

Google Now is for the Extraverted, Cortana for the Introverted: Investigating the Influence of Personality on IPA Preference.

Patrick Ehrenbrink
Quality and Usability Lab
Technische Universität Berlin
Ernst-Reuter-Platz 7
10587 Berlin
patrick.ehrenbrink@tu-berlin.de

Seif Osman
Quality and Usability Lab
Technische Universität Berlin
Ernst-Reuter-Platz 7
10587 Berlin
seif.osman@campus.tu-berlin.de

Sebastian Möller
Quality and Usability Lab
Technische Universität Berlin
Ernst-Reuter-Platz 7
10587 Berlin
sebastian.moeller@tu-berlin.de

ABSTRACT

¹ Intelligent Personal Assistants (IPAs) have become a mainstream product and often show a high level of variance in terms of their interaction philosophy. For example, Google's Now has no personality at all, whereas personality plays a strong part in the advertisement of Apple's Siri. We have assessed the personality profile of users and their preference for either Apple's Siri, Google's Now or Microsoft's Cortana, based on attractiveness and psychological state reactivity. Analysis revealed how the preference for an IPA depends on a person's character traits. Preferences of individual traits are discussed and average profiles for devotees of different IPAs are given. The results can be used to recommend IPAs specifically to a user's personality profile. The work concludes with a number of recommendations for the design of IPAs to address specific personality traits of users.

CCS CONCEPTS

• **Ubiquitous and mobile computing** → **Ubiquitous and mobile devices**; *Personal digital assistants* • **User characteristics**

KEYWORDS

Intelligent Personal Assistants, IPA, Siri, Cortana, Now, Big Five

ACM Reference format:

P. Ehrenbrink, S. Osman and S. Möller. 2017. Google Now is for the Extraverted, Cortana for the Introverted: Investigating the Influence of Personality on IPA Preference. In *Proceedings of the 29th Australian Conference on Human-Computer Interaction, Brisbane, Texas Australia, December 2017 (OZCHI'17)*, 8 pages. <https://doi.org/10.1145/3152771.3152799>

1 INTRODUCTION

Intelligent Personal Assistants (IPAs) are becoming more popular in mobile contexts.

For example, Microsoft's Cortana is available not only on Windows 10 Mobile smartphones, but also on Android, iOS and

Windows 10 for desktop computers [22]. The three big IPAs, Microsoft's Cortana [22], Apple's Siri [2], Amazon's Alexa [1] and Google's Now (followed by Assistant) [13] are all able to control third party devices such as smart home devices and offer hands-free interaction for, e.g., usage while driving cars. The manufacturers of the IPAs follow different approaches to make them appeal to their customers and to represent the desired company brand. On one hand, Siri and Cortana clearly try to imitate or appear as having a personality by offering non-productive interaction, such as telling jokes or responding emotionally to personal questions. On the other hand, Now does not do this. Interaction with Now resembles conventional human-computer interaction more closely. The user states a question and then receives either a neutral, unemotional answer or no answer. Besides underlining the desired brand image of the developing company, such design decisions may also influence the acceptance of the respective IPA for different users, based on the users' personal preferences and personality. Naturally, it is impossible to equally appeal to all users simultaneously, as some aspects might please one user but repel another.

A solution for this problem is user-specific adaptation. As IPAs have a large set of functions that is constantly growing, it is hardly possible for users to effectively compare and choose IPAs on the basis of their functionality. Also, IPAs are meant for frequent interaction which makes their ease of use particularly important. Adapting interaction behaviour to fit the individual user's preferences as close as possible can therefore be of high importance for each IPA developer in order to influence the user's decision towards choosing a particular IPA over the other available ones. The current study investigates which attributes of people's personalities correspond to preferences of IPA usage. To accomplish this, six different personality traits were measured among the test population. Also, each participant of the study interacted with the three IPAs Siri, Cortana and Now and indicated whether he or she liked each IPA according to self-report scales. This study is meant to provide a starting point for adapting IPAs specifically to users or user groups according to their personality profile.

There is a large number of IPAs with different interaction modalities, functions and market share. For this study, the IPAs Siri, Cortana and Now were selected because they have a similar target audience and are well integrated in their respective mobile operating systems, even though Cortana is not exclusive to Windows 10 Mobile. Amazon's Alexa and Google's Assistants

¹ OzCHI '17, November 28-December 1, 2017, Brisbane, QLD, Australia
© 2017 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-5379-3/17/11.
<https://doi.org/10.1145/3152771.3152799>

were not part of this study because at the time of the experiment the integration of German (the study was conducted in Germany) as interaction language was not fully functional.

1.1 Big Five

The Big Five personality traits have been identified by various research groups independently and are thus highly reliable constructs [20][12]. The Big Five can be regarded as higher level factors, because they consists of other personality traits that are less generalizable [12]. With small variations in naming across different authors, the Big Five include: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism. Each factor will be briefly explained in the following.

