Credibility and Chatbots – Investigating the influence of chatbot credibility on perceived usefulness among users.

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

ABSTRACT

Overall, the use of AI is on the rise. A segment that might profit from specific AI technologies is B2B marketing. Chatbots are frequently used as marketing and diagnosis tools, and their added value is to be investigated. As research on chatbots for idea evaluation shows that credibility is an influential factor that might affect the perceived usefulness of the feedback provided by the chatbot. For this study, a survey experiment was designed to examine the association between credibility disclosure and measures of perceived usefulness. The experiment encompasses 2x2 conditions, i.e., credibility disclosed or no statements about credibility and positive or negative feedback. To answer the research questions, only those who received a positive evaluation were included. In total, 30 responses complied with the inclusion criteria, mostly from Germany and Bulgaria. It was explored how useful the feedback provided by the chatbot was perceived among the sample and if there is an association between credibility disclosure and the perceived usefulness. It was found that the respondents perceived the chatbot as neither useful nor useless, while there was no association between credibility and perceived usefulness found in this study. However, this contradicts with several findings in the literature that indicate a relationship between usefulness and credibility disclosure, as well as positive effects of the non-human feedback on usefulness measures.

Graduation Committee members:

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

Keywords

AI, chatbots, perceived usefulness, credibility, evaluation apprehension

---

1. INTRODUCTION

1.1 Academic relevance

Artificial intelligence (AI) is currently among the most rapidly growing technologies that are based on data and enjoys increasing popularity (Lin, 2022). It is commonly defined as "the ability of a machine to display human-like capabilities such as reasoning, learning, planning and creativity" (European Parliament, 2021). Areas of application are numerous and incrementally increasing. Inter alia, it is an integral element of healthcare, transportation, manufacturing, and administration (European Parliament, 2021). Furthermore, it is central in the business context, since it offers a variety of benefits, e.g., in terms of performance, efficiency, and accuracy (Wamba-Taguimdje et al., 2020).

More specifically, marketing is an area of application that heavily relies on AI and its various functions. Marketing can be described as "the activity, set of institutions, and processes for creating, communicating, delivering, and exchanging offerings that have value for customers, clients, partners, and society at large" (American Marketing Association, 2022). From that, it becomes clear that value creation and delivery are the main objectives of marketing, which can be achieved through different means. Even though human intelligence is still crucial and frequently employed to attain these objectives, AI applications are widely used where they outperform humans, e.g., in terms of time or veracity. Essentially, the following operations are commonly used in marketing: voice processing, text processing, image recognition/processing, decision making, and autonomous robots/vehicles (Jarek & Mazurek, 2019).

Conversational agents, also called chatbots, are a well-known example of artificial intelligence in marketing. They are primarily used to automate customer interaction and to facilitate the customer experience (Selig, 2022). Presumably, chatbots are largely recognised by the general population as tools that are predominantly used for e-commerce (Brandtzaeg & Følstad, 2017). However, AI in marketing is more than just chatbots that make shopping smoother and more convenient. Generally, markets are divided into business-to-customer (B2C) and business-to-business (B2B), while chatbots for customer service in e-commerce are an example of tools used for the former market type (Janssen et al., 2021). Chatbots are nowadays also utilised within B2B organisations to support and supplement human labour (Han et al. 2021).

Due to chatbots' capability to engage in creative tasks, they can be used in innovative processes. In the business context, innovation is central to improve current practices and deliver new products and services (Cefis & Marsili, 2006). Through these improvements and developments, companies create value for customers and business partners, while aiming at return on investment. Thus, the generation of novel ideas is important for innovation. Moreover, idea generation is tightly related to idea evaluation, which is needed to determine in an early stage if pursuing the idea is practically possible and potentially profitable (Siemon, 2022).

According to Følstad et al. (2021), the research base on the impact of chatbots is limited, be it considering individuals, certain groups, or society at large. Even though there is research that deals with, inter alia, customer experience, recommendation system, algorithm aversion, and relative advantage of chatbots over human interaction, there is a need to investigate the specific value of chatbots for idea evaluation (Benbasat & Wang, 2005; Kushwaha et al., 2021; Logg et al., 2019; Prahl & Van Swol, 2017). For example, it might be interesting to investigate how the target group perceives the chatbot, e.g., in terms of its usefulness. Moreover, the role of trust was shown to be important within the literature, which is tightly related to credibility (Følstad et al., 2018). Thus, in the context of business idea evaluation within organisations, the influence of credibility disclosure seems to be a crucial factor to investigate.

1.2 Research questions

Based on that, the central research questions of this study are:

1. *What is the perceived usefulness of the feedback provided by the chatbot when the user receiving positive feedback?*
2. *Does credibility disclosure of chatbot influence how users perceive the usefulness of the feedback provided by the chatbot in case of positive feedback?*

2. THEORETICAL BACKGROUND

2.1 Evaluation apprehension

A central aspect of innovation and idea evaluation is evaluation apprehension, which is defined as "uneasiness or worry about being judged by others, especially worry experienced by participants in an experiment as a result of their desire to be evaluated favorably by the experimenter or by others observing their behavior" (American Psychological Association, n.d.). Basically, people might be restrained by their fear of being judged negatively, which could ultimately harm their reputation. In the business context, such an impaired reputation does not only have an effect on how people perceive an individual person, but also on how credible and competent they are. For example, after presenting an idea that is largely rejected by other team members, the person could be seen as having a peculiar way

of thinking. As a result, they might be considered less in future idea generation processes. Since evaluation apprehension constitutes a threat to innovation and idea generation, inter alia by harming creativity and decreasing the chances of ideas being brought up, it is an important concept to examine innovatively. For example, Siemon (2022) stated that artificial intelligence is nowadays used to counteract some of these negative effects by providing a preliminary non-human evaluation. By doing that, immature ideas can be discarded or contrived, people receive a first feedback on the quality of their idea, and they might be strengthened in their ambition to present the idea to others. Although the value of chatbots for counteracting idea evaluation was shown in recent literature, there is a need to replicate these findings to build a solid knowledge base. Hence, the following hypothesis deals with the value of AI-based conversational agents to counteract evaluation apprehension.

