Convey: Exploring the Use of a Context View for Chatbots

Mohit Jain[+*], Ramachandra Kota[+†], Pratyush Kumar[+], Shwetak Patel[*]
[+]IBM Research, India. mohitjain@in.ibm.com, pratyushkpanda@gmail.com
[†]Realtor.com, Vancouver, Canada. ramachandra.kota@move.com
[*]Computer Science & Engineering, University of Washington, Seattle, USA. shwetak@cs.washington.edu

ABSTRACT
Text messaging-based conversational systems, popularly called *chatbots*, have seen massive growth lately. Recent work on evaluating chatbots has found that there exists a mismatch between the chatbot's state of understanding (also called context) and the user's perception of the chatbot's understanding. Users found it difficult to use chatbots for complex tasks as the users were uncertain of the chatbots' intelligence level and contextual state. In this work, we propose *Convey* (CONtext View), a window added to the chatbot interface, displaying the conversational context and providing interactions with the context values. We conducted a usability evaluation of *Convey* with 16 participants. Participants preferred using chatbot with *Convey* and found it to be easier to use, less mentally demanding, faster, and more intuitive compared to a default chatbot without *Convey*. The paper concludes with a discussion of the design implications offered by *Convey*.

ACM Classification Keywords
H.5.2. Information Interfaces and Presentation (e.g. HCI): User Interfaces

Author Keywords
Chatbots; conversational system; dialog; shopping bot; context; user interface; design; evaluation.

INTRODUCTION
In 2015, text-messaging became the most popular class of smartphone applications, overtaking social networking [10]. This strong indication of users' preference for messaging-based, real-time conversations motivated the growth of chatbots. *Chatbots* refer to messaging-based conversational agents. Chatbots received significant attention in 2016 [4] with the expectation that users can 'text' intelligent agents of businesses, just as they text their friends and family. Over 34,000 chatbots have been developed on Facebook's Messenger Bot platform alone [5] within 6 months of its release in 2015.

Recent works by several researchers evaluating users' experience while interacting with chatbots have discovered a gulf between experience and expectation with respect to both intelligence and the user interface of chatbots [13, 14, 17]. They found that although users enjoy chatbots that can continue a conversation specifically by retaining *conversational context*, there is a mismatch between the chatbot's real context versus the user's perception of the chatbot context, *i.e.*, there is a difference between their mental models [12]. This is even more problematic in the case of lengthy, complex conversations. Similarly, users were found to be apprehensive in using conversational systems (*e.g.,* Siri, Cortana) for complex tasks, as the users were not certain of the system's intelligence level and had a poor mental model of its contextual state [12, 14, 15]. Moreover, certain chatbot assumptions are not evident to the user, further exacerbating this issue. The importance of explicitly providing contextual information in a GUI communication channel has been well established [6, 20], especially in text messaging domain [9, 11].

Against this background, in this paper, we propose *Convey* (CONtext View), a window added to the chatbot interface that displays the (inferred and assumed) context of the conversation to the user (Figure 2). It also provides intuitive interactions on the context values, enabling users to modify them in a simple and efficient manner. The *Convey* content gets updated as the conversation proceeds, thus always showing the latest understanding of the chatbot. To evaluate the effectiveness and usability of the proposed design, we conducted a 16 participant user study centered around a chatbot for buying shoes. The results show that participants preferred using chatbot with *Convey* and found it to be easier to use, less mentally demanding, intuitive, and faster compared to the default chatbot without *Convey* (seen in Figure 1). We conclude the paper with a discussion on the implications of *Convey* on future chatbots.

A BRIEF HISTORY OF CHATBOTS
Research in chatbots started with *chatterbots*, whose sole purpose was to maintain a conversation with a human user. The first chatbot, called ELIZA [18], emerged in 1966 from MIT. ELIZA worked on simple declarative rules: if a certain keyword was identified in the user text, it responded with one or more pre-defined outputs. Subsequently, in the latter chatbots, the rules used for both natural language understanding and natural language generation were enriched. Ontologies were used to represent word meanings, reasoning was used to identify user intent, and memory was used to continue a contextual conversation [16, 19]. The development of chatterbots has remained research-driven and not yet adopted by industry.