1.1.1 Openness. Openness describes the level of openness for new experiences. Persons that show a high score of in this trait often seek new experiences and the unfamiliar [23]. A stereotypical high scorer of this trait would be a performance artist, a stereotypical low scorer would be a clerk.

1.1.2 Conscientiousness. Conscientiousness describes a person’s character in terms of organization and goal-orientation [23]. Highly Conscientious people are usually well organized and reliable and may show a high level of discipline. People with low scores in conscientiousness will probably be more hedonistic and careless.

1.1.3 Extraversion. Extraversion describes the extend of a person’s sociableness. People that show a high score on extraversion are usually sociable and talkative, whereas people that show a low score on extraversion are rather reserved and quiet [23].

1.1.4 Agreeableness. Agreeableness is a trait that assesses the general attitude of a person towards others in terms of antagonism or compassion. People with low scores in agreeableness are typically rude or cynical, whereas people with high scores in agreeableness are warm-hearted and trusting [23].

1.1.5 Neuroticism. The trait neuroticism describes the emotional stability of a person. A person who achieves a high score in neuroticism is nervous and insecure and prone to worrying. A person with a low score in neuroticism can be regarded as emotionally stable and relaxed [23].

1.2 Psychological Reactance

Reactance is a concept from social psychology that was introduced by Brehm [5][6]. There are two manifestations of reactance. State reactance and trait reactance, that will be briefly introduced in the following.

1.2.1 State Reactance. State reactance is a motivational or affective state that that can bear negative consequences for the opinion of a person. A person that enters a reactant state often also engages in reactant behavior, namely negative re-evaluation and acting against the (ascribed) wishes of the entity that caused the reactance. For example, heightened reactance has been shown to correlate with diminished acceptability of a Smart TV [10], also, sometimes public health campaigns cause reactance which then results in the opposite behavior of the addressed

people [9]. State reactance is thought to be triggered mainly by a perceived loss of personal freedom of choice or control [5][6]. However, also other factors are known to influence state reactance. Roubroeks et al. found that the level of social agency of a recommender system can moderate state reactance [24]. According to an evaluation of different models by Dillard and Shen [9], state reactance is a higher level construct that is comprised of anger and negative cognitions.

1.2.2 Trait Reactance. Trait reactance is a personality trait and can be regarded as a person’s proneness to enter a reactant state. State reactance could first be assessed by the “Fragebogen zur Messung der psychologischen Reaktanz” [21]. This questionnaire was later translated and slightly modified via several iterations by Hong [17][15][16].

2 RESEARCH QUESTION

Personality traits influence behaviour and their manifestation can be very different between persons. It is unclear whether they also influence whether someone likes an IPA or not. Therefore, the research question is: Do personality traits influence whether a person likes or dislikes certain IPAs?

3 METHODS

3.1 Participants

Twenty-four persons participated in the current study. All participants were fluent German speakers. The age ranged between 18 and 35 with a mean age of 25.29 years. The gender was roughly balanced with 13 females and 11 males who took part in the study. The former experience with intelligent personal assistants was unbalanced between the IPAs. While 11 out of 24 participants had used Siri before, only one participant had used Cortana before. An overview of the previous experiences of the participants with the three IPAs is given in Tab. 1. An average personality profile of all participants can be viewed in Fig. 1.

3.2 Personality Assessment

Personality was assessed using standardized questionnaires. The Big Five personality traits were assessed with the Ten-Item Personality Inventory [14]. It measures the Big Five personality traits with a set of ten items on a Likert-type [19] scale.

Trait reactance was assessed with the revised version of Hong’s Psychological Reactance Scale [17][15][16]. It measures trait reactance based on eleven Likert-type items. Following the

Table 1: Prior experience with Siri, Cortana and Now.
N=24

	knew it before	used it before
Siri	22	11
Cortana	14	1
Now	20	9

Table 2: Average ratings of each IPA on Attractiveness (Att.), State Reactance (S.R.) and Approval Score (A.S.).

	Att.	S. R.	A.S.	
Siri	4.81	-0.27	0,43	relatively positive
Cortana	4.25	0.62	-0,83	average
Now	4.66	-0.35	0,40	
All	4.57	0	0	relatively negative

suggestion of other authors, this scale was regarded as unidimensional [18][25].

3.3 IPA Rating

Each IPA was rated by the participants with the AttrakDiff Mini questionnaire [8]. The AttrakDiff Mini is comprised of ten semantic differentials that correspond to the three dimensions Pragmatic Quality, Hedonistic Quality and Attractiveness. Only Attractiveness was considered in this work, as it represents a global judgment of a product. This dimension is represented via two 7-point semantic differentials within the AttrakDiff Mini.

