The Role of Dialogue User Data in the Information Interaction Design of Conversational Systems

Heloisa Candello[(✉)] and Claudio Pinhanez[(✉)]

IBM Research, São Paulo, Brazil
{heloisacandello, csantosp}@br.ibm.com

Abstract. Designers face several challenges when designing information for conversational systems. In this paper, we discuss those challenges in the context of Information Interaction Design and the role of conversational data to address them. Using an actual study performed prior to the development of a financial chatbots adviser, we identified a set of common issues which led to 18 design recommendations. We categorize those recommendations according to the disciplines they are related (Information Design, Interaction Design, and Sensorial Design). The guidelines were employed by the actual developers of the system, simplifying considerably the development, and show the importance of actual user conversational data in the design process of conversational systems.

Keywords: Conversational interfaces · Dialogue systems · Chatbots
User experience

1 Introduction

The rapid progress of natural language processing in the recent years, fueled by notable advances in Artificial Intelligence and Neural Networks, made the construction and deployment of conversational systems and machines not only possible, but a reality on the fingertip of any smartphone user. However, understanding how to design compelling and useful experiences for such systems is still in its infancy, with very limited theoretical frameworks to support it. In this paper, we explore some of those design issues under the perspective of Information Design.

Information design investigates the organization and presentation of data to result in meaningful information. Shedroff [19] claims that while Information Design primarily focuses on the representation of data and its presentation, the emphasis in Interaction Design is on the creation of compelling experiences. Those experiences are shaped by people perceptions in context, making the sensorial design also a valuable topic of tailoring people's communication. *Information Interaction Design* is the intersection of the disciplines of Information Design, Interaction Design and Sensorial Design [19]. Those three disciplines are essential to design compelling user experiences with conversational systems. We expect that conversational interface design also may benefit from the intersection of those disciplines.

We define conversational interfaces as computational systems which interact with humans through dialogues. Such machines or systems are known by various names such as virtual personal assistants, intelligent assistants, chatbots, and cognitive advisers. There are two primary types of conversational interfaces, the ones where the user input is through audio and the ones where it is via text.

In this work, we focus on the challenges of designing interactive information for text-based conversational systems. First, we show technical challenges and unveil how conversational systems are usually built. Then, we discuss the design of the verbal messages supported by users studies and data collected before the development of prototypes of the actual system. We show that the user data highlights issues and not only provide design recommendations for conversational interfaces creators but also structures the programming of such systems. The challenges we present here are tailored by the three disciplines: Information Design, Interaction Design and, Sensorial Design. We consider essential those perspectives to shape effective conversational experiences with conversational machines (Fig. 1).
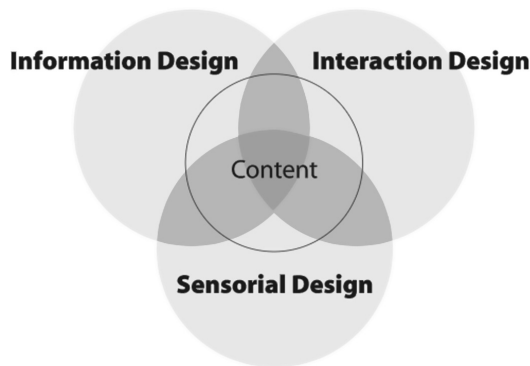

Fig. 1. Information Interaction Design. [19] illustrates the combination of the three categories that compose Interface Design. Illustration source [19].

2   Technical Limitations in the Design of Conversational Systems

Information interactive design for current conversational systems presents many challenges. Challenges such as bots being too responsive, not understanding utterances, lack of personality consistency, and giving too many options in the form of buttons for users are some of the issues affecting negatively the user experience. Some of those challenges are connected to the limitations and particularities of the technology currently employed. In most cases, the information flow of chatbots is tailored manually by developers using framework tools which structure the conversation quite rigidly. Those frameworks are used by designers and developers to build the content and structure of a possible conversation in detail, often using rule-based paradigms. Examples of commercial conversational frameworks and API platforms are Facebook's

wit.ai, *IBM's Watson Conversational Services, Microsoft's LUIS, Amazon's Lex, Google's Conversation APIs*, and as well as an increasing number of systems provided by startups.

