The difference in UX between open and closed questions for novice users in conversational search

WYBE PIETERSE, University of Twente, The Netherlands

In human-computer interaction, natural language is important to improve the user experience. An attribute of natural dialogue is vague language, but vague language is hard to interpret for systems as meanings differ per user and context. So, to have a usable system with conversational search, it needs to decipher a user's need by understanding vague language. This can be done by asking the user clarifying questions to probe for more specifications and narrow the scope of their requirements, but what clarifying question should be asked differs per user. An expert user who knows the domain is better at naming the requirements outright than a novice user. That is why this research conducted a qualitative user experiment with novice users. In this experiment, the difference in user experience between open and closed questions was analysed. It was expected that closed questions would be an improvement for novice users due to the added context. From this, the results show that novice users who were asked closed questions felt more confident and more able to answer the questions in comparison to novice users with open questions.

Additional Key Words and Phrases: clarifyingj questions; chatbot; human-computer interaction; user experience; UX; open and closed questions; conversational search; eCommerce; multiple-choice; novice users

1 INTRODUCTION

While looking for products on the web, users often search for terms like *"cheap laptop"* or *"big screen"*. Terms like 'big' and 'cheap' are vague as what they entail differs per user [6]. The consequence of this is that users often get inaccurate results due to the system not being able to interpret correctly what the user's needs are. This results in a drop in user satisfaction as the system does not return what is expected [9]. To combat this, the system needs to adapt so it understands the user's needs better.
Currently, this can be done using conversational search with clarifying questions so that the system builds a better understanding of a user's need. This can be seen on the internet where many chatbots have been developed over time to engage in a natural language dialogue with the user [15].

An example of this would be when a user asks a chatbot for laptop recommendations and the system responds with "How much RAM do you want your laptop to have?". We argue that this splits users into two groups, as expert users would know what RAM is and how much they need, but novice users would not. Previous literature shows that these novices get frustrated by a question they do not know the answer to [11]. A novice user would not know how much performance their usage requires as the domain is unknown to them. Resulting in that they are not able to respond with relevant information, which then means that expert users who do know the domain need to be asked different questions than novices. Also, in other literature, it is found that novice users focus more on ease of use than expert users [13].
The existence of such a system is currently unknown to us. Resulting in that for these conversations all users are treated the same and asked the same type of clarifying questions. Which kind of clarifying questions should be asked to which user is unknown but it influences user experience, which in the domain of eCommerce is critical [3].

Open questions give more space for vague language and a broader input as the user has the freedom to choose their own vocabulary. This makes it harder for a system to interpret the meaning of the user as the answer given is not restricted within a scope. Whereas with a closed question the scope would be restricted. The advantage of this is that the user can express themselves however they want and can formulate their needs in their own words. In comparison, closed questions force the user to choose out of pre-set answers and stay within that strict scope pre-set by the system. A benefit of this is that closed questions give users more information, as possible answers are already predetermined and given. The fact that closed questions have a limited scope also makes it easier for a system to interpret, as the number of outcomes is limited. However, closed questions do have a greater risk of incurring bias [5] as the user is given more information which could sway a user into a certain direction.

To address this issue for novice users, the system needs to be adapted in such a way that it can uncover the needs of a user. As they are a novice, it is likely that this user does not know what their requirements explicitly are [11]. This research focuses on determining what the difference in effect is between different kinds of clarifying questions. The focus will be on open versus closed questions, because of the difference between the two types and it is expected to illustrate a difference in user experience. As open questions give a broad scope and closed questions narrow the scope for the user. Therefore, we formulate the following research question:

> *"What is the difference in user experience between closed questions and open questions for novice users?"*

To answer the research question, we decided to do a qualitative user study with novice users. By conducting this study with novice users, we seek to gain a deeper understanding of how users perceive and interact with these question types. The study utilises an existing system that employs open questions and extends it to support closed questions. From open interviews, we were able to obtain data and answer the research question. The closed questions are presented as multiple-choice options based on relevant use cases. This is to give the users proper context for how they would utilise their laptop and what they need.

TScIT 39, July 7, 2023, Enschede, The Netherlands

The open interviews revealed that users who received open-ended questions from the system demonstrated that they had little confidence in answering them. This stemmed from their lack of knowledge about RAM, which led to them wanting more examples and context. The data also showed that the users had problems indicating what display size they wanted as they could not translate how big a measurement is in practice. They lacked reference to answer this question properly.
In comparison the participants who received the closed questions from the system felt more confident, as they were able to obtain context from the options that the system showed them. From this, they were able to formulate an answer to the question. This was again evident in the technical question as they felt a lack of knowledge but expressed that the added context helped them. Even though these users felt more confident some still expressed that these options still did not give them enough reference to properly answer the question about display size.
This research's most significant contribution is its addition of knowledge on user experience, which provides valuable insights into how novice users perceive closed-ended and open-end questions.