The participants also rated the IPAs in terms of state reactance. State reactance can be regarded as a counterpart to acceptance [10]. It was measured using a mixed method approach where its component anger was measured with a four-item Likert-type questionnaire with a 5-point scale. The number of negative cognitions towards the stimulus (Siri, Cortana or Now) were counted in a thought-listing task [9], where a user writes down all thoughts that come to her mind and afterwards rate if each thought was negative, neutral or positive. Only negative thoughts are counted. Afterwards, the results from the anger questionnaire and the number of negative thoughts are both z-transformed [4] and added, forming a score for state reactance. The z-transformation [4] allows to add or compare data from different scales because it transforms a dataset into a state where it has a mean value of zero and a standard deviation of 1.

3.4 Approval Score

An approval score was calculated to approximate user's acceptance for each IPA. There is no consensus on how to measure acceptance. Therefore, it was decided to include two measures that are supposed to approximate acceptance. The first measure is the attractiveness dimension of the AttrakDiff Mini [8] questionnaire. The second measure is state reactance as induced by the IPA.

An approval score for each IPA was calculated for each person, resulting in three scores per person. The z-transformation was calculated using the overall average and standard deviation over all IPAs and persons, which ensures a fair comparison among the IPAs.

The values of attractiveness and state reactance were z-transformed. Since attractiveness has a positive valence and state

reactance has a negative valence, state reactance (z-transformed) was subtracted from attractiveness (z-transformed). To achieve a better readability, the result was again z-transformed, because this results in a mean value of zero over the sample [3]. An overview of the ratings in attractiveness and state reactance and the resulting approval scores can be viewed in Tab. 2.

3.4.1 Interpretation. The z-transformed approval score will be used to determine if a certain personality likes an IPA or not. Since the average of the z-transformed score is zero, a positive value can be regarded as liking, whereas a negative value can be regarded as disliking of an IPA, relative to the average population sample.

3.5 Procedure

The experiment took 60 minutes on average. Before the participants interacted with the IPAs, they filled out questionnaires to assess their personality profile. Those included Hong's Psychological Reactance Scale [17][15][16] and the Ten-Item Personality Inventory [14].

An alternative to this scale would have been a longer Big Five questionnaire, such as the NeoFFI [3]. The shorter questionnaire was used to reduce the amount of questions that the participants had to answer during the experiment. This also provided the opportunity of longer interaction times with the IPAs. Basic demographic data was assessed, including age, gender and previous experiences or knowledge about Siri, Cortana and Now.

Afterwards, all participants received a short introduction on how to operate an IPA. They then started the main part of the experiment, in which they were given a set of 24 tasks that they had to perform with each of the IPAs. The tasks included simple queries like "When was Einstein born?", but also complex task-cascades like searching for nearby restaurants, then select the second on the list, then find out opening hours and navigate there. The tasks were selected to include a wide variety of functions that the IPAs can perform. Smart Home functions like turning on light or increase room temperature were not included.

The tasks were chosen in a way that they could be completed by all IPAs. A problem in IPA operation by novices is a lack of knowledge regarding the appropriate commands. This often results in long sessions of trying out different commands and might cause frustration. To counteract this and since the goal of the interaction was to give the participants a realistic impression of the functionality of each IPA from an expert user perspective, the exact commands that the participants could use in order to perform a task were handed out to them on a sheet of paper. As a result, most of the participants were able to solve all tasks. After the participants had completed the tasks with one IPA, they were asked to fill out the AttrakDiff Mini [8], the anger questionnaire and then had to perform a thought listing with the instructor [9]. The participants then proceeded with the next IPA.

Table 3: Correlations between the measured personality traits. Significant correlations are indicated with *, highly significant correlations are indicated with **. O=Openness, C= Conscientiousness, E=Extraversion, A=Agreeableness, N=Neuroticism.

	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Trait Reactance
O	-0.29	-0.18	-0.54**	0.34	-0.59**
C		0.43*	0.45*	-0.22	0.37
E			0.42*	-0.43*	0.17
A				-0.33	0.18
N					-0.17

In order to avoid sequence effects, the sequence of the IPAs was changed with each participant and balanced so that each of the six possible sequences was performed by four participants in total.

4 RESULTS

4.1 IPA Ratings

4.1.1 T-Tests. Paired samples T-Tests with the data of all participants were conducted to give an overview of the general rating of the IPAs. The tests did show a significant (alpha < 0.05) effect on the approval score between Cortana and Siri, where Siri was rated better with $t(23)=-2.240, p=0.030$. Also, Now was rated significantly better than Cortana in the approval score with $t(23)=-2.125, p=0.039$. No significant differences in the approval score ratings could be found between Siri and Now. See [Tab. 2](#) for the descriptive results.

4.1.2 Individual IPA Preference. The preferred IPA of each person was determined, by comparing the approval scores of the three IPAs. The IPA with the highest approval score was regarded as the preferred IPA of that person. According to this method, Siri was preferred by nine persons, Now was preferred by eight persons and Cortana was preferred by six person. One person achieved identical approval scores for all three IPAs. A further investigation of the individual data points that led to the approval scores of that person revealed no obvious irregularities. The average personality profiles of the participants who preferred either Siri, Cortana or Now were calculated and are shown in [Fig. 1](#). Also, the average personality profile of all participants can be viewed there.