In many ways, the technology behind most of today's commercial deployments is not too different from the one used by the early chatbots [24]. But the first key dimension in differentiating capabilities of different technologies and platforms refers to whether the dialogue is driven by the user, known as *user-initiative*, by the computer, known as *system-initiative*, or by both, or *mixed-initiative* [27]. It is therefore essential that the needs of the application and the interface, regarding the initiative, match the capabilities of the platform.

But independently of whether the user or the system has the initiative, most conversational systems today are built using an *intent-action* approach. The system is created by defining a basic set of user and the systems utterances and how they should match to each other. In user-initiative systems (for example, typical Q&A systems), groups of questions from the user are mapped into a single answer from the systems, sometimes with some variations. The term *intent* is used to describe the goal of the group of questions, so the basic task of the conversational platform is to identify the intent of a given question written or spoken by the user, and then output its associated answer or *action* (Fig. 2).
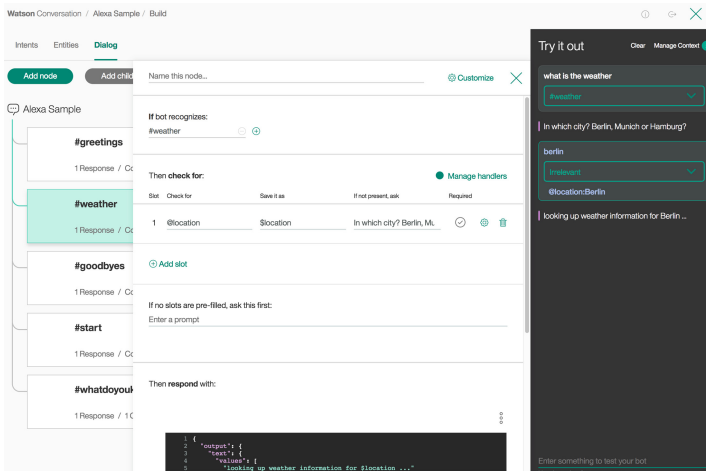

Fig. 2. An example of template-based system. IBM Watson Conversation Service. Designers of chatbots create a list of user intents, which are associated with an answer. Nowadays, designers may test the intent matching in real time and fix any possible discrepancy.

In system-initiative systems, the designers and developers of the conversational system have to provide sets of typical user answers for each question the system is going to make. Based on the *intent* of the answer of the user, an *action* is often produced with the help of primary natural language parsing technology to help extract by the system needed information such as numbers, choices, etc. Notice that in both

cases, as well as in mixed-initiative systems, the users' utterances always have to go through a process of matching with a set of available examples (often mistakenly called as "intents") for that context. The *intent matching* is often the most important source of problems in the development of conversational systems, due to the complexity and difficulty of analyzing natural language.

Many different technologies and platforms can be used for intent matching. A conventional approach is to use *template-based* systems in which the intent is determined by the presence of manually defined terms or groups of words in the user utterance. When the template matches the utterance, the associated intent is identified. Template-based systems, although often the simplest way to start developing a conversational system, suffer from two critical problems. First, it is hard to capture in simple templates the many nuances of human language. Second, as the number of templates increases, typically beyond one hundred or more, it becomes complicated to track the source of errors and debug the system successfully.

An alternative approach which is gaining increasing popularity is to use *machine learning-based* intent recognizers. In this approach, the conversational system developers provide a set of examples in natural language for each intent and use machine learning (ML) techniques to train an automatic classifier that is used in the run-time. Different types of classifiers can be used, such as *Bayesian networks*, *support vector machines (SVM)*, and the currently popular *deep neural networks (DNN)*. The main differences, advantages, and properties of those technologies are beyond the scope of this paper. Suffice to say that often the critical element of success when using machine-learning based intent recognizers is the quality and comprehensiveness of the data set provided to the ML classifier. Designers should not underestimate the importance of the often time-consuming task of collecting and organizing the training dataset.

An alternative method of creating conversational systems, still in the research stage, bypasses the definition of the intent-action sets and uses a large corpus of real dialogues to learn from the scratch how the conversation system should behave [21, 18, 25]. Notice that some of those works use corpora with hundreds of millions of conversations samples, or, conversely, very narrow domains, so this is an approach can only be applied with current technology to particular cases which meet those constraints.

