Informing the Design of a News Chatbot

Zhirun Zhang
Department of Computer Science
Hong Kong Baptist University
Hong Kong, China
zrzhang@comp.hkbu.edu.hk

Xinzhi Zhang
Department of Journalism
Hong Kong Baptist University
Hong Kong, China
xzzhang2@hkbu.edu.hk

Li Chen
Department of Computer Science
Hong Kong Baptist University
Hong Kong, China
lichen@comp.hkbu.edu.hk

ABSTRACT

Chatbots use conversational interfaces to simulate human communication and recently have been applied to different domains due to advancing techniques of natural language understanding and generation. In particular, in the domain of digital journalism, chatbots provide a new channel for audiences to engage with news. However, most chatbots operated by news organizations have so far failed to achieve business growth. In this paper, we have conducted a qualitative user study to better understand users' attitudes towards news chatbots. Specifically, 15 participants were asked to interact with several news chatbots implemented on Facebook Messenger by large international news organizations (e.g., ABC News, NBC News, and BBC News), by issuing a set of 22 sample questions covering various search and recommendation goals related to COVID-19. Then the participants expressed their expectations of news chatbots and the advantages/disadvantages perceived in using them. From these findings, we derive several design guidelines on effectiveness, informativeness, efficiency, humanization, and facility, suggesting developments to news chatbots that may better serve users' needs.

CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in interaction design**; **Empirical studies in HCI**.

KEYWORDS

News chatbots, qualitative user study, design guidelines, conversational agents

1 INTRODUCTION

Chatbots, also known as conversational agents, are a type of intelligent virtual agents that can converse with users in natural language via text, speech, or both. Many news organization implemented

chatbots to promote, disseminate, and recommend news via direct interactions with audiences in a conversational format. The effectiveness of implementing news chatbots for interacting with the audience, however, has remained mixed and under researched. Some scholars suggest that journalists should communicate with audiences to understand and meet their needs [31]; whereas counterarguments and evidence documenting failed implementations of chatbots in the newsrooms indicate that the logic of chatbot development is different with that of journalism [4]. The former is more experimental and focusing on efficiency and technical advancement, while the latter mainly focuses on the news values in the news production (i.e., news formats and the autonomy of human editors and journalists) [4]. Therefore, whether or not existing news chatbots can offer users the experience they desire still remains undetermined in the previous literature.

When existing studies on news chatbots have focused on interviewing journalists or editors, the current paper fills a gap by implementing a user study to yield more insights into users' perceptions and needs, as well as the limitations of existing chatbots. The current research has the following objectives:

(1) To understand users' opinions on news chatbots after they have interacted with news chatbots offered by mainstream news organizations;
(2) To summarize users' perspectives on the advantages and limitations of these current news chatbots;
(3) To develop and inform the design guidelines for news chatbots that can help serve users' needs.

A qualitative user study was conducted in June 2020. We focused on young adults because they are not only regarded as digital natives who are savvy in digital media technologies, but also because they are the main target users of the news chatbot [10, 19]. Fifteen young adults were recruited to interact via Facebook Messenger with operating, workable news chatbots implemented by several international large news organizations. Qualitative data were obtained from interviews with these users via an online questionnaire after they had used these news chatbots to perform a number of information searching and recommendation tasks. The results aim to inform the design guidelines of news chatbots that better meet users' needs and thus pave the way for chatbot developers and media practitioners to deploy news chatbots.

2 RELATED WORK

2.1 News Chatbots

Applying chatbots in the journalism domain is motivated by two impetuses: the idea of conversational journalism and the advancement of automatic journalism. Chatbots can disseminate and recommend news to audiences and increase their engagement [19].

The idea of conversational journalism [6, 26] argues that journalists are "real, flesh-and-blood people" [26] and audiences are social, friendly, informal. "Being conversational" suggests an informal tone of interactions between the news providers and users [10].

Another impetus of news chatbots is the development of conversational agents and natural language processing (NLP) techniques. The evolutionary techniques, such as word embedding, enable more accurate interpretations of user inputs [39]. Methods for information retrieval, such as knowledge graph, generate more elaborate responses [2]. In addition, more commercial and open-source platforms reduce the barrier of developing a chatbot [1].

Exploration of news chatbots by several large news organizations— such as BBC, the New York Times, and the Guardian—started in the mid-2010s [4, 19]. Since 2018, however, most news organizations either closed or stopped maintaining their chatbots because these initiatives encountered "lack of interest, functionality, or investment" [4]. Solving this puzzle requires a better understanding of users' needs when using the news chatbots.

2.2 User Experience of Chatbots

Conducting user studies may help chatbot designers better understand users [24, 40]. In the healthcare domain, the user feedback could reveal the effect of chatbots on assessing mental state [21], relieving mental issues [15], and promoting health behaviors [32]. In e-commence, chatbots facilitate user perception of social support and increase purchase behavior [7]. In the fashion domain, a study analyzed the conversation of stylist chatbots with users to understand the interaction pattern [37]. Another study focused on understanding user experiences of chatbots in several domains (e.g., travel, entertainment, and chit-chat), and identified that the chatbot interface is more suitable for domains that require interactive communication [17].

In the news domain, a study on user experiences of the news chatbot (developed by the Australian Broadcasting Corporation (ABC)) found that the news chatbot can help media and audiences form an intimate relationship [10]. ABC's news chatbot was regarded as an intimate friend of users that could provide personalized experience for them. It encouraged users to give responses and feedback and regarded users' feedback as "cultivating a closer relationship between the ABC and their audiences than is provided by other services on offer" [10]. For the users, they appreciated the "non-intimidating, non-threatening" [10] tone of the news chatbot.