4.2 Correlation of personality traits

A Pearson correlation was calculated to explore correlations among the personality traits that were used in this study.

Table 4: Average approval scores calculated for different grouping of the participants according to their personality traits. "All" represents the average of Siri, Cortana and Now.

IPA	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Trait Reactance	Trait Score
Siri	0,38	1,19	-0,92	0,37	0,98	-0,01	high
	0,47	0,42	0,67	0,45	0,44	0,91	medium
	0,40	-0,23	0,00	0,43	0,02	0,26	low
Cortana	0,36	-1,56	-2,87	-3,59	0,12	-2,34	high
	0,33	-0,73	-0,88	-0,41	-1,54	-0,21	medium
	-2,78	-0,30	0,43	0,56	-1,19	-0,20	low
Now	1,12	1,37	1,78	1,58	0,34	0,33	high
	0,13	0,71	0,37	0,41	1,30	1,06	medium
	0,41	-0,79	-0,15	-0,59	0,04	-0,27	low
All	0,62	0,33	-0,67	-0,55	0,48	-0,67	high
	0,31	0,13	0,05	0,15	0,06	0,59	medium
	-0,66	-0,44	0,09	0,13	-0,38	-0,07	low

The results can be viewed in [Tab. 3](#). There is a highly significant (alpha < 0.01), medium, negative correlation between trait reactance and openness. Also, a highly significant, medium, negative correlation between Agreeableness and Openness was observed. Several other positive and negative correlations that reached significance and were of medium strength were found, too.

4.3 Personality specific results

The six personality traits of each participant were assessed during the experiment. Afterwards, the median of each personality trait over all participants was calculated. The participants were then graded as either high, medium or low for each personality trait.

When assigning the participants to personality groups according their personality traits, it was tried to create equally large groups for each level of the traits (low, medium, high). However, as the Ten-Item Personality Inventory only provides two items per personality trait and a five-point Likert scale was used, the resolution of the trait scores was not always sufficient to be able to discriminate three equally sized groups. In such cases, more participants were assigned to the medium group. The medium approval scores for each personality trait and its value-group were calculated for Siri, Cortana and Now. The results can be viewed in [Tab. 4](#).

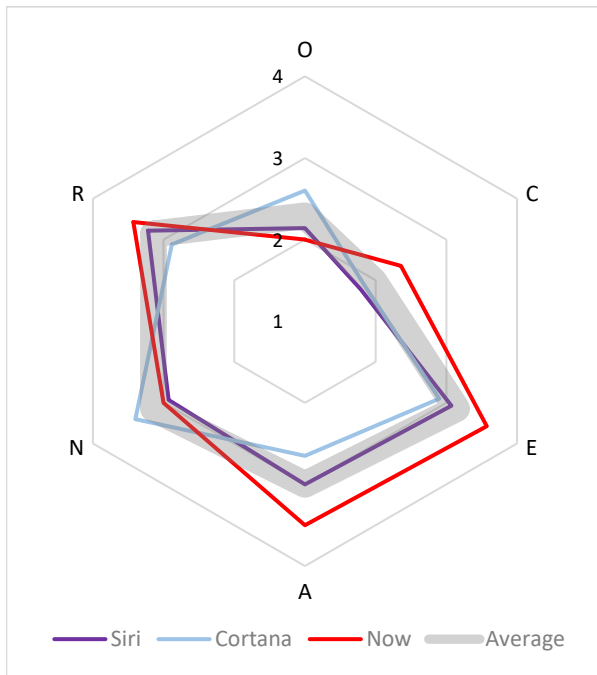


Figure 1: Average personality profiles of users that preferred either Siri (9), Cortana (6) or Now (8). The grey line represents the average personality profile of the whole test population. O=Openness, C=Conscientiousness, E=Extraversion, A=Agreeableness, N=Neuroticism, R=Trait Reactance.

Paired samples T-Tests were performed to investigate effects between IPAs within the different levels of each personality trait (low, medium and high).

4.3.1 Openness. Paired samples T-Tests show a highly significant effect between Cortana and Siri with $t(8)=-4.831$, $p=0.001$ and between Cortana and Now with $t(8)=-3.988$, $p=0.004$ in the low openness group, where Cortanas ratings were lower than Siri or Now. Medium or high openness showed no effect on IPA preference.

4.3.2 Conscientiousness. Analysis showed a significant difference between Cortana and Siri with $t(6)=-2.876$, $p=0.028$. Also, the difference between Cortana and Now fell just short on significance with $t(6)=-2.026$, $p=0.089$ in the high conscientiousness group. Again, Cortana was rated lower than Siri or Now. Low or medium conscientiousness showed no significant effect on IPA preference.

