Are users who are confident in their high-quality ideas & ability to formulate high-quality ideas affected by the type of feedback received by a chatbot?

Author: Karim Duerrschnabel
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands

ABSTRACT
Over the past few years' human confidence and trust in artificial intelligence has changed completely based on advancement and improvement of these technologies. This paper analyses if users that are confident in their high-quality ideas and their ability to formulate are affected by feedback from chatbots. To answer this question, we conducted a survey with an integrated chatbot that provided feedback and used respondents' answers as data to test the different components we created through analysing the data with a total amount of 110 respondents. Data was used to conduct test based on different hypothesis mentioned in the paper. Our results showed that based on the different variables we had selected based on the components of the survey, we found that a respondents trust in chatbot advice and a respondent's perceived usefulness of the chatbot advice effected the confidence a user believes he has in his high-quality ideas and his ability to formulate them. While variables like trust in AI algorithms and trust in technologies had little to no effect of perceived confidence a respondent had. Based on our results we can conclude that some variables effect the respondents perceived confidence level in their ability to formulate ideas and the level of quality of the ideas. This study emphasizes the need to better connect and bridge the gap between human and AI when it comes to depending on and trusting AI response to utilize AI technologies better.

Graduation Committee members:
First supervisor – dr. M.L. Ehrenhard (Michel)
Second supervisor – F. Koefer MSc (Franziska)

Keywords
Artificial Intelligence, Confidence, Trust, Technology, Chatbot Feedback, Algorithm Aversion

1. INTRODUCTION

Artificial intelligence is on the rise in many aspects of life and sectors of business, ranging from machine learning, chatbots, sentiment analysis, and many more different types of artificial intelligence that could benefit us in our day to day lives as individuals and employees. Big data and analytics are increasingly becoming more and more important in the modern landscape of the business world. "Artificial intelligence (AI) in marketing is currently gaining importance, due to increasing computing power, lower computing costs, the availability of big data, and the advance of machine learning algorithms and models" (Huang et al., 2020). We can see an increase in market size for chatbots and similar services from $250 million to over $1.3 billion in recent years (Pise, 2018 cited in Luo et al., 2019). The perceived usefulness of such technologies and data collection has been drastically increasing over the past decade as they gradually improve more and more. Artificial Intelligence has made its way into many sectors in business operations. Artificial intelligence in marketing link together data collection and new advanced technologies for businesses to understand and create profile for their target customers and segment they wish to market and advertise in. AI allows marketers to use this data to develop the right strategies with which to target the right customer segments and position themselves ahead of competition (Overgoor et al., 2019). With such new technologies and new developments arise challenges for us as humans in using these technologies.

With AI usage arises issues pertaining to confidence of AI judgement and low confidence in AI systems, "AI chatbots are computer programs that simulate human conversations through voice commands or text chats and serve as virtual assistants to users" (Luo et al., 2019). Different questions arise with AI in marketing, how can machine learning understand different words, slang, feeling, etc. that can only be understood by humans? How can we program *machine learning to feel* when we still have not developed such a code for it to understand feelings to be able to provide feedback? We look at Chatbots and how they can provide services for humans. Chatbots are found everywhere around us, and they provide many services to us. "Customer service chat and commercial social media interactions are increasingly managed by intelligent agents, many of which have been developed with human identities and even personalities" (Simonite, 2017). But, even though chatbots and automated services are increasingly used in different areas of business practices and customer interactions, "despite seeking advice from automation, decision makers frequently discount advice obtained from it, especially when compared to advice from a human advisor" (Onkal, Goodwin, Thomson, Gonül, &Pollock, 2009 as cited in Prahl et al., 2017).

According to recent surveys, 42% of participants lack general trust in AI, and 49% of participants could not name any AI product they trust (Dujmovic, 2017). Surveys help portray and visualize user trust in AI, and we can deduct that still several people have not experience AI in a positive manner to the extent where they still distrust AI and not use it to or give it a chance. The purpose of looking at trust is to see how forgiving and unforgiving humans can be to AI error as humans are prone to error, "Recent literature points to inappropriate trust as the reason for under- or over-relying on AI (Bansal et al., 2019a, 2019b, Dzindolet et al., 2003; Hoffman et al., 2013; Lee & See, 2004; Parasuraman & Riley, 1997; Siau & Wang, 2018; Zhang et al., 2020 cited in Chong et al., 2021)" showing that humans trust AI when they shouldn't have trusted AI suggestion or advice based on the fact that the technology is only as trustworthy as it is made to be. AI can mostly outperform human judgement, but humans are highly unforgiving to AI error, in turn this creates distrust in AI regardless of its quality (Alvarado-Valencia & Barrero, 2014; Dietvorst et al., 2015).

The following sections will explore the different aspects of the mentioned research to review needed literature, formulate the hypothesis, and test our data to conclude and understand the variables that we wish to test.

1.1 Research Objective

The goal of this research is to determine *how users who believe that they are confident in their high-quality ideas or their ability to formulate high quality ideas are more or less likely to change their perception in confidence after interacting with a chatbot.* We view this change in confidence by looking at how the different variables that we chose for our hypothesis affect the causal relationship they have on confidence for either negative feedback or positive feedback. We conduct tests on users that received negative or positive advice to understand users that believe in their confidence to formulate high quality ideas can be affected from the responded advice based on chosen variables.

1.2 Research Question

Based on what we have explained in the research objective we formulated the following research question for this paper:

"Are users who are confident in their high-quality ideas & ability to formulate high-quality ideas affected by the negative or positive feedback received by a chatbot based on their trust in AI and its perceived usefulness?"

2. THEORETICAL BACKGROUND

The following looks at the existing literature on AI chatbots and AI technologies. The following literature look at concepts that help understand the concepts of confidence, lay beliefs, algorithm aversion, human judgement, and AI decision making in marketing. The mentioned literature will help build foundation and allow for hypothesis formulation and the ability to build a conceptual framework.

2.1 Human Confidence in Artificial Intelligence and Themselves

"Confidence in AI is formed from trustor's perception of trustee's (in this case an AI) ability to perform a given task, while self-confidence contributes to the trustor's willingness to rely on trustee. Confidence in AI and self-confidence respectively provide insight into two antecedents of trust proposed by Mayer et al.: perception of trustee's attributes such as ability, and a propensity to trust (related to a personal disposition to trust) (Mayer et al., 1995; Rousseau et al., 1998)" (Chong et al., 2021).

Chong et al. here makes an important distinction between confidence in oneself and in AI. We see that the trust a user has in the AI system is in the actual ability it must perform and conclude results and suggestions that are in fact correct, and the confidence the user has is in the extent the user trusts the AI systems. These two factors allow us to understand at how confidence in oneself and confidence in AI can affect decision making. This paper explores the preconceptions and thoughts of users towards AI before making decisions and receiving suggestions. Exploring user confidence in oneself and in AI are crucial factors to understand how confidence in oneself can develop trust in decision making and how that confidence in AI contrasts to oneself also when making decisions.

We see that "good decision-makers uniquely display a positive correlation between self-confidence and probability of accepting AI suggestions; they accept the AI when they are confident in themselves and reject the AI when they are not." (Chong et al., 2021) The research that Chong et al. provided allows for insights into how we can understand confidence in oneself and in AI, Chong et al. concluded that good decision makers show that they are confident in their own ideas, but also confident in the ability of AI to suggest good ideas which in turn as shown in their research positively correlates together. We can see that having confidence in themselves means they are confident in their ideas and are able to generate good ideas but can also mean that it is not necessary that because they have good ideas they will not explore or use the suggestion of AI. Users with high quality ideas might be confident in themselves but it might not be the case that they reject AI because of those factors. We explore these factors when conducting our research to understand the effect of confidence towards oneself and AI.