2 RELATED WORK

It has been shown in previous work that natural dialogue is important to a natural conversation and is needed to improve human-computer interaction [10]. In a natural conversation, humans use vague language which is often seen as a reduction in precision and clarity but vague language can be more effective and more productive [7]. An example of this from *Jucker (2003)* was that in a casual chat, very precise language might be off-putting. Apart from this, it also has and is not limited to social functions like nuancing statements and personal evaluation [7]. So although vague language is important to natural dialogue, computer systems have trouble interpreting them as it can be hard to quantify [6]. This is due to the fact that users use different vocabularies [9] and search engines do generally not handle these differences in keywords well, which results in poor performance [9, 4].
In the research domain of conversational search, numerous papers dive into clarifying questions to improve human-computer dialogue. These papers are focused on multiple facets of conversational search like information retrieval [18] and evaluate the effect of clarifying questions on the performance of these systems [1]. On the other hand also what effect clarifying questions have on the user [4] and how willingly users are to answer them [19, 12]. However, no literature was found on what the effect is on users between different clarifying questions and what they prefer. In *Bondarenko, 2022* [4] it was shown that the majority of participants in that study enjoyed the system with a clarification component and found it helpful to find satisfactory answers.

eCommerce is a domain where conversational search is of importance. As users need to find products they are looking for using search queries to complete their purchase and let their visit be successful for the business [8]. As a result, chatbots have been developed to improve upon these search queries. [15]. In this domain, there is research into the different kinds of users and their different needs, an example of this is that novice users focus more on ease of use than expert users would [13]. This difference is of importance as technical jargon can irritate these novice users which has the consequence of lower user satisfaction [11].

The discussion of open versus closed questions has been held for decades, the bigger differences are that closed questions are quicker and easier to encode for research but can bias the response. While open questions allow for more user freedom reducing the risk of bias however also increasing the work of coding the replies and requiring more work from the user [5]. There is also a paper that reasons to use open questions as these can have a higher information gain. Resulting in a reduction in the number of questions that a user has to answer in comparison to only using closed questions [18]. Another study shows that users are willing to answer only a limited number of questions, supporting the idea that reducing the number of questions is important. [19].

3 APPROACH TAKEN

For this research to compare open versus closed, we chose to go for scenarios as it limits the scope of the user as to what they can answer but at the same time abstracts the specific information into direct use cases. These scenarios were chosen because previous work in conversational search asked for open questions [2], and in this research, we want to focus on what the effect is if closed questions are asked to novice users. As closed questions can give more information and users might not know exactly what they need [2]. So it provides context to their needs. The abstracting of technical information could benefit novice users as they get irritated by jargon they do not understand [11]. To make this comparison with scenarios a qualitative user study was done to understand how users perceive their interaction with the system.

3.1 Participants

We recruited 12 participants for this experiment. The gender and age distribution of these participants can be found in Table 1. All these participants were related to the University of Twente, meaning that they currently work or study there or have studied there. They were approached by the researcher, asking if they wanted to help by doing a small experiment and interview. All of the participants were friends or acquaintances of the researcher. As this research focuses on novice users, it had to be determined beforehand if these participants classify as novice users. This was done by asking them the question *"How do you classify your knowledge of laptops?"* on a scale from one to seven. Where *"1"* meant *little knowledge* and *"7"* meant *expert level.* The researcher excluded participants with a self-assessed knowledge score of 5 or higher. The distribution of these participants among the knowledge levels can be seen in Table 2.

3.2 Interactive system

To conduct this experiment, we used an existing chatbot and extended it. This existing chatbot was designed to help users find a new laptop using conversational search and clarifying questions.