4.3.3 Extraversion. Paired samples T-Tests showed a significant effect in the medium extraversion group. Cortana's approval score is significantly lower compared to Siri's with $t(17)=-2.620$, $p=0.018$. Also, Cortana's approval score is lower than that of Now, but fell just short of significance with $t(17)=-1.974$, $p=0.065$. Also, there were no significant effects of low or high extraversion on IPA preference.

4.3.4 Agreeableness. Significant effects between Cortana and Siri with $t(4)=-4.438$, $p=0.011$, and highly significant effects between Cortana and Now with $t(4)=-10.843$, $p<0.001$ were found

with paired samples T-Tests in the high agreeableness group with Cortana being rated lower than Siri and Now. There were no significant effects of high or medium agreeableness on IPA preference.

4.3.5 Neuroticism. Paired samples T-Tests revealed no significant effects in any of the three neuroticism groups. However, the difference between Cortana and Siri fell just short of significance with $t(4)=-2.234$, $p=0.089$ in the low neuroticism group. Also, the difference between Cortana and Now fell just short of significance with $t(4)=-2.404$, $p=0.074$ in the low neuroticism group. Again, Cortana was rated lower than Siri or Now in this group.

4.3.6 Trait Reactance. There was a significant effect between Cortana and Siri in the high trait reactance group with $t(6)=-2.578$, $p=0.042$. The Difference between Cortana and Now fell just short of significance with $t(6)=-2.248$, $p=0.066$ in the same group. Cortana was rated lower than Siri and Now in this Group. No significant effects were observed in the low or medium trait reactance groups.

5 DISCUSSION

5.1 Trait Reactance as an Addition to the Big Five

The Big Five are constructs that are regarded as the universal traits that cover more fine-grained personality traits, which are not always reproducible or generalizable [12]. This also implies that trait reactance can be represented as a function of the Big Five in some way (this should be possible for other fine-grained personality traits).

Since state reactance was an important dependent variable in the presented study, it was decided to include trait reactance, anyway. The analysis of correlations among the six personality traits showed a medium correlation between trait reactance and openness. However, the correlation between openness and agreeableness was similarly high. Also, there were some other correlations of medium strength among the Big Five traits. Therefore, we regard trait reactance as not redundant among the six measured traits.

5.2 Traits and Approval Score

5.2.1 Average Profiles. We build average profiles of the participants that rated either Cortana, Siri or Now highest. The resulting profiles are shown in Fig. 1. It is evident, that the group that prefers Siri matches the average personality profile of all participants very closely. This also explains the group that preferred Siri over the other IPAs was largest. The Now group shows in general more highly expressed personality traits towards the valence that was used in this study, except for openness, where it shows the lowest expression and neuroticism. The average profile of the Cortana group on the differs quite greatly from the other groups. People that prefer Cortana are on average more neurotic, less agreeable and less extroverted than members of the other groups. These differences are especially salient compared to the Now profile.