2.2 Lay Beliefs About AI

Lay beliefs are subjective and informal ideas people have about various things revolving around them. We wish to understand the factors that lead to lay belief's users have for AI and machine learning to help grasp a better concept of these beliefs' users are subjected to, when conducting the experiment. "Research has found that people regard someone to be intelligent if she or he possesses analytical abilities such as solving number problems, processing of new information, and logical reasoning, whereas other types of skills (e.g., socio-emotional skills) are less strongly associated with an intelligent person (Furnham, 2001)." (Walter et al., 2021) Users can be seen as "intelligent" based on such attributes that they possess, even if AI possesses the same attributes the trust in AI is less. We use these attributes in our own experiment to measure if these "intelligent users" are less influenced by AI. We look at how such users with attributes like these can affect themselves and others before receiving a response from AI based on their lay beliefs towards AI. "When making decisions under uncertainty, individuals tend to be receptive to the advice of others and allow it to inform their own choices." (Gunaratne et al., 2018). User beliefs in AI capability and how they view AI is not always based on actual understanding of the AI but mistrust and ignorance towards technologies (Walter et al., 2021). Mistrust in AI can come from various reasons, the factors that lead to mistrust could be a wide array of things that we have to explore.

Lay beliefs of AI may exist in individuals because of the complex systems of AI that not many individuals may understand. "Consumers who believe that AI is higher than human intelligence may feel that they receive more accurate advice. In contrast, when consumers perceive a low level of complexity, lay beliefs about AI may not exert a similar impact." (Walter et al., 2021) When individuals understand complexity of systems, they can set aside beliefs pertaining to AI systems. It is evident that perceived complexity of these systems and lay belief have a positive correlation based on Walter et al. paper we can see how users perceived complexity affects their lay beliefs towards AI. With the individuals that will conduct our own research we set in mind the lay beliefs that exist within individuals and use this to lay beliefs based on complexity or to what degree they understand the AI systems that will be granting them advice. This can be used to understand how the beliefs and complexity correlate with a user's trust and confidence in their own ideas and in the AI systems advice and suggestions. Complexity and beliefs of these systems pre use are factors that can be set to help understand to what degree they are confident in themselves and how much they trust their own decision-making skills.

2.3 Overcoming and Understanding Algorithm Aversion in Users

"Despite the preponderance of evidence demonstrating the superiority of algorithmic judgment, decision makers are often averse to using algorithms, opting instead for the less accurate judgments of humans." (Dietvorst, 2015) We need to look at the factors that put decision makers in a position where they tend to lean towards human judgement rather than AI or machine learning judgement. "Forecasts made by evidence-based algorithms are more accurate than forecasts made by humans." (Dietvorst, 2015) When such systems can make better decisions than humans, why do humans still rely on their own judgement rather than follow the judgement of Ai systems that provide better solutions? "Although evidence-based algorithms consistently outperform human forecasters, people often fail to use them after learning that they are imperfect, a phenomenon known as algorithm aversion." (Dietvorst, 2015) Algorithm aversion is what we call the preference of human judgement to AI suggestions after understanding algorithm-based systems can also be imperfect. This concept challenges confidence and trust people have in themselves and AI. We understand the concept of algorithm aversion to use as a factor to help conceptualize and use as a factor to help understand the confidence and trust people have in their own ideas and in algorithmic systems.

Understanding algorithm aversion is the first step to be able to overcome it. "Despite seeking advice from automation, decision makers frequently discount advice obtained from it, especially when compared to advice from a human advisor (Önkal, Goodwin, Thomson, G.nül, & Pollock, 2009)" (As Cited in Prahl & Van Swol, 2016) we see the level of algorithm aversion apparent and to the extent that automated advice is discounted as opposed to human advice even though human advice is not as error free or might not be as error free as AI. This comes down to trust in automation and credibility. We determine the credibility of users through their competence, "The relationship between competence and advice utilization is so strong that it is common for advice literature to use a two-dimensional model of advice utilization, which differentiates advisor competence from all other factors such as intentions and integrity (Jodlbauer & Jonas, 2011; Schrah, Dalal, & Sniezek, 2006).” (As cited in Prahl & Van Swol, 2016) Understanding the credibility of users is an essential factor to be able to understand the confidence in their own idea to be able to discount AI suggestions as this could influence their thought process to block AI suggestions. Users’ high-quality ideas needed to be tested against their credibility at providing those high-quality ideas to understand if that is the case in discounting AI suggestions. Credibility of users and ideas is a factor that needs to be incorporated in understanding users’ preconceptions to of AI advice, we use the credibility factor to understand algorithm aversion and users’ ability to be confident in their own ideas or trust themselves against AI.

2.5 Human Trust in AI

Trust is defined as “the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party.” (Mayer et al., 1995) Trust is needed as a factor to understand the extent of vulnerability, reliability, and transparency a user has when using AI and using the suggestions or outcomes of AI. Depending on the level of trust a user has in AI and the level of the AI technology itself can birth various outcomes for users like low trust in highly capable AI would result in misuse and the opposite with high trust in AI with a low capability would also result in misuse. “Trust can predict the level of reliance on technology, while the level of correspondence between user’s trust and the technology’s capabilities, known as calibration, can influence the actual outcomes of technology use.” (Glikson & Woolley, 2020) understanding the levels of trust and reliance help determine a user’s capability in using technologies, the factors mentioned help assess the likelihood of a user adopting Ai suggestions based on high reliability and trust in AI technology. Our research targets trust users have in technology to see how they rely on such technology and if there could be mistrust in AI based on how the much the user is confident in his own idea or vice versa. “When researchers examine cognitive trust in AI, they measure it as a function of whether users are willing to take factual information or advice and act on it, as well as whether they see the technology as helpful, competent, or useful.” (Glikson & Woolley, 2020) Our research looks at trust to understand the extent of reliability or the vulnerability a user gives to the AI. Looking at this factor it helps better understand where the confidence a user has in his own ideas could develop. Trust can be used as a factor to measure the willingness of users to reject or accept the AI suggestion based off what they thought of the suggestion before receiving it to understand if they trust themselves more or the AI choice which in turn can also show the confidence of a user in their own ideas and in AI.

2.6 Hypothesis

Based on the previous literature and findings we develop hypothesis to help visualize a conceptual framework for our research.

