Do Make me Think! How CUIs Can Support Cognitive Processes

Leon Reicherts
UCL Interaction Centre
University College London
London, UK
l.reicherts.17@ucl.ac.uk

Yvonne Rogers
UCL Interaction Centre
University College London
London, UK
y.rogers@ucl.ac.uk

CHI 2020 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
DOI: https://doi.org/10.1145/3334480.XXXXXXX
*update the above block & DOI per your rightsreview confirmation (provided after acceptance)

Abstract
Traditionally, a key concern of HCI has been to design interfaces that should not make the user think. While this is – and will continue to be – desirable for most systems, there are also situations in which a system that prompts and questions the user may be more appropriate. In educational systems for instance, tasks are often intentionally made more challenging to enable "deeper" thinking and more thorough learning. Although conversational interfaces are still relatively limited in their capabilities, they seem very promising for contexts where questioning is needed, such as learning, analytics or sensemaking as well as safety-critical systems. Overly simple interactions – when the user can just *tap* or *drag and drop* – may not be beneficial in this context or may even be risky. In this position paper, we discuss previous work as well as opportunities where questioning users through conversation can be beneficial, based on insights from our own research.

Author Keywords
Conversational user interfaces; scaffolding; cognition; slow interaction; sensemaking, questioning; analytics.

CSS Concepts
• Human-centered computing~Human computer interaction (HCI)~Interaction paradigms~Natural language interfaces

Introduction

Traditionally, the focus of HCI and usability has been on making interactions as easy as possible. The user should not even have to think – *Don't make me think* by Steve Krug [5]. Indeed, it is desirable for most tools, devices and applications to be as easy to use as possible. However, here we argue that this can be different depending on what the user is trying to achieve at the interface, such as improving their activities of sensemaking, problem-solving and learning – or performing a safety-critical task. In such situations, prompting and questioning at appropriate/opportune points in the interaction may lead to more meaningful outcomes [7]. Although such prompts may sometimes be annoying, they can also help users think about certain actions, choices or conclusions.

An example of an 'easy-to-click' interface is *Tableau* [12]. The program makes it simple for the user to filter and visualize complex datasets. It enables lay users to conduct analyses, which only data analysts could do a couple of years ago. With just a few clicks, it allows users to generate almost any type of visualization that is 'generatable' based on the variables, data types and scales of the given dataset. However, choosing an appropriate visualization and understanding it generally requires certain methodological and domain-related knowledge. It is often the case that the easier an interface, the less users need to think about their actions. While it may be appropriate to make a task less cognitively demanding in many cases, there are also situations where the opposite is true. For example, how does a lay user know whether a treemap or a stacked bar chart is the more appropriate visualization of a certain dataset? Asking users certain questions about what they are hoping to discover at the interface may help them in their decision.

Natural language interfaces or Conversational User Interfaces (CUI) offer great potential in situations like the one described above, since they can prompt, guide and scaffold users' thinking when doing a task. More generally, conversational interaction seems suitable for tasks that are new to the user, ill-defined or exploratory. In such scenarios, users may not know what they are looking for and a conversational agent can help them to keep on track.

Graphical User Interfaces (GUIs), on the other hand, are ideally suited for direct manipulation (e.g. filtering, selecting, zooming and scanning). When users know what they are looking for or if they are doing a familiar task, GUIs may be most suitable. Of course, GUIs and CUIs are not mutually exclusive, particularly when thinking of a chatbot, which may be integrated into an application's GUI. For example, users can interact with the *Amazon Echo Show* [13] through voice as well as through a touchscreen interface. Furthermore, users can be verbally or textually prompted within a GUI through pop-up windows, for example.

The goal of this paper is not to argue about the benefits of CUIs over GUIs, but to consider the potential of CUIs in being able to scaffold users' thinking when simple GUI prompts, such as pop-ups, may not be sufficient or appropriate. To begin to explore which scenarios these are can help to inform design choices as to which interface type should be used to best support different activities and tasks – which ultimately means how CUIs and GUIs can complement each other best.

Background

Guiding Users and Scaffolding their Thinking

A benefit of using conversational agents at the interface is that they can provide suggestions to the user concerning what to do next or what else to consider. For example, Tegos et al. [10] demonstrated that conversational agents can trigger dialogs between students by intervening in a conversation, which substantially improved both individual and group learning outcomes. Moreover, Winkler et al. [11] showed that an Alexa-based tutor can have a positive effect on task outcome and collaboration among users in a problem-solving task.