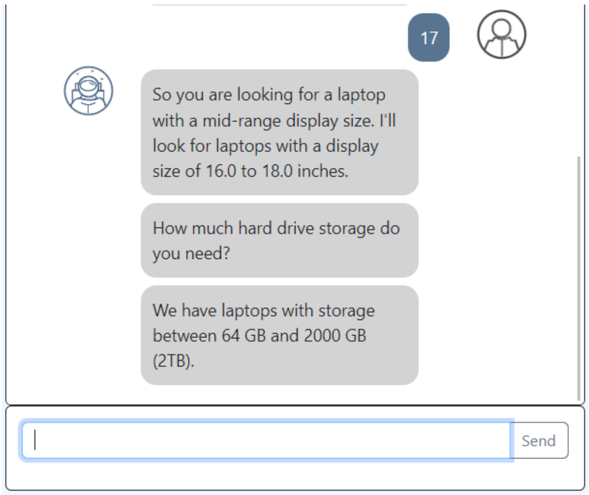

Fig. 1. The original chatbot with open answers

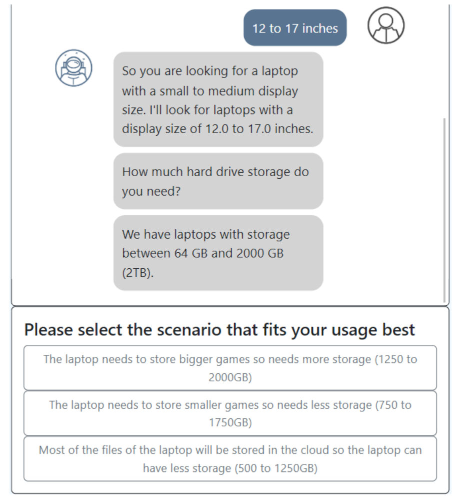

Fig. 2. The extended chatbot with scenarios

	Age Group		
	18-24	25-31	32-40
Male	1	1	1
Female	8	1	0

Table 1. Distribution of participants over the gender and age groups

User knowledge rating (1-7)	1	2	3	4
Participant count	1	6	3	2

Table 2. Distribution of participants across knowledge ratings

The chatbot asks the user about the following subjects: *purpose, price, display size, RAM size, storage size,* and *battery duration*. The original existing chatbot would ask questions about these subjects and have the users give open answers. To make a comparison between open and closed answers, this chatbot was modified. It was chosen to replace the open answer for *display size, RAM size, storage size, and battery duration* with closed answers. We did not replace the open question about *purpose*, the reason for this is that it is used as an introductory question. It is a very broad question and if users want a laptop they know why they want it. This means they can answer this question, by leaving this open you leave the user a lot of freedom. We also did not replace *price* as this depends on the financial situation of the user which the chatbot cannot help figure out. So, giving extra information to the user for these questions is not needed which is the point of the closed questions.
For the options of the closed questions, scenarios were implemented. This means that the user got a multiple-choice question with different scenarios, each representing a use case. Each use case had its own technical specifications assigned to it. For example, in a use case where the laptop is always connected to an external power source the specification is only a few hours of battery life. Whereas in a use case where the laptop is travelled with a lot it needs a longer battery life. Examples of this are shown in figure 1 and 2.

Moreover, there were also different kinds of scenarios for different types of users, because someone who needs a basic laptop is not in need of 32GB of RAM. With this method, fewer scenarios or more specific scenarios for the user can be shown. To determine the type of user the very first question that the user is asked about their *"purpose"* for their laptop was used. The answer that is given by the user is analysed. From this, the user is put into one of three classifications;

Basic: *A user who does not need fast hardware for compute-intensive tasks*
Advanced: *A user who does use their laptop for compute-intensive tasks*
Gaming: *A user who uses their laptop for gaming*

Examples of how these classifications impact the scenarios for the user can be found in figure 6. When presented with a multiple-choice option the users were asked to pick the scenario that fits their use case *"best"*. As it could that the perfect scenario for them is not included as an option.
When constructing the scenarios it was chosen to have at least three options to choose from and at most five. This was to not limit the user with too few options and at the same time not tire the user with too many options as that increases reading time and complexity.

3.3 Procedure

The experiments were conducted in a private space wherever it suited the participants, sometimes at home and other times in a meeting room at the university. For this experiment, the participants were split into two groups. One group used the original chatbot with open answers, and the other group used the extended chatbot with multiple-choice scenarios. The participants were equally split between each group, meaning six participants per group. They were handed a laptop with the chatbot running and requested to answer all the questions that the chatbot asked. Examples of the chatbot

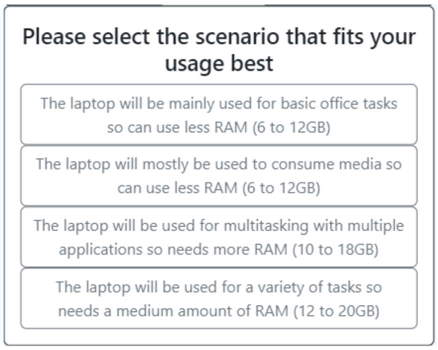

Fig. 3. Options for a user with a basic purpose

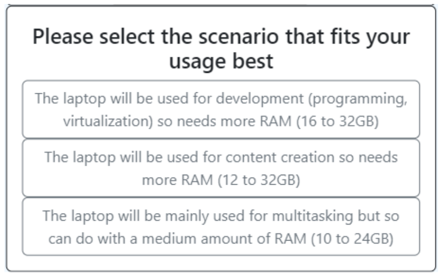

Fig. 4. Options for a user with an advanced purpose

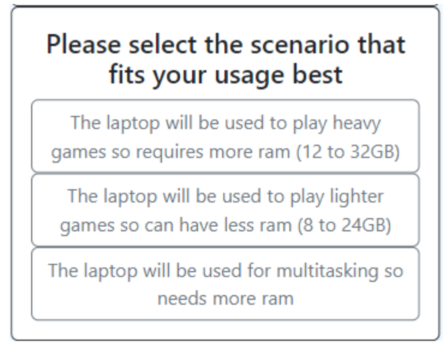

Fig. 5. Options for a user with a gaming purpose

Fig. 6. Screenshots of different multiple-choice options based on the purpose the laptop will be used for

are shown in Figure 1 and 2. Before they started the interaction, they were asked to think out loud and share all their thoughts and opinions during the interaction. The experiment was recorded for later reference and transcribed for data analysis. During the participants' interaction with the chatbot, observations were noted and any questions the user had about the interaction were answered. Technical questions like *"What is RAM?"* were not answered to not give one participant more information than the other. The observations that were made were later on used for follow-up questions in the interview.