H1: Receiving the chatbot evaluation decreases people's tendency to experience evaluation apprehension when receiving positive feedback.

2.2 Creativity

As stated above, fear of negative judgement can affect creativity negatively, which is an integral part of innovation. Innovation is hardly possible without creative input and processing; both are considered progress drivers (Hughes et al., 2018). The relationship between the two concepts gets clearer when looking at sophisticated definitions that were created in the business context. Hughes et al. (2018) defined workplace creativity, in consideration with other available definitions, as: "the cognitive and behavioral processes applied when attempting to generate novel ideas", whereas workplace innovation is said to be: "the processes applied when attempting to implement new ideas" (p. 551). Due to the interconnection between the two concepts, it leads to the assumption that impaired creativity hinders innovation (Roskes, 2014). Moreover, Siemon (2022) highlighted the influence of evaluation apprehension on creative potential since the former can affect the latter negatively. Based on the premise that chatbots can have a positive influence on fear of negative evaluation, the following hypothesis concerns the capability of chatbot evaluation to increase creativity through decreasing evaluation apprehension.

H2: Receiving the chatbot evaluation enhances people's creative potential when receiving a positive evaluation.

2.3 Confidence

Moreover, being confident about one's own thoughts and ideas is important. For example, Matulessy and Hikmah (2022) stated that people who lack confidence tend to hold back their ideas and wait a long time before presenting them to others. However, limited confidence is not fixed, but it can be increased by internal and external factors. For example, trying to understand oneself better and engaging in positive thinking can increase overall confidence, but only relying on internal factors might not be enough to feel encouraged to express a premature idea (Matulessy & Hikmah, 2022). This is where chatbots come into play, as they constitute a means to facilitate confidence externally. As chatbots for idea evaluation are hypothesised to potentially counteract evaluation apprehension, people might at the same time feel more confident about their idea, in case of positive feedback or even input to refine the idea. As a result of increased confidence, people's beliefs in their idea as being accurate and persuasive might be strengthened. Even more, employees might be motivated to express their idea to others, which is critical for further idea development and, hence, innovation (Siemsen et al., 2009). Based on this, the following hypothesis relates to the effect of chatbot evaluation on confidence to present the idea to others.

H3: Receiving the chatbot evaluation increases people's confidence in their own idea in case of positive feedback.

2.4 Feeling encouraged

However, being confident about one's own idea might not be enough to express it. Since confidence is defined as: "being certain of your abilities or having trust in people, plans, or the future", it can be seen that it rather relates to feelings and cognitive processes (Cambridge Dictionary, n.d.). Even though some people might already start to express their idea as a result of increased confidence, others might need some more motivation to speak up. In contrast to being confident in one's idea, feeling encouraged to express it might be necessary to act on this confidence. To ensure a comprehensive examination of the concepts of chatbot usability, evaluation apprehension, and increased courage to present ideas, the concept of feeling encouraged as a result of preliminary non-human feedback will be integrated in addition to simply feeling more confident. Thus, the following hypothesis relates to the feeling of encouragement as a consequence of receiving feedback from a conversational agent.

H4: Receiving the chatbot evaluation leads people to feel more encouraged to express their idea in case of positive feedback.

2.5 Perceived usefulness

Using chatbots for idea evaluation has many advantages for employees, firms, and innovation management in general. One way to investigate the concrete practicality of using AI for idea evaluation is to look at the perceived usefulness. Whereas the term useful relates to being "capable of being used advantageously", perceived usefulness as defined by

Davis (1989) in the context of technology acceptance relates to "the degree to which a person believes that using a particular system would enhance his or her job performance" (p. 320). However, investigating the perceived usefulness of a specific technology in a specific business context should not be done by simply assessing general usefulness, but rather by looking at specific aspects of the chatbot's usefulness within the context of idea evaluation. Thus, in this case, it might be useful in that it decreases evaluation apprehension, fosters creativity, or increases confidence and feelings of encouragement. Based on all these aspects of innovation and idea generation, the following hypothesis relates to the perceived usefulness of the chatbot by taking these factors into account.

H5: Receiving the chatbot evaluation is perceived as useful by the participants in case of positive feedback.

2.6 Credibility

Having trust in the chatbot's ability to evaluate the idea in an accurate and professional manner is important as doubt about the credibility of the chatbot potentially impacts the perceived usefulness negatively (Følstad et al., 2018). According to Corritore et al. (2003), trust can be divided into three concepts, i.e., credibility, risk perception, and ease of use. Moreover, research shows that only if the output provided by the chatbot is credible, people are inclined to act in accordance with it. Otherwise, the chatbot might not have the intended effect on people's willingness to share their idea with co-workers (Shin, 2021). Hence, if there are doubts about the credibility of the chatbot, users might perceive the AI technology as less useful. Based on that, the following hypothesis relates to the effect of credibility disclosure on the perceived usefulness of the chatbot evaluation.

H6: Highlighting the credibility of the chatbot will increase people's perception of the usefulness of the feedback provided in case of positive feedback.

H7: Non-disclosure of the chatbot's credibility will decrease people's perception of the usefulness of the feedback provided in case of positive feedback.

2.7 Framework

The framework of the study consists of the above-mentioned factors, including perceived usefulness, evaluation apprehension, creativity, confidence, feeling encouraged, and chatbot credibility (Figure 1). First, the perceived usefulness of the chatbot evaluation will be investigated, taking the aspects of evaluation apprehension, creativity, confidence, and feeling encouraged into account. Basically, the mentioned aspects will be examined individually and in relation to the perceived usefulness that arises due to the effects of the chatbot evaluation on th ese factors. Moreover, it will be checked whether credibility disclosure affects the perception of usefulness, which constitutes the second major aim of this study.