When previous studies have shed light on user perception of news chatbots, they only relied on a single case (i.e., one media organization), and they did not summarize principles to inform a better design of news chatbots. In light of the above reviews, we focus on performing a qualitative user study in order to gain an in-depth understanding of users' perceived advantages and disadvantages of using news chatbots to acquire news information. Hence we seek to devise a set of design principles to support both developers and journalistic practitioners.

3 METHOD

3.1 Participants

We recruited 15 participants (13 females and 2 males) who were postgraduate students majored in digital media and technologies

Table 1: The five news chatbots used in the experiment

News Organization	Country	The URL of the Chatbot
ABC News	Australia	m.me/abcnews.au
BBC News	United Kingdom	m.me/bbcnews
La Jornada	Mexico	m.me/lajornadaonline
Le Figaro	France	m.me/lefigaro
NBC News	United States	m.me/nbcnews

in a public university in Hong Kong through purposive sampling. Compared to a random sample, they have knowledge of journalism and technology, enabling them to better observe the function of news chatbots and express their expectation more accurately. All participants are Chinese and fluent in English. The average age was 22 years old. The younger generation is regarded as the digital natives, active users of social media, and the major target audience of news chatbots [10, 19]. The selection of those participants is suitable for us to understand users' experience of news chatbots.

3.2 Materials: News Chatbots on Facebook Messenger

We selected the Facebook Messenger because it has been widely used by news organizations all over the world and also the most-downloaded apps in the 2010s [30]. In 2016, Facebook announced a bot platform for Facebook Messenger, which allows developers to build their own chatbots on Facebook Messenger and provide audiences "anything from automated subscription content like weather and traffic updates, to customized communications like receipts, shipping notifications" [27]. More than 300,000 chatbots have been developed on Facebook Messenger [5]. Most news organizations have also began to develop chatbots based on Facebook Messenger.

We select five chatbots (see Table 1) on Facebook Messenger (see some screenshots of ABC News chatbot in Figure 1), which have been discussed and analysed in the digital journalism literature as case studies [10, 19, 41]. These chatbots also represent news organizations in different media systems according to Hallin and Mancini's framework [13]. For example, the cases in U.S. and Australia represent the liberal model; the case in U.K. represents the liberal model but with a focus on public service journalism; the case in France represents the polarized pluralist model; and the case in Mexico represents the transitional democracy.

3.3 Procedure and Data Analysis

The user study was implemented in June 2020. To assure their familiarity with news chatbots, participants were asked to interact with the five news chatbots (in a random order) based on 22 sample questions (see Table 2). The questions are related to COVID-19, which is a recent news event involving a global public health emergency. We selected such a news event with a global impact and a global-level concern, rather than locally salient events, to ensure a better comparison among the different chatbots. These 22 questions were designed based on major information searching and recommendation goals from the previous works [20, 33]. With such theory-driven interaction rather than random chit-chats, the participants could have a comprehensive understanding of news chatbots and better evaluate their functionality. The pre-defined
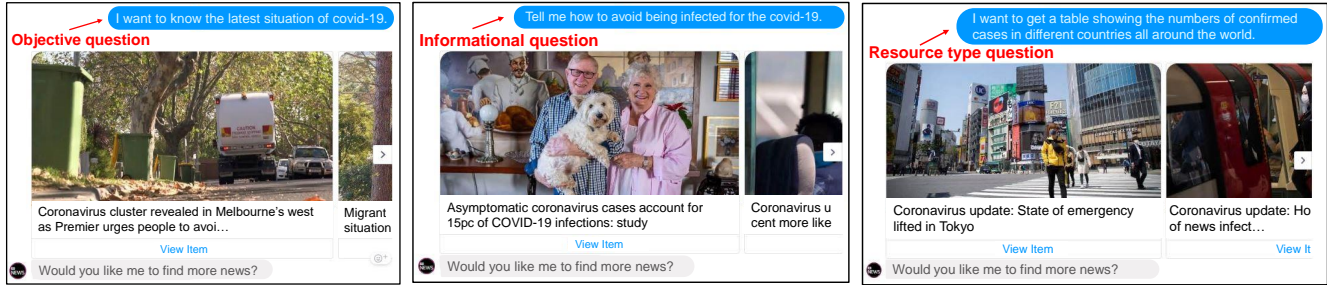
Figure 1: Screenshots of ABC News chatbot responding to three types of user questions (enlarged fonts to improve readability).

questions could also decrease the possible language mistakes made by non-native English speakers.

During the experiment, participants should ask all the 22 questions. If participants did not receive expected response, they could use keywords as input to avoid underestimating the chatbot because some chatbots may only have the search-and-retrieval function. To give participants enough time to experience the news chatbots, they could perform the task at any time and any place they wanted within a two-week period.

We have asked the participants to document the time they had used to interact with the chatbots. According to their self-reported duration, the participants in general spent 4 hours (M = 4.46, SD = 1.92). Once they completed the interaction with all the news chatbots, they were invited to complete an online questionnaire to share their experiences and expectations. The questions included:

Q1: *Among the news chatbot you have used, please indicate which one you most preferred and give an explanation as to why.*

Q2: *Please list three **advantages** and three **disadvantages** of using news chatbots to acquire news information, compared to using traditional news websites/papers.*

Q3: *Please express your **expectations** about what you think a good news chatbot should be, and then follow up with examples supporting your expectations.*

Q4: *Rate those 22 questions according to your own experience as to how likely you are going to ask them when using a news chatbot (on 5-point Likert scale).*

All participants submitted answers to the above four questions. After collecting their replies, the first author identified the related content and initially coded the content. The other two authors then discussed the results from the first author, refined the codes, and summarized them into categories. Finally, the first author reviewed the content based on the categories. In addition, we counted the number of participants and the percentage among all participants for each category.