Interfaces for Learning, Analytics and Sensemaking

Russell et al. [8] considered sensemaking as *"the process of searching for a representation and encoding data in that representation to answer task-specific questions"*. However, it can be difficult to see patterns, make inferences and understand what the representations mean in the context in which they are presented. Having an interface, which can guide and facilitate human reasoning through conversational interactions, may be highly desirable. For example, Subramonyam et al. [9] developed a dialog tool, which facilitated collaborative decision-making through enabling data-driven conversations. In the context of learning/educational science, questioning and problematizing is used by tutors/teachers to scaffold learners' problem solving [4]. For this purpose, software tools such as *ExplanationConstructor* by Reiser [7] have been designed, which allow users to articulate research questions and explanations as well as managing and structuring evidence. An interface that scaffolds, probes and guides users when exploring datasets might play a central role in other cognitive tasks, such as data analysis and decision-making. Therefore, combining proactivity with suitable ways of prompting for the task at hand, such as exploratory data analysis, is promising.

Existing Analytics Tools Incorporating CUIs

Iris by Fast et al. [1] provides a chatbot interface which supports data scientists in open-ended modelling tasks. Fast and colleagues mention in their paper that the 'structural guidance' provided by *Iris* was found particularly useful in their user study. Another example is *Ava* by John et al. [3], which allows data scientists to assemble data analytics pipelines, using a chatbot interface. Overall, this line of research suggests that the use of natural language as a modality of interaction can scaffold complex (data analytics) tasks.

Our Approach

To further investigate the potential of CUIs, which prompt the user and scaffold their thinking, we have developed an interface prototype for a data analytics scenario. As part of the system, an assistant asks users questions about the data being visualized. The questions aim to draw the users' attention to differences and trends within the dataset, which should help users to think about the data more thoroughly and to articulate hypotheses about why there are certain patterns. The chosen visualizations were time series graphs on increasing obesity levels in different countries from 1990-2013. They were based on Marinez [6] and the underlying dataset taken from the Global Burden of Disease Study [2]. These visualizations were chosen, as they were simple enough to understand but also as they contain relatively nuanced differences that need to be teased out (e.g. subtle changes in growth rates), which often require more detailed analysis.

A study was conducted in which participant pairs were asked to look at the visualizations and to discuss and hypothesize about patterns they saw in the data. There were two reasons for choosing pairs instead of single users. First, in a single user scenario, it can feel unnatural for someone to speak to a system. Second and most importantly, running the study in pairs provides opportunities to examine the kinds of conversations that would take place. Furthermore, our rationale was that talking to each other about the data could trigger further reasoning and thinking about the data.

The aim of the study was to examine the sensemaking that occurs when prompted by an interface. The prompts, which were provided through a virtual assistant, were aimed at stimulating the participants' discussion. A *Wizard of Oz* experiment was conducted, hence the prompts were controlled by a human, who pretended to be an intelligent virtual assistant. Participants were told that they could, but did not have to, respond to the assistant's suggestions. The assistant was designed to ask a number of predefined questions, such as "*Would you say that the increase is slowing down for all four groups?"* or *"Did you see this pattern elsewhere?"*. There were two rules for triggering a question, (a) that there was silence for more than three seconds and (b) that participants have not discussed a topic directly related to that question. After the participants had explored the dataset, they were asked about their experience of interacting with the system and the assistant in a semi-structured interview. In the following section, we provide an overview of some of our insights from these interviews.

Insights From our User Study

Overall, most participants found the assistant's prompts particularly useful for making them think about the given dataset from different perspectives. For example:

> [Participant 1 - Pair 5] *"It would tell us to think about stuff we didn't see at first, but they were really interesting to think about."*

Several pairs found that the assistant made them do the analysis task more slowly than if they had done the task without it. Many participants also found that this "slowing down effect" had benefits, for example, in situations where users are trying to better understand a certain topic/dataset to get a new perspective:

> [P1-P18] *"I think it is good [to use this system] if you have time and you are trying to figure out things."*

> [P2-P11] *"I mean it was more time consuming than traditional tools but that also has benefits if you are not in a rush."*

Some individuals also mentioned that when users become more familiar with the dataset, they may prefer the system to become faster.

> [P1-P6] *"If you are looking at the same data for an extended period of time, you mostly want it to be very fast to get data out. This isn't exactly fast. I guess this is more suitable if you are introducing a new topic or you are trying to get a new perspective on the same data."*

Similarly, many participants mentioned that they would rather not use the system when doing a specific task and know what to look for.

> [P1-P8] *"If you have a lot of variables and you are not really sure what you are looking for or if you are training someone it might be a good thing to use. If I know what I am looking for, I probably won't use it. (…) I would use it to generate hypotheses instead of testing my hypotheses."*

However, several participants pointed out that the assistant helped them to not get lost or stuck on a particular data visualization. It also allowed them to find additional differences or trends in the data when they thought that they had already discovered everything, for example:

> [P2-P7] *"For complex datasets this would be very useful, because when there are so many parameters (…) you might get lost in the data – like where you have started and where you are ending it (…) It could give me a starting point when I am confused."*

> [P1-P8] *"I think one thing that helped was that when we were kind of stuck and we were not saying anything, it would just generate a suggestion. I found that useful."*

> [P2-P10] *"I like the fact the questions were about finding more in the data. By looking at the question, you would think about the question and you would think about why (…) this is more steady than the other, which wouldn't happen without the assistant."*