CHI 2018, April 21-26, 2018, Montreal, QC, Canada

ACM 978-1-4503-5620-6/18/04 $15.00
https://doi.org/10.1145/3173574.3174042

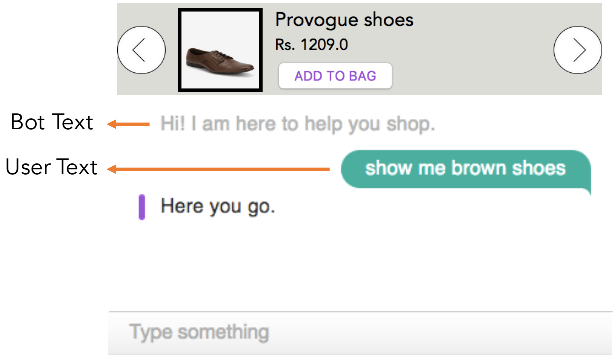

Figure 1: Shoe shopping default chatbot.

Instead, the tech industry has mainly been devoting its efforts towards 'utility-driven' chatbots - those designed to provide specific and limited services to the user (*e.g.*, Dominos chatbot for ordering pizzas). Facebook Messenger, Skype, Slack, Kik, Telegram, *etc*. together host more than a million chatbots [2], with use-cases ranging from food delivery (Domino's) to exploratory shopping (Burberry), from connecting like-minded humans (Chatible) to flight booking support (Kayak), and from casual conversation (Pandorabots) to reading news (CNN). The primary focus of these chatbots is not to mimic human conversation but to enable tasks through the ease of conversation. Anthropomorphism in these chatbots, when it exists, seeks to augment the efficiency of the task-solving process. This paper mainly focuses on advancing such utility-driven chatbots.

In spite of the growing industry adoption and the advancements in AI to make chatbots 'smarter' and more 'easy-to-use', the user interfaces of chatbots have not evolved much. They still closely resemble a messaging interface, wherein a user or a bot response results in a message bubble. While some chatbot platforms may have a few multimedia and interactive elements (such as buttons, hyperlinks, carousels, gifs, videos, and so on) to enhance interactivity, the essence of chatbot interfaces has remained unchanged. The main benefit of persisting with such an interface is that it is highly flexible and familiar to anyone who has used a messaging app before. In contrast, each website/app has its own interface, thus incurring a small learning curve [7]. In this work, we augment the familiar chatbot user interface with *Convey*. The aim of *Convey* is to enhance the effectiveness of using chatbots without losing the flexibility afforded by the messaging interface. We evaluate usefulness of *Convey* through a user study.

DESIGN OF CONVEY

In this section, we start with an overview of the basics of a conversation system, and then discuss the design of *Convey*.

Basics of a Conversation System

A conversation system identifies *intents* and *entities* from user's input, to understand the meaning of user text. The user's *intent* is the current goal or purpose of their interaction with the chatbot. The *entities* add value to that purpose and narrow it further to make it specific. For any chatbot, the intent and entity types are defined by the chatbot designer

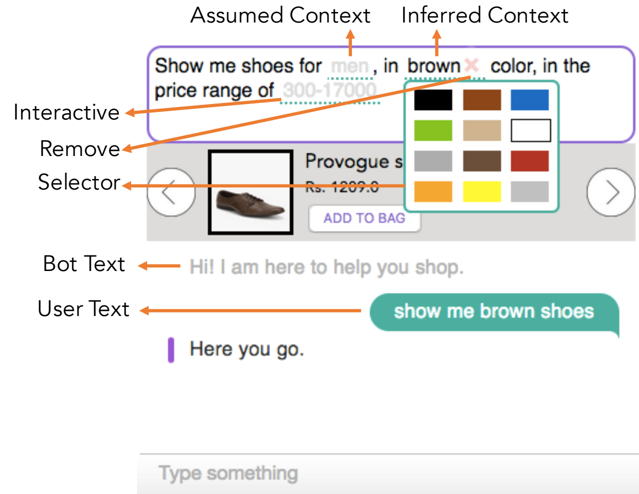

Figure 2: Shoe shopping chatbot with *Convey* at the top. Screenshot of the interface used in the user study.