4 RESULTS

4.1 User Preference for News Chatbot

For Question 1, most of participants (10 out of 15, 66.7%) preferred ABC News chatbot, 3 participants (20%) preferred NBC News, and 1 participant (6.7%) favored BBC News, while 1 participant did not indicate preference.

The preference for ABC News is mainly because it provided news relevant to their questions (indicated by 8 participants). For instance, "*It provided content related to my queries.*" (P2). "*ABC's*

chatbot could detect the key words in my sentences and phrases, and then search out the relevant news reports from its database for me.*" (P3). Another reason is attributed to its interactivity (indicated by 6 participants), e.g., "*Its interactive way is varying.*" (P13). "*I think it is the most interactive chatbot. Its reply is richer and more anthropomorphic [...]*" (P7). Some participants (5) were also impressed by its instant response, e.g., "*always fast*" (P1), "*ABC News's chatbot reply is instant.*" (P7).

The preference for NBC News was similarly owing to its quick response (indicated by 3 participants), provision of relevant news (2), or display of news in the form of pictures and videos (1). For example, "*It can reply immediately at any time [...]*" (P15). "*NBC recommends news in the form of pictures and videos, which is more attractive than simply giving the URL, and can see at first glance whether it is what I want to see.*" (P8).

As for BBC News, one user liked it most, because "*its chatbot is more intelligent than others.*" (P3). Some participants also expressed suggestions for further improving the news chatbots. Because the comments overlap with their feedback on the news chatbots' disadvantages, we discuss them in Section 4.3.

4.2 Advantages of News Chatbot

For Question 2, participants first listed the advantages of using news chatbots compared to traditional news websites or newspapers. Investigating the advantages of news chatbots from the users' perspective can better understand their motivation for using a news chatbot. We summarize their answers into six categories as follows.

4.2.1 Easy to Search for News. The first advantage of news chatbots is its ability of allowing users to easily search for news information. Ten participants (66.7%) believed that it is easier or more convenient to search for news using a chatbot than using traditional news websites and newspapers. For instance, P8 stated that a news chatbot is "*easy to search for news.*" P12 wrote "*[The news chatbot] makes searching for news easier; we can send a sentence to find the news we want. It can make the process of searching for news more intelligent, and perhaps the robots can better understand the news we want to search.*" P13 made a similar opinion: "*More convenient to find news*".

It is hence interesting to further identify what kind of information users tend to search/seek for. According to participants' ratings on how likely they are going to ask those 22 sample questions when using a news chatbot (in response to Question 4; see Table 2), we may see that participants are most likely to seek for news based on **objective** goals (e.g., "*I want to know the latest situation of Covid-19*"; mean rating = 3.90 out of 5). They are also likely to ask news

Table 2: Twenty-two sample user questions (derived from related papers about general web searching or obtaining recommendations in natural language [20, 33]) for participants to interact with the news chatbots and the result of how likely participants are going to ask these questions when using a news chatbot (on a 5-point Likert scale) in the post-task questionnaire.

Type	Goal	Sample Question	Mean	SD
Objective	News genre	1. *I want to know the latest situation of Covid-19.*	4.60	0.51
	Language	2. 你会说中文吗? *(Do you speak Chinese?)*	4.33	1.05
	Release date	3. *I want to know some news stories on Covid-19 on 12 Mar 2020.*	3.87	1.06
	Region	4. *I want to know some news stories on Covid-19 in [the capital city of the country involved].*	3.73	1.10
	Deep features of news	5. *I want to read an investigative report on the immigrants of Covid-19 prevention.*	3.47	1.30
	People in news	6. *I want to know how [national leader] discusses the Covid-19.*	3.40	1.30
			3.90	**1.14**
Informational	Advice	7. *Tell me how to avoid being infected for the Covid-19.*	4.33	1.05
	Undirected	8. *Tell me about the quarantine policy in [the largest city of the country]*	3.87	1.19
	Directed-closed	9. *What is the current quarantine policy in [the largest city of the country]?*	3.80	1.15
	Locate	10. *Tell me where to get masks!*	3.73	1.28
	List	11. *Show me the list of restaurants still open.*	3.53	1.46
	Directed-open	12. *Why the [the largest city of the country] would have such a crazy current quarantine policy?*	3.40	1.18
			3.78	**1.22**
Subjective	Emotion	13. *Do you have some good news on Covid-19 to share?*	3.67	1.40
	Quality	14. *Do you have some credible news stories on Covid-19 to share with me?*	3.53	1.36
	Relationship to another news (sports)	15. *Do you have some news on Covid-19 that is related to the Tokyo Olympics?*	3.47	1.06
	Relationship to another news (politics)	16. *Do you have some news on Covid-19 that is related to the 2020 US Election?*	3.33	1.18
	Relationship to another news (finance)	17. *Do you have some news on Covid-19 that is related to the Dow Jones Index?*	2.93	1.28
			3.39	**1.25**
Resource	Obtain	18. *I want to get a table showing the numbers of confirmed cases in different countries all around the world.*	4.07	0.88
	Interact	19. *Can I contact the news editor of this news story?*	3.67	1.29
	Download	20. *I want to download the podcast audio files.*	2.93	0.96
	Entertainment	21. *Do you have some music to play?*	2.87	1.36
			3.38	**1.22**
Navigational		22. *I want to go to the webpage of the New York Times.*	3.00	1.31

chatbots some **informational** questions (e.g., "*Tell me how to avoid being infected for the Covid-19*"; mean rating = 3.78). Relatively, the likelihood of asking *subjective* and *resource* type questions is low (mean ratings are 3.39 and 3.38 respectively), except a high demand for **obtaining resource** under "Resource" type (i.e., "*I want to get a table showing the numbers of confirmed cases in different countries all around the world*"; mean rating = 4.07). The chance of asking *navigational* question is the lowest (e.g., "*I want to go to the webpage of the New York Times*"; mean rating = 3.00).