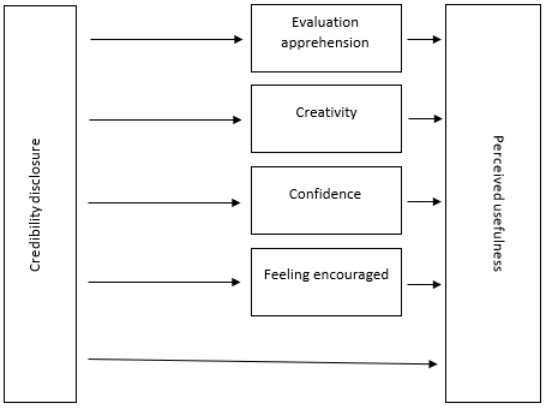


Figure 1. Framework

3. METHODS

3.1 Research Design

A between-groups design was chosen to investigate how useful the chatbot feedback is perceived as well as to examine whether credibility disclosure influences people's perception of the chatbot's usefulness. More specifically, a survey experiment was constructed for the purpose of this study. To test the hypotheses, an experiment is most suitable since different conditions can be compared (Webster & Sell, 2014). In this case, the factor of chatbot credibility could be manipulated. Based on that, a reasonable assessment of the predictions can be made, in contrast to a simple questionnaire that requires the respondents to simply imagine using the technology. Moreover, it was shown to be especially useful to investigate new technologies. As said, when using a survey experiment, the results rely on imagination and self-report. By actually using the technology in the experimental setting, respondents can make a more precise indications about the technology as a whole or particular features (Webster & Sell, 2014).

To create and distribute the survey, the software Qualtrics was used. A pre-and post-questionnaire were included to assess relevant information as well as the actual experiment. The experiment design includes 2x2 conditions with he variables credibility and type of feedback. The experimental group is introduced to the chatbot called "EVA" by stating it is a UT tool, which is already used by startups and has shown to be effective in the past. Additionally, it is briefly explained how the chatbot works. In contrast, the control group is introduced to "EVA" by only merely explaining its

functionality. The control group receives no statements concerning the chatbot's credibility. The type of feedback provided is either positive or negative, together with reasoning in order to make it seem more valid for the participants.

First, the participants were asked to fill out the pre-evaluation questionnaire. Next, they were introduced to the topic of AI and chatbots, and were randomly assigned to two groups afterwards. Subsequently, "EVA" was introduced, either with a statement regarding its credibility (experimental group) or without (control group). Afterwards, the participants were asked to think of a business idea and formulate it so that it can be digitally evaluated. This was done in several steps in order to increase the feeling of a real evaluation. After providing the business idea, participants were informed about the process of evaluation that was performed by the chatbot before they received an indication of the business idea quality in terms of a good or bad idea. The evaluation was presented to the participants realistically and validly, although the feedback of the chatbot was randomly allocated to the participants and did not reflect the actual quality of their idea.

3.2 Designing questions

The pre- and post-questionnaire were designed in English, Dutch and German language in collaboration with fellow researchers (Appendix B-G).

3.2.1 Pre-questionnaire

To ensure replicability of the survey experiment, respondent characteristics were assessed in the beginning. Afterwards, before being led to the experiment, respondents were asked two questions assessing evaluation apprehension. As can be seen from the definition of evaluation apprehension, two central aspects are worrying about negative evaluation and uneasiness. Thus, two questions, each focused on one of the two aspects were posed using a 5-point Likert scale, which can be seen in Appendix A. An example of the questions is: "I would feel at ease when having to present an idea to others." Evaluation apprehension was measured twice, once in the pre-questionnaire and once in the post-questionnaire after the experiment. This way, it can be investigated if the individual participants indicate feeling less evaluation apprehension after being provided with non-human feedback. Only when comparing general evaluation apprehension and evaluation apprehension after getting some feedback before having to present an idea to others, it can be concluded that using the chatbot possibly affects fear of negative evaluation.

3.2.2 Post-questionnaire

In the post-questionnaire, six items are included with the intention to measure perceived usefulness, measured on a 5-point Likert scale, which can be seen in Appendix G. The first item is simply focused on usefulness, which is inspired by the technology acceptance model (Davis, 1989). The following two items are similar to the questions assessing evaluation apprehension in the pre-questionnaire, but they are posed in relation to the chatbot evaluation. For example, one item is as follows: "The evaluation provided by the chatbot would help me to feel at ease when presenting my idea to others." This way, it can be observed if the chatbot evaluation has an effect on the level of evaluation apprehension experienced. Moreover, as the role of creativity was highlighted above, an item was included to assess if the chatbot evaluation can foster decreased creativity. As creativity is important for innovation, and evaluation apprehension was found to have a negative impact on measures of creativity, the concerning item was included on the basis of the added value of the chatbot's feedback on creativity levels (Siemon, 2022). The last two items relate to the concepts of confidence and feeling encouraged, as was shown that being confident about one's idea is crucial to presenting it to others, whereas feeling encouraged goes one step further and relates to the actual intention to act on this confidence. An example of the items would be: "The evaluation provided by the chatbot would help me to have more confidence in my idea."

3.3 Data collection

For the purpose of this research, it was decided to design an online survey experiment since most research activities take place in an online environment at the moment, presumably a consequence of the still ongoing pandemic. This way, it was ensured that as many people as possible could be invited to participate while keeping the effort needed from the side of the participants as low as possible. Ethical approval was granted prior to data collection by the BSM Ethics Committee of the University of Twente (request number: 220945). Afterwards, the link to the survey experiment was distributed using convenience and snowball sampling, which means that the researchers reached out to their respective social networks to gather respondents, who were in turn permitted to forward the link to others who might be interested in taking part. To distribute the survey, social media applications like WhatsApp and Instagram were used. The gathered data was handled in an anonymous manner, meaning that no data was collected that could possibly lead to identifying the participants. Moreover, it was clarified that participation was voluntary and could have been terminated at any time.