based on the purpose of the chatbot. As the conversation involves multiple back-and-forth rounds between the user and the chatbot, the conversation system maintains *context* to keep track what the user and the chatbot have been discussing. The context values comprise of a combination of intents and entities. Therefore, without context, a user's new input would be analysed completely oblivious of their previous inputs.

As an example, here is a typical conversation with a delivery-ordering chatbot:

> Human: *i want to order a hawaiian pizza*
> Bot: *ok, anything else?*
> Human: *yeah make that medium size, and add a coke*

From the first message by the user, the chatbot recognizes that their intent is to *order food*, with the entity being '*hawaiian pizza*'. As the conversation proceeds, the chatbot maintains the context of pizza ordering so that '*medium size*' can be related to '*hawaiian pizza*'. Without the propagation of context, '*medium size*' is just another entity which is not attached to any intent. Thus, maintaining appropriate context of the conversation is crucial to a chatbot's success [12, 15].

Our proposed design (the *Convey* box) explicitly displays the context (including assumptions) of the conversation system to the user and provides a way to efficiently interact with the context values. We now discuss the primary features of the proposed *Convey* (a sample can be seen in Figure 2).

Convey: Showing Context

Context can be of two types: *inferred* and *assumed*. Inferred contexts are extracted from the conversation between the user and the chatbot. In the example shown in Figure 2, the user typed '*show me brown shoes*', so '*brown*' is an inferred context value. Additionally, a chatbot may typically assume a few context values based on the input. For example, on asking for '*brown shoes*', the chatbot might automatically assume that the user is looking for '*male*' shoes (perhaps based on user history). Depending on the chatbot design, even the price range of the shoes can be assumed based on the buying history of the specific user.

Convey shows both these contexts differently such that it is clear to the user whether the context was inferred or assumed. In *Convey*, inferred contexts are shown in black, while assumed contexts are shown in gray (Figure 2). The context values in *Convey* get updated in real-time as the conversation proceeds. Moreover, all displayed contexts are interactive, as indicated using a dotted underline. When a participant updates an assumed context, either by interacting with *Convey* (*e.g.*, by clicking '*men*' and selecting '*women*' from the drop-down list), or by stating the updated value as part of the conversation (*e.g.*, by texting '*looking for female shoes*'), *Convey* converts the assumed context into an inferred context. Alternately, based on chatbot design, a user confirmation can be attached to assumed context values, (*e.g.*, the chatbot asking the user explicitly, '*are you looking for male or female shoes?*').

Convey: Interaction

As stated earlier, the displayed context values in *Convey* are interactive in nature. The user can perform three actions: *confirm context*, *modify context*, and *remove context*. Confirming assumed context is as discussed in the previous section. Another way to confirm assumed context is by long pressing (*i.e.*, holding one's finger/mouse over an item for more than 0.5s) it. Note that this long-press feature is not visible to the user, so it can add to the learning curve. For this reason, the long-press feature was not included in *Convey*'s evaluation.

The user can modify a context by clicking on it. Each context value has a specific UI element associated with it. The element is populated with domain-specific options extracted from the chatbots' catalog in the database. Clicking the context value in *Convey* shows the UI with options as a *selector* pop-up. For example, clicking on the '*brown*' color context value in *Convey* pops up a color palette showing colors available in the catalog for male shoes (Figure 2). Similarly, clicking on the price range '*Rs 300-17000*' shows a slider-based price selector, while clicking on the gender '*male*' shows a drop-down menu with two gender options (*male* and *female*) to choose from.