4.2.2 Interactivity. Ten participants (66.7%) mentioned the advantage of providing two-way interaction between users and news chatbots. These participants believed that the interactive form makes news chatbots more "*fun*" (P15), "*highly interactive*" (P11), or different from the "*boring one-way communication in the traditional media*" (P14).

4.2.3 Personalized Content. Seven participants (46.7%) praised the personalized content offered by news chatbots. For example, P1 believed that the news provided by news chatbots is "*based on my interest.*" P4 also noted that the content of news is "*more targeted and personalized.*" P11 stated that news chatbots "*give the feeling of a private customized service.*"

4.2.4 Quick Response. Six participants (40%) mentioned the chatbot's quick responses to their questions. It can help improve the efficiency of news organizations when responding to user requests. For example, P1 and P9 used "*fast*" to describe news chatbots. P2 regarded the "*instant response*" as an advantage. P11 said that "*query service is fast and can quickly solve readers' problems.*"

4.2.5 Latest News. Three participants (20%) believed that news chatbot can provide the latest news. P1 wrote the word "*latest*" as an advantage. P6 said: "*As the content is updated, users can see the most cutting-edge news.*" P10 also believed that "*users can know the recent events in time.*"

4.2.6 Humanized Response. Three participants (20%) thought the reason why news chatbots are interesting to users is because they have human-like features, e.g., "*feel like chatting with a person*" (P3), "*[...] users have the experience of talking to real people*" (P15), or making the interaction process "*more fun*" (P6).

4.3 Disadvantages of News Chatbot

When being asked the disadvantages of using news chatbots compared with traditional news websites or newspapers, participants mentioned several *addressable* limitations based on their experience of using the existing immature chatbots.

4.3.1 Limited Effectiveness. Most of the participants (12 out of 15, 80%) were dissatisfied with the effectiveness of news chatbots they used. Specifically, four (26.7%) complained about their limited understanding, e.g., "*Sometimes they can't understand what you're talking about, and it takes a lot of effort to explain.*" (P12). Nine (60%) were dissatisfied with the inaccurate responses, e.g., "*So many responses cannot meet my needs. Ineffective communication wastes a lot of time.*" (P3). "*The robot's algorithm is not yet mature, and sometimes the desired response cannot be obtained accurately.*" (P14).

4.3.2 Limited Information. The limited news information, or information cocoons, was mentioned by nine participants (60%). For example, "*It still provides very little reports, which may lead to information cocoons.*" (P3). "*Perhaps the content of the news prevented users from getting other related content.*" (P15). A participant discussed the balance between the diversity and relevance of news, i.e., "*When users search actively, they can only see the topics they are interested in, and they will have a one-sided understanding of the news. If chatbot pushes [too much] other news, it will disturb the user's life and make the user feel harassed. The balance between the two needs to be examined.*" (P6).

4.3.3 Limited Efficiency. Six participants (40%) pointed out the slow response of some chatbots in their experience, e.g., "*slow response sometimes*" (P9). A participant mentioned the possible consequence of slow response: "*If the response speed is not timely, it is easy for users to lose patience.*" (P10).

4.3.4 Lack of Humanness. Three participants (20%) complained that the news chatbots they used lacked humanness, which can be interpreted as the lack of human characteristics in the chatbot's responses, such as the "*lack of emotion*" (P2) and the "*lack of human touch*" (P5).

4.3.5 Inadequate Facility. Three participants (20%) expressed dissatisfaction with the facilities of the current news chatbots, whose inadequacies include limited filtering support, no multi-lingual support, and no note-taking function. E.g, "*When we search for news in a chatbots conversation, we have no way of knowing its criteria for filtering news. On traditional news sites, we enter keywords and then we can filter the content we are interested in ourselves.*" (P11).

These disadvantages might possibly be addressed in the implementation. However, some intrinsic disadvantages of news chatbots were also indicated by participants. For example, one participant pointed out that it is hard for users to receive a deeper understanding of news contents from a conversation with a news chatbot, i.e., "*Readers tend to lose their focus in conversations with chatbots, making it difficult to engage in deeper news reading in conversations.*" (P11). Another participant mentioned the "*need for internet and electricity*" (P1), in comparison to reading newspapers.

4.4 User Expectations of News Chatbot

From participants' replies to Question 3 (see Section 3.3), we summarize the following five major types of expectations.

4.4.1 Accuracy. Almost all participants (14 out of 15, 93.3%) mentioned accuracy, which implies two aspects (or steps): **understanding** and **relevance**. Five participants (33.3%) explicitly discussed the importance of understanding (i.e., understanding users' inputs).

For example, "*First, a good news chatbot should understand well what I said and what I really want, then it could provide me with the reports which meet my needs. This is the most important thing for a news chatbot.*" (P3). "*I hope a good news chatbot can understand my questions accurately [...].*" (P8).

In the examples given by participants to support this expectation, some users preferred ABC News because "*compared to other chatbots that fail to recognize user questions, it can recognize the keyword in my query and understand roughly.*" (P3). Another participant mentioned NBC News because "*in most cases, it can understand the keywords in my question and give recommended news.*" (P8). It can thus be seen that understanding capability is a precondition for the chatbot to provide relevant information to users.