Most of the communication expressions between conversational machines and humans consist of words and statements verbalized or in written mode. Conversational systems, popularly named chatbots, are in their essence verbal-based and supported by visual elements such as emoticons, pictures, and sometimes even action buttons. Most often than not, each chatbot's response is manually created by the designer or developer. They face an array of challenges not only to design of the conversational verbo-visual elements which reflect the personalized expression of the machine but also how to do this while obeying the constraints of the implemented algorithms and classifiers and handling the on-the-fly identification of context and user intentions. Understanding how those platforms work help and limit designers to create the conversation flow is important, since in practice the generated conversations are characterized by "moment-by-moment management of the distribution of turns" [22].

Those are some of the central concerns designers should keep in mind to battle the very common problem of conversation breakdowns. Conversational systems built

using intent-action rules have, by the way they are built, clear limits on their ability complex dialogues and utterances beyond the programmed scope. On the one hand, humans are experts to understand implicit interactions in a certain domain while machines struggle to keep up with it [9]. Ideally, understanding the conversation context and real-world knowledge should be part of chatbots repertories, but the reality of today's technology is that most conversational systems are very limited and easily broken. This challenge is even more prominent when more than one bot and a person is part of the same conversation. Turning-taking and governance of the conversation are vital elements to group conversation flows, and the same might be true when machines take part in group conversations. General rules have been created by designers and developers to guide chatbots behavior inspired by human conversation social rules [16, 7, 17], but they are still quite limited.

In the rest of the paper we discuss how some of those design challenges can be more effectively tackled with the help of collecting user conversational data before the system is developed. Based on the utterances and patterns detected, it is possible to simplify the design and deployment process considerably. Notice that performing user studies prior to system's development is a practice rarely used in the field.

3   User Data as the Design Guide of a Conversation System

Collecting natural human conversation data to understand the conversation dynamics is a way to build machines which reflect what users expect. Conversation logs can be used as a resource to simulate conversations and can be grounded in fieldwork studies. Although the main concentration of bot expression is textual, it is paramount to understand contextual clues (voice tone, facial expressions, environmental sounds, and utterances) which are only able to be collected in the field. As humans, the context where those conversations are situated matter and the actors in the same discussion affect the way people react with information. Enriching user studies with video and semi-structured interviews help to understand why specific questions were made to machines. Several design activities may help on collecting user data to understand the dynamic of the conversation. Some techniques such as *Wizard of Oz* technique [1, 3, 12, 13, 26], *Roleplaying* [8, 15, 23], and *Magic thing* [5, 9] may assist designers in investigating user interaction with conversational systems.

In this section, we describe a Wizard of Oz study which highlighted important challenges for designers to consider. Based on this study we exemplify challenges related to specific user issues and in the next section provide some design recommendations for overcoming those.

This study belongs to a series of user studies [1] with potential users with the limited financial knowledge. Those studies belong to a wealth management project where multiple chatbots are governed to provide investment advice to users. [2].

We conducted a Wizard of Oz study to understand human-machine conversation patterns, to map typical user's reactions to financial assistant answers, and to collect data with real user questions to build the first system corpus. Fifteen participants were invited to test a "the first version" of an intelligent financial adviser dialogue which could answer questions related to two kinds of investments: savings accounts and a

fixed-income investment called CDB (Bank Deposit Certificate). Following a typical Wizard of OZ protocol, the participants believed they were interacting with a functional system. The participants were remote and each session took approximately 30 min. The main data gathered were notes and audio and video recordings (screen captures). Participants were young adults (26 to 43 years old), highly educated and high-income bracket. All the participants described themselves as not interested or not keen on finances, particularly investments. All the participants answered positively to the consent form document, allowing us to use the data gathered.