Any of the context values can be removed by the user. Apart from users wishing to modify their preference, context may have been wrongly inferred or assumed by the chatbot, which also necessitates deletion by the user. Deleting context has been found to be an issue with current chatbots [14]; either the chatbot does not support deleting context, or it is hard for users to specify the deletion request in text so that the chatbot is able to correctly understand it. In our example, typing '*show all colored shoes*' results in removal of the '*brown*' color context value. *Convey* makes the deletion task much easier by allowing users to click on the cross ('x') button next to the context value (as shown in Figure 2 next to '*brown*'). The *cross* button for deletion, along with the *selector* pop-up for modifying context, only appear after clicking a particular context value. Deleting context is subject to the consistency of *Convey* after their removal.

Finally, the *Convey* design ensures symmetry between the two user modalities: typing and clicking to interact with context. Any interaction with a context value in *Convey* is logged as an equivalent message on the messaging window, which helps the user recognize exactly what happened and also learn additional phrases to message the bot. Both user modalities have equivalent capabilities, and interacting with either of them updates both the *Convey* window and the messaging window. Thus, the two modalities complement each other and can be used interchangeably.

STUDY DESIGN

In this section, we present the study design by describing the participants, the systems used and the study procedure.

Participants

Sixteen participants (11 male and 5 female, mean age = 32.5 years, sd = 7.4 years) were recruited for the study by emailing employees of a local IT company and snowball sampling. Fourteen of them had an engineering background, and the remaining two were from non-technical backgrounds (finance and social sciences). All participants held a Bachelor's or higher degree. Although none of the participants were native English speakers, all rated themselves fluent in English. Five of them reported using Facebook Messenger every hour of the day, while the rest reported using Messenger at least every four hours daily. All participants understood chatbots at a conceptual level, while five had prior experience interacting with chatbots on the Facebook Messenger platform. Two participants stated that "*proper context understanding*" was one of the major difficulties they faced while interacting with chatbots in the past.

System Description

For the user study, we developed a chatbot using IBM Conversation platform [3] with functionalities similar to an e-commerce chatbot for buying footwear. We used the shoes catalog data from jabong.com [1], an e-commerce website. The chatbot was designed to understand and filter shoes based on several features, including price, color, material, style, and brand, to help participants in their decision process. The user can click the shoe image to view a zoomed version of the image. Clicking on '*ADD TO BAG*' (Figures 1 & 2) results in placing an order for the shoe .

Procedure

We conducted a within-subject user study with two interfaces: default chatbot and the same default chatbot with the added *Convey* feature (also referred as *Convey* chatbot). The ordering of the interfaces was randomized across participants to counter ordering effects. With each interface, participants were required to perform one of these two tasks: (a) Select party footwear for yourself, and (b) Select a pair of sports shoes for the opposite gender. Half of the participants had to select a party footwear for themselves using the default chatbot and select sports shoes for the opposite gender using *Convey* chatbot, while the other half had to select a party footwear for themselves using *Convey* chatbot, followed by selecting sports shoes for the opposite gender using the default chatbot. For the tasks (a) and (b), the combined budget was 3000 INR (45 USD). To motivate the participants, the reward for participation was that a randomly-selected participant would receive

Data type	Default Bot	*Convey* Bot
Time taken	7.6±2.1	8.3±2.0
# of shoes viewed	71.5±21.5	77.6±22.7
# of shoes zoomed into	9.8±5.3	10.6±9.2
# of typed messages	9.8±4.4	6.3±3.7
# of words/message	33.9±13.3	26.1±11.2
# of chars/message	174.4±65.8	136.6±50.9
# of *Convey* interactions		7.8±6.5

Table 1: Results from Log Data, mean±std (bold with $p<0.05$)

his/her selected shoes as a free gift. No other rewards were given for participation.