As for another aspect of accuracy, i.e., *relevance*, 12 participants (80.0%) mentioned its importance, e.g., "*provide effective and targeted answers rather than broad and unhelpful information*" (P2). "*[...] in response to user question about Trump's remarks about Covid-19, the final result should include both Trump and Covid-19, rather than just including one of them, which can reflect the chatbot's accuracy and intelligence.*" (P15). As supporting examples, one participant mentioned BBC News, "*the answers given by BBC News' chatbot are more abundant, and many questions can be answered accurately.*" (P12). Another participant preferred ABC News because "*most answers are related to the questions.*" (P4). It hence reveals the significance of returning relevant responses to user requests from news chatbots.

4.4.2 Quick Response. The ability of quickly responding to user requests was mentioned by eight participants (53.3%), e.g., "*Reply immediately after the user sends a message, even if it is an automatic reply with an exchange button set in advance.*" (P14). The importance of quick response is also reflected in their preferences for a particular news chatbot. For example, one participant liked NBC News because "*it responds very quickly. Even if it is an automatic reply, it also makes me feel that my inquiry is echoed.*" (P8). Some participants preferred ABC News also because of its fast response, e.g., "*the answers are always fast.*" (P1). "*It responds the most instantly, regardless of whether it matches what the user wants to find.*" (P14).

4.4.3 Interactivity. Eleven participants (73.3%) expected the interactivity of news chatbots on conversational content and presentation, which can be concretely reflected on the chatbot's ***flexibility*** (or called *responsiveness*) and ***humanization***. Flexibility refers to providing flexible responses in different situations, as mentioned by nine participants (60%). For instance, "*Flexible. Even if it can't answer the question, it can provide other options.*" (P4). "*It is able to give some explanation and elaboration of the questions raised, rather than providing me with template-based answers aimed at different questions.*" (P12). Still, this expectation was supported by their experience of using a specific news chatbot such as ABC News, i.e., "*If the chatbot can't recognize the question that users asks, it will ask the users if they want to search for some news, and also give some feedback.*" (P9). "*Each time ABC News chatbot gives me an answer to my query, it continues to ask me if I need more information.*" (P11).

In addition, five participants (33.3%) expected for humanization of a news chatbot, i.e., expecting the chatbot to act like a human. For instance, "*Humanized. I think the language it used is also an important issue. First, I hope it could speak like an ordinary person. Second, I also hope it can reply with more diverse sentences.*" (P3). "*I*

think interesting speaking style will not make users get bored when they interact with a chatbot. For example, chatbot can make some jokes when users ask them question." (P9). One participant even mentioned humanization as her/his major reason behind preferring ABC News, i.e., "*It is the most interactive chatbot; its reply is richer and more anthropomorphic.*" (P7). It can thus be seen that conversational style, variety of response, and sense of humor can make users perceive the chatbot as more human.

4.4.4 Informativeness. In our study, several participants expressed expectation of informativeness on news chatbots. Concretely, six participants (40%) expected that news chatbots can provide smart recommendations, e.g., "*Smart recommendation function. Chatbot can provide users with more information according to the question they ask.*" (P9). One participant stated that "*Although ABC News chatbot is not completely problem-solving, compared to other news chatbots, its recommended news matches my query well.*" (P7).

Moreover, three participants stressed the importance of news chatbots responding with concise information, e.g., "*concise reply*" (P10) and "*clear expression*" (P11). They expected news chatbots to have a clear and brief response, e.g., "*I will perceive the news organization to be more professional if its news chatbot will produce less meaningless messages.*" (P10).

4.4.5 Other Functions. Four participants (26.6%) expected other functions to be provided in a news chatbot. To be specific, a participant proposed a "*fact-check service*" (P11), which can verify whether a returned news is fake or real. Another participant suggested to allow users to rate the news by "*choosing to like or dislike, or agree and disagree. At the same time, I hope to see other people's feeling about the news.*" (P8). Yet another participant suggested that the news chatbot could "*guide readers to human service if it cannot handle readers' queries*" (P2). One participant suggested that the news chatbots can send real-time updates and reminders, e.g., about "*weather, traffic, or something else*" (P13).

5 DISCUSSION

The responses of our participants to this qualitative user study pointed out several aspects to improve news chatbots. In the following, we conclude five major design guidelines.

5.1 Providing *relevant* news information

The capability of providing relevant news information reflects the news chatbot's **effectiveness**, which means that it can understand users' aims and provide appropriate feedback. In the literature, some researchers also used the term "intelligence," "smart," or "wit" to define a chatbot's effectiveness [16, 19, 29]. In our study, the majority of participants expressed the limitation of existing news chatbots on effectiveness. One possible reason might be due to the limited technologies of chatbots they used [19]. Consistent with this limitation, participants overwhelmingly mentioned *accuracy* in the expectation of news chatbots, including understanding their requests and responding to them accurately. These results indicate that effectiveness should be imperative for news chatbots in user expectation. Previous research also indicated the expectation of understanding on general chatbots [17, 36] and the importance of providing relevant and accurate answers [40]. A previous study

found that understanding and answering questions is the essential precondition for the widespread use of chatbots, which may substitute for other traditional channels [41]. To achieve this goal, we believe that news chatbots should accommodate more advanced natural language understanding and generation techniques (e.g., Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) models, and Transformer) [14, 28, 38], so as to more effectively respond to user questions especially *objective* and *informational* type questions that users are more likely to ask when they use a news chatbot as reflected in our findings.

In addition, the utility of responses delivered news chatbots could be improved, especially in the case that they cannot fully recognize and understand users' requests. As suggested by our participants, when no suitable results are available, the chatbot may provide other options for users. Previous research recommended admitting the problem and clarifying the capability of the chatbot [24, 35]. We hence suggest that the news chatbot can first express its capability limitation and then provide some options to users, such as guiding users to some specific news topics.