3.1 Procedure

Participants were recruited by a snowball sampling and invited to be part of the remote study. The sessions started with demographic questions and questions of their financial investment experiences. Following that, they shared their screen with the researcher and started interacting with the chat mock up, the supposed Intelligent Financial adviser (Fig. 3).
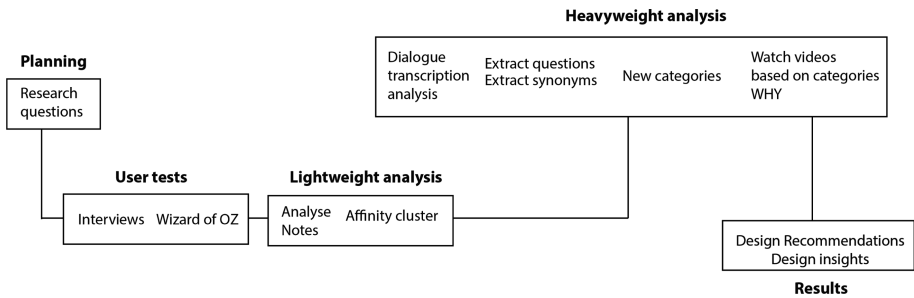

Fig. 3. Experiment procedures.

A human operator, who was not the same as the researcher facilitating the user session, answered their questions using a protocol. The human operator used a small table of content to answer the questions. The table was composed of 36 small paragraphs extracted from popular financial websites. The content relied on investment definitions, pros (return) and cons (risk) of two types of investments. Every table cell had a label (e.g. interest, safety, minimum value) to help the operator find the questions quickly during the sessions. The human operator could use sample answers in case she did not have an answer (*1. I don't know; 2. Ask again please; 3. I don't have enough information*). In the end of the session, the facilitator asked the participants to give their impressions about the system and disclosed the true identity of the intelligent system.

3.2 Data Gathering and Analysis

Lightweight and heavyweight analyses were the approaches to analyze the data [6]. The lightweight analysis consisted in an affinity cluster extracted from notes and offered guidelines for the main categories to look for in the audio transcriptions. The

main categories emerged from the data were: User reactions; Investment questions; Improvements; Technical issues; Communication issues; Conversation flow break-downs. In the heavyweight analysis, the *Nvivo* software was used to analyze the data. Notes, chat transcriptions and videos were analyzed. Categories from the affinity cluster phase were used as a base to analyze chat transcriptions and video transcriptions. Videos were mainly a source for understanding why people asked some questions to the financial adviser, allowing us to investigate how users structured their interaction during the study. For example, sometimes people repeated a question, or rephrased a question before writing the question and not always they typed what they wanted to know. Some reactions and contextual information were only possible to gather from watching some sessions again (Fig. 4).
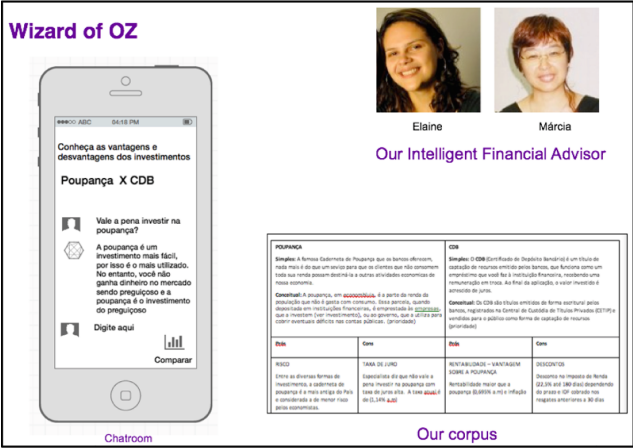

Fig. 4. This figure illustrates the components of a Wizard of Oz technique applied to collect user questions in the finance context. A chatroom environment, two researchers simulated the agent and a preliminary corpus in the table. Researchers simulated the agent and only used the content of the table to answer users.

3.3   Findings

In a previous paper [1], we described shortly the design process and how the results of this experiment impacted the development team. Here we present the main findings uncovered by this study which show typical issues designers of chatbots face. We classify those findings into 5 main categories, illustrating typical results obtaining from running user studies before system development starts. Those categories guided the elaboration of design recommendations, discussed in the next section, which were used by the development team for the actual construction of the conversational system.