H1: *Users who are confident in their high-quality ideas are less influenced by chatbot advice*

H2: *Users who are confident in their high-quality ideas are less trusting of chatbot advice*

H3: *Users who are confident in their high-quality ideas are more likely to think AI is useless*

H4: *Users who have high quality ideas feel that they are less likely to trust AI and technologies before using it*

2.7 Conceptual Framework

The following figure visualizes and represents the hypothesis connections and show how they represent the outcome of the research

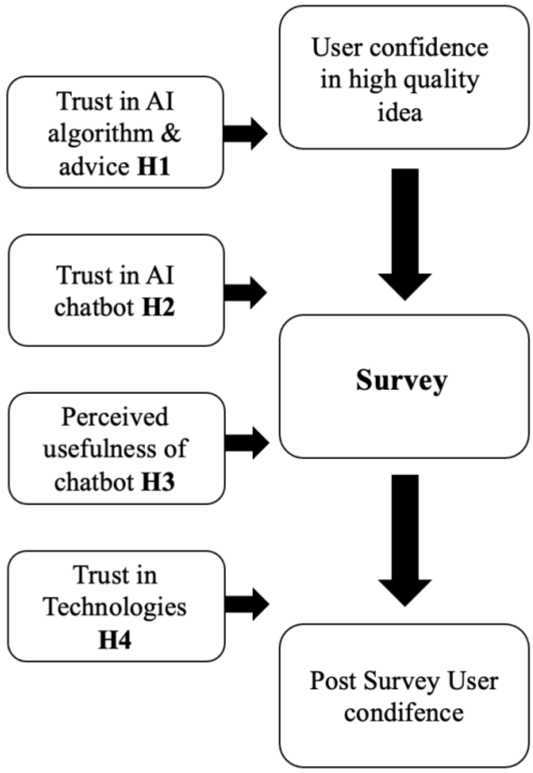


Figure 1. Conceptual Framework

3. METHODS

3.1 Research Design

For the research design of our experiment, we are conducting a between groups experiment, a control group, and a treatment group. Our research design is a two-by-two matrix design shown in *figure 2.* The first dimension of the matrix design informs the participants of the experiment whether the chatbot is credible or not. The other dimension of the experiment informs the participants of the outcome or result of the business idea suggestion. As mentioned, the experiment is split into 2 groups, the control group, and the treatment group. The control group will receive no information on the credibility of the chatbot, while the treatment group will receive full disclosure of the chatbot’s credibility.

Our experiment is accompanied by 2 survey's the pre-experiment questionnaire and the post-experiment questionnaire. Participants will answer questions before conducting the survey about background and other information for data collection and will conduct the experiment and interact with the chatbot, and after interaction and completion of the experiment they will fill out a questionnaire to understand different aspects of the experiment.

The experiment we will conduct aims at understanding whether disclosing the credibility of a chatbot has any effect in a participant's confidence or trust in AI chatbots and their given suggestion or outcome of the AI. Data collection will be made based on users answers to surveys and interaction with chatbot which will be used to test hypothesis created.

Credibility / Evaluation	Chatbot no info (= control)	Chatbot credible
Positive	Control group 2a	Treatment group 1a
Negative	Control group 2b	Treatment group 1b

Figure 2. 2x2 Matrix design

3.2 Survey Design

The pre-experiment survey and post-experiment survey were created in collaboration with other researchers that look at different aspects of this experiment to collect data. In total there are in total X questions. We look at the design of the pre and post experiment questions in the following. The experiments are conducted in English, Dutch, and German.

3.2.1 Pre-Experiment Survey

The questions developed in the pre-experiment survey at targeted at the characteristics of the individuals participating in the experiment. After developing the background of the individual, the pre-experiment questions are divided into 5 parts, trust in technologies, familiarity with AI, trust in AI, feelings about being judged on your idea, and confidence in ability to formulate idea. Each sections contains questions that characterize a person and show different factors that are considered before participating in the experiment. The questions were written, and all have a scale, they use the Likert scale (1-5) ranging from strongly agree to strongly disagree. The Likert scale measures attitudes, knowledge, perception, values, and behavioral changes (Vogt, 1999). Both groups will ben answering the same questions and provide data to understand the nature of each participant before the experiment is conducted.

3.2.2 Post-Experiment Survey

The questions developed in the post-experiment survey are targeted at the experience that each different group had after conducting the experiment. The questions are split into 4 parts, nature of the evaluation, trust in the AI, advice utilization, and perceived usefulness. These questions also are measured through the Likert scale (1-5) ranging from strongly agree to strongly disagree. The questions are designed to understand if the participants use the advice or not. It is also designed to understand and gather data on the different factors mentioned in the questionnaire like trust, confidence, usefulness, etc. The post-experiment survey gathers data to understand the difference between both groups and see if knowing credibility effects participants thoughts on the AI and the advice.

3.3 Data Collection

To collect the needed data for this experiment it was decided upon between the researchers of the different subtopics to conduct an online survey. The experiment would be conducted fully online, the chatbot and survey would be sent together to the different participants of the experiment. Conducting the experiment online would allow us to reach a wider and more diverse audience of people of all backgrounds. Researchers would use their social networks to send and share he link to the experiment. This would be done with the consent of individuals as the experiment states in the beginning a form to collect data anonymously from users. As mentioned, the experiment is conducted in a confidential matter, users would just have to agree to terms and conditions stated on the first page. Users participating would participate voluntarily, all data collected will be used just for the purpose of this research. The experiment also will be approved by the BMS ethics committee before beginning the experiment.

3.4 Sampling

The sample size of the experiment reached a total of 110 participants. Sampling was done randomly with 110 users participating we found that 46 participants were male, and 58 participants were female with different ranges of age. The ages ranged from 19 to 60 years old. To conduct the experiment, it was required that the people in the sample were able to either fluently speak English, German, or Dutch. Users that rushed the experiment because their time was less than 3 minutes will be removed from the data pool. Users who have not fully completed the survey will also be removed from the data pool as a non-completed survey cannot grant useful data for us. Data will only be accepted from users that fit our criteria and have completed the experiment truthfully and answered the questions consistently and properly to allow for the best results for our research. Samples of individuals will be from many backgrounds and ages as conducting the experiment online allows for a diverse data pool.

The final sample size for the experiment after collecting the data was in total 60 respondents after removing users that have answered in a false manner, or in an inappropriate manner compared to the standards needed. Of those, 28 were male (46.7%) and 32 (53.3%) were female. We had 28.3% of respondents from Germany, 15% of respondents were from Bulgaria, and 8.3% of respondents from The Netherlands.

To Filter and make the data set better we conducted a check to see if participants were aware of the results the chatbot gave. We do this to check if participants were focused during the experiment, or survey. We do this through a cross-table in spss through checking the actual advice given to the advice they thought the chatbot gave them. In the below figure we show the cross-table to filter out the cases, and after that add a filter variable to remove the cases of participants that were wrong with what the advice, they thought they had against what they had.

What type of evaluation have you received from the chatbot? * Advice Combined Crosstabulation

			Advice Combined Positive	Advice Combined Negative	Total
What type of evaluation have you received from the chatbot?	Positive	Count	29	3	32
		% of Total	48.3%	5.0%	53.3%
	Negative	Count	1	27	28
		% of Total	1.7%	45.0%	46.7%
Total		Count	30	30	60
		% of Total	50.0%	50.0%	100.0%

Figure 3. Crosstab evaluation pre filter

			AdvComb = 2 (FILTER) Selected	Total
What type of evaluation have you received from the chatbot?	Positive	Count	3	3
		% of Total	10.0%	10.0%
	Negative	Count	27	27
		% of Total	90.0%	90.0%
Total		Count	30	30
		% of Total	100.0%	100.0%

Figure 4. Post Filter Negative Advice

			AdvComb = 1 (FILTER) Selected	Total
What type of evaluation have you received from the chatbot?	Positive	Count	29	29
		% of Total	96.7%	96.7%
	Negative	Count	1	1
		% of Total	3.3%	3.3%
Total		Count	30	30
		% of Total	100.0%	100.0%

Figure 5. Post Filter Positive Advice

After removal of 4 selected cases from the data set, we are left with 56 respondents that we can use for our random sample to test our hypothesis in the next sections. We will firstly analyze the data, conducting a factor analysis of the variables and components, then conducting a test to look at Cronbach's alpha for the selected variables and components, and after compiling these analyses we look at testing for normality to understand what tests we will use to test our hypothesis.