After participants finished the interaction with the chatbot, an open interview followed. The purpose of this open interview was to obtain data about how users perceived the interaction. This interview used open-ended questions with extended probing in order to obtain a holistic understanding of the subject [17]. Meaning that whenever the user made a statement about their experience with the chatbot, follow-up questions were asked, in order to obtain a deeper understanding of why a user experienced this. The earlier made observations during the interactions were used to support these questions. The total length of the interaction with the chatbot and open interview lasted between 10 and 20 minutes.

3.4 Analysis

To analyse the data that was retrieved from the experiment, the first step taken was to transcribe all the recordings. Whereafter, we then used inductive coding to develop a conceptual understanding of the experiences of participants. We chose inductive coding as it is a systematic used process to gather reliable and valid findings for qualitative data analysis [16]. The inductive coding was done in three stages. Firstly, observations were made, for example, a user had trouble answering the question about RAM. Then we sought patterns and finally, we developed a general conclusion. During the coding process, we identified two broad themes with smaller sub-themes.

4 RESULTS

This research aims to answer the research question:

> *"What is the difference in user experience between open and closed questions for novice users?*

As the premise is that asking technical closed questions to a novice user could be better than asking open questions. Due to the fact that closed questions can give more information and thus context to the user.

As outlined in section 3 an experiment was conducted with two user groups, both consisting of six participants. This section will outline the results that were found from this experiment. It will split into the themes that were found during the analysis as outlined in section 3.4 with subsections detailing each user group.

4.1 Reference

4.1.1 Open Questions. The first question that the chatbot asks the users is; *"How big should the display be?"*. During the experiment, two users asked the researcher *"How big is this display?"*, referring to the laptop that the interview was conducted on. When asked why they asked this question in the follow-up interview they said:

> *"I have no clue how big an inch is, so I could use this as a reference"*
> *"It is hard to imagine a size in inches"*

Other users who did not ask this question during the interaction with the system also indicated similar issues with understanding how big they wanted their display to be. They said:

> *"Like for the display I have no clue how big it is. So maybe some pictures to get an idea of how large it is"*
> *"I had trouble with inches of a display as I cannot imagine that very well"*

So, all of these users had issues with imagining how big they want their display to be, they lack a reference to how big a display is in inches. This experiment was done in The Netherlands where the metric system is used. As display sizes are generally specified in inches this could be the reason for this problem. However, when these users were asked if they could specify their display-size requirement in centimetres they answered *"probably not actually"* or *"I don't think so"*.

One of the participants of the open-question group said they did not have trouble answering this question. When asked why this was they said *"I already know the laptop size I currently have and want the same"*.

4.1.2 Closed Questions. During the interaction with the chatbot, the group that used the version with multiple-choice options issues with reference also arose. Two users asked the researcher for the size of the laptop the interview was conducted on, in order for them to get a reference for size. Another user said:

> *"It would be nice if I could see more, for example, the size of the monitor if I could have a comparison with what I am currently using."*

One user did find that the scenarios helped him choose the display size and gave him a reference:

> *"I think the closed questions were good because then you can choose more the direction you want it and it gave me extra information on what I need. For example, the screen size I have no grasp on how big a 12-inch display actually is."*

As said in section 4.1.1 here also it could be an issue that the display size was in inches while the interviews were conducted in The Netherlands. Just like the other group, these participants were also asked if they could express it in centimetres and answered negatively.

4.2 Context

The theme of participants wanting a reference to answer the question can be connected with another theme that arose from the analysis. That is context. This was the most visible during the question about what RAM size the user wanted, which is the most technical question that the users were asked.

4.2.1 Open Questions. During the interaction, multiple users asked the researcher what RAM was or how much is *"normal"*. The researcher did not answer these questions as explained in section 3.3. Some users also tried to ask the bot *"What is RAM"* or *"How much do you recommend?"*. To which the bot replied that it did not understand or that the average user uses 4 to 8GB of RAM.
When the users were asked about their experience with the RAM question in the interview it was evident that they struggled:

> *"Absolutely no idea, an indication of something would be nice"*
> *"No clue what RAM is and how much I would need"*
> *"The RAM question especially as I did not what I needed and the recommendation for the system was not very helpful"*

From these excerpts, it can be seen that the users struggled. The other participants did as well and asked for examples or recommendations for a good answer:

> *"I had trouble with the RAM, maybe you can give examples"*
> *"I needed more explanation for the RAM as I do not know how much I needed"*
> *"I do not know what RAM is if I could get more of an explanation of how much RAM I would need that would be nice"*

These are answers from all six participants from the group, it can be seen that they all lack the knowledge to answer the question *"What RAM size do you need?"*. This is understandable as they all are novice users with little knowledge about laptops. Multiple users indicate that they want *"examples"* or an *"explanation"* to help them answer the question, they lack context to answer.
For the other questions only one participant said they struggled with storage as *"I do not know how much storage I need"* and the rest did not indicate they struggled with storage or battery life when asked.

