Towards the Design of a Student-Centered Question-Answering System in Educational Settings

Thiemo Wambsganss  
*University of St. Gallen*, thiemo.wambsganss@unisg.ch

Leonie Haas  
*University of St. Gallen*, leonie.haas@student.unisg.ch

Matthias Soellner  
*University of Kassel*, soellner@uni-kassel.de

Follow this and additional works at: [https://aisel.aisnet.org/ecis2021_rip](https://aisel.aisnet.org/ecis2021_rip)

**Recommended Citation**

[https://aisel.aisnet.org/ecis2021_rip/7](https://aisel.aisnet.org/ecis2021_rip/7)

This material is brought to you by the ECIS 2021 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2021 Research-in-Progress Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
TOWARDS THE DESIGN OF A STUDENT-CENTERED QUESTION-ANSWERING SYSTEM IN EDUCATIONAL SETTINGS

Research in Progress

Thiemo Wambsganss, University of St.Gallen, St.Gallen, Switzerland, thiemo.wambsganss@unisg.ch
Leonie Haas, University of St.Gallen, St.Gallen, Switzerland, leonie.haas@student.unisg.ch
Matthias Söllner, University of Kassel, Kassel, Germany, soellner@uni-kassel.de

Abstract

Enrollments in distance-learning scenarios have been tremendously rising. Here, the ability of students to receive answers to questions is hindered due to an uneven educator-student ratio. Students often do not receive quick answers to simple questions, and educators feel stressed by answering the same questions repeatedly. However, advances in Natural-Language-Processing and Machine Learning bear the opportunity to design new forms of human-computer interaction by embedding question-answering (Q&A) models in conversational agents. Such a system enables students to receive personalized answers independent of an instructor, time, and location. This paper presents the first steps of our design science research project on designing a student-centered Q&A system that helps learners receive personalized answers in large-scale settings. Based on social response theory and user interviews, we propose five design principles for the design of a conversational Q&A system. Furthermore, we instantiate those principles as design features in a natively built prototype.

Keywords: Question-Answering Systems, Conversational Agents in Education, Knowledge Management in Education, Design Science Research

1 Introduction

Enrollment in so-called Massive Open Online Courses (MOOCs) has been grown rapidly (Seaman, Allen and Seaman, 2018). In 2017, 33.1% of learners worldwide took at least one course online, compared to 24.8% in 2012 (Lederman, 2018). Not only in the distance-learning scenarios of MOOCs but also in traditional large-scale lectures at universities, students are confronted with uneven educator-student ratios. The standard default at public universities is often 100 students per educator, whereas in MOOCs the relationship between educator and students can be decoupled entirely in time and location, quickly building ratios of 10,000 students per educator (R. Winkler and Söllner, 2018; Meyer von Wolff, Nörtemann, Hobert and Schumann, 2020). Individualized support provided by educators is nearly impossible, and students are thus unable to engage in effective learning (Brinton et al., 2015). The ability of students to receive answers to simple and personal questions in their learning process is naturally hindered, and educators are often confronted with answering the same questions repeatedly. Several studies have revealed that this lack of individualized support leads to weak learning outcomes, high dropout rates, and dissatisfaction (Eom, Wen and Ashill, 2006; Brinton et al., 2015; Hone and El Said, 2016).