3.5 Data Analysis

The following analysis of data will look at various aspects of the data set to understand our set better regarding our selected variables and components. These analyses will be made through spss.

First, we will conduct a factor analysis on the different components of the survey. For each of the questions in the survey we used a Likert scale from 1-5. The Likert scale ranges from strongly agree to strongly disagree, we made 6 components consisting of the different Likert scales questions that we used in the survey. We selected these components because they were of the most significant for our research. For the factor analysis we first look at the Kaiser-Meyer-Olkin test to see if our sample size is sufficient for the test, any sample size below 0.5 would not be sufficient but ideally, we look for a score above 0.7.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.522
Bartlett's Test of Sphericity	Approx. Chi-Square	40.844
	df	15
	Sig.	.000

Figure 6. KMO test

As we can see we received a score of 0.522 which is just about okay for our sample meaning data sample size is sufficient to use a factor analysis. In this analysis we also have Bartlett's test of sphericity meaning we look for an adequate number of correlations. This is represented by a significance level less than alpha (0.5) which is present ($p<0.001$). This means we have enough correlations for factor analysis, the rules are satisfied to create a factor analysis.

Rotated Component Matrix[a]

	Component 1	Component 2	Component 3
Perceived usefulness of chatbot	.854		
Trust in AI chatbot	.822	.319	
Confidence in ability to form ideas		.801	
Familiarity with AI and AI chatbots		.716	
Trust in Technologies			.846
Trust in AI algorithms & its advice			.724

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 4 iterations.

Figure 7. Factor analysis

Based on the above analysis we can see that the questions that we selected were split into 3 components. These components also almost match the split in components that we had initially made with the survey. For the factor loading of the variables, you can see all variables except one are above 0.7 which is close to ideal. One variable, which is present in component 1 and 2 has a score of .822, and .319, it would be better to include this variable only in component 1 and remove it from component 2 when analyzing data. Our factor analysis shows that data is sufficient when it comes to variance within variables. We next conducted a reliability analysis to see our Cronbach's alpha level.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.909	.901	28

Figure 8. Cronbach's Alpha

After conducting a reliability analysis, we receive a Cronbach's alpha of 0.909, this is because we used the components we have, and included all survey questions within the test, this means we have high internal consistency between items. This Cronbach's alpha was conducted to test as mentioned the reliability, or the internal consistency of the scale items we have for our Likert scale questions. We can see with 28 items in our scale we have a high internal consistency with an alpha of 0.909. Lastly for data analysis we conduct a Shapiro-Wilks test for normality. We conduct this test to understand if our data is normally distributed or not, this will reveal which tests we can use.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Trust in Technologies	.263	56	.000	.872	56	.000
Familiarity with AI and AI chatbots	.258	56	.000	.887	56	.000
Trust in AI algorithms & its advice	.236	56	.000	.874	56	.000
Confidence in ability to form ideas	.180	56	.000	.916	56	.001
Trust in AI chatbot	.125	56	.030	.939	56	.007
Perceived usefulness of chatbot	.177	56	.000	.888	56	.000

Figure 9. Shapiro-Wilks Test

Based on the above figure, we can say that our data is not normally distributed, that means we can only conduct non-parametric tests, for the figure you can see that for all components you can see ($p<0.05$) meaning that we reject the null hypothesis of normal population distributions. For the following results of the analysis of our hypothesis we will conducts tests for association and correlation between two or more means.

4. RESULTS

For the following analysis of the results, we look at each of our 4 hypothesis and conduct tests to understand association and correlation between the selected means. Our independent variable, "*Confidence in ability to form ideas*" is the variable that will be tested against the other dependent variables to understand association, correlation, and the regression the variables have on each other to analyze our hypothesis for pre and post survey variables.

4.1.1 Respondents who are confident in their high-quality ideas are less affected by chatbot advice

For hypothesis H1: "*Users who are confident in their high-quality ideas are less influenced by chatbot advice"* We formulate this as follows; H0: User trust in AI algorithms is not affected by chatbot advice.

H1: User trust in AI algorithms is affected by chatbot advice. We look at the variables "confidence in ability to form high-quality ideas" and "Trust in AI algorithms & its advice". To measure this, we conducted an independent samples t-test between the users that received negative advice and users that received positive advice. This was done to understand and determine whether there is statistical significance between these two groups based on their trust in AI algorithms and advice. We look for difference in both advice groups in our dependent variable, trust in AI algorithms and then we test it against confidence in ideas for linear regression.

Group Statistics

	Advice Combined	N	Mean	Std. Deviation	Std. Error Mean
Trust in AI algorithms & its advice	Positive	29	3.2069	1.04810	.19463
	Negative	27	3.1111	.89156	.17158

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Trust in AI algorithms & its advice	Equal variances assumed	1.669	.202	.367	54	.715	.09579	.26098	-.42744	.61902
	Equal variances not assumed			.369	53.581	.713	.09579	.25946	-.42449	.61607

Figure 11. Independent Samples T-test H1

Based on the above figure, an independent samples t-test was run to determine if there were differences in trust in AI algorithms and its advice between users that received positive feedback and users that received negative feedback. We look at the users with positive feedback (N=29, M=3.2) against users with negative feedback (N=27, M=3.1), looking at the results of the t-test we find that there was not a significance in data (t (54) =0.367, p=0.715). These results suggest that feedback does not affect whether a user's trust in AI and algorithms is affected. With $p>0.05$ we accept our null hypothesis.

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.123	.681		4.588	.000
	Confidence in ability to form ideas	.010	.184	.008	.057	.955

a. Dependent Variable: Trust in AI algorithms & its advice

Figure 12. Linear regression H1

Based on the above figure we can conclude that with ($p>0.05$) p = 0.955 This shows data is not significant as $p>0.05$. From conducting an independent samples t-test to look at statistical difference between both positive and negative feedback groups and regression we can find that both these variables have data that is not statistically significant, we accepted the null hypothesis in both cases. This means we can reject hypothesis H1 based on the data we found.

4.1.2 Respondents who are confident in their high-quality ideas are less trusting of chatbot advice

For hypothesis H2: "*Users who are confident in their high-quality ideas are less trusting of chatbot advice"* We formulate this as follows; H0: User trust in chatbot is not affected by chatbot advice.

H2: User trust in chatbot is affected by chatbot advice. We look at our independent variable "confidence in ability to form high-quality ideas" and the variable "trust in AI chatbot". This looks at the post advice survey component to understand how confidence in a respondent's own ideas and ability to formulate ideas is affected and associated post survey. To measure this, we conducted an independent samples t-test between the users that received negative advice and users that received positive advice. This was done to understand and determine whether there is statistical significance between these two groups based on their trust in AI chatbots post survey. We look for difference in both advice groups in our dependent variable, trust in AI chatbots and then we test it against confidence in ideas for linear regression.