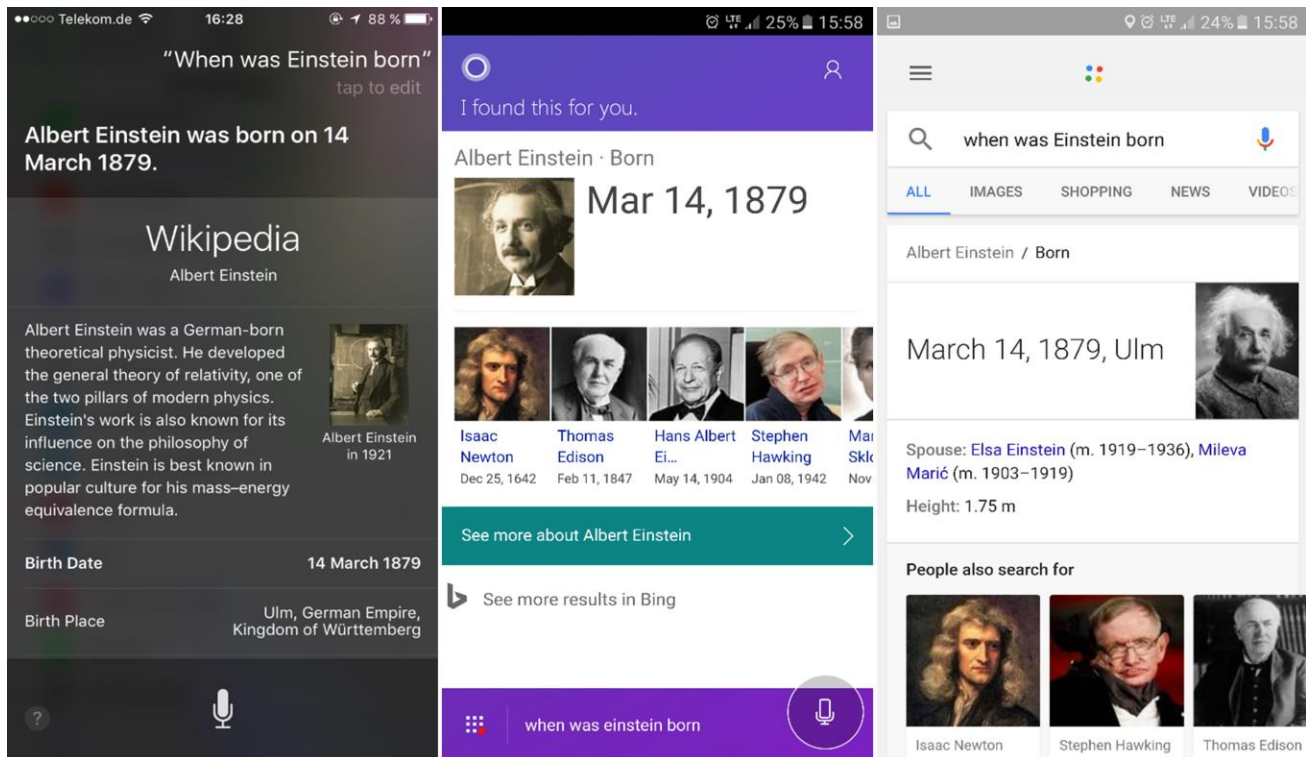


Figure 2: Example query for Siri, Cortana and Now.

5.2.2 *Openness*. Tab. 4 shows that people who have a low score on openness dislike Cortana, compared to Siri and Now. The T-Tests show a highly significant effect between Cortana and Siri and between Cortana and Now. People who score low on openness are usually less open to new experiences. Tab. 1 shows that Cortana is the least known and by far the least used IPA in the current population of participants. The lack of experience with this IPA could explain the low approval score of Cortana: state reactance is a component of the approval score. In order to complete the experiment, participants had to interact with all three IPAs. If participants had been reluctant to interact with the formerly unknown Cortana against their intention, this could have been regarded as a freedom threat, causing state reactance and hence lowering the approval score. As stated before, people that show a low level of openness may be reluctant to try out unknown services like IPAs. It is therefore advisable to provide a strong incentive for such people to get to know a potentially new service thoroughly and to overcome that reluctance. Also, low openness correlates with high trait reactance. Highly reactant persons are more sensitive to threats to their freedom of choice. Therefore, the incentive has to be a positive one and the introduction method to the new service should be visibly cancellable at all times.

5.2.3 *Conscientiousness*. Tab. 4 shows that highly conscientious people like Now and Siri, but rather dislike Cortana. A statistical test showed that the difference was significant between Cortana and Siri and fell just short of significance between Cortana and Now. One explanation for this

could be the amount of information that is given by each IPA as a response to the task (see Fig. 2). While Cortana gives the correct answer but no additional information to questions like “When was Einstein born?”, Siri and Now both enrich their answer with additional information such as the birth place (Siri and Now), a short biography drawn from Wikipedia (Siri) or other additional information (spouses and height, Now). Very conscientious people may value additional information higher than less conscientious people, because highly conscientious people are very conscious on potentially beneficial technology (or information) and therefore show a preference for such technology (or information) [7]. Additional information in the system answer might appear beneficial for future use, which could explain the preference for Siri and Now, compared to Cortana, by highly conscientious people.

5.2.3 *Extraversion*. Extraversion showed a low variance of values in the medium level. As a result, only two participants could be classified as highly extraverted and only four participants could be classified as lowly extraverted. The advantage for Now compared to Cortana and Siri in the highly extraverted group can therefore be regarded as unreliable and will not be interpreted at this point.

5.2.4 *Agreeableness*. Agreeableness showed significant differences between Cortana and Siri and between Cortana and Now in the group of highly agreeable persons. Cortana received the lowest approval score in this group, Now the highest. Devaraj et al. proposed that highly agreeable personalities might be especially sensitive to how technology might influence one’s

appearance towards others [7]. The approval scores of the three IPAs differ quite greatly.

While Siri is rated similarly over low, medium and highly agreeable personalities, Now was rated better by highly agreeable personalities, compared to personalities that show low agreeability. On the other hand, a contradictory effect is observed with Cortana, where highly agreeable personalities show lower approval scores than lowly agreeable personalities. Therefore, Now and Cortana are probably at the two poles of a factor that influences approval by highly or lowly agreeable personalities. Since Devaraj et al. suggested that highly agreeable personalities are very conscientious about how they are perceived by others [7], the perceived reputation of the IPA or its manufacturer could be a candidate for such a factor.

5.2.5 Neuroticism. Devaraj et al. hypothesize that highly neurotic personalities are generally more negative towards technology and technological advances [7]. The data presented here does not reflect this.

There was no significant (or non-significant) effect that would support this. Arguably, IPAs represent a recent advancement in technology but the average approval score over all IPAs even indicates that highly neurotic personalities like IPAs more, compared to persons that show a lower level of neuroticism.

There are some non-significant effects in the medium and in the low neuroticism groups, however. Medium level neurotic persons seem to prefer Now, especially compared to Cortana, whereas highly neurotic personalities seem to prefer Siri. In the group of people showing a low level of neuroticism, data hints that people rather prefer Now over Cortana. Again, none of the effects of Neuroticism reached significant levels and thus, should be investigated further to confirm or reject that the observed differences are just the result of chance.