At the start of interacting with each interface, a one-minute tutorial video (screen-cast with no audio) was played to showcase the capabilities of that interface. At the end, participants were asked to rate their experience on a 5-point Likert scale on several metrics, including ease of use, fun, and frustration [8] and also provide subjective feedback regarding the interface by typing their responses in an online form. After interacting with both the interfaces, participants were asked to compare the two interfaces, and specify which one they preferred and why. Every participant was asked to use their personal laptops/phones for the study with the URL provided by the study facilitator. Participants were not primed to use the *Convey* chatbot in any particular way. All input events were logged and saved on the server for analysis. The study took place in an IT office, and on an average, it took 45 minutes.

RESULTS

In general, participants enjoyed their experience interacting with a chatbot for buying shoes, as a majority of them (11) were interacting with a chatbot for the first time: "*It was a fun exercise... got to know how to use chatbots.*" - P_1, "*... was able to try lots of custom queries*" - P_{10}. Also, 9 participants liked that the chatbot was "*very responsive*" and "*prompt*". Seven participants stated that it was "*easy to use*", and five appreciated the "*enormous catalog*". Out of the 16 participants, 7 used their phone for the study, while remaining used their laptop/computer.

Log Data

On average, participants viewed 77.6±22.7 shoes with the *Convey* chatbot and zoomed into 10.6±9.2 of them before adding a shoe to the cart, while with the default chatbot, participants viewed 71.5±21.5 shoes and zoomed into 9.8±5.3 of them (Table 1). Participants took an average time of 8.3±2.0 mins to complete the task with the *Convey* chatbot, while with the default chatbot, they took 7.6±2.1 mins. This hints that the participants spent enough effort and time in shoe selection.

We conducted paired t-tests between the two interfaces on several parameters, including time taken to complete the task, total number of words input by the user, and total number of shoes browsed and zoomed into. Except for the total number of words input, we did not find any significant difference between the two interfaces. This might be attributed to the fact that the study task was not a performance-measurement task, rather it was a subjective decision-making task. It could also be due to the small sample size. As expected, participants

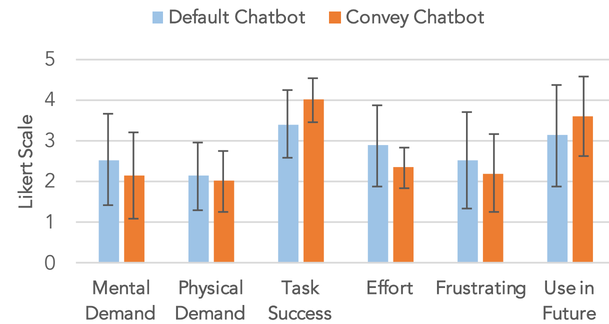

Figure 3: Likert-scale rating by the participants (with standard deviation shown as error bars)

typed significantly more text messages in the default chatbot interface (9.8±4.4 messages) compared to *Convey* chatbot interface (6.3±3.7 messages), with $t_{15}=1.9$, $p<0.05$. Instead of typing, participants interacted with the context values in *Convey*. Moreover, participants also typed longer messages with default chatbot (174.4±65.8 characters/message) compared to *Convey* chatbot (136.6±50.9 characters/message) with $p<0.05$, which was mostly attributed to the text messages for updating the price range. Overall, participants interacted with elements in *Convey* 124 times, using a combination of drop-down menus (67 times to select a brand, change gender, modify shoe type, *etc*.), range sliders (26 times to choose the price range), and button menus (21 times to select the shoe color). Also, the remove option on *Convey* was used 10 times.

Ratings

Participants rated both the interfaces on a 5-point Likert scale rating [8] (Figure 3). Note that for all metrics other than *Task Success* and *Use in Future*, a lower score is better. We conducted a paired t-test analysis and found the *Convey* chatbot to be significantly better than the default chatbot, with respect to perceived success in performing the task ($t_{15}=3.0$, $p=0.01$), and potentially using it in future ($t_{15}=3.1$, $p=0.01$). The *Convey* chatbot also outperformed the default chatbot in the effort required to achieve the participants' level of performance ($t_{15}=-2.4$, $p=0.05$) and mental demand of the task ($t_{15}=-2.3$, $p=0.05$). The ratings clearly show that participants preferred the *Convey* chatbot over the default chatbot.