4.2.2 Closed Questions. The expectation for the multiple-choice questions is that they give extra information. Resulting that the users have more context and it is easier to answer the question. In section 4.1.2 there is a quote from a user that indicates that *"it was nice that there were examples of what the size was"*, meaning the extra reference and context helped him choose. Another user also said something similar about their general experience with the chatbot;

> *"Worked quite well, it was clear for me to fill in due to the examples. I was able to understand it due to that"*

Specifically for the RAM question the multiple-choice mentioned the positive influence of the scenarios:

> *"Storage I understand but RAM is a bit too complicated so the examples made it very easy for me."*
> *"Unfortunately it doesn't explain what RAM exactly is but the scenarios gave me some context which was nice"*
> *"It was nice because you do not have to know anything about it since the scenarios really helped. To give you a good idea of which answers you should give and it was pretty fast. I like to know what I am choosing so more explanation would be nice"*
> *"Not entirely, I didn't know what RAM was but the scenarios helped"*

All the users that did say that the RAM question was hard to answer did find that the scenarios helped them. Two users did not mention RAM when asked which questions were hard to answer. However, these also made positive notes about the context that the multiple-choice options gave them:

> *"The scenarios were good, it gave me more context and what I need for what I use it for"*
> *"Rather easy, the scenarios helped and gave me context"*

All users in this group did find that the scenarios had a positive effect on the amount of context they had and some indicated that this improved their ability to answer the questions. None of the users of this group indicated that they found it hard to answer the storage size or battery life questions.

4.3 Overall experience

During the interview and interaction, the participants also made comments regarding their overall experience with the chatbot. Whereas the other two sections focused more on the results that followed from the user's experience with the questions. First, this section will look at the confidence the users had during the interaction.

4.3.1 Confidence. Both groups were asked how they felt when answering the question and how confident they were. Between the two groups, there was a stark difference. The open-answer group felt not very confident in their ability to answer:

> *"I was not very confident as I don't really know what the options are and are good for me if the bot had a more thorough understanding of my needs I would feel more confident"*
> *"Some I was confident as I know how much storage my computer has, but for the things I didn't know I was not very confident. It would be nice if the bot gave even more help and more handles"*
> *"I would be more confident if the system gave more information"*

From this, it is apparent that these users are lacking information and these are some answers that the closed-question user group gave:

> *"Quite confident, the scenarios helped me narrow it down to a nice range"*
> *"Yes I would say I feel more confident with scenarios as now I know what are the considerations"*
> *"I think it is mostly a good thing as it gives you more comfort I guess. Due to the fact that you have a bit of an idea of what you are saying"*
> *"It was fine, especially with the sentences you can choose between as it helped me make a choice"*

In comparison, the closed-question users had more confidence in answering the question and were more comfortable. A user of the closed question group said that the reason was that they trusted the multiple-choice answers as they expected some expert to have thought about them;

> *"Positively because I think someone has thought about these examples and they can guide me to what my needs are."*

4.3.2 Overall comments.

Open questions. As outlined in the previous sections it became clear that the users of this group felt that the system did not provide enough information. The users did indicate that they were interested in using such a chatbot. However, they were put off by the fact that they could not ask it questions about things they did not understand. As they felt there was a lack of recommendations:

> *"It seemed very helpful but when I asked it questions it couldn't explain stuff. I felt like it could give me good advice."*
> *"It was okay but gave some really weird answers from time to time. The system didn't give me any recommendations"*
> *"It did not answer the questions I have, it did not explain what ram size actually meant and how important it was for me."*

The users did however have some positive notes about their interaction with the chatbot and liked how the chatbot reacted to them if they stayed within the scope of the bot:

> *"It was pretty clear, like the way the answers of the bot were structured in small pieces of text. "*
> *"The chatbot was really straightforward and easy to navigate."*

Closed questions. This user group was more positive about their experience with the chatbot and was interested in using such a chatbot:

> *"Yeah, as I usually use my dad so this would be pretty nice"*
> *"Yes, I feel it could give me a good suggestion and I could start from there"*
> *"Yes, if it works well and I'm looking for a laptop I'm interested"*

This indicates that the users are willing to use such a chatbot if it were to be available. One user also mentioned that they liked that the interaction was shorter as they could just press buttons; *"Nice and easy to use, it was nice that I didn't have to specify it but could just simply answer quickly."*. As mentioned in earlier sections the users did like the multiple-choice options that were given however one user did feel that they were limited by the options; *"Maybe more options would be nice, sometimes I wanted to select both so this limited me in my choice."*. In a further clarification, they added: *"I could not express what I found more important"*.