3.4 Sampling

In total, 110 people participated in the survey experiment. The responses were filtered according to the following

inclusion criteria: (1) being at least 18 years old, (2) completion of the survey including the pre-/post-questionnaire and the experiment, (3) receiving a positive evaluation by the chatbot and (4) passing the attention check. The attention check involved a question concerning which type of feedback was provided by the chatbot to test if respondents did the survey experiment in a serious and attentive manner. Thus, if respondents did not indicate the right answer to this question, they were excluded as well. After the filtering process, 30 responses were considered eligible to include for analysis. Respondent's age ranges from 21 to 60, although the majority indicated to be between 21 and 24 years old (63.3%) and were female (76.7%) Most respondents were from Germany (26.7%) and Bulgaria (16.7%). Moreover, the average respondent has experience with AI (M = 3.8, sd = .91) as well as with chatbots (M = 3.8, sd = .7). Lastly, respondents showed slightly increased levels of evaluation apprehension in general (M = 3.5, sd = .87).

3.5 Validation

To analyse the data, SPSS statistics (version 26) was used. As a first step, the validity of the items was investigated by running a factor analysis. By doing that, the number of constructs that are measured with the items within the post-questionnaire can be determined. It was revealed that all six items measure one construct, as only one item displays an Eigenvalue > 1. Thus, the measurement instruments that were constructed for the purpose of this study seem to be valid.

Additionally, the internal consistency of the items assessing perceived usefulness of the feedback provided by the chatbot was assessed by calculating Cronbach's alpha. The scale demonstrates excellent internal consistency, as shown by an alpha of .942. Based on that, it was decided to transform the six items into one variable that is intended to measure various aspects of perceived usefulness in the context of chatbots and idea evaluation.

4. RESULTS

4.1 Normality of distribution

First, the dataset was checked in terms of normality of distribution by using the Shapiro-Wilk test. Only the combination of the two items assessing evaluation apprehension in the pre-questionnaire was shown to be normally distributed. In contrast, all other individual items, the new variable assessing evaluation apprehension in the pre-questionnaire as well as the scale for perceived usefulness were demonstrated to be non-normally distributed (Table 1). Based on that, non-parametric tests will be utilised as an alternative to account for this.

Table 1
Test of normality (Shapiro-Wilk)

	W	p
Feeling at ease (pre-evaluation)	0.886	.004
Worrying (pre-evaluation)	0.842	<.001
Evaluation apprehension (pre-evaluation)	0.937	.074
Perceived usefulness	0.929	.046
Feeling at ease (post-evaluation)	0.892	.005
Worrying (post-evaluation)	0.820	<.001
Evaluation apprehension (post-evaluation)	0.906	.012
Creativity	0.890	.005
Confidence	0.852	.001
Feeling encouraged	0.882	.003

4.2 Evaluation Apprehension

To test the first hypothesis "*Receiving the chatbot evaluation decreases people's tendency to experience evaluation apprehension",* the Wilcoxon signed-rank test was employed. Since the conditions for normal distribution were not met, this non-parametric test was chosen to test whether significant differences were observable between evaluation apprehension in the pre- and post-questionnaire. To calculate the median (IQR) for evaluation apprehension, the two questions assessing the concept were used to compute a new variable, each for the pre- and post-questionnaire. Following, the two variables "pre-evaluation apprehension" and "post-evaluation apprehension" were created to compare their medians (IQR) with the help of the Wilcoxon signed-rank test. The results indicate no significant differences between evaluation apprehension ratings in the pre-questionnaire (Mdn = 3.5, IQR = 1) and post-questionnaire (Mdn = 3.75, IQR = 2) (Z = -0.858, p = .396). Thus, the feedback provided by the chatbot does not seem to decrease evaluation apprehension in the sample.

4.3 Creativity

In an attempt to investigate the second hypothesis "*Receiving the chatbot evaluation enhances people's creative potential"*, a one-sample sign test was chosen as an

alternative to the one-sample t-test due to non-normality of the data. Since running a one-sample sign test using SPSS statistics is not possible, the item assessing creativity was recoded in order to utilise a binomial test. The recoding was executed as follows: all scores below 3 were given a value of 1, whereas all scores above 3 were given a value of 2. Scores of 3 were excluded. Hence, the respondents' creativity score could be compared to the scale score of 3, which represents the value "neutral". After that, the binomial test was run to test for significant difference between the median (*Mdn = 3*, *IQR* = 2) and the value of 3. The test revealed that the median score does not significantly differs from the value of 3 (p = .446). Based on that, it is reasonable to assume that the positive evaluation provided by the chatbot does not influence perceptions of evaluation apprehension positively within the sample.

4.4 Confidence

To test the following hypothesis *"Receiving the chatbot evaluation increases people's confidence in their own idea in case of positive feedback"*, the same procedure as for testing the second hypothesis was employed with the item assessing confidence in one's own idea. Again, the item regarding confidence was recoded as follows: all scores below 3 were given a value of 1, whereas all scores above 3 were given a value of 2. Scores of 3 were excluded. Afterwards, a binomial test was executed to compare respondents' confidence rating with the scale score of 3 (="neutral"). The results indicate that the median score (*Mdn* = 4, *IQR* = 2) is not significantly different from 3 (p = .054). Thus, receiving positive feedback by the chatbot does not seem to increase feelings of confidence to present one's idea among the sample.