Comparison

When asked to choose between the two chatbots for shopping in future, all 16 participants preferred chatbot with *Convey*. These positive comments about *Convey* summarize the participants' response: "*It* (*Convey* chatbot) *was more like a shopping experience, the other one was more like an exam!*" - P_{13}, "*it felt good interacting with this* (*Convey*) *chatbot.*" - P_4, and "*Well, just keep the GUI at the top, plz, it helps!*" - P_{12}.

Seven participants mentioned that the *Convey* chatbot was easier to use ("*easier to find products with different combinations" - P_7, "easier to narrow down products*" - P_5, "*very intuitive*" - P_{15}), and five participants stated that the *Convey* chatbot was faster than the default chatbot. The *Convey* chatbot was perceived to be faster as it "*saves typing effort*" - P_1,P_5, and helped in providing precise input, "*I gave less false inputs to* *the chatbots in case with top interactive part* (*Convey*)"- P $_{11}$. False inputs have been reported as a major barrier to adoption of conversational systems in general [14].

Interestingly, six participants liked the *Convey* chatbot because it showed context; "*it* (*Convey* chatbot) *can keep track of what we are searching currently*" - P$_3$, "*easy to see what are the choices made and edit them*" - P$_{15}$, "*maintains effective cumulative history*" - P$_6$, "*I was not clear on what filters were getting applied* (with default chatbot). *I had to go through the old chats to figure that out.*" - P$_{12}$. By showing context, *Convey* reduced confusion; "*It* (*Convey*) *showed what the bot understood so there weren't any misunderstandings... It* (*Convey* chatbot) *is WYSIWYG of chatbots!*" - P$_4$.

Five participants mentioned that *Convey* guided them by "*showing what options are available to choose from.*" - P$_8$. For brands and colors in particular, participants were not sure of the available options in the default chatbot interface. Only one participant asked for the options by typing "*show all available brands*", and received a list of brands in response, but this was not obvious to other participants. Two participants pointed at the efficiency of the price range slider, as entering the price range using text was "*almost impossible*" for them; it requires typing "*more than Rs 1000 and less than Rs 2000*". Compared to the default chatbot, one of the participants commented that he was "*easily able to remove preferences once selected* (with *Convey*)" - P$_7$. Also, three participants mentioned liking the fact that even with *Convey*, they can always use the default text mode, if needed. None of the participants complained that *Convey* took up space at the top of the chatbot, even with the limited screen space of a mobile device.

DISCUSSION AND DESIGN IMPLICATIONS

In our study, participants preferred the *Convey* chatbot, and found it to be easy to use, intuitive, less mentally demanding, and faster compared to the default chatbot. Interestingly, even though they interacted with the *Convey* interface for a short period of time, participants were cognizant of the benefits offered by *Convey* and appreciated them. Next, we briefly discuss design implications for chatbots, as derived from the positive comments by users while interacting with *Convey*

Summary and Persistent View: Participants perceived *Convey* as showing a summary of the conversation between the human and chatbot so far, which gets updated after every message turn. The default chatbot interface is non-persistent, *i.e.*, as the conversation proceeds, the text messages are eventually removed from the messaging window. Although user can always scroll up to view past messages, it quickly becomes cumbersome as conversation proceeds. A persistent summary of the conversation not only adds to the usability and but also helps ensure that the humans and the chatbot have the same mental model. This helps counter an important drawback of chatbots, as seen in earlier studies [12, 14, 15], that users lose track of the chatbot's contextual state.