One solution avenue to provide students with accurate personal answers independent of an instructor, time, and location is to leverage recent advantages in Natural Language Processing (NLP) and Machine Learning (ML) (Winkler and Söllner, 2018; Wambsganss, Küng, Söllner and Leimeister, 2021).
Researchers train question-answering (Q&A) systems to redefine human-computer interaction (HCI) within educational information systems by changing how humans access and retrieve information (Kratzwald and Feuerriegel, 2019). Embedded in conversational agents (CAs), this novel technology replaces traditional information retrieval systems (Simmons, 1965), such as users using keywords to retrieve answers in matching documents (Belkin, 1993). For large-scale educational settings, Q&A systems bear several advantages. Research has shown that dialog-based question-answering is more natural to humans than traditional keyword searches, especially for students with poor technical skills (Vodanovich, Sundaram and Myers, 2010). Conversational Q&A systems can, therefore, significantly contribute to the ease of use (Radev et al., 2005) and user acceptance rates (Giboney, Brown, Lowry and Nunamaker, 2015). Moreover, Q&A systems are not only available at all times independent of educators or location but also promise to speed up the information search process, as students directly obtain the correct answer to their questions without searching through a collection of possible answers themselves (Roussinov and Robles-Flores, 2007). Lastly, Q&A systems can be embedded in different electronic hardware such as in Smart Personal Assistants (e.g., Amazon’s Alexa) (Rainer Winkler et al., 2020) or in wearables and thus make novel human-interaction scenarios possible. The embedding of conversational agents in MOOCs, for example, not only enhanced students’ engagement and participant levels but also decreased the risk of dropouts by up to 50% (Guarino and Welty, 2004; Meyer von Wolff et al., 2020)

Therefore, Q&A systems have been intensively investigated in various domains in past research, for example, in the field of medicine (e.g., Cao et al., 2011), general question-answering (Kratzwald and Feuerriegel, 2019), or IT security (Roussinov and Robles-Flores, 2007). Also, in educational scenarios, various studies exist that investigated the technical feasibility of Q&A systems (e.g., Ranoliya, Raghuwanshi and Singh, 2017; Ch’ng, Yeong and Ang, 2019; Mikic-Fonte, Llamas-Nistal and Caeiro-Rodriguez, 2019). However, only a few studies approach educational Q&A systems from a user-centered interaction design perceptive. Thus, as Winkler and Söllner (2018) and Hobert and Wolff (2019) stated, the literature on conversational embedded Q&A systems in educational settings falls relatively short of providing a holistic design perspective with principles and proof on how to design a Q&A bot in a particular pedagogical scenario. Therefore, we aim to contribute to design science and IS research by answering the following research question:

**RQ:** What are the design principles for a user-centered conversational question-answering system that helps students to receive personalized answers in large-scale pedagogical scenarios?

To answer the stated research question, we follow the design science research (DSR) approach by Hevner 2007. As stated above, there is a lack of design knowledge for Q&A systems in educational settings from a user-centered and theory-based perspective. The DSR approach is particularly suited to address such research gaps. We intend to iteratively design and evaluate a conversational Q&A system on the baseline of existing theory (social response theory by Nass, Steuer and Tauber, 1994) informing the artifact design (Hevner, March, Park and Ram, 2004). We believe that social response theory could explain why a user-centered design of conversational Q&A systems in educational settings might lead to positive user acceptance and usage rates. To the best of our knowledge, no study rigorously derives requirements from scientific literature and potential users to develop an educational Q&A system based on this theory. With a user-centered conversational Q&A system, we implicate a conversational agent that provides students with answers to natural language questions independent of an instructor, time, and place during, after, or before a large-scale learning setting. In this paper, we aim to present the preliminary results of our DSR project's first steps. In the following, we will first introduce the reader to the necessary theoretical background. Afterwards, we present our methodological approach for developing the artifact following the three-cycle view of Hevner 2007. Finally, our preliminary results of the first four steps are presented, followed by an outline of the subsequent steps and the expected implications once our research is completed.
2 Related Work

2.1 Traditional Question-Answering Systems in Education

In educational settings, the natural default is that students communicate with educators by telephone, e-mail or offline conversations to get answers to university-related questions (Sandoval, 2018). As a result, educators of large-scale settings often receive a high number of questions, which can become a burden for those responsible (Ch’ng et al., 2019). The problem also affects the students themselves, who sometimes must wait a long time for answers to relatively simple questions (Sumikawa, Fujiyoshi, Hatakeyama and Nagai, 2019). Alternatively, students can search for their questions on multiple university websites or in a list of frequently asked questions (FAQs). However, often, these are confusing due to the amount of information they contain, which makes the process inefficient (Sandoval, 2018). Hence, several authors have stated the importance of providing students with an effective information system to enable them to obtain information (e.g., Heo and Lee, 2019). This is especially significant since the lack of support leads to dissatisfaction, weak learning outcomes, or even high dropout rates (Brinton et al., 2015; Hone and El Said, 2016; Ikhsan et al., 2019). A solution to these challenges would be to employ a Q&A system in which students can receive an answer in natural language independent of educators, time, and location. A conversational agent could be precisely tailored to the respective Q&As and therefore provide precise answers to frequently asked questions (Shawar and Atwell, 2007). This would allow students to not only find information quickly and immediately in an anonymous way (Heo and Lee, 2019) but would also come with further advantages of conversational agents, such as 24-hour availability and reduced service costs (Meyer von Wolff et al., 2020).