4.5 Feeling encouraged

To investigate if *"Receiving the chatbot evaluation leads people to feel encouraged to express their idea in case of positive feedback"*, the item assessing feelings of being encouraged were taken into account by running the same test as for the former hypotheses, i.e., a binomial test. The same recoding process was conducted to enable the investigation of differences between the respondents' feeling encouraged ratings and the "neutral" scale score 3. Apparently, the median score (*Mdn* = 4, *IQR* = 2) does not significantly differ from 3 (p = .212). Based on that, the positive feedback provided by the chatbot does not seem to influence feelings of being encouraged to present one's idea positively among the sample.

4.6 Perceived usefulness

To test the fifth hypothesis *"Receiving the chatbot evaluation is perceived as useful by the participants"* and simultaneously answer the first research question *"What is the perceived usefulness of the feedback provided by the chatbot?"*, the six items intended to assess perceived usefulness of the feedback provided by the chatbot were used to create a scale. Since Cronbach's alpha was shown to be excellent for the scale (α = 0.942), all six items were included. Afterwards, the median scale scores were recoded as stated above (scores < 3 = 1, scores > 3 = 2) in order to run a binominal test. The output shows that the median scale score (*Mdn* = 4, *IQR* = 2) does not significantly differ from the scale score of 3, representing the value "neutral" (p = .339). Following, it is reasonable to assume that receiving a positive evaluation of the chatbot is not perceived as useful among the sample.

4.7 Influence of credibility disclosure on perceived usefulness

Lastly, the following hypotheses and research question were investigated:

H6: *"Highlighting the credibility of the chatbot will increase people's perception of the usefulness of the feedback provided."*

H7: *"Non-disclosure of the chatbot's credibility will decrease people's perception of the usefulness of the feedback provided."*

RQ2: "*Does credibility disclosure of chatbot influence how people perceive the usefulness of the feedback provided by the chatbot?"*

To do that, the Mann-Whitney U test was utilised as a non-parametric alternative to the independent samples t-test that would have been employed in case of a normal distribution. The analysis was conducted to reveal if there are significant differences in terms of perceived usefulness of the chatbot between those who received information about its credibility compared to those who did not receive any credibility-related information. The results revealed no significant differences in terms of perceived usefulness between the conditions credibility disclosure (*Mdn* = 3.3) and no credibility statements (*Mdn* = 3.5) (U = 89.5, p = .345). Hence, credibility disclosure has seemingly no significant effect on ratings of the chatbot's perceived usefulness among the sample (Table 2).

Table 2

Mann Whitney U test results to investigate differences in terms of perceived usefulness between credibility disclosure conditions

Variable	n	Mean Rank	Sum of Ranks	U	p

Disclosed credibility	15	13.89	209.50	89.5	.345
Non-disclosed credibility	15	17.03	255.50		

5. DISCUSSION

Chabot are increasingly used within businesses for various reasons. One aspect that might be facilitated by the use of chatbots is business idea evaluation in the context of innovation management (Siemon, 2022). To ensure a good fit between the chatbot's design and future users, achieving insight into users' needs, wishes, and preferences concerning the chatbot is crucial. First, trust and credibility were shown to play a vital role in how people perceive the chatbot (Følstad et al., 2018). Moreover, Davis & Venkatesh (1996) acknowledged the importance of perceived usefulness, as intention to use a certain technology are rather low in case the perceived usefulness is lacking. Based on this, a survey experiment was designed to investigate the two aspects of chatbot credibility disclosure and perceived usefulness of the feedback provided by the chatbot.

5.1 Evaluation apprehension

No significant effects of the positive feedback provided by the chatbot were found in terms of evaluation apprehension within the sample. Thus, receiving the evaluation apparently did not lead people to feel more at ease (or worry less) when presenting their business idea to others. Siemon (2022) highlighted the role of evaluation apprehension in the context of idea evaluation and found significant differences between chatbot evaluation and human evaluation in terms of evaluation apprehension. Based on his results, people tend to experience less evaluation apprehension when receiving a AI-based evaluation compared to a human evaluation. In contrast, Wieland et al. (2022) did not find any significant effects of receiving non-human feedback on evaluation apprehension. As a result of these mixed findings, the added value of chatbots to counteract evaluation apprehension in the context of idea evaluation is still ambiguous.

5.2 Creativity

Moreover, no significant effects of receiving the chatbot evaluation were observed in terms of creativity. Seemingly, participants did not perceive the non-human evaluation to increase their creative potential, as suggested by Siemon (2022). Previous research demonstrated that the quantity and quality of ideas that were created by humans were significantly increased when designed in collaboration with chatbots (Hwang & Won, 2021). Furthermore, increased diversity was observed when brainstorming with non-human agents, which suggests that chatbots have the potential to enhance creative potential (Wieland et al., 2022). Thus, the results of this study seem to contradict with the findings regarding chatbot collaboration and creativity presented in the literature.

5.3 Confidence

Regarding the influence of positive chatbot feedback on confidence rating among participants, no significant effects were found. Presumably, being provided with AI-based feedback does not influence people's level of confidence in their own idea. This contrasts with the work of Siemsen et al. (2009), which deals with psychological safety and knowledge sharing. Knowledge sharing in this context might be compared to sharing innovative ideas since both concepts involve human interaction and being confident in one's own thoughts and ideas. Generally, psychological safety is theorised to influence knowledge sharing. However, it was shown that increasing confidence diminishes the effects of psychological safety on knowledge sharing (Siemsen et al., 2009). Hence, if greater confidence influences expressing one's idea positively, the chatbot's feedback should be designed to enhance people's confidence. Unfortunately, this relationship was not observed in this study.

5.4 Feeling encouraged

Again, no significant effects of receiving feedback by a chatbot on participants ratings of feeling encouraged were found as the median scores does not significantly differ from the "neutral" scale score. Based on that, the chatbot does not seem to be able to encourage the participants to present their positively evaluated idea. In other contexts, e.g., experience sharing or self-disclosure concerning mental health, chatbots were said to be encouraging and engaging (Finch et al., 2020; Lee et al., 2020). However, the results of the current study are rather inconclusive regarding the context of idea evaluation specifically.