Question Categories and Intent Definition. The study highlighted the main topics and questions potential users expected to be answered by a conversational investment adviser. The questions collected were organized in topics and illustrated by a visual taxonomy for each investment type (Fig. 5). Emerged topics from the study helped

classifiers to recognize the user questions grouped into user intents and connect suitable answers to user intents. Overall 125 questions were gathered which provided the first corpus for the financial adviser (see examples in Table 1).
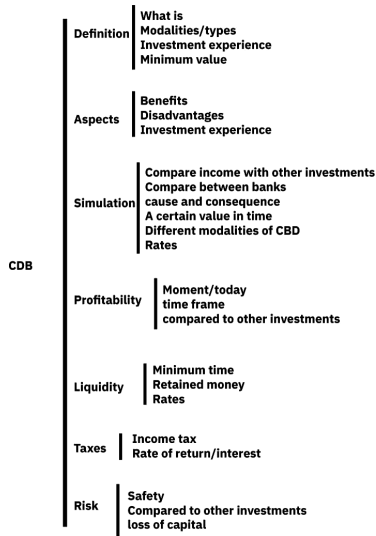

Fig. 5. Example of a taxonomy for CDB investment.

Essential Non-answered User Questions. It was also possible to detect information which participants expected from the system. Essential information not answered relied on participants asking for real-time value investment calculation; meanings of acronyms; a comparison between investments; and the system not answering questions about itself: *(P8) If it was you, which investment would you invest? (P5) Did you understand my question?* and generic questions such as *(P7) What is the best choice for investment?* Rarely participants asked out of scope questions; it might be because of the presence of a researcher facilitator observing their interactions.

User Perceptions of the Nature of the Chatbot Expression. The set of questions used by the human operator in the experiment had a neutral tone, which helped to identify in which degree participants expected the system biased the answers. Several participants verbalized their concern on how to communicate with this machine. Should it be formal or informal tone the questions? Should it understand acronyms of not? Should it understand punctuation? Showing transparence of how to communicate was essential for participants. It might be that some of the participants' concerns could be minimized by personalizing the chatbot answers, for example the tone of the utterances.

Identification of Context-Free and Contextual Issues. While performing the lightweight and heavyweight analysis we noticed issues that were connected to investment decisions and content specific and other context-free issues. We defined context-free

Table 1. Examples of categories and questions extracted from the study.

Categories	Sample question
Definition	What are the types of CDB?
Advantages and disadvantages	What are the advantages to invest on CDB?

conversation what might be extracted as ordered phenomena from conversational materials which would not turn out to require reference to one or another aspect of situatedness, identities, particularities of content, or context. [16]. Both types of issues are described in Tables 2 and 3.

Table 2. Context-free design recommendations classified by aspects of Information Design (ID), Interactive Design (IXD) and Sensorial Design (SD).

	Context-free - Information Interaction design recommendations
01	**The system cannot appear to be hiding information.** (ID) Participants felt the system showed what it knows by short pieces of textual information, and they wanted a complete information (pros and cons) to decide
02	**The system must be prepared to be tested. Repeated questions may have a more complete answer and new information.** (SD) Participants asked repeatedly for more information because they were not satisfied with the first answer, some thought the system did not understand them, others because they want to test the system.
03	**The system could help users to ask new questions.** (IXD) Participants did not know if the machine would understand them, they had trouble formulating sentences. Others did not have a basic knowledge of investments that would give ideas of what to ask.
04	**Basic information is expected.** (ID)(SD) Participants expect the system to know what they know at least.
05	**The system must engage the participant in the conversation**. (SD) (IXD) After moments of silence, or after the system answers a question not always the users kept asking. If the system asks: *Any more questions?* – and repeats prompt phrases of engagement, it would improve the flow of the conversation
06	**The system should identify the context of the question.** (IXD) (SD) Participants did not always include the name of the investment in the question, especially when the previous questions were related to an investment
07	**People expect the system to respond as a person even knowing that it is not a person** (SD) Questions allocated to the system were not answered to participants. The system reacted as *I don't have enough information*
08	**The system should introduce itself as being a machine**. (ID) (IXD) Participants were informed that they were talking to a machine by the researcher. Participants verbalized many concerns of how the system understand utterances and punctuations
09	**Consider the ambiguity of words.** (ID) (SD) Words may have double meanings. For example: rate is related to profitability rates or interest rates

Table 3. Contextual design recommendations classified by aspects of Information Design (ID), Interactive Design (IXD) and Sensorial Design (SD).