5.2 Providing *diverse* and *up-to-date* news information

Compared with other types of chatbots, news chatbots mainly aim to deliver news that is essentially a kind of information. A prior study showed that *informing* is the primary goal of news [22], and an investigation of the situation of existing news bots on Twitter also indicated that *informing* is their foremost function [23]. Therefore, we believe that **informativeness** should be an important aspect of improvement to news chatbots. Based on the feedback of our participants, the existing news chatbots are limited in terms of their information diversity, which may cause the problem of information cocoons [12]. A previous study on general chatbots expressed similar concern on the risk of biased personalized information [41]. Another study proposed a chatbot that stimulates users to receive diverse sources of news for developing their critical thinking [9]. These observations suggest that information diversity of news content should be considered in the design process.

Moreover, some participants mentioned the **timeliness** of information as they thought the news provided by a news chatbot should be more timely and related to recent events. Related work also discussed the importance and necessity of information timeliness for news chatbots [17, 25]. We hence believe that providing up-to-date information should be important to a news chatbot development.

5.3 Responding to user requests *quickly*

The **efficiency** of a news chatbot is mainly about how quickly it could respond to user requests. In our study, participants strongly expressed their expectation of a quick response. They also pointed out that slow response may cause poor experience and decrease their intention of using the news chatbot. Consistent with the feedback, a study reported the slow responses of current newsbots [18]. Another study regarded time- and cost-savings as two advantages of chatbots to engage users [41]. In addition, the results are in agreement with a case study of a news chatbot during COVID-19, which found that timely information is significant for the public

[25]. Therefore, we think news chatbots should better respond to user requests in an efficient manner.

5.4 Responding *like a human*

Previous studies on general chatbots have paid a lot of attention to the aspect of humanization [10, 19, 29, 36]. **Humanization** means that a chatbot is given human characteristics like personality and emotion. In our study, some participants expressed dissatisfaction with the inadequacy of humanization of their used news chatbots, and expectation of the chatbot to act like a person. Several studies also discussed what personality of a chatbot users may like. For example, a study indicated that a user can have different preferences for chatbot personality in different situations [29]. Another study showed that humanization is a challenge for news chatbots because it may conflict with the authoritative nature of news delivery [10]. It was believed that messages from news chatbots should be personal, but they still questioned the conversational style and tone of a news chatbot [10]. Overall, these studies acknowledge the importance of humanization and provide a direction for in-depth exploration of users' perceptions of humanization particularly in news chatbots.

To achieve humanization, our participants suggested that the news chatbot can adopt multiple forms to answer user questions, e.g., providing various responses to the same request. This expectation is consistent with a recent study on BBC News where the BBC Mundo bot uses different expressions to increase its humanness [19]. Another suggestion by our participants is to increase the chatbot's response humor, which was also recommended in other work in relation to improving users' impression of chatbots [17, 29]. However, some studies [24, 35] found that humor may raise users' expectations of chatbots and in consequence make them more likely to be disappointed. Therefore, we suggest that humor should be carefully used, and be better considered when the news chatbot reaches a certain level of accuracy and intelligence.

5.5 Including other functions

Our participants proposed to add new functions on top of the news chatbots they used. For instance, a participant suggested having a "fact-check service" to verify the facts of the returned news. Some participants suggested that the chatbots can allow users to provide feedback (e.g., like/dislike) on the news, which was also discussed in a related work that reported ABC News chatbot collects user feedback via thumbs up/down [10]. As for the multilingual support mentioned by our participants, a related study pointed out that language barriers often lead to low performance in chatbots [40]. For the news organizations serving multilingual audiences, providing language support should be considered. Furthermore, one participant expected news chatbots to guide users to a human service when they cannot handle the user's request. A previous study concurred that chatbots need a customer consultant when they cannot understand or answer most of the user's questions [41]. Therefore, we believe that news chatbots could better provide a channel for users to access the human service. Overall, these suggestions provide some considerations for enriching the facilities of current news chatbots in the design process, but how to implement them still needs to be further explored.

6 LIMITATIONS AND FUTURE WORK

The present study has several limitations to be addressed in the future studies. First, we acknowledged the limitation of homogeneous sample and small number of participants. Participants were mainly from media and communication faculties and lack of gender balance because there are more female students in Journalism [3, 34]. We will conduct a more gender-balanced design with more participants from different backgrounds in the future studies. Second, we only focused on the Facebook Messenger as the deployment platform. Future studies can compare the news chatbots operated on different platforms, since the cross-platform comparison may offer more design insights. Third, although English is the global language and a common language shared by the present study's focused news organizations and the users, some news chatbots may be less advanced in English but more compatible in their native language. It would hence be interesting to involve them in our studies to identify influences potentially caused by language difference.

Fourth, the event topic for sample questions is limited to COVID-19. The merit of such choice is to control the topic and make the results more comparable, but in the real-world settings, people may interact with the chatbots with different topics. Fifth, the current study was primarily through qualitative survey to acquire user feedback and comments. Future studies can implement other experimental approaches, such as ethnography observations [11] to understand how news chatbots are embedded in the dynamic communication process in different social settings, and longitudinal studies [8] to in-depth understand users' changing needs over time.