5.5 Perceived usefulness

To test whether the positive feedback provided by the chatbot is overall perceived as useful, taking the aspects of evaluation apprehension, creativity, confidence, and feeling encouraged into account, it was assessed if participants scale score significantly differs from the value of "neutral". Since no significant differences were observed, it might be assumed that participants considered the chatbot's feedback neither useful nor useless. As stated above, several studies found that people consider chatbots for idea evaluation in respect to evaluation apprehension and creativity as generally useful (Hwang & Won, 2021; Siemon, 2022). However, the results of the current study are rather

inconclusive with regard to the perceived usefulness in terms of evaluation apprehension, creativity, confidence, and feeling encouraged.

5.6 Credibility disclosure

Lastly, the influence of credibility disclosure on perceived usefulness was shown to be insignificant, which suggests that making statements about the chatbot's credibility does not affect the perceived usefulness among the sample. When chatbots are used to give recommendations, trust and credibility were shown to be of importance (Følstad et al., 2018). However, the results of the current study do not correspond to these findings as the perceived usefulness rating was not significantly different between the credibility disclosure subgroups.

5.7 Theoretical implications

As described above, the findings of the study seem to contradict the results presented in the literature. First, the value of chatbots in regard to evaluation apprehension varied, depending on which sources were consulted (Siemon, 2022; Wieland et al., 2022). However, in terms of creativity, the chatbot was shown to be useful within other studies (Hwang & Won, 2021; Wieland et al., 2022). Despite effects on measures of confidence and feeling encouraged were found in other contexts, the results were not observed in this research context (Finch et al., 2020; Lee et al., 2020; Siemsen et al., 2009). Even though using chatbots for idea evaluation was found to be perceived as useful in different respects (e.g., to decrease evaluation apprehension), it does not seem to be the case within this study. Also, the credibility disclosure was considered based on existing literature that shows an association, but the findings could not have been replicated (Følstad et al., 2018). Overall, the findings are rather inconclusive, especially taking the contradicting research base into account.

5.8 Practical implications

Relating the results to the practical context, it seems like the designed chatbot does not provide added value to facilitate idea evaluation and innovation management. It is neither said to decrease evaluation apprehension, nor does it influence creative potential, confidence, or feeling encouraged positively. Overall, the perceived usefulness of the provided feedback does not seem to be perceived as useful among the sample. This rises the question if the same results can be found in practice. In case the perceptions of the general population, which was studied in this research, align with those of employees in the context of business idea evaluation, the added value of the chatbot evaluation is questionable. Generally, chatbots with the purpose to assist in evaluation and idea sharing should be designed to enhance confidence and to have an encouraging effect. Expressing one's thoughts and ideas is central in this context, which is why the chatbot should be designed to facilitate these aspects. As perceived usefulness, evaluation apprehension, confidence, and feeling encouraged were shown to play a major role, the design of the chatbot should be selected by taking these aspects into account (Davis & Venkatesh, 1996; Siemon, 2022; Siemsen et al., 2009). For example, the feedback could be structured in a more motivating manner and directly reassuring phrases might be helpful.

6. CONCLUSION

This study was designed to answer the following research question: *Does credibility disclosure of chatbot influence how people perceive the usefulness of the feedback provided by the chatbot?*

To arrive at an informed conclusion, another research question and seven hypotheses were designed. The results revealed no significant effects of credibility disclosure on the perceived usefulness of the provided feedback. Moreover, the chatbot does not seem to provide an advantage in terms of perceived usefulness in general. Thus, all hypotheses were rejected. Neither does the provided feedback seem to decrease evaluation apprehension, nor does it show to increase creativity, confidence, or feeling encouraged.

7. LIMITATIONS AND FURTHER RESEARCH

7.1 Limitations

Some limitations need to be addressed, which potentially implicate the conclusions drawn from the study. First, and most importantly, the study is underpowered due to the small sample size (N = 30). A sample size of 100 is commonly considered the minimum to find any statistically significant results. Thus, the fact that all hypotheses had to be rejected might arise due to the limited sample size instead of actual absence of effects. Thus, the findings of the current study should be interpreted cautiously, taking the limited power to arrive at meaningful results into account. Moreover, the chatbot was not completely realistic as the provided feedback was randomly distributed and did not represent the actual quality of the idea. That might have impacted the perceptions of the chatbot among participants. Moreover, a more meaningful and concise interaction could have been designed to encourage people to provide a sincere business idea. Lastly, the studied sample was rather representative of the general public or young adults instead of people that work in the context of innovation management and idea evaluation. Hence, it is debatable to which extent the results of this study apply to the perceptions of the actual target group of the chatbot.

7.2 Further research

Taking the above-mentioned limitations into account, future research should continue to investigate perceptions of chatbots for this specific purpose and the role of credibility disclosure. Especially due to the central role of AI and chatbots in the media, literature, and society as a whole, further investigations are necessary to improve the current technologies and enhance their uptake. First, a larger sample needs to be consulted to ensure arriving at meaningful results. As shown by this study, well-designed items are of limited use when the sample size restricts the conclusions drawn. To ensure that as many participants as possible are reached and can be included for analysis, only positive feedback should be provided. With the current study design, numerous people had to be excluded from the dataset because they received negative feedback, although they filled out the survey and executed the experiment in a sincere manner. Moreover, as the feedback was kept rather simple, a more encouraging wording should be used in order to influence people's confidence in their idea. Only when the chatbot is designed in a promoting and encouraging manner, it will potentially be perceived as useful. Moreover, the target group should be shifted from the general population to people from the business context, ideally a group of people who are used to generate and evaluate ideas. Perhaps, a specific focus on those who tend to experience high levels of evaluation apprehension might be interesting as well. Lastly, the experiment could be executed in person, instead of online. This would have the advantage of observing participants, while creating an atmosphere that is more serious compared to executing the survey experiment alone. When checking the business ideas provided, it was evident that many people did not take the task of generating a business idea serious. Conducting the experiment offline could potentially prevent this.