	Contextual (Finance) - Information Interaction design recommendations
10	**Showing evidence of real time value calculation** (IXD) (SD) Only four participants did not ask the system for a simulation with numbers
11	**Savings account should be used as a baseline for comparison.** (SD)(ID) (IXD) Participants had previous experience with Savings accounts, and often asked for comparing Savings to other investments
12	**The system should provide better information than bank managers.** (SD)(ID) People do not like to ask managers and often do not understand what the manager is talking about
13	**The system should have a visualization to compare the data.** (ID) Participants felt lack of visual presentations and tables to compare investments
14	**The system should give priority or weight to information related to loss.** (SD) Participants find more decisive the information that showed how much money will be lost than other kind of information such as: Minimum Application or Income

Reflections on Verbal Design Interaction. From the design activity, several reflections on how to conduct the next stage of the design emerged. Those are displayed as self-reflection questions: (a) What could be strategies to open spacesfor collaboration between man-machine in the decision making? (b) How a chatbots using text expression can help people to ask what they really want to ask? (we often saw that participants in the experiment rephrased or changed the questions verbalized to the researcher before typing); (c) How to better present and compare investments using dialogue?

4   Recommendations for Design of the Conversational System

Eighteen (19) design recommendations emerged from this study. Those were classified by strength of confidence and evidence accompanied by issues that occurred in the experiment. This rating scale was inspired by the scale applied in [20] work. The list of design recommendations accompanied by correlated issues and strengths was made available for the development team as an internal wiki page.

The recommendations were rated according to the level of confidence and organized into two groups: *Context-free* (Table 2) and *Contextual Design* (Table 3) recommendations. The Information Interaction Design recommendations are described in bold, followed by user issues observed in the study. Fourteen design recommendations for conversational interfaces were classified by aspects of three disciplines: Information Design (ID), Interactive Design (IXD) and Sensorial Design (SD).

We also classified each recommendation accordingly by the perspective of our collected data from the experiment participants. For instance, Sensorial aspects were classified in respect to verbalized participant feelings - expectation, frustration, confusion, satisfaction [19]. Information design aspects were the ones considering organization, presentation and/or text structure [14]. Interaction design aspects were

concerned of action, control, feedback, learning, balancing, engagement and conversing [4]. Some of the issues and recommendation fall into more than one of those three disciplines.

Context-free design recommendations posts challenges for Information designers, claiming legibility and transparence of information (01); literacy (04); personalization (08) and understanding (09). From the Interaction design perspective engagement with the system (03, 05), contextual reference (06), and interaction actions (02, 08). Participants also shaped their interaction in sensorial ways felling compelling to repeat their questions when not satisfied with answers or not felt understood by the system (02, 09). Expectation of machine-like behavior and lack of engagement were identified as important issues to considerate too. (05, 07). We expect that those design recommendations be useful for other intelligent conversational advisers in other areas.

Information Interaction design recommendations for similar financial advisers rely all 5 into the aspects of Sensorial design. (10, 11, 12, 13) were connected to expectations and previous experiences with financial products or bank manages. Moreover, the user sense of comparison with well-known investments, such as Savings, may assist to shape trust and reliability on this context. The last recommendation (14) is also supported by the loss aversion element present in the Prospect Theory [11]. The participants need and expectation for real time simulation and comparison of the investments influenced on the interaction design experience (10). Information designers should consider how to shape utterances based on previous user knowledge and experience (11, 12), with the aim of effectiveness and clarity of information. Visualizations and pictures are supported by conversational systems nowadays and are more suitable representations for comparing data than text (13).

5   Final Remarks

In this paper, we described technical and information interaction design challenges to create conversational interfaces and how using conversational data collected from users can help designers to face them. We exemplify with data captured from an experiment with real users which enriched and gave strength to design recommendations used by a development team of a conversational financial adviser. Fourteen design recommendations pointed out issues designers of conversational systems should consider also in other contexts. The lens provided by the three disciplines Information Design, Interaction Design and Sensorial design helped to shape and unveil our findings. We hope other designers may benefit from this study and apply the Information Interaction Design recommendations described here on similar conversational projects.