Form-based UI: Most existing chatbots do not provide value over alternatives such as search engines, webpages and native mobile apps [12]. Current chatbots do not allow previous messages to be edited. In certain scenarios, such as flight booking, changing one of the parameters (*e.g.*, departure date) is easier on a website due to the form-based UI. *Convey*, in a way, combines the benefits of form-based UI with the flexibility of a text-based chat interface. However, unlike a typical form-based UI, a chatbot should not be dense, it should not show all the options available to the user all the time. Instead, showing only those options that pertain to the context explicitly mentioned by the user or assumed by the system, makes it easier to interact.

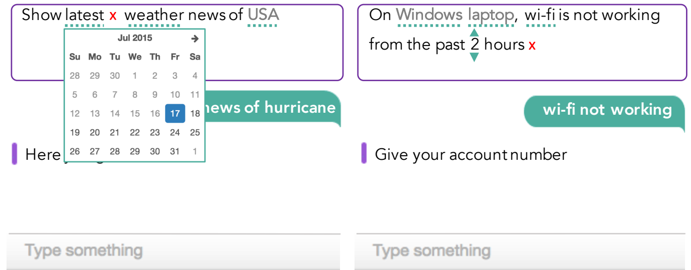

Figure 4: Other *Convey* use cases - Left: News; Right: Support

Precise Input: Participants enjoyed the fact that they could specify precise inputs with *Convey*, especially the price range selector. This is necessary at times, as text might be too cumbersome to type, resulting in lengthy chats to reach the desired outcome, and/or the chatbot might not be intelligent enough to understand complex input text. To elaborate, natural language input to chatbots is highly flexible as anything can be expressed. However, it has a low bit-rate since it requires time for users to type and intelligence for bots to understand. Future chatbots should combine natural language with standard UI elements to enhance the interaction medium between humans and computers in order to combine high flexibility with high bit-rate.

Finally, in this paper, we centered our study on a shopping bot. However, *Convey* can be adapted for chatbots in other domains such as IT support, travel booking, news, movie booking, *etc.*, (see Figure 4), as the concept of 'context' remains consistent across utility-driven conversational systems. The results of the study should also be generalizable to other domains as *Convey* can provide a way for precise input, along with providing a persistent view summarizing the conversation. Adding these capabilities to the current-day chatbots will help in making them more user-friendly and bridging the gap between user experiences and expectations [13, 14, 17].

Limitations

The shopping chatbot used for the study had limited capabilities, which participants pointed out. Participants suggested adding more items to the catalog (in particular, more brands), improving the understanding capability (NLP) of the chatbot, enabling viewing of multiple shoes in a carousel, adding images of the same shoe from different angles, enabling an option to maintain a list of shortlisted shoes, auto-correcting spelling mistakes, and providing user reviews and ratings. However, as the purpose of the study was to understand the usability of *Convey*, a chatbot without such advanced features sufficed. In fact, the capabilities of the chatbot used in this study is similar to most of the existing utility-driven chatbots.

CONCLUSION

In this work, we added a context view called *Convey* to the top of the chatbot interface to help users have an understanding of the mental-state of the chatbot during the conversation (helping users and chatbot be on the same page) while sustaining the familiarity of the text-based messaging interface. Moreover, *Convey* adds the benefits of a form-based user interface by enabling entry of precise input through the interactive elements. The results from a 16-participant user study demonstrated that participants perceived chatbot with *Convey* to be faster and easier to use. *Convey* is generalizable to chatbots in any domain, and in future, we expect *Convey* to be integrated and offered by many chatbot-hosting platforms.