2.2 Conversational Agents in Education

CAs are software applications designed to communicate with users through natural language interaction interfaces (Shawar and Atwell, 2005; Rubin, Chen and Thorimbert, 2010). In today’s world, conversational interfaces, such as Amazon’s Alexa, Google’s Assistant, or Apple’s Siri, are ubiquitous, with their popularity having steadily been growing over the past few years (eMarketer, 2017; Krassmann et al., 2018). They are implemented in various areas, such as customer service (Hu et al., 2018; Zierau et al., 2020), counseling (Cameron et al., 2017; Fitzpatrick, Darcy and Vierhile, 2017), healthcare (Kowatsch et al., 2017; Launer, Maier and Gubler, 2019), or education (Wambgsans, Söllner and Leimeister, 2020; Wambgsans, Winkler, Schmid and Söllner, 2020). Hobert and Wolff (2019) define CAs used in education as a particular form of learning application that interacts with learners individually. The development of CAs in education goes back to the 1970s research stream of Intelligent Tutoring Systems (ITS) (e.g., Atkinson and Shiffrin 1968; Suppes and Morningstar 1969). Like a human tutor, these systems can present instructions, ask questions, and provide immediate feedback (Kulik and Fletcher, 2016). ITS evolved from abstract entities with limited technological possibilities to systems that can interact with learners using multiple communication channels, exhibit social skills, and perform different roles, such as tutors (Payr, 2003), motivators, or learning companions (Kim and Baylor, 2008). While existing research on CAs in education has mainly focused on providing task-based learning support for students (Song, Oh and Rice, 2017; Hobert and Wolff, 2019), several authors have started to investigate the potential of available question-answering systems in higher education (e.g., Mikic-Fonte, Llamas-Nistal and Caeiro-Rodriguez, 2019; Oliveira et al., 2019). Nevertheless, past research has mainly focused on a technical use-case assessment. User-centered and theory-grounded design knowledge on embedding a conversational agent in a question-answer system is rather less.

2.3 Conversational Question Answering Systems in Education

Past research has mainly investigated the technical feasibility of Q&A systems in education (e.g., Ranoliya, Raghuwanshi and Singh, 2017; Ch’ng, Yeong and Ang, 2019; Mikic-Fonte, Llamas-Nistal
The studies of Sandoval (2018) investigated the practical understanding of conversational Q&A systems in education. They developed a chatbot for higher education settings using qualitative analysis to improve their FAQ bots. However, as also described in Santos et al. (2018), Ch’ng et al. (2019) and Sumikawa et al. (2019), past research has mainly focused on the functionality of Q&A bots. Santos et al. (2018) programmed a Q&A bot for students on Dialogflow using a forwarding chaining ontology. Ch’ng, Yeong, and Ang (2019) described how to develop a Q&A bot with Dialogflow displayed on Telegram and how to improve the functionality. Sumikawa et al. (2019) tested a new framework for supporting dataset creation to collect the Q&A data. However, as (Ch’ng et al., 2019; Sumikawa et al., 2019; Meyer von Wolff et al., 2020) noted, there is less investigation on the interaction design, the perception, and the acceptance of Q&A chatbot services for students. This lack of literature on design sciences was also identified as a literature gap by Polatidis (2014) and in recent studies of Meyer von Wolff et al. (2020). Therefore, we aim to address this research gap by deriving user-centered and theory-motivated design principles for the interaction design of Q&A bots in large-scale educational settings. In the long-term, we aim to provide not only requirements but also empirical investigations of our design principles on students’ behaviour and user perception.