5.2.6 Trait Reactance. In the group of highly reactant personalities, Cortana received a significantly lower approval score than Siri. Also, Cortana's score was lower than Now, but did not reach significance. Highly reactant personalities are often members of the low openness group. As discussed earlier, such persons can be less comfortable with using an IPA that they did not know before (Tab. 1).

If they have to use that IPA because they participate in a study, this could also induce a perceived loss of control and thereby state reactance. Therefore, the low approval score of Cortana by persons that are highly reactant personalities (and thereby probably less open personalities) could be explained by a reluctance to use an unknown IPA. The pressure of still having to use that IPA (in an experimental situation) then induces state reactance (as indicated by the highest state reactance score for Cortana in Tab. 2).

5.3 Ratings of Cortana

In terms of absolute numbers, Cortana received significantly worse ratings in the approval score than Siri and Now (see Tab. 2 for a comparison). However, when comparing the individual preferences of persons between the three IPAs, the result is not

so obvious. Among the 24 participants were six individuals who preferred Cortana above Now and Siri, according to the respective approval scores. This can lead to the assumption that Cortana is currently the preferred option for people who usually do not like IPAs. Another possibility is that Cortana profited from a novelty effect. Only one participant had used Cortana prior to the experiment (Tab. 1).

5.4 Shortcomings and Weaknesses

5.4.1 Approval Score. A weakness of the current study is the approval score. During the experiment, there was no separate task that specifically asked to rank the IPAs by preference, because it was assumed that such a task would be highly biased by the person's opinion of the respective company. Therefore, a scoring was needed that referred to the individual IPA in a more analytical manner. The score is comprised of attractiveness and state reactance. Attractiveness as assessed with the AttrakDiff Mini questionnaire is already described as a global judgement [8]. However, the items of the questionnaire are all addressing cognitive aspects of people's opinion, only. The currently used measure of state reactance also includes affective aspects. The authors assume that the combination of attractiveness and state reactance as used in this work may provide a more reliable approximation of real world acceptance than any of those two measures alone.

5.4.2 Sample Size. Only 24 persons participated in the study. The small number of datasets might have prevented some effects from reaching significance in the statistical analysis, especially after separating the participants in three subgroups for each character trait. Still, a high number of effects reached significant levels and were in line with theory, indicating their validity.

6 CONCLUSION AND FUTURE WORK

6.1 Conclusion

The current study provides detailed data on how different character traits influence IPA preference. We measured six different character traits with three levels each and were able to show a variety of significant differences in IPA ratings and preference between these traits. Furthermore, we provided average personality profiles of users that prefer one of the IPAs over the others. This data can already help to recommend a certain IPA to users, especially in the case of Cortana, which is also available for Android and iOS smartphones.

The findings can act as a starting point for adapting or recommending IPAs to individual users on the basis of their personality profile. However, the study does not provide specific design recommendations for individual character traits. Follow up studies are intended to use the repertory grid technique [11] to assess the relevant differences between a variety of IPAs to allow for conclusions on personality-specific design guidelines.

6.2 Research Question

The research question of this paper is: Do personality traits influence whether a person likes or dislikes certain IPAs?

The current paper investigates differences in the ratings of three different IPAs between different personality types. It could be shown that significant differences of the ratings of individual IPAs between different expressions of personality traits exist.

For example, highly extraverted persons liked Now but disliked Cortana. It is therefore concluded that personality traits indeed influence or at least hint at the type of IPA that a person likes or dislikes.

6.2 Recommending IPAs

Based on the approval scores that the IPAs were rated with a specific IPA can be recommended to persons with corresponding expressions of a personality trait.

The preferred IPA for each of the assessed personality traits

Table 5: Recommendation for IPAs according to personality traits. N=Now, S=Siri, C=Cortana, based on approval score.

Trait Score	high	medium	low
Openness	N	S	N & S
Conscientiousness	N	N	S & C
Extraversion	N	S	C
Agreeableness	N	S & N	C
Neuroticism	S	N	N & S
Trait Reactance	N	N	S

and its level of expression can be viewed in Tab. 5. If the difference in the approval score of two IPAs was smaller than 0.1, both IPAs were added to the table.

6.2 Recommendations for IPA Design

Based on the findings from literature [7] and the current study, some recommendations for IPA design can be given:

6.2.1 Openness. People that show low scores on openness should be persuaded into trying out a service via positive incentives.

6.2.1 Conscientiousness. People that show high conscientiousness should be provided with additional information or references to such information. This holds for graphical user interfaces where it does not take additional time to present such information. It does not necessarily hold for audio interfaces.

6.2.3 Trait Reactance. Reactance is triggered by freedom threats or a perceived loss of control [5][6]. Devices should therefore refrain from forcing or pressure users into any actions.