A few participants also found that discussing the data in in the context of the assistant's questions helped them to better remember certain aspects of the data:

> [P1-P18] *"I am really impressed by what we all remember from that, so maybe it is also a good thing for remembering data by talking about it and having some kind of facilitator."*

Overall, the system and its assistant were perceived as helpful for the exploratory, open-ended task in our user study. Prompting via questions, which were aimed to help users to discover (nuanced) differences, was found to be useful in most cases. Many participants also found that the prompts had a stimulating effect on their discussion. The fact that the system sometimes slowed users down was generally perceived as beneficial. However, in line with our expectations, participants commented that it might not be so desirable to be slowed down when completing a familiar task or topic, where users know what they are looking for – unless they would like to approach it from a new angle. These comments highlight the trade-off between 'speed and scaffolding', where one enables the user to get their task done efficiently and the other can lead to deeper thinking.

Topics to Discuss in the Workshop

Based on these initial insights, we are interested in exploring where, when and how CUIs could be used to good effect to slow users down. Therefore, the key questions, which we would like to discuss in our workshop are:

- What are the situations and use cases where CUIs, which prompt and question users, could be used to augment and enhance cognitive processes?
- What are the key challenges of using CUIs for this purpose along with or instead of GUI-based interaction?
- Which design strategies could be used for combining the advantages of 'fast' graphical interfaces with 'slow' conversational interactions in hybrid/multi-modal systems?

Conclusion

Findings form our preliminary research suggest that questioning user actions through conversational interfaces has great potential, especially in scenarios where other types of interfaces may not be able to scaffold users thinking and provide sufficient guidance. However, the effect of slowing someone's thinking down in this way can increase the cognitive effort required. This may be appropriate for certain types of tasks and scenarios but not for others.

References

[1] Ethan Fast, Binbin Chen, Julia Mendelsohn, Jonathan Bassen, and Michael S. Bernstein. 2018. Iris: A Conversational Agent for Complex Tasks. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18, 1–12. https://doi.org/10.1145/3173574.3174047

[2] Global Burden of Disease Collaborative Network. 2014. Global Burden of Disease Study 2013 (GBD 2013) Obesity Prevalence 1990-2013. Institute for Health Metrics and Evaluation (IHME), Seattle. Retrieved from http://ghdx.healthdata.org/record/global-burden-disease-study-2013-gbd-2013-obesity-prevalence-1990-2013

[3] Rogers Jeffrey Leo John, Navneet Potti, and Jignesh M. Patel. 2017. Ava: From Data to Insights Through Conversations. In CIDR.

[4] Minchi C. Kim and Michael J. Hannafin. 2011. Scaffolding problem solving in technology-enhanced learning environments (TELEs): Bridging research and theory with practice. Computers & Education 56, 2: 403–417. https://doi.org/10.1016/j.compedu.2010.08.024

[5] Steve Krug. 2005. Don't Make Me Think: A Common Sense Approach to the Web (2nd Edition). New Riders Publishing, USA.

[6] Ramon Marinez. 2015. Level and Trends of Overweight and Obesity. Tableau Public. Retrieved September 17, 2019 from https://public.tableau.com/profile/ramon.martinez#!/vizhome/LevelandTrendsofOverweightandObesity/Overweightandobesitylevel

[7] Brian J. Reiser. 2004. Scaffolding Complex Learning: The Mechanisms of Structuring and Problematizing Student Work.

[8] Daniel M. Russell, Mark J. Stefik, Peter Pirolli, and Stuart K. Card. 1993. The Cost Structure of Sensemaking. In Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI '93), 269–276. https://doi.org/10.1145/169059.169209

[9] Hariharan Subramonyam, Bongshin Lee, Sile O'Modhrain, and Eytan Adar. 2017. Data Dialog: Facilitating Collaborative Decision Making Through Data-driven Conversations. In Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare

(PervasiveHealth '17), 440–443. https://doi.org/10.1145/3154862.3154921

[10] Stergios Tegos and Stavros N. Demetriadis. 2017. Conversational Agents Improve Peer Learning through Building on Prior Knowledge. Educational Technology & Society 20: 99–111.

[11] Rainer Winkler, Matthias Söllner, Maya Lisa Neuweiler, Flavia Conti Rossini, and Jan Marco Leimeister. 2019. Alexa, Can You Help Us Solve This Problem?: How Conversations With Smart Personal Assistant Tutors Increase Task Group Outcomes. In Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems - CHI EA '19, 1–6. https://doi.org/10.1145/3290607.3313090

[12] Tableau: Business intelligence and analytics software. Tableau Software. Retrieved February 9, 2020 from https://www.tableau.com/en-gb

[13] Echo Show 2nd Generation. Retrieved September 20, 2019 from https://www.amazon.co.uk/Amazon-Echo-Show-With-Screen-2nd-Generation-Alexa/dp/B0793G9T6T