2.4 Social Response Theory as Kernel Theory

Our design approach is anchored in social response theory. According to this theory, humans tend to respond socially to IS that display characteristics similar to humans (e.g., animals or technologies)(Moon, 2000). Behavioral clues and social signals from computers, such as interacting with others, using natural language, or playing social roles, subconsciously trigger humans' responses, no matter how rudimentary those clues or signals are (Nass et al., 1994; Nass and Moon, 2000). Following the “Computers are Social Actors” (CASA) paradigm, existing research has examined different social clues and their influence on HCI (e.g., Feine et al. 2019). Especially for the user-centered design of CAs, social response theory has been successfully used in IS research to embed CAs in different educational settings (e.g., Gnewuch, Morana, Adam and Maedche, 2018; Diederich, Brendel and Kolbe, 2020; Wambsganss, Winkler, Schmid and Söllner, 2020). However, to the best of our knowledge, the potential of embedding anthropomorphic clues in a Q&A system for educational purposes has been rarely investigated (Winkler and Söllner, 2018). By addressing the stated potential from a design science oriented perspective, we aim to contribute to better user acceptance, user experience and user-centered design for conversational Q&A systems in large-scale educational settings according to social response theory (Nass and Moon 2000).

3 Methodology

![Figure 1. Three-cycle design science process according to Hevner (2007)](image)

To answer our research question, we follow a DSR approach (Hevner, 2007). We decided to follow this methodology, as it allows us to solve a set of practical problems and to contribute to the existing body of knowledge by deriving and evaluating new design knowledge based on a sound understanding of the
current knowledge base and user perceptions of a recent technological phenomenon. Figure 1 shows the steps that are being carried out. In this research-in-progress paper, we report on the preliminary findings of the first four steps (see highlighted circles). Overall, our research project aims to contribute to research with a nascent design theory that gives explicit prescriptions for conversational Q&A systems in large-scale or distance-learning scenarios (Gregor and Hevner, 2013). We followed a theory-driven design approach by grounding our research on social response theory (Nass et al., 1994; Nass and Moon, 2000). The first step of the DSR process includes the problem formulation. The relevance of the practical problem was therefore described in the introduction of this work. In the second step, we derived a set of requirements in the form of literature issues (LIs) from the current state of scientific literature for the design of conversational Q&A systems. Therefore, we conducted a systematic literature review in the fields of Human-Computer Interaction (HCI) and Information Systems (IS) design. Next, we conducted thirteen semi-structured interviews with master’s students, using the expert interview method by Gläser and Laudel (2010). Based on the interviews, we gathered user stories (USs) as user requirements for the design of a user-centered educational Q&A system. In the fourth step, we derived preliminary design principles (DPs) addressing the LIs and USs from the previous steps using the structure suggested by Gregor, Chandra Kruse and Seidel (2020). We built an initial version as a first instantiation of these principles based on twelve carefully derived design features. Thus, we leveraged a first collection of around 50 questions and answers from one of our large-scale lectures at our university from the last three years. We trained a Q&A system based on pre-trained vector models and embedded the model in our natively developed CA, utilizing the frameworks spacy.io, flask, TensorFlow and ChatterBot in Python according to our design principles. A screenshot is displayed in Figure 3.