REFERENCES

1. 2016. jabong. (2016). Retrieved Dec 1, 2017 from https://www.jabong.com/
2. 2017. Chatbots. (2017). Retrieved January 11, 2017 from https://chatbottle.co/
3. 2017. IBM Watson Conversation. (2017). Retrieved March 1, 2017 from https://www.ibm.com/watson/services/conversation/
4. Kathleen Chaykowski. 2016. More Than 11,000 Bots Are Now On Facebook Messenger. (2016). Retrieved December 28, 2016 from http://www.forbes.com/sites/kathleenchaykowski/2016/07/01/more-than-11000-bots-are-now-on-facebook-messenger/
5. O' Brien Chris. 2016. Facebook Messenger chief says platform's 34,000 chatbots are finally improving user experience. (2016). Retrieved February 7, 2017 from http://venturebeat.com/2016/11/11/facebook-messenger-chief-says-platforms-34000/-chatbots-are-finally-improving-user-experience/
6. Alan Dix, Tiziana Catarci, Benjamin Habegger, Yannis Ioannidis, Azrina Kamaruddin, Akrivi Katifori, Giorgos Lepouras, Antonella Poggi, and Devina Ramduny-Ellis. 2006. Intelligent Context-sensitive Interactions on Desktop and the Web. In *Proceedings of the International Workshop in Conjunction with AVI 2006 on Context in Advanced Interfaces (CAI '06)*. 23–27.
7. Craig Elimeliah. 2016. Why chatbots are replacing apps. (2016). Retrieved January 20, 2017 from http://venturebeat.com/2016/08/02/why-chatbots-are-replacing-apps/
8. Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988), 139–183.
9. Mariam Hassib, Daniel Buschek, Pawel W. Wozniak, and Florian Alt. 2017. HeartChat: Heart Rate Augmented Mobile Chat to Support Empathy and Awareness. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. 2239–2251.
10. BI Intelligence. 2016. Messaging apps are now bigger than social networks. (2016). Retrieved February 7, 2017 from http://www.businessinsider.com/the-messaging-app-report-2015-11?IR=T
11. Joonhwan Lee, Soojin Jun, Jodi Forlizzi, and Scott E. Hudson. 2006. Using Kinetic Typography to Convey Emotion in Text-based Interpersonal Communication. In *Proceedings of the 6th Conference on Designing Interactive Systems (DIS '06)*. 41–49.
12. Vera Q. Liao, Matthew Davis, Werner Geyer, Michael Muller, and N. Sadat Shami. 2016. What Can You Do?: Studying Social-Agent Orientation and Agent Proactive Interactions with an Agent for Employees. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems (DIS '16)*. 264–275.
13. Vera Q. Liao, Muhammed Masud Hussain, Praveen Chandar, Matthew Davis, Marco Crasso, Dakuo Wang, Michael Muller, Sadat N. Shami, and Werner Geyer. 2018. All Work and no Play? Conversations with a Question-and-Answer Chatbot in the Wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 13.
14. Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf Between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5286–5297.
15. Susan Robinson, David Traum, Midhun Ittycheriah, and Joe Henderer. 2008. What would you ask a conversational agent? Observations of Human-Agent dialogues in a museum setting. In *Language Resources and Evaluation Conference (LREC)*. Marrakech (Morocco).
16. Ronald Rosenfeld, Dan Olsen, and Alex Rudnicky. 2001. Universal Speech Interfaces. *Interactions* 8, 6 (2001), 34–44.
17. Indrani M Thies, Nandita Menon, Sneha Magapu, Manisha Subramony, and Jacki O'Neill. 2017. How do you want your chatbot? An exploratory Wizard-of-Oz study with young, urban Indians. In *Proceedings of the International Conference on Human-Computer Interaction (HCI) (INTERACT '17)*. IFIP, 20.
18. Joseph Weizenbaum. 1966. ELIZA - A computer program for the study of natural language communication between man and machine. *Commun. ACM* 9, 1 (1966), 36–45.
19. Tsung-Hsien Wen, Milica Gasic, Nikola Mrksic, Pei-hao Su, David Vandyke, and Steve J. Young. 2015. Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. *CoRR* abs/1508.01745 (2015).
20. Tom Yeh, Tsung-Hsiang Chang, Bo Xie, Greg Walsh, Ivan Watkins, Krist Wongsuphasawat, Man Huang, Larry S. Davis, and Benjamin B. Bederson. 2011. Creating Contextual Help for GUIs Using Screenshots. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. 145–154.