8. ACKNOWLEDGEMENTS

First, I would like to thank my supervisors Michel Ehrenhard and Franziska Koefer for supporting me and for setting up the experiment together with our group. Also, I would like to thank Tina Hormann for supporting me with my statistical questions. Finally, I would like to thank my bachelor thesis group that supported me and helped me with all my questions while writing this thesis.

REFERENCES

American Marketing Association. (2022, March 31). Definitions of marketing. American Marketing Association. Retrieved April 7, 2022, from https://www.ama.org/the-definition-of-marketing-what-is-marketing/

American Psychological Association. (n.d.). *Apa Dictionary of Psychology*. American Psychological Association. Retrieved May 29, 2022, from https://dictionary.apa.org/evaluation-apprehension

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

Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. *Internet Science*, 377–392. https://doi.org/10.1007/978-3-319-70284-1_30

Cambridge Dictionary. (n.d.). *Confident*. CONFIDENT | Bedeutung im Cambridge Englisch Wörterbuch. Retrieved May 30, 2022, from https://dictionary.cambridge.org/de/worterbuch/englisch/confident

Cefis, E., & Marsili, O. (2006). Survivor: The role of innovation in firms' survival. *Research Policy*, *35*(5), 626–641. https://doi.org/10.1016/j.respol.2006.02.006

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-line trust: Concepts, evolving themes, a model. *International Journal of Human-Computer Studies*, *58*(6), 737–758. https://doi.org/10.1016/s1071-5819(03)00041-7

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

Davis, F. D., & Venkatesh, V. (1996). A critical assessment of potential measurement biases in the technology acceptance model: Three experiments. *International Journal of Human-Computer Studies*, *45*(1), 19–45. https://doi.org/10.1006/ijhc.1996.0040

European Parliament. (2021, March 29). *What is artificial intelligence and how is it used?* EuropeanParliament. Retrieved April 7, 2022, from https://www.europarl.europa.eu/news/en/headlines/society/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used

Finch, S., E., Finch, J. D., Ahmadvand, A., Choi, I. J., Dong, X., Qi, R., Sahijwani, H., Volokhin, S., Wang, Z., Wang, Z., & Choi J. D. (2020). Emora: An Inquisitive Social Chatbot Who Cares For You. *3rd Proceedings of Alexa Prize*. https://doi.org/10.48550/arXiv.2009.04617

Følstad, A., Araujo, T., Law, E. L.-C., Brandtzaeg, P. B., Papadopoulos, S., Reis, L., Baez, M., Laban, G., McAllister, P., Ischen, C., Wald, R., Catania, F., Meyer von Wolff, R., Hobert, S., & Luger, E. (2021). Future Directions for Chatbot Research: An interdisciplinary research agenda. *Computing*, *103*(12), 2915–2942. https://doi.org/10.1007/s00607-021-01016-7

Følstad, A., Nordheim, C. B., & Bjørkli, C. A. (2018). What makes users trust a chatbot for customer service? an exploratory interview study. *Internet Science*, 194–208. https://doi.org/10.1007/978-3-030-01437-7_16

Han, R., Lam, H. K. S., Zhan, Y., Wang, Y., Dwivedi, Y. K., & Tan, K. H. (2021). Artificial Intelligence in business-to-business marketing: A bibliometric analysis of current research status, development and Future Directions. *Industrial Management & Data Systems*, *121*(12), 2467–2497. https://doi.org/10.1108/imds-05-2021-0300

Hughes, D. J., Lee, A., Tian, A. W., Newman, A., & Legood, A. (2018). Leadership, creativity, and innovation: A critical review and practical recommendations. *The Leadership Quarterly*, *29*(5), 549–569. https://doi.org/10.1016/j.leaqua.2018.03.001

Hwang, A. H.-C., & Won, A. S. (2021). IdeaBot: Investigating social facilitation in human-machine team creativity. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3411764.3445270

Janssen, A., Rodríguez Cardona, D., & Breitner, M. H. (2021). More than FAQ! chatbot taxonomy for business-to-business customer services. *Chatbot Research and Design*, 175–189. https://doi.org/10.1007/978-3-030-68288-0_12

Jarek, K., & Mazurek, G. (2019). Marketing and artificial intelligence. *Central European Business Review*, *8*(2), 46–55. https://doi.org/10.18267/j.cebr.213

Kushwaha, A. K., Kumar, P., & Kar, A. K. (2021). What impacts customer experience for B2B enterprises on using AI-enabled Chatbots? insights from big data analytics. *Industrial Marketing Management*, *98*, 207–221. https://doi.org/10.1016/j.indmarman.2021.08.011

Lee, Y.-C., Yamashita, N., & Huang, Y. (2020). Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW1), 1–27. https://doi.org/10.1145/3392836

Lin, Y. (2022, March 14). *10 artificial intelligence statistics you need to know in 2022* . Oberlo. Retrieved April 7, 2022, from https://www.oberlo.com/blog/artificial-intelligence-statistics

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

Matulessy, A., & Hikmah, U. R. (2022). Students' confidence in expressing opinions at a senior high school in Pademawu. *Proceedings of the 2nd International Conference on Social Science, Humanity and Public Health (ICOSHIP 2021)*. https://doi.org/10.2991/assehr.k.220207.008

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

Roskes, M. (2014). Constraints that help or hinder creative performance: A motivational approach. *Creativity and Innovation Management*, *24*(2), 197–206. https://doi.org/10.1111/caim.12086