In our future research, we aim to evaluate this initial version based on the evaluation framework proposed by Venable, Pries-Heje and Baskerville (2016). They suggest four evaluation strategies from which we aim to use the human risk and effectiveness strategy since our research focuses on a user-centered artifact that needs to prove its utility and benefit in a real-world context (e.g., application in a large-scale lecture). Thus, we will first evaluate our prototype in a formative and artificial setting (i.e., lab experiment) to assess if all design principles are fulfilled. In this experiment, students are asked to use the developed prototype to answer preselected questions about a particular lecture to simulate a natural Q&A setting in a common pedagogical scenario. Subsequently, user perceptions will be captured with a questionnaire. Then, we will refine our design principles based on the findings from this evaluation before designing a second version, which can then be tested in a more extensive evaluation in a natural setting in a large-scale learning scenario. We will rely on one control group (participants will use an alternative FAQ document) and one treatment group (participants will use our Q&A system based on our design principles). Participants in the treatment group will use our designed CA to receive personalized answers independent of an educator, time, and location. Participants in the control group will receive the same question-and-answer pairs but displayed in a traditional static FAQ document. We aim to measure different behavior characteristics (such as search time, number of searched questions, number of correctly found answers) and user perceptions at the end of the study. Finally, we want to contribute with an evaluated natively designed Q&A system, which can be embedded in large-scale and distance-learning scenarios. The evaluation findings will be summarized as a nascent design theory (Gregor and Hevner, 2013) for student-centered Q&A systems.

4 Designing a Conversational Q&A System for Education

This section will describe and discuss how we gathered the preliminary requirements and derived the preliminary DPs. The problem formulation (step one), described in the introduction, serves as the foundation for deriving the requirements from literature and users. The main insights are illustrated in Figure 2.
4.1 Deriving Meta-Requirements from Scientific Literature

To derive requirements from scientific literature, a systematic literature search was conducted using the methodological approaches of Cooper (1988) and vom Brocke et al. (2015). We initially focused our research on studies that demonstrate the successful implementation of Q&A systems in education. Two broad areas for deriving requirements were identified: human–computer interaction and information systems. We only included literature that deals with or contributes to a conversational knowledge retrieval system in education. On this basis, we selected 35 papers for more intensive analysis. We have summarized similar topics of these contributions as literature issues (LIs) and formed broad areas for deriving requirements were identified:

For humans, the emotional intelligence of computers is important because they interact with users in socially meaningful ways (Creed & Beale, 2000; Nass & Moon, 2000). Accurate information retrieval systems should reduce the risk of failure by the user and thereby increase their self-confidence (Othman, 2006). Wilson (1999) pointed out that information retrieval systems should reduce the risk of failure and thereby increase the user’s self-confidence (LI 1). This is also reflected in the social responsibility theory of Nass and Moon (2000). Accordingly, users want to interact with computers that incorporate social characteristics. Thus, a Q&A system should comprise emotional intelligence to be accepted by the user and a social role adapted to the purpose (MR 1). Wilson (1999) and Othman (2006) point out that information retrieval systems should reduce the risk of failure and thereby increase the user’s self-confidence (LI 1). Therefore, a Q&A system should be easy to use and improve the search for information to reduce the risk of failure for students (MR 2). In the theory of Hemp (2009), information overload leads to difficulties for users to obtain relevant and valuable information (LI 3). Therefore, a Q&A system must be well trained based on the latest advances in NLP and ML so that it filters existing data to provide the user with useful and relevant information (MR 3).

4.2 Deriving Requirements from Expert Interviews

Based on the derived literature issues and meta requirements, we conducted thirteen semi-structured interviews according to Gläser and Laudel (2010). The interviewees were randomly selected students from our university, and each interview lasted around 20 to 35 minutes. The participants were between 23 and 28 years old; ten were female, three were male. Our interview guideline consisted of 23 questions about the interviewees’ experience with the use of chatbots, information retrieval in educational settings.
and particularly about requirements and expectations for a conversational Q&A system in educational settings. Additionally, we asked the participants to draw the design and interaction of a Q&A chatbot in an academic setting. Thereby, we further captured preferences in terms of design, interaction, and functionality. To evaluate the interviews, a qualitative content analysis was conducted. Based on the interviewees’ answers, we formed several clusters about requirements for a conversational Q&A system and gathered user stories (US) and derived user requirements (UR), following Cohn (2004). 80% of students mentioned that they would like to use a Q&A system that immediately provides clear and correct answers in natural language (US 1), which we incorporated in UR 1. Eight of the thirteen students stated that they would like to use a chatbot for either administrative or organizational questions (US 2) to receive quick answers for frequently asked questions (UR 2). From the responses of most students (60%), we derived the following user story (US 3): “As a student, I want a Q&A-system that can respond to my needs and provide me with the best possible support in answering my questions so that I am not left on my own when looking for information”. This implies that the system needs to understand the individual needs of each user and needs to be able to help answer the questions (UR 3). Another requirement (UR 4), which we derived from US 4, is that the Q&A system should apply to different educational topics and different semester levels. Five students expressed their desire for a Q&A system that is easily accessible and usable on different devices (US 5), which we incorporated in UR 5. All students would like a Q&A system that is professional and friendly at the same time (US 6). Thus, the conversational agent should be professional, social, and personal in character and design. On top of that, it should be tailored to the cooperate design of the educational institution (UR 6).