7 CONCLUSION

When news organizations are attempting to deploy news chatbots as a type of intelligent virtual agents to interact with their audiences in an interactive manner, less is known about how users are perceiving the advantages and disadvantages of using news chatbots, and their expectations of a more effective news chatbot. The present study implements a qualitative user study of news chatbots on Facabook Messenger operated by several international news organizations (such as ABC News, NBC News, and BBC News). We found that news chatbots could have several advantages, such as allowing users to easily search for news information, and providing interactive interface and personalized content to engage users. However, the existing news chatbots are still limited in terms of their effectiveness, informativeness, efficiency, humanization, and facility. The participants also expressed their expectations for high level of accuracy, quick response, flexibility, and informativeness of a news chatbot. Based on the results, we finally conclude with a set of five guidelines that may help enhance the design of news chatbots to better serve both users' needs and journalistic purpose.

REFERENCES

[1] Eleni Adamopoulou and Lefteris Moussiades. 2020. *An Overview of Chatbot Technology*. Vol. 584 IFIP. Springer International Publishing. 373–383 pages. https://doi.org/10.1007/978-3-030-49186-4_31

[2] Ram G. Athreya, Axel-Cyrille Ngonga Ngomo, and Ricardo Usbeck. 2018. Enhancing Community Interactions with Data-Driven Chatbots–The DBpedia Chatbot. In *Companion Proceedings of the The Web Conference 2018* (Lyon, France) *(WWW '18)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 143–146. https://doi.org/10.1145/3184558.3186964

[3] Lyn Barnes. 2017. An inexplicable gap: Journalism and gender in New Zealand. *Journalism* 18, 6 (2017), 736–753. https://doi.org/10.1177/1464884915620231

[4] Valerie Belair-Gagnon, Seth C Lewis, and Colin Agur. 2020. Failure to Launch: Competing Institutional Logics, Intrapreneurship, and the Case of Chatbots. *Journal of Computer-Mediated Communication* 25, 4 (jul 2020), 291–306. https://doi.org/10.1093/jcmc/zmaa008

[5] Facebook Business. 2018. *F8 2018: David Marcus' Keynote.* Retrieved January 18, 2021 from https://www.facebook.com/business/news/david-marcus-f8-keynote-2018

[6] James W. Carey. 2019. The press and the public discourse. *Kettering Review (winter)* (2019), 9–22.

[7] Veena Chattaraman, Wi-Suk Kwon, and Juan E. Gilbert. 2012. Virtual agents in retail web sites: Benefits of simulated social interaction for older users. *Computers in Human Behavior* 28, 6 (nov 2012), 2055–2066. https://doi.org/10.1016/j.chb.2012.06.009

[8] Deborah Compeau, Christopher A. Higgins, and Sid Huff. 1999. Social Cognitive Theory and Individual Reactions to Computing Technology: A Longitudinal Study. *MIS quarterly* 23, 2 (1999), 145–158. https://doi.org/10.2307/249749

[9] Mateusz Dubiel, Alessandra Cervone, and Giuseppe Riccardi. 2019. Inquisitive Mind: A Conversational News Companion. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 6, 3 pages. https://doi.org/10.1145/3342775.3342802

[10] Heather Ford and Jonathon Hutchinson. 2019. Newsbots That Mediate Journalist and Audience Relationships. *Digital Journalism* 7, 8 (2019), 1013–1031. https://doi.org/10.1080/21670811.2019.1626752

[11] Giampietro Gobo. 2008. *Doing ethnography.* SAGE, Los Angeles, Calif. ; London.

[12] Cèdric Gossart. 2014. Can digital technologies threaten democracy by creating information cocoons? In *Transforming Politics and Policy in the Digital Age*, Jonathan Bishop (Ed.). IGI Global, Chapter 10, 145–154. https://doi.org/10.4018/978-1-4666-6038-0.ch010

[13] Daniel C. Hallin and Paolo Mancini. 2004. *Comparing Media Systems: Three Models of Media and Politics.* Cambridge University Press. https://doi.org/10.1017/CBO9780511790867

[14] Shafquat Hussain, Omid Ameri Sianaki, and Nedal Ababneh. 2019. A Survey on Conversational Agents/Chatbots Classification and Design Techniques. In *Advances in Intelligent Systems and Computing.* Vol. 927. Springer International Publishing, 946–956. https://doi.org/10.1007/978-3-030-15035-8_93

[15] Becky Inkster, Shubhankar Sarda, and Vinod Subramanian. 2018. An Empathy-Driven, Conversational Artificial Intelligence Agent (Wysa) for Digital Mental Well-Being: Real-World Data Evaluation Mixed-Methods Study. *JMIR mHealth and uHealth* 6, 11 (nov 2018), e12106. https://doi.org/10.2196/12106

[16] Mohit Jain, Ramachandra Kota, Pratyush Kumar, and Shwetak Patel. 2018. Convey: Exploring the use of a context view for chatbots. In *Conference on Human Factors in Computing Systems - Proceedings*, Vol. 2018-April. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3173574.3174042

[17] Mohit Jain, Pratyush Kumar, Ramachandra Kota, and Shwetak N. Patel. 2018. Evaluating and Informing the Design of Chatbots. In *Proceedings of the 2018 Designing Interactive Systems Conference* (Hong Kong, China) *(DIS '18)*. Association for Computing Machinery, New York, NY, USA, 895–906. https://doi.org/10.1145/3196709.3196735

[18] Al Johri, Sam Eui-Hong, and Mehta Dhrumil. 2016. Domain Specific Newsbots - Live Automated Reporting Systems involving Natural Language Communication. *Standord Computation Journalism Symposium* (2016), 1–5.