Selig, J. (2022, March 24). *The power of Chatbots explained*. Expert.ai. Retrieved April 7, 2022, from https://www.expert.ai/blog/chatbot/

Shin, D. (2021). How do people judge the credibility of algorithmic sources? *AI & SOCIETY*, *37*(1), 81–96. https://doi.org/10.1007/s00146-021-01158-4

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

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

Wamba-Taguimdje, S.-L., Fosso Wamba, S., Kala Kamdjoug, J. R., & Tchatchouang Wanko, C. E. (2020). Influence of Artificial Intelligence (AI) on firm performance: The Business Value of AI-based transformation projects. *Business Process Management Journal*, *26*(7), 1893–1924. https://doi.org/10.1108/bpmj-10-2019-0411

Webster, M., & Sell, J. (2014). Why do experiments? *Laboratory Experiments in the Social Sciences*, 5–21. https://doi.org/10.1016/b978-0-12-404681-8.00001-7

Wieland, B., de Wit, J., & de Rooij, A. (2022). Electronic Brainstorming with a Chatbot Partner: A Good Idea Due to Increased Productivity and Idea Diversity. *Frontiers in Artificial Intelligence*. https://doi.org/10.3389/frai.2022.880673

APPENDIX

Appendix A – Variables, items, and their corresponding sources
Appendix B – Pre-interaction survey
Appendix C – Task description + video
Appendix D – Credibility disclosure scenarios
Appendix E – Questions asked by the chatbot
Appendix F – Chatbot's advice depending on the scenario
Appendix G – Post-interaction survey

Appendix A

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 personalization agents, virtual assistants, etc)**.	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
	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*

	I don't believe that my confidence in my high-quality idea will be affected by a machine response.	*in Human Behavior, 127, 107018. https://doi.org/10.1016/j.chb.2021.107018*
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

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 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	○	○	○	○	○

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.

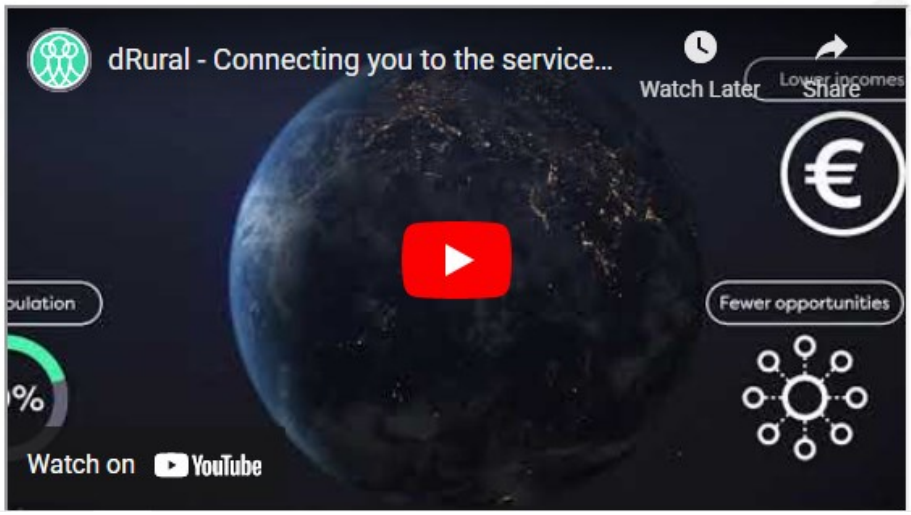
	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
If I needed to, I would feel at ease when presenting an idea to others	○	○	○	○	○
I tend to worry about being judged by others when presenting an idea	○	○	○	○	○

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

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.



https://www.youtube.com/watch?v=RCQc24UYfeI

Appendix D

Scenario 1 & 2

Hello, I am **EVA**, the Business Idea Diagnosis Tool I am programmed to help you evaluate the potential of your business ideas. Based on my research and analysis, you will be provided with a concrete evaluation of whether the idea has a potential for further development.

If you wish to proceed using my services please click on the button below.

Scenario 3&4

Hello, I am **EVA**, the Business Idea Diagnosis Tool! I am programmed to help you evaluate the potential of your business ideas. I specialize in business ideas' diagnosis and development, and my evaluations are based on analyses of thousands of business ideas over the past four years. My evaluations have already helped with the founding of quite a few startups, including some highly successful ones**.** According to previous users, I am known for providing clear questions that help in finetuning your idea and contribute to your success. I was created by experienced researchers at a reputable technical university in the Netherlands, in collaboration with a startup consultancy, both of which are known for their success in creating societal impact.

For the analysis of your business ideas, I use IT algorithms, external sources and databases to which I compare the information that is given. Based on my research and analysis, you will be provided with a concrete evaluation of whether the idea has a potential for further development.

For the successful analysis you will be asked a number of questions, some of which contain a certain amount of personal information. However, I would like to assure you that we treat such information with utmost caution. You may read our privacy policy. for more detailed information on that matter.

It will take you around 10 to 15 minutes to successfully complete the tool.

If you wish to proceed using my services please click on the button below and let's go!! :)

Appendix E

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

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 F

I am ready! To receive your advice, please click on the button below.

Scenarios 1&3

Your business idea has great potential as it is highly innovative and/or in high demand in the target market. Given your experience, the likelihood of success is relatively high. Based on the information you have provided me your idea promises to yield ROI in the long run. I recommend you to develop a detailed investment and financial plan.

Scenarios 2&4

Your business idea has low potential to be successful in the target market. There is low demand for your product/service and the potential for your idea to be disrupting the market is low. Given your experience you may struggle as an entrepreneur. Based on the information you have provided me your business idea may be too expensive and/or indifferent for it to enter the target market

Thank you for using my services and I hope they are being of use to you! I wish you big success and I hope to see you again soon!

Appendix G

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.	○	○	○	○	○