg, J.; Schmidt, J.; Wittmann, W.W.; Nübling, R. Psychometrische Überprüfung einer deutschsprachigen version des relationship scales questionnaire (RSQ). In *Störungsspezifische Therapieansätze – Konzepte und Ergebnisse*; Bassler, M., Ed.; Psychosozial Verlag: Gießen, Germany, 2001; pp. 320–342.
94. Kaplan, A.D.; Sanders, T.; Hancock, P.A. The relationship between extroversion and the tendency to anthropomorphize robots: A Bayesian analysis. *Front. Robot. AI* **2019**, *5*, 135. [CrossRef]
95. Kashdan, T.B.; McKnight, P.E. The darker side of social anxiety: When aggressive impulsivity prevails over shy inhibition. *Curr. Dir. Psychol. Sci.* **2010**, *19*, 47–50. [CrossRef]
96. Nestler, S.; Back, M.D.; Egloff, B. Psychometrische Eigenschaften zweier Skalen zur Erfassung interindividueller Unterschiede in der Präferenz zum Alleinsein. *Diagnostica* **2011**, *57*, 57–67. [CrossRef]
97. Stokes, J.P. The relation of social network and individual difference variables to loneliness. *J. Personal. Soc. Psychol.* **1985**, *48*, 981–990. [CrossRef]
98. Abdellaoui, A.; Chen, H.Y.; Willemsen, G.; Ehli, E.A.; Davies, G.E.; Verweij, K.J.; Nivard, M.G.; de Geus, E.J.; Boomsma, D.I.; Cacioppo, J.T. Associations between loneliness and personality are mostly driven by a genetic association with neuroticism. *J. Personal.* **2019**, *87*, 386–397. [CrossRef]
99. Cacioppo, J.T.; Hawkley, L.C. Perceived social isolation and cognition. *Trends Cogn. Sci.* **2009**, *13*, 447–454. [CrossRef]
100. Cacioppo, J.T.; Hughes, M.E.; Waite, L.J.; Hawkley, L.C.; Thisted, R.A. Loneliness as a specific risk factor for depressive symptoms: Cross-sectional and longitudinal analyses. *Psychol. Aging* **2006**, *21*, 140–151. [CrossRef]
101. Epley, N.; Waytz, A.; Akalis, S.; Cacioppo, J.T. When we need a human: Motivational determinants of anthropomorphism. *Soc. Cogn.* **2008**, *26*, 143–155. [CrossRef]
102. Kim, A.; Cho, M.; Ahn, J.; Sung, Y. Effects of gender and relationship type on the response to artificial intelligence. *Cyberpsychology Behav. Soc. Netw.* **2019**, *22*, 249–253. [CrossRef]
103. Roy, R.; Naidoo, V. Enhancing chatbot effectiveness: The role of anthropomorphic conversational styles and time orientation. *J. Bus. Res.* **2021**, *126*, 23–34. [CrossRef]
104. Luger, E.; Sellen, A. "Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 7–12 May 7 2016; pp. 5286–5297.
105. OpenAI. GPT-4 Technical Report. *arXiv* **2023**, arXiv:2303.08774.
106. Shafeeg, A.; Shazhaev, I.; Mihaylov, D.; Tularov, A.; Shazhaev, I. Voice Assistant Integrated with Chat GPT. *Indones. J. Comput. Sci.* **2023**, *12*, 1. [CrossRef]
107. Rzepka, C.; Berger, B.; Hess, T. Why another customer channel? Consumers' perceived benefits and costs of voice commerce. In Proceedings of the 53rd Hawaii International Conference on System Sciences, Honolulu, HI, USA, 7–10 January 2020.
108. Choi, T.R.; Drumwright, M.E. "OK, Google, why do I use you?" Motivations, post-consumption evaluations, and perceptions of voice AI assistants. *Telemat. Inform.* **2021**, *62*, 101628. [CrossRef]
109. Narynov, S.; Zhumanov, Z.; Gumar, A.; Khassanova, M.; Omarov, B. Chatbots and Conversational Agents in Mental Health: A Literature Review. In Proceedings of the 2021 21st International Conference on Control, Automation and Systems (ICCAS), Jeju, Republic of Korea, 12–15 October 2021; pp. 353–358.
110. Striegl, J.; Gotthardt, M.; Loitsch, C.; Weber, G. *Investigating the Usability of Voice Assistant-Based CBT for Age-Related Depression*; Springer International Publishing: Cham, Switzerland, 2022; pp. 432–441.
111. Bordin, E.S. The generalizability of the psychoanalytic concept of the working alliance. *Psychother. Theory Res. Pract.* **1979**, *16*, 252–260. [CrossRef]
112. Beutler, L.E.; Harwood, T.M. *Prescriptive Psychotherapy: A Practical Guide to Systematic Treatment Selection*; Oxford University Press: New York, NY, USA, 2000.
113. Horvath, A.O. Research on the alliance. In *The Working Alliance: Theory, Research, and Practice*; Horvath, A.O., Greenberg, L.S., Eds.; John Wiley & Sons: New York, NY, USA, 1994.
114. Baker, J.; Hudson, J. Friendship quality predicts treatment outcome in children with anxiety disorders. *Behav. Res. Ther.* **2013**, *51*, 31–36. [CrossRef] [PubMed]