Group Statistics

	Advice Combined	N	Mean	Std. Deviation	Std. Error Mean
Trust in AI chatbot	Positive	29	3.2414	1.16205	.21579
	Negative	27	2.4074	1.06552	.20506

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Trust in AI chatbot	Equal variances assumed	.195	.660	2.793	54	.007	.83397	.29862	.23528	1.43267
	Equal variances not assumed			2.802	53.990	.007	.83397	.29768	.23716	1.43079

Figure 13. Spearman's Rho H2

Based on the above figure, an independent samples t-test was run to determine if there were differences in trust in AI chatbots post survey and its advice between users that received positive feedback and users that received negative feedback. We look at the users with positive feedback (N=29, M=3.2) against users with negative feedback (N=27, M=2.4), looking at the results of the t-test we find that there was not a significance in data (t (54) =2.793, p=0.007). These results suggest that in fact the feedback received by users whether positive or negative influences users trust in chatbot post survey. We look at the $p<0.05$ with p = 0.007 we find that data is in fact significant making it possible to reject the null hypothesis we have stated. This means that most probably user trust in chatbots is affected by chatbot advice. We conduct linear regression.

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	1.944	.823		2.363	.022
	Confidence in ability to form ideas	.246	.222	.149	1.108	.273

a. Dependent Variable: Trust in AI chatbot

Figure 14. Linear regression H2

The above figure depicts the linear regression and correlation of both the chosen variables. We can conclude from the above figure that with (p>0.05) p = .273 that our data is not significant. Based on our independent samples t-test data is significant, so we reject our null hypothesis and accept our hypothesis that in fact user trust in chatbot is affected by chatbot advice. Based on our regression output we find that p>0.05. With p<0.05 in our t-test we reject the null hypothesis and accept our alternate hypothesis. We accept hypothesis H2

4.1.3 Respondents who are confident in their high-quality ideas are likely to view AI useless

For hypothesis H3: "*Users who are confident in their high-quality ideas are more likely to think AI is useless"* we formulate this as follows; H0: Users perceived usefulness of AI is not affected by chatbot advice

H3: Users perceived usefulness of AI is affected by chatbot advice. We conduct the test using the two variables "confidence in ability to form ideas" and "perceived usefulness of chatbot". To measure this, we conducted an independent samples t-test between the users that received negative advice and users that received positive advice. This was done to understand and determine whether there is statistical significance between these two groups based on their trust in AI algorithms and advice. We look for difference in both advice groups in our dependent variable, trust in AI algorithms and then we test it against confidence in ideas for linear regression.

Group Statistics

	Advice Combined	N	Mean	Std. Deviation	Std. Error Mean
Perceived usefulness of chatbot	Positive	29	3.0862	1.24692	.23155
	Negative	27	2.1852	1.14479	.22031

Independent Samples Test

		Levene's Test for Equality of Variances F	Sig.	t-test for Equality of Means t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Perceived usefulness of chatbot	Equal variances assumed	.470	.496	2.810	54	.007	.90102	.32061	.25825	1.54380
	Equal variances not assumed			2.819	53.991	.007	.90102	.31961	.26023	1.54181

Figure 15. Spearman's Rho H3

Based on the above figure, we can conclude an independent samples t-test was run to determine if there were differences in perceived usefulness of AI and its advice between users that received positive feedback and users that received negative feedback. We look at the users with positive feedback (N=29, M=3.0) against users with negative feedback (N=27, M=2.1), looking at the results of the t-test we find that there was not a significance in data (t (54) =2.810, p=0.007). These results suggest that in fact the feedback received by users whether positive or negative influences users perceived usefulness of AI. We look at the p<0.05 with p = 0.007 we find that data is in fact significant making it possible to reject the null hypothesis we have stated. This means that most probably user's perceived usefulness of AI is affected by chatbot advice. We conduct linear regression.

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.088	.892		3.461	.001
	Confidence in ability to form ideas	-.120	.241	-.068	-.498	.621

a. Dependent Variable: Perceived usefulness of chatbot

Figure 16. Linear regression H3

Based on the above figure we can conclude that with (p>0.05) p = 0.621 that data is not significant based on a linear regression analysis. Based on our independent samples t-test data is significant, so we reject our null hypothesis and accept our hypothesis that in fact user trust in chatbot is affected by chatbot advice. Based on our regression output we find that p>0.05. With p<0.05 in our t-test we reject the null hypothesis and accept our alternate hypothesis We accept hypothesis H3

4.1.4 Respondents who are confident in their high-quality ideas don't trust technology & AI pre survey

For our final hypothesis test we look at H4: "*Users who have high quality ideas feel that they are less likely to trust AI and technologies before using it"*

We formulate this as follows; H0: User trust in technologies is not affected by chatbot advice.

H1: User trust in technologies is affected by chatbot advice. We look at the variables "confidence in ability to form high-quality ideas" and "Trust in technologies". To measure this, we conducted an independent samples t-test between the users that received negative advice and users that received positive advice. This was done to understand and determine whether there is statistical significance between these two groups based on their trust in AI algorithms and advice. We look for difference in both advice groups in our dependent variable, trust in AI algorithms and then we test it against confidence in ideas for linear regression.

Group Statistics

	Advice Combined	N	Mean	Std. Deviation	Std. Error Mean
Trust in Technologies	Positive	29	3.4138	1.05279	.19550
	Negative	27	3.1852	.87868	.16910

Independent Samples Test

		Levene's Test for Equality of Variances F	Sig.	t-test for Equality of Means t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
Trust in Technologies	Equal variances assumed	2.426	.125	.879	54	.383	.22861	.26017	-.29301	.75022
	Equal variances not assumed			.884	53.388	.380	.22861	.25849	-.28976	.74698

Figure 17. Spearman's Rho H4

Based on the above figure we can see that an independent samples t-test was run to determine if there were differences in trust in AI algorithms and its advice between users that received positive feedback and users that received negative feedback. We look at the users with positive feedback (N=29, M=3.4) against users with negative feedback (N=27, M=3.1), looking at the results of the t-test we find that there was not a significance in data (t (54) =0.879, p=0.383). These results suggest that feedback does not affect whether a user's trust in technologies is affected. With p>0.05 we accept our null hypothesis.

Coefficients[a]

Model		Unstandardized Coefficients B	Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	3.275	.683		4.798	.000
	Confidence in ability to form ideas	.008	.184	.006	.043	.966

a. Dependent Variable: Trust in Technologies

Figure 18. Linear regression H4

Based on the above figure we can conclude that with (p>0.05) p = 0.966 This shows data is not significant as p>0.05. From conducting an independent samples t-test to look at statistical difference between both positive and negative feedback groups and regression we can find that both these variables have data that is not statistically significant, we accepted the null hypothesis in both cases. This means we can reject hypothesis H4 based on the data we found.