5 DISCUSSION

The data collected in this experiment suggests multiple things. Firstly, the results indicate that users want reference and context. Reference means that it is hard for them to imagine text and measurements in the real world, so they want to compare it to something they do know. Context means that the users are lacking the knowledge to answer questions about what their requirements are. In the experiment, some users of the multiple-choice group found that the closed questions did help them with referencing, as it gave them a feel for which use cases use what kind of laptop size. But some users still needed more reference to answer.
The experiment did indicate that the closed question did help with the context, especially for the more technical question. This continued in that the data from the user group with closed questions suggested that they were more confident than the open questions participants. Overall, the closed-question participants were more positive and interested in potentially using such a chatbot.
Secondly, there was one user who felt that the closed questions limited their ability to choose. They felt that they could not express all their needs due to the simplicity of the closed question.

5.1 Interpretations

As was expected, novice users do indeed benefit from closed questions in this technical context. It was found that this is due to the lack of context and reference that the users have, closed questions improved this by giving examples. This way the users felt more confident in answering the questions which then in turn improved their overall user experience. Although, users with closed questions still found some questions hard to answer they got help from the options given. For the screen size, some users indicated that they still needed more reference as they found it hard to imagine inches in real life. A solution with a direct comparison with what the user is currently using might solve this. This does mean that although the closed-question approach that was taken for this research was not perfect it is an improvement over asking open questions.
Closed questions limit the scope of the user to what is set by the system. As there is no additional input from the user where they can express nuances or opinions. One user expressed that this was indeed limiting for them. That is one of the key drawbacks of using closed questions and should be looked at with care. This is why in section 6.3 we suggest that future research looks into a hybrid version.

5.2 Implications

The result is in line with previous literature that novice users have issues with technical jargon and that it negatively impacts user experience [11]. As it was shown that when users got the open question they were confused and tried to gain more knowledge. Whereas this was not the case with the closed questions. So, although previous literature found that open questions can have a higher information gain [18] closed questions should not be overlooked, because if the user does not know how to answer then the system will not receive a quality answer. It also should not be forgotten that users are only willing to answer a limited number of questions [19] and open questions could prove helpful in this. So these closed questions should be carefully constructed to be effective for the user. In combination with this it could be there are different user groups who would benefit from different kinds of questions. As was utilised in this research by splitting the users into *basic, advanced, and gaming*. Meaning that the application of these closed questions needs a thorough understanding of their users.
Users did show a positive experience with the chatbot and that they were willing to use it. This corresponds with what was shown in earlier studies that users enjoy systems with a clarification component and find it helpful to find satisfactory answers [4].

5.3 Limitations

5.3.1 Participants. As explained in section 3.1 the participants were collected from the University of Twente, which is a technical university. With that, these users also are intensive users of their laptops for their studies and work. Those factors could influence the self-assessed knowledge score as their environment is highly technical. On top of this, as illustrated in table 1 the participants were 75% in the age group of 18-24 and 75% female. Meaning that these participants do not represent an equal spread across all genders and age groups which could affect results.

5.3.2 Technical context. This experiment was conducted in a technical context where users are searching for a new laptop. Closed questions helped in this instance to answer questions. It was not researched if this is applicable to different contexts.

6 FUTURE WORK

6.1 Kind of closed answers

For this research, the closed answers were multiple-choice representing use cases. In this study, it helped novice users but it could be that different kinds of options would work better, this relates to the content of the options. Another subject that should be researched is a different kind of closed question. For example, a user could rank the multiple-choice options to what they like the most. Or select the option which they like the most and the least and from there the system could calculate an answer. This gives more freedom to the user and could be an improvement but also increases complexity, further research is needed.

6.2 Broader scope

The context of this experiment was technically focused and here it helped novice users. But it could be that this is also the case in other environments, less technically inclined. Novice users in different disciplines could maybe also benefit from more context and reference. That is why research should be done into what the effect is in other aspects of closed versus open questions.