References

1. Candello, H., Pinhanez, C., Millen, D., Andrade, B.D.: Shaping the experience of a cognitive investment adviser. In: Marcus, A., Wang, W. (eds.) DUXU 2017. LNCS, vol. 10290, pp. 594–613. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58640-3_43
2. de Bayser, M.G., Cavalin, P., Souza, R., Braz, A., Candello, H., Pinhanez, C., Briot, J.-P.: A Hybrid Architecture for Multi-Party Conversational Systems. arXiv:1705.01214 [cs] (2017)
3. Dow, S.P., Mehta, M., MacIntyre, B., Mateas, M.: Eliza meets the Wizard-of-Oz: blending machine and human control of embodied characters. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 547–556. ACM (2010)
4. Dubberly, H., Pangaro, P., Haque, U.: ON MODELING What is interaction?: are there different types? Interactions **16**, 69–75 (2009)
5. Forlizzi, J., Battarbee, K.: Understanding experience in interactive systems. In: Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, pp. 261–268. ACM, Cambridge (2004)
6. Goodman, E., Kuniavsky, M., Moed, A.: Observing the User Experience: A Practitioner's Guide to User Research. Elsevier, Waltham (2012)
7. Grice, H.P.: Logic and Conversation, pp. 41–58 (1975)
8. Iacucci, G., Iacucci, C., Kuutti, K.: Imagining and experiencing in design, the role of performances. In: Proceedings of the Second Nordic Conference on Human-Computer Interaction. pp. 167–176. ACM, Aarhus (2002)
9. Iacucci, G., Kuutti, K., Ranta, M.: On the move with a magic thing: role playing in concept design of mobile services and devices. In: Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, pp. 193–202. ACM (2000)
10. Ju, W.: The design of implicit interactions. Synth. Lect. Hum. Centered Inform. **8**, 1–93 (2015)
11. Kahneman, D.: Thinking, Fast and Slow. Macmillan (2011)
12. Kerly, A., Bull, S.: The potential for chatbots in negotiated learner modelling: a Wizard-of-Oz study. In: Ikeda, M., Ashley, Kevin D., Chan, T.-W. (eds.) ITS 2006. LNCS, vol. 4053, pp. 443–452. Springer, Heidelberg (2006). https://doi.org/10.1007/11774303_44
13. Martelaro, N., Ju, W.: WoZ way: enabling real-time remote interaction prototyping & observation in on-road vehicles. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 169–182. ACM, New York (2017)
14. Pettersson, R.: Information design theories. J. Vis. Literacy **33**, 1–96 (2014)
15. Pinhanez, C.: Computer Interfaces to Organizations: Perspectives on Borg-Human Interaction Design. arXiv:1712.03012 [cs] (2017)
16. Sacks, H., Schegloff, E.A., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language **50**, 696–735 (1974)
17. Searle, J.R.: Speech Acts: An Essay in the Philosophy of Language. Cambridge University Press, Cambridge (1969)
18. Serban, I. V. S., Sordoni, A., Bengio, Y., Courville, A., Pineau, J.: Building end-to-end dialogue systems using generative hierarchical neural network models. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence (AAAI-2016) (2016)
19. Shedroff, N.: Information Interaction Design: A Unified Field Theory of Design. Information Design. MIT Press, Cambridge (2000)
20. Shneiderman, B., Leavitt, M.: Research-Based Web Design and Usability Guidelines. Department of Health and Human Services, Washington, D.C. (2006)

21. Sordoni, A., Galley, M., Auli, M., Brockett, C., Ji, Y., Mitchell, M., Nie, J.-Y., Gao, J., Dolan, B.: A neural network approach to context-sensitive generation of conversational responses. In: Proceedings of Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL (2015)
22. Suchman, L.: Human-Machine Reconfigurations: Plans and Situated Actions. Cambridge University Press, New York (2007)
23. Wagner, G.: Exploratory and Participatory Simulation. In: Proceedings of the 2013 Winter Simulation Conference: Simulation: Making Decisions in a Complex World. pp. 1327–1334. IEEE Press, Piscataway (2013)
24. Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM **9**, 36–45 (1966)
25. Weston, J. Dialog-based Language Learning. arXiv preprint arXiv:1604.06045 (2016)
26. Yu, Z., Nicolich-Henkin, L., Black, A.W., Rudnicky, A.: A Wizard-of-Oz study on a non-task-oriented dialog systems that reacts to user engagement. In: Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pp. 55–63 (2016)
27. Zue, V.W., Glass, J.R.: Conversational interfaces: advances and challenges. Proc. IEEE **88**, 1166–1180 (2000)