5.DISCUSSION AND IMPLICATIONS

This research looked at the confidence users have in their ability to formulate ideas and how that is affected by the selected variables that we selected. We selected these variables to find how users feel pre and post survey when it comes to trust and believing AI. Confidence in AI is formed from trustor's perception of trustee's (in this case an AI) ability to perform a given task, while self-confidence contributes to the trustor's willingness to rely on trustee. (Chong et al., 2022) we looked at that confidence opposed to our variables, confidence in AI is essential and understanding confidence in human ability to formulate high quality ideas can relate to situations that are high risk or need quick and proper decision making. Humans often remain responsible for the final decisions due to ethical concerns; therefore, these teams can only reach their collaborative potential when human decision-makers appropriately accept or reject AI input (Zhang et al., 2020) to appropriately accept or reject trust in AI comes from confidence in AI and users that have more confidence in their ideas appear based on results to not have much trust in AI and its decisions. We find that much research explores and connects the idea of confidence and trust in AI and we believe through looking at the research we can understand why humans either trust or don't trust AI and how we can use this to increase confidence in AI ability to maximize efficiency when it comes to critical decision making, Not only this, the performance of Watson for Oncology, IBM's cancer treatment recommendation system, varies greatly depending on the population and the type of cancer (Strickland, 2019). If doctors fail to reject Watson's faulty recommendations, patients can receive inappropriate treatment for their cancer. (Chong et al., 2022)

5.1 Practical implications

Based on the results we found in this research we can see the relationship present with users who believe they have high quality ideas and their trust and feelings towards AI and AI advice. "That humans accept or reject AI suggestions when they should not because their trust for the AI does not match the AI's trustworthiness." (Chong et al., 2022) The data shows that humans who are confident in their ability to formulate high quality ideas indeed are confident in that ability and have less trust and belief in advice and feedback from AI and chatbots.

5.2 Theoretical Implications

An interesting theoretical implication of this research is that respondents of the survey that were confident in their ability to formulate high quality ideas feel like they can't trust AI advice and perceive AI advice to be low in usefulness to them. This cannot be ignored as mentioned in previous parts that AI is important for critical and fast decision making because AI provides a needed view on different decision making and laying trust into AI that is trustworthy is essential. If this research shows that mostly users and confident in themselves rather than AI we have to explore the psychological effect of this and understand how to improve the relationship between AI and human when it comes to decision making.

5.3 Practical Relevance

This study aims at understanding if users with high quality ideas or individuals described as "intelligent" are confident in their ability to formulate ideas and trust their ideas more than the AI suggestions, we wish to explore if those individuals are more/less effected by chatbot response and how much their trust in AI can change their confidence in their own ability and trusting their own ideas. AI and machine learning are on the rise and have proven to give companies advantages if implemented properly. We wish to see how AI can change the perspective of users with low confidence and trust in chatbot evaluation or response, we want to understand the variables that effect individuals confidence/trust in AI and see how to maintain a more positive relationship between users and AI. This analysis will look at the proposed experimentation model to understand user idea confidence/trust in relation to machine learning before they receive feedback from AI or the chatbot we will implement. We use this to understand how much better AI can be in relation to human thought processes and understand how to change human behaviour towards AI or chatbots by understanding affecting factors.

5.4 Academic Relevance

Relevance for such research can be seen as research in value co-creation using artificial intelligence is lacking (Kaartemo et al., 2018). Value co-creation is defined as "allowing companies and customers to create value through interaction." (Galvagno et al., 2014) This is better described for us as the creation of value using AI or chatbots when it comes to idea generation. Our paper and experiment explore the influence of chatbot or AI response for individuals during idea generation processes. Such research will help look at a field of value co-creation through a different scope. "AI and robots in a beneficiary's value co-creation processes and well-being remain a nearly untouched territory in marketing and service research." (Kaartemo et al., 2018) There is a lack of research on the usefulness of AI and chatbots around value co-creation. Our paper is relevant to add to the academic field of understanding the added value of AI for individuals and businesses.

6. CONCLUSION

The research conducted was carried out to find the answer to the following research question provided in the beginning: *"Are users who are confident in their high-quality ideas & ability to formulate high-quality ideas affected by the negative or positive feedback received by a chatbot based on their trust in AI and its perceived usefulness?"* we formulated 4 hypothesis that were looked at in our results to be able to answer our research question. In the previous figures found under results you can find the results computed for each of our hypothesis and the concluded outcomes based on these results.

Testing these hypotheses, we look at our independent variable, "confidence in ability to formulate ideas" this variable was important to test against the other variables to find the conclusions for our stated hypothesis. What we found was that for each hypothesis based on the outcomes we rejected H1 and H4, we accepted H2 and H3 every hypothesis was based on independent sample t-tests and association through linear regression tests. We found that for H1 we asked if a user's trust in AI algorithms can be affected by the feedback it receives from the chatbot and based on our results what we found was that that in fact user trust most likely will not be affected by the chatbot feedback, we rejected the first hypothesis. For H2, we found that users trust in chatbot as indeed affected by the advice the users receive post survey, looking at the data we found that this was in fact the case and we accepted this hypothesis. For H3, we found that the perceived usefulness users had of AI was in fact affected by the feedback users received from the chatbot. Data showed that it was in fact mostly true that perceived usefulness was affected by chatbot feedback, we accepted our hypothesis for H3. For H4 we looked at user trust in technologies can be affected by chatbot feedback and in fact found that this is not true based on the data, so we rejected the H4 and accepted the null hypothesis.

We cannot fully conclude based on the outcomes of the hypothesis that users who are confident in their high-quality ideas and their ability to formulate those ideas are truly affected by the feedback of a chatbot. But what we can conclude is that there are certain variables like the ones mentioned in our hypothesis that can affect the confidence of a user in their ability to formulate high-quality ideas and their confidence in following them.

7. LIMITATIONS AND FUTURE RESEARCH

In the following we look at the limitations of this research and what could be done in future research to improve this research.

7.1 Limitations

A few limitations and issues present themselves to us in this research. Firstly, the design of this research experiment took much time to have finished which hindered the final number of responses which could have been larger and provided more insight into our research. More data could have influenced our results and helped give a different insight into our research. The second limitation to this research was the users that we had that responded to our research. We could not guarantee proper response from users as we time was constrained, and user responses could have been varied from serious to not serious as some users could be doing surveys as a favor. Respondents could limit research as sometimes the target group is not met so we do not have the proper respondents that we want. The next limitation are the variables, the variables are well defined, but they survey could use more components to formulate a better variable for certain research. We cannot be so sure that users are confident in their ability to formulate high quality ideas, maybe they state that they are confident but just say that without understanding the question or state that because they feel that might be true, but we need another question or component to challenge that in users that take the quiz. Another limitation that could be seen in this research is that the chatbot used in the research was not a chatbot developed by us but used of a program to use which could limit tis actual ability to give proper advice. This could deviate data and give different responses that might be true or false but can still limit responses given to users which in turn can deviate data. The last limitation we can mention is the usage of scales for each component, the survey or research could use different scales to have different types of data for different types of tests to be conducted.

7.2 Future Research

Looking at what we mentioned previously we can see that for future research we need to have more time to compile and collect proper data which might show different outcome to our research. This research also needs to have an improved survey with a proper chatbot developed to give users correct advise. It is also advised to change or add new validated scales in the next research to be able to collect different data types to for more varied research and be able to explore more avenues of research and data research. For future research we need to add more questions to challenge what users believe their answers were previously. This is to confirm actual confidence levels, trust levels, and other levels pertaining to AI beliefs. For future research we need to target a group of people that are more homogenous rather than such a random sample to have proper results when it comes to replies in our survey, this will also help create a better image of respondents and see the different relationships users will have within this specific group which can open more research and filter out outliers found within responses.

8. ACKKNOWLEDGMENTS

After concluding this research, I would like to acknowledge the different individuals that helped me complete and conclude this paper. First, I would like to thank my supervisor; dr. Michel Ehrenhard, and MSc Franziska Koefer for the support and dedication to create such an experiment to help understand AI and its applicability in marketing, furthermore their support and feedback to help continuously improve our research and gain more knowledge into different matters. Second, I would like to acknowledge and thank the wonderful individuals that are in my thesis circle, our thesis circle was very helpful to one another and helped together get through rough patches which made this research much less strenuous and easier for us as a group to complete together. Lastly, I would like to offer my thanks to friends and family for being there to support and help me during my thesis, any help received allowed me to continuously improve and better myself and my research.