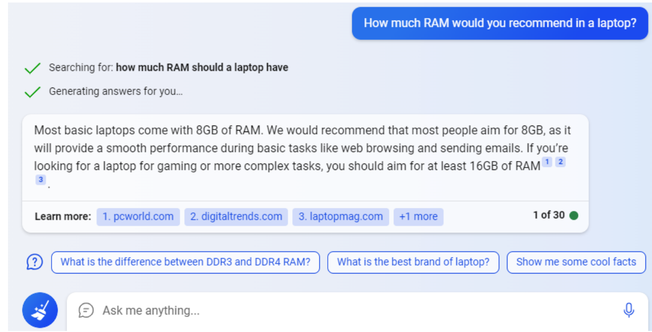

Fig. 7. An example of Bing Chat utilising a hybrid approach between open answers and suggesting answers

6.3 Hybrid solution

This research showed that closed questions can help novice users, but it also showed that users still want to be able to answer differently than what the options are if they have questions for example. And ideally, you want the same user interface for expert users as you have for novice users. As this reduces the complexity of the system which increases if two systems needed to be maintained. This experiment made a hard comparison between open and closed answers, it was one or the other, however, a hybrid solution could maybe combine the best of both worlds. An example of this format being used already is "Bing Chat" from Microsoft of which an example can be seen in Figure 7. They use GPT-4, which is a large language model, for their chatbot and also generate follow-up answers for users with that [14]. Extra research should be conducted on what the best solution would be.

7 CONCLUSION

This research aimed to give insights into how novice users perceive the difference between open and closed questions. Based on the analysis of the qualitative study that was done, it can be concluded that closed questions can improve the user experience of novice users in comparison to open questions. The results show that when users are asked questions users need to have enough context to formulate a proper answer to the question. This study shows that closed questions can provide the context that is necessary. This research meets the expectations that were set by previous works that had shown that novice users need more information. It does raise the question of what the ideal kind of question is for which user. For this further research is needed.
To conclude, this research contributes to the field of user experience by providing new valuable insights. From this better user experiences can be created and there exists a better understanding of how users perceive questions.

ACKNOWLEDGMENTS

Firstly, I would like to majorly thank Andrea Papenmeier, my supervisor, for giving me thorough feedback plenty of times even though there were some doubts along the way. Then, Mariët Theune and all my fellow students for all the track meetings and advice that was given. Last but not least all the friends who participated in my study and were willing to take time out of their day to help me.

REFERENCES

[1] Mohammad Aliannejadi et al. "Asking Clarifying Questions in Open-Domain Information-Seeking Conversations". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '19: The 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. Paris France: ACM, July 18, 2019, pp. 475–484. ISBN: 978-1-4503-6172-9. DOI: 10.1145/3331184.3331265. URL: https://dl.acm.org/doi/10.1145/3331184.3331265 (visited on 04/27/2023).

[2] Keping Bi et al. "Conversational Product Search Based on Negative Feedback". In: *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*. CIKM '19: The 28th ACM International Conference on Information and Knowledge Management. Beijing China: ACM, Nov. 3, 2019, pp. 359–368. ISBN: 978-1-4503-6976-3. DOI: 10.1145/3357384.3357939. URL: https://dl.acm.org/doi/10.1145/3357384.3357939 (visited on 04/27/2023).

[3] Laia Bonastre and Toni Granollers. "A Set Of Heuristics for User Experience Evaluation in E-commerce Websites". In: *ACHI 2014 : The Seventh International Conference on Advances in Computer-Human Interactions* (2014). ISBN: 978-1-61208-325-4. URL: https://www.academia.edu/download/44494283/achi_2014_2_10_20126_1.pdf.

[4] Alexander Bondarenko, Ekaterina Shirshakova, and Matthias Hagen. "A User Study on Clarifying Comparative Questions". In: *ACM SIGIR Conference on Human Information Interaction and Retrieval*. CHIIR '22: ACM SIGIR Conference on Human Information Interaction and Retrieval. Regensburg Germany: ACM, Mar. 14, 2022, pp. 254–258. ISBN: 978-1-4503-9186-3. DOI: 10.1145/3498366.3505832. URL: https://dl.acm.org/doi/10.1145/3498366.3505832 (visited on 04/27/2023).

[5] Saoirse Connor Desai and Stian Reimers. "Comparing the use of open and closed questions for Web-based measures of the continued-influence effect". In: *Behavior Research Methods* 51.3 (June 2019), pp. 1426–1440. ISSN: 1554-3528. DOI: 10.3758/s13428-018-1066-z. URL: http://link.springer.com/10.3758/s13428-018-1066-z (visited on 05/04/2023).

[6] G. W. Furnas et al. "The vocabulary problem in human-system communication". In: *Communications of the ACM* 30.11 (Nov. 1987), pp. 964–971. ISSN: 0001-0782, 1557-7317. DOI: 10.1145/32206.32212. URL: https://dl.acm.org/doi/10.1145/32206.32212 (visited on 04/27/2023).

[7] Andreas H. Jucker, Sara W. Smith, and Tanja Lüdge. "Interactive aspects of vagueness in conversation". In: *Journal of Pragmatics* 35.12 (Dec. 2003), pp. 1737–1769. ISSN: 03782166. DOI: 10.1016/S0378-2166(02)00188-1. URL: https://linkinghub.elsevier.com/retrieve/pii/S0378216602001881 (visited on 04/27/2023).