6.3 Future Work

When the experiment was conducted, Windows 10 was not common on desktop or laptop computers. Cortana is a part of Windows 10 now and can thereby be used by a much larger portion of the population. It would be interesting to know if the relative amount of people who prefer Cortana to Siri or Now

changed after this. This study compared ratings of three different IPAs according to the user personality. The result provides some suggestions on what personality type prefers which IPA, however the reasons for those preferences remain unclear. For example, Now seems to be especially appealing to personalities with high levels of the personality traits that were discussed here. We are currently analysing qualitative data from experiments with IPAs to identify factors that can be used to rate IPAs according to discriminative features, instead of a simple overall rating.

ACKNOWLEDGEMENTS

This work was supported by BMBF, Software Campus, grand no. 01IS12056, Sozialpsychologische Aspekte von Smart Homes (SPASH).

REFERENCES

- [1] Amazon.com. 2017. Alexa. Retrieved January 27, 2017 from <https://developer.amazon.com/alexa>
- [2] Apple Inc. 2016. Siri. Retrieved December 23, 2016 from <http://www.apple.com/ios/siri/>
- [3] Peter Borkenau, and Fritz Ostendorf. 1993. *NEO-Fünf-Faktoren Inventar:(NEO-FFI); nach Costa und McCrae*. Hogrefe.
- [4] Jürgen Bortz. 2006. *Statistik: Für Human-und Sozialwissenschaftler*. Springer-Verlag.
- [5] Jack W. Brehm. 1966. *A Theory of Psychological Reactance*, Academic Press.
- [6] Sharon S. Brehm, and Jack W. Brehm. 1981. *Psychological reactance: A theory of freedom and control*. Academic Press.
- [7] S. Devaraj, R. F. Easley, & J. M. Crant. 2008. Research note—how does personality matter? Relating the five-factor model to technology acceptance and use. *Information Systems Research*, 19(1), 93-105.
- [8] Sara Diefenbach, & Marc Hassenzahl, M. 2010. *Handbuch zur Fun-ni Toolbox—User Experience Evaluation auf drei Ebenen*. Retrieved March 14, 1916 from: http://fun-ni.org/wpcontent/uploads/Diefenbach+Hassenzahl_2010_HandbuchFun-niToolbox.pdf.
- [9] James P. Dillard, & Lijiang Shen. 2005. On the nature of reactance and its role in persuasive health communication. *Communication Monographs*, 72(2), 144-168.
- [10] P. Ehrenbrink, X. G. Gong, & S. Möller. 2016. Implications of different feedback types on error perception and psychological reactance. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (pp. 358-362). ACM.
- [11] F. Fransella, R. Bell, & D. Bannister. 2004. *A manual for repertory grid technique*. John Wiley & Sons.
- [12] Lewis R. Goldberg. 1990. An alternative "description of personality": the big-five factor structure. *Journal of personality and social psychology*, 59(6), 1216.
- [13] Google Inc. 2016. Now. Retrieved December 23, 2016 from <https://google.com/landing/now/>
- [14] S. D. Gosling, P. J. Rentfrow, & W. B. Swann. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in personality*, 37(6), 504-528.
- [15] Sung M. Hong. 1992. Hong's Psychological Reactance Scale: A further factor analytic validation. *Psychological Reports*, 70(2), 512-514.
- [16] Sung M. Hong, & Salvatora Faedda. 1996. Refinement of the Hong psychological reactance scale. *Educational and Psychological Measurement*, 56(1), 173-182.
- [17] Sung M. Hong, & Sandra Page. 1989. A psychological reactance scale: Development, factor structure and reliability. *Psychological Reports*, 64(3), 1323-1326.
- [18] Peter K. Jonason, & Heather M. Knowles. 2006. A unidimensional measure of Hong's psychological reactance scale. *Psychological reports*, 98(2), 569-579.
- [19] Rensis Likert. 1932. A technique for the measurement of attitudes. *Archives of psychology*.
- [20] Robert R. McCrae, & Paul T. Costa. 1987. Validation of the five-factor model of personality across instruments and observers. *Journal of personality and social psychology*, 52(1), 81.
- [21] Jürgen Merz. 1983. Fragebogen zur Messung der psychologischen Reaktanz. *Diagnostica*.
- [22] Microsoft. 2016. Cortana. Retrieved December 23, 2016 from <https://www.microsoft.com/windows/cortana/>

- [23] Lawrence A. Pervin, & Oliver P. John. 1999. *Handbook of personality: Theory and research*. Elsevier. Handbook of personality: Theory and research. Elsevier.
- [24] M. Roubroeks, J. Ham, & C. Midden. 2011. When artificial social agents try to persuade people: The role of social agency on the occurrence of psychological reactance. *International Journal of Social Robotics*, 3(2), 155-165.
- [25] Lijiang Shen, & James P. Dillard. 2010. Psychometric properties of the Hong psychological reactance scale. *Journal of personality assessment*, 85(1), 74-81