9. REFERENCES

Alvarado-Valencia, J. A., & Barrero, L. H. (2014). Reliance, trust and heuristics in judgmental forecasting. *Computers in Human Behavior*, *36*, 102–113. https://doi.org/10.1016/j.chb.2014.03.047

Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. *Computers in Human Behavior*, *127*, 107018. https://doi.org/10.1016/j.chb.2021.107018

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2014). Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2466040

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them. *Management Science*, *64*(3), 1155–1170. https://doi.org/10.1287/mnsc.2016.2643

Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*, *48*, 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *Academy of Management Annals*, *14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Gunaratne, J., Zalmanson, L., & Nov, O. (2018). The Persuasive Power of Algorithmic and Crowdsourced Advice. *Journal of Management Information Systems*, *35*(4), 1092–1120. https://doi.org/10.1080/07421222.2018.1523534

Huang, M. H., & Rust, R. T. (2020). A strategic framework for artificial intelligence in marketing. *Journal of the Academy of Marketing Science*, *49*(1), 30–50. https://doi.org/10.1007/s11747-020-00749-9

Jodlbauer, B., & Jonas, E. (2011). Forecasting clients' reactions: How does the perception of strategic behavior influence the acceptance of advice? *International Journal of Forecasting*, *27*(1), 121–133. https://doi.org/10.1016/j.ijforecast.2010.05.008

Kaartemo, V., & Helkkula, A. (2018). A Systematic Review of Artificial Intelligence and Robots in Value Co-creation: Current Status and Future Research Avenues. *Journal of Creating Value*, *4*(2), 211–228. https://doi.org/10.1177/2394964318805625

Lee, M. D., & Dry, M. J. (2006). Decision Making and Confidence Given Uncertain Advice. *Cognitive Science*,*30*(6),1081–1095. https://doi.org/10.1207/s15516709cog0000_71

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. Humans: The Impact of Artificial Intelligence Chatbot Disclosure on Customer Purchases. *Marketing Science*. https://doi.org/10.1287/mksc.2019.1192

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An Integrative Model Of Organizational Trust. *Academy of Management Review*, *20*(3), 709–734. https://doi.org/10.5465/amr.1995.9508080335

MIT Technology Review. (2017). *404 - Page not found*. https://www.technologyreview.com/s/603895/customerservice-chatbots-are-about-to-become-frighteningly-realistic/

Önkal, D., Goodwin, P., Thomson, M., Gönül, S., & Pollock, A. (2009). The relative influence of advice from human experts and statistical methods on forecast adjustments. *Journal of Behavioral Decision Making*, *22*(4), 390–409. https://doi.org/10.1002/bdm.637

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the Computers Take Over: Using AI to Solve Marketing Problems. *California Management Review*, *61*(4), 156–185. https://doi.org/10.1177/0008125619859318

Prahl, A., & van Swol, L. (2017). Understanding algorithm aversion: When is advice from automation discounted? *Journal of Forecasting*, *36*(6), 691–702. https://doi.org/10.1002/for.2464

Schrah, G. E., Dalal, R. S., & Sniezek, J. A. (2006). No decision-maker is an Island: integrating expert advice with information acquisition. *Journal of Behavioral Decision Making*, *19*(1), 43–60. https://doi.org/10.1002/bdm.514

von Walter, B., Kremmel, D., & Jäger, B. (2021). The impact of lay beliefs about AI on adoption of algorithmic advice. *Marketing Letters*, *33*(1), 143–155. https://doi.org/10.1007/s11002-021-09589-1

Zhang, G., Raina, A., Cagan, J., & McComb, C. (2021). A cautionary tale about the impact of AI on human design teams. *Design Studies*,*72*,100990. https://doi.org/10.1016/j.destud.2021.100990

10. APPENDICES

Appendix A- Variable items & Scales

Variable	***Item***	***Source***
Trust in Technologies	**My typical approach is to trust new technologies until they prove me that I shouldn't**	Chi, O. H., Jia, S., Li, Y., & Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (AI) social robots in service delivery. *Computers in Human Behavior*, *118*, 106700. https://doi.org/10.1016/j.chb.2021.106700
	I generally give a technology the benefit of the doubt when I first use it	
	I usually trust a technology until it gives me a reason not to trust it	
Familiarity with AI and AI chatbots	**I am familiar with AI**	Gillath, O., Ai, T., Branicky, M. S., Keshmiri, S., Davison, R. B., & Spaulding, R. (2021). Attachment and trust in artificial intelligence. Computers in Human Behavior, 115, 106607. https://doi.org/10.1016/j.chb.2020.106607
	I am familiar with AI chatbots	Chi, O. H., Jia, S., Li, Y., & Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (AI) social robots in service delivery. Computers in Human Behavior, 118, 106700. https://doi.org/10.1016/j.chb.2021.106700
	I have much knowledge about AI chatbots	
	I am more familiar than the average person regarding AI chatbots	
	I know how to interact with AI chatbots	
Trust in AI algorithms and its advice	**I trust the recommendations by algorithms-driven services (chatbots, predictive**	Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*,

	personalization agents, virtual assistants, etc).	102551. https://doi.org/10.1016/j.ijhcs.2020.102551
	Recommended items through algorithmic processes are trustworthy.	
	I believe that the algorithm service results are reliable.	
Feelings about being judged by others when telling them about an idea you recently had.	**If I needed to, I would feel at ease when presenting an idea to others**	Siemon, D. (2022). Let the computer evaluate your idea: evaluation apprehension in human-computer collaboration. *Behaviour & Information Technology, 1–19.* *https://doi.org/10.1080/0144929x.2021.2023638*
	I tend to worry about being judged by others when presenting an idea	
Confidence in ability to formulate ideas	**I'm confident in my ability to formulate high quality ideas.**	Chong, L., Zhang, G., Goucher-Lambert, K., Kotovsky, K., & Cagan, J. (2022). Human confidence in artificial intelligence and in themselves: The evolution and impact of confidence on adoption of AI advice. *Computers in Human Behavior, 127, 107018.* *https://doi.org/10.1016/j.chb.2021.107018*
	I don't believe that my confidence in my high-quality idea will be affected by a machine response.	
Trust in the AI chatbot	**I trust the advice the chatbot provided me with.**	Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, 102551. https://doi.org/10.1016/j.ijhcs.2020.102551
	I find the chatbot's advice to be trustworthy.	
	I believe that the chatbot's advice is reliable.	

	I believe that the chatbot was credible during our conversation.	Toader, D. C., Boca, G., Toader, R., Măcelaru, M., Toader, C., Ighian, D., & Rădulescu, A. T. (2019). The Effect of Social Presence and Chatbot Errors on Trust. *Sustainability, 12(1), 256.* https://doi.org/10.3390/su12010256
Advice utilization	**I am willing to let this chatbot assist me in deciding whether or not to develop my business idea**	Benbasat, I., & Wang, W. (2005). Trust In and Adoption of Online Recommendation Agents. *Journal of the Association for Information Systems, 6(3), 72–101.* https://doi.org/10.17705/1jais.00065
	I am willing to use this chatbot as an aid to help with developing my business idea.	
	I am willing to use this chatbot's advice recommendations.	
Perceived usefulness of the chatbot	**The evaluation provided by the chatbot would be useful to me.**	Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13(3), 319.* https://doi.org/10.2307/249008
	The evaluation provided by the chatbot would help me to feel at ease when presenting my idea to others.	Siemon, D. (2022). Let the computer evaluate your idea: evaluation apprehension in human-computer collaboration. *Behaviour & Information Technology, 1–19.* https://doi.org/10.1080/0144929x.2021.2023638
	The evaluation provided by the chatbot would help me to worry less about being judged by others when I present my idea.	
	The evaluation provided by the chatbot would help me to be more creative.	