[8] Aarno Lehtola, Johannes Heinecke, and Catherine Bounsaythip. "Intelligent Human Language Query Processing in Mkbeem". In: *C. Stephanidis (ed), Universal Access in HCI, Vol. 4 of the Proceedings of HCI International, Lawrence Erlbaum Associates Publishers, Mahwah, New Jersey, 2003, pp. 750-754* ().

[9] Subhadeep Maji et al. "Addressing Vocabulary Gap in E-commerce Search". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '19: The 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval. Paris France: ACM, July 18, 2019, pp. 1073–1076. ISBN: 978-1-4503-6172-9. DOI: 10.1145/3331184.3331323. URL: https://dl.acm.org/doi/10.1145/3331184.3331323 (visited on 05/04/2023).

[10] Erinc Merdivan et al. "Dialogue Systems for Intelligent Human Computer Interactions". In: *Electronic Notes in Theoretical Computer Science* 343 (May 2019), pp. 57–71. ISSN: 15710661. DOI: 10.1016/j.entcs.2019.04.010. URL: https://linkinghub.elsevier.com/retrieve/pii/S1571066119300106 (visited on 05/05/2023).

[11] IM Moczarny, de Villiers, and van Biljon. "How can usability contribute to user experience?: a study in the domain of e-commerce". In: *SAICSIT '12: Proceedings of the South African Institute for Computer Scientists and Information Technologists Conference* (). DOI: 10.1145/2389836.2389862. URL: https://dl.acm.org/doi/10.1145/2389836.2389862.

[12] Paloma Ochi et al. "Predictors of user perceptions of web recommender systems: How the basis for generating experience and search product recommendations affects user responses". In: *International Journal of Human-Computer Studies* 68.8 (Aug. 2010), pp. 472–482. ISSN: 10715819. DOI: 10.1016/j.ijhcs.2009.10.005. URL: https://linkinghub.elsevier.com/retrieve/pii/S1071581909001591 (visited on 04/28/2023).

[13] Sandeep Kumar Panda, Santosh Kumar Swain, and Rajib Mall. "An Investigation into Usability Aspects of E-Commerce Websites". In: *ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 1, No.13 , January 2015* 4.1 (2015). ISSN: 2322-5157. URL: https://www.academia.edu/download/36497207/ACSIJ-2014-4-1-631.pdf.

[14] Jay Peters. "Microsoft's new AI-powered Bing chatbot has been relying on the newly announced GPT-4 model all along". In: (Mar. 14, 2023). URL: https://www.theverge.com/2023/3/14/23639928/microsoft-bing-chatbot-ai-gpt-4-llm (visited on 06/25/2023).

[15] Filip Radlinski and Nick Craswell. "A Theoretical Framework for Conversational Search". In: *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval*. CHIIR '17: Conference on Human Information Interaction and Retrieval. Oslo Norway: ACM, Mar. 7, 2017, pp. 117–126. ISBN: 978-1-4503-4677-1. DOI: 10.1145/3020165.3020183. URL: https://dl.acm.org/doi/10.1145/3020165.3020183 (visited on 05/04/2023).

[16] David R. Thomas. "A General Inductive Approach for Analyzing Qualitative Evaluation Data". In: *American Journal of Evaluation* 27.2 (June 2006), pp. 237–246. ISSN: 1098-2140, 1557-0878. DOI: 10.1177/1098214005283748. URL: http://journals.sagepub.com/doi/10.1177/1098214005283748 (visited on 05/07/2023).

[17] Susan C. Weller et al. "Open-ended interview questions and saturation". In: *PLOS ONE* 13.6 (June 20, 2018). Ed. by Andrew Soundy, e0198606. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0198606. URL: https://dx.plos.org/10.1371/journal.pone.0198606 (visited on 05/05/2023).

[18] Yunfeng Zhang, Q. Vera Liao, and Biplav Srivastava. "Towards an Optimal Dialog Strategy for Information Retrieval Using Both Open- and Close-ended Questions". In: *23rd International Conference on Intelligent User Interfaces*. IUI'18: 23rd International Conference on Intelligent User Interfaces. Tokyo Japan: ACM, Mar. 5, 2018, pp. 365–369. ISBN: 978-1-4503-4945-1. DOI: 10.1145/3172944.3172998. URL: https://dl.acm.org/doi/10.1145/3172944.3172998 (visited on 05/04/2023).

[19] Jie Zou, Evangelos Kanoulas, and Yiqun Liu. "An Empirical Study on Clarifying Question-Based Systems". In: *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*. CIKM '20: The 29th ACM International Conference on Information and Knowledge Management. Virtual Event Ireland: ACM, Oct. 19, 2020, pp. 2361–2364. ISBN: 978-1-4503-6859-9. DOI: 10.1145/3340531.3412094. URL: https://dl.acm.org/doi/10.1145/3340531.3412094 (visited on 04/27/2023).