[19] Bronwyn Jones and Rhianne Jones. 2019. Public Service Chatbots: Automating Conversation with BBC News. *Digital Journalism* 7, 8 (2019), 1032–1053. https://doi.org/10.1080/21670811.2019.1609371 arXiv:https://doi.org/10.1080/21670811.2019.1609371

[20] Jie Kang, Kyle Condiff, Shuo Chang, Joseph A. Konstan, Loren Terveen, and F. Maxwell Harper. 2017. Understanding How People Use Natural Language to Ask for Recommendations. In *Proceedings of the Eleventh ACM Conference on Recommender Systems* (Como, Italy) *(RecSys '17)*. Association for Computing Machinery, New York, NY, USA, 229–237. https://doi.org/10.1145/3109859.3109873

[21] Masamune Kawasaki, Naomi Yamashita, Yi-Chieh Lee, and Kayoko Nohara. 2020. Assessing Users' Mental Status from Their Journaling Behavior through Chatbots. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (Virtual Event, Scotland, UK) *(IVA '20)*. Association for Computing Machinery, New York, NY, USA, Article 32, 8 pages. https://doi.org/10.1145/3383652.3423870

[22] Bill Kovach and Tom Rosenstiel. 2007. *The Elements of Journalism: What Newspeople Should Know and the Public Should Expect.* Three Rivers Press.

[23] Tetyana Lokot and Nicholas Diakopoulos. 2016. News Bots: Automating news and information dissemination on Twitter. *Digital Journalism* 4, 6 (2016), 682–699.

https://doi.org/10.1080/21670811.2015.1081822

[24] Ewa Luger and Abigail Sellen. 2016. "Like having a really bad PA": The gulf between user expectation and experience of conversational agents. *Conference on Human Factors in Computing Systems - Proceedings* (2016), 5286–5297. https://doi.org/10.1145/2858036.2858288

[25] Theodora A. Maniou and Andreas Veglis. 2020. Employing a chatbot for news dissemination during crisis: Design, implementation and evaluation. *Future Internet* 12, 12 (2020), 1–14. https://doi.org/10.3390/FI12070109

[26] Doreen Marchionni. 2015. Online story commenting: An experimental test of conversational journalism and trust. *Journalism Practice* 9, 2 (mar 2015), 230–249. https://doi.org/10.1080/17512786.2014.938943

[27] David Marcus. 2016. *Messenger Platform at F8.* Retrieved February 25, 2021 from https://messengernews.fb.com/2016/04/12/messenger-platform-at-f8/

[28] Wari Maroengsit, Thanarath Piyakulpinyo, Korawat Phonyiam, Suporn Pongnumkul, Pimwadee Chaovalit, and Thanaruk Theeramunkong. 2019. A Survey on Evaluation Methods for Chatbots. In *Proceedings of the 2019 7th International Conference on Information and Education Technology - ICIET 2019*, Vol. 29. ACM Press, New York, New York, USA, 111–119. https://doi.org/10.1145/3323771.3323824

[29] Indrani Medhi 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. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10513 LNCS (2017), 441–459. https://doi.org/10.1007/978-3-319-67744-6_28

[30] Chance Miller. 2019. *These Were the Most-Downloaded Apps and Games of the Decade.* Retrieved January 18, 2021 from https://9to5mac.com/2019/12/16/apps-and-games-of-the-decade/

[31] Jacob L. Nelson. 2019. The next media regime: The pursuit of 'audience engagement' in journalism. *Journalism* (jul 2019), 1–18. https://doi.org/10.1177/1464884919862375

[32] Stefan Olafsson, Teresa K. O'Leary, and Timothy W. Bickmore. 2020. Motivating Health Behavior Change with Humorous Virtual Agents. In *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents* (Virtual Event, Scotland, UK) *(IVA '20)*. Association for Computing Machinery, New York, NY, USA, Article 42, 8 pages. https://doi.org/10.1145/3383652.3423915

[33] Daniel E. Rose and Danny Levinson. 2004. Understanding User Goals in Web Search. In *Proceedings of the 13th International Conference on World Wide Web* (New York, NY, USA) *(WWW '04)*. Association for Computing Machinery, New York, NY, USA, 13–19. https://doi.org/10.1145/988672.988675

[34] Cynthia Ross, Karen; Carter. 2011. Women and news: A long and winding road. *Media, Culture Society* 33, 8 (2011), 1148–1165. https://doi.org/10.1177/0163443711418272

[35] Nina Svenningsson and Montathar Faraon. 2019. Artificial Intelligence in Conversational Agents: A Study of Factors Related to Perceived Humanness in Chatbots. *Dl.Acm.Org* (2019), 151–161. https://doi.org/10.1145/3375959.3375973

[36] Ekaterina Svikhnushina and Pearl Pu. 2020. Social and Emotional Etiquette of Chatbots: A Qualitative Approach to Understanding User Needs and Expectations. (2020). arXiv:2006.13883 http://arxiv.org/abs/2006.13883

[37] Kristen Vaccaro, Tanvi Agarwalla, Sunaya Shivakumar, and Ranjitha Kumar. 2018. Designing the Future of Personal Fashion. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–11. https://doi.org/10.1145/3173574.3174201

[38] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), Vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

[39] Anbang Xu, Zhe Liu, Yufan Guo, Vibha Sinha, and Rama Akkiraju. 2017. A New Chatbot for Customer Service on Social Media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3506–3510. https://doi.org/10.1145/3025453.3025496

[40] Jennifer Zamora. 2017. I'm Sorry, Dave, I'm Afraid I Can't Do That. In *Proceedings of the 5th International Conference on Human Agent Interaction*. ACM, New York, NY, USA, 253–260. https://doi.org/10.1145/3125739.3125766

[41] Darius Zumstein and Sophie Hundertmark. 2018. Chatbots - An Interactive Technology for Personalized Communication, Transactions and Services. *IADIS International Journal on WWW/Internet* 15 (2018), 96–109.