	The evaluation provided by the chatbot would help me to feel encouraged to present my idea to others.	Siemsen, E., Roth, A. V., Balasubramanian, S., & Anand, G. (2009). The Influence of Psychological Safety and Confidence in Knowledge on Employee Knowledge Sharing. *Manufacturing & Service Operations Management, 11(3), 429–447.* https://doi.org/10.1287/msom.1080.0233
	The evaluation provided by the chatbot would help me to have more confidence in my idea.	

Appendix B- Pre-chatbot Survey

Please indicate your level of agreement with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I am familiar with Artificial Intelligence (AI)	○	○	○	○	○
I am familiar with AI chatbots	○	○	○	○	○
I have much knowledge about AI chatbots	○	○	○	○	○
I am more familiar than the average person regarding AI chatbots	○	○	○	○	○
I know how to interact with AI chatbots	○	○	○	○	○

What is your year of birth?

Where do you come from?

What gender do you identify as?

- Male
- Female
- Non-binary / third gender
- Prefer not to say

According to Duan et al., (2019), Artificial Intelligence refers to “the ability of a machine to learn from experience, adjust to new inputs and perform human-like tasks”. And a chatbot is a tool that is designed to mimic human-like conversations (Kushwaha et al., 2021). The purpose of this new source of advice (Logg et al., 2019) is to provide people with information they can use in their decision-making (Klaus & Zaichkowsky, 2020 as cited in Kushwaha et al., 2021).

Please indicate your level of agreement with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
My typical approach is to trust new technologies until they prove me that I shouldn't	○	○	○	○	○
I generally give a technology the benefit of the doubt when I first use it	○	○	○	○	○
I usually trust a technology until it gives me a reason not to trust it	○	○	○	○	○

Please indicate your level of agreement with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I trust the recommendations by algorithms-driven services (chatbots, predictive personalization agents, virtual assistants, etc)	○	○	○	○	○
Recommended items through algorithmic processes are trustworthy.	○	○	○	○	○
I believe that the algorithm service results are reliable.	○	○	○	○	○

Please indicate your level of agreement with the following statements.

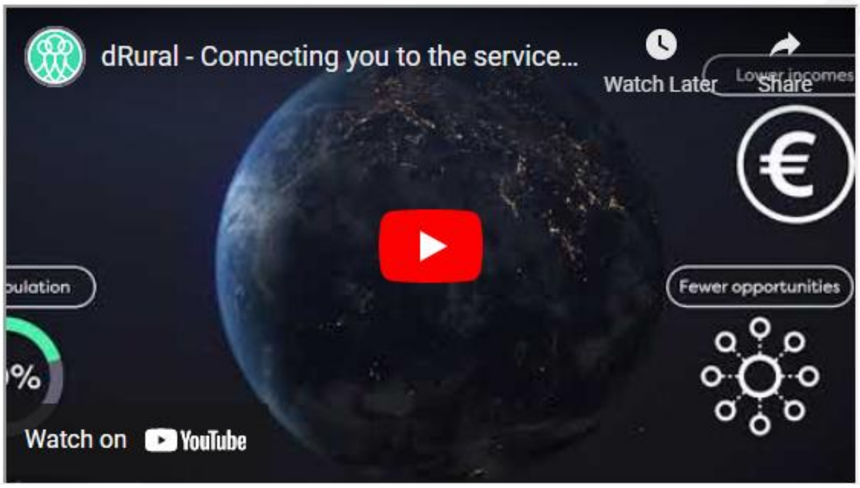
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
If I needed to, I would feel at ease when presenting an idea to others	○	○	○	○	○
I tend to worry about being judged by others when presenting an idea	○	○	○	○	○

Please indicate your level of agreement with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I'm confident in my ability to formulate high quality ideas.	○	○	○	○	○
I don't believe that my confidence in my high quality idea will be affected by a machine response.	○	○	○	○	○

Appendix C- Task Description & Video

In this study you are asked to imagine that you would like to start a new business on a digital services platform. Below, you will watch a video about a new digital platform for rural services in Europe. Please watch the video carefully and think of potential business ideas for services that could be offered via this digital services platform – try to be as innovative and creative as you can. Next, please choose the business idea that in your opinion would be the most viable, which means that there is a clear customer base that is willing to pay for your service. After the press release, a chatbot will help you evaluate various aspects of your business idea and then provide an overall assessment of the quality of your business idea.



Appendix D- Questions Asked by chatbot

But first, what is your first name? (This question is optional)

[ ]

Nice to meet you! As I already said I am EVA and I am here to help you evaluate your business idea. So, to begin I would like to know your current occupation, please?

Great! Can you please give me an indication how much relevant business experience you have had? (e.g. 3 months, 3 years etc).

Lovely! Now, please answer the questions below so I can get familiar with your business idea.

Please briefly introduce to me your product or service (nature of your idea).

What problem(s) will your business idea solve? Please explain it to me very briefly.

Thank you for all your input so far! To better understand your idea can you please briefly describe who your target customers will be.

Great! Why do you believe your target customers are interested to buy your product/service?

In what country or region are you planning to establish/sell your product or service?

How will your business idea generate revenue? Please briefly describe your strategy, e.g. your pricing strategy

Fantastic! You already provided a lot of informative details about your business idea which will help me to compare it with existing databases.

My final question: What do you think will be the competitive advantage of your business idea in your target region? (for example, the price, uniqueness of the product/service, high social or environmental impact)

Amazing! Thank you! Now please give me a minute to make the evaluation and I will get back to you with my advice! You can move forward.

Appendix E- Post Survey Questions

What type of evaluation have you received from the chatbot?

- ○ Positive
- ○ Negative

Please indicate your level of agreement with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I trust the advice the chatbot provided me with.	○	○	○	○	○
I find the chatbot's advice to be trustworthy.	○	○	○	○	○
I believe that the chatbot's advice is reliable.	○	○	○	○	○
I believe that the online agent was credible during our conversation.	○	○	○	○	○

Please indicate your level of agreement with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I am willing to let this chatbot assist me in deciding whether or not to develop my business idea	○	○	○	○	○
I am willing to use this chatbot as an aid to help with developing my business idea.	○	○	○	○	○
I am willing to use this chatbot's advice recommendations.	○	○	○	○	○

Please indicate your level of agreement with the following statements.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
The evaluation provided by the chatbot would be useful to me.					
The evaluation provided by the chatbot would help me to feel at ease when presenting my idea to others.					
The evaluation provided by the chatbot would help me to worry less about being judged by others when I present my idea					
The evaluation provided by the chatbot would help me to be more creative.					
The evaluation provided by the chatbot would help me to feel encouraged to present my idea to others.					
The evaluation provided by the chatbot would help me to have more confidence in my idea.					