### 4.3 Deriving Design Principles for Initial Version

Based on the requirements, we derived five design principles (DP) for a conversational Q&A system in educational settings following the structure of Gregor, Kruse and Seidel, 2020. We illustrated them in Figure 2. Based on these design principles, we derived and formulated twelve design features (DF) as an instantiation of the DPs. The design features are incorporated in our initial version called “Herbie”. We leveraged a first collection of around 50 questions and answers from one of our large-scale lectures at our university over the last three years. We trained a Q&A system based on pre-trained vector models and embedded the model in our natively developed conversational agent, utilizing the frameworks spacy.io, flask and chatterbot in Python according to our design principle.

Following the design features, we instantiated the conversational Q&A system with a friendly, human-like avatar (DF 1) that uses the name of the user without asking, retrieved from a “Single-Sign-On” login information from our university (DF 2). Furthermore, we provided buttons to highlight the most popular topics. By clicking, the user can quickly find the information he or she is looking for (DF 3). After successfully answering a question, the bot displays similar questions to choose from (DF 4). Moreover, the CA provides links or PDFs for further information (DF 5). Furthermore, the CA is embedded in a commonly used website (the learning management system of our university), so it is easily accessible (DF 6). Also, the CA allows the user to choose how high the confidence level of the retrieved answers of the chatbot is (DF 7). Besides, the user is able to provide feedback to the CA with a bar of emojis (DF 8). After using the Q&A system, the user has the possibility to receive the entire chat via mail (DF 9). Furthermore, the interface displays a button on the top right to enable the user to connect to the educator in case a question is not appropriately answered (DF 10). Besides, the CA incorporates a large text input field, so the user can type long sentences and is still able to see the entire input (DF 11). Moreover, the interface has a mix of neutral colors and the cooperate design of our educational institution for a seamless experience (DF 12). In Figure 3 all the design features are displayed in our first prototype.
5 Next Steps and Expected Contributions

We follow the design science research (DSR) approach by Hevner 2007 to contribute to the research question what are the design principles for a user-centered conversational question-answering system that helps students to receive personalized answers in large-scale pedagogical scenarios? As outlined, there is a lack of design knowledge for Q&A systems in educational settings from a user-centered and theory-based perspective. Therefore, we aim to iteratively design and evaluate a conversational Q&A system on the baseline of existing theory (social response theory by Nass, Steuer and Tauber, 1994) informing the artifact design (Hevner, March, Park and Ram, 2004). We believe that social response theory could explain why a user-centered design of conversational Q&A systems in educational settings might lead to positive user acceptance and usage rates. In this paper, we discussed three literature issues and six user stories on how to design an educational Q&A setting for large-scale educational settings. We presented three preliminary meta requirements and six user requirements from thirteen interviews as well as five design principles that address them. We presented an initial version based on twelve instantiated design features as an instantiation of these design principles.

Next, we will evaluate our initial version in a formative and artificial setting (Venable et al., 2016), revise the design principles, and analyze the impact of the instantiated conversational Q&A system on students' perception and behaviour in a large-scale educational setting. We aim to measure different behavior characteristics (such as search time, number of searched questions, number of correctly found answers) and user perceptions to provide empirical insights of our design principles on the user. Thereby, we want to contribute to research with an evaluated natively designed Q&A system that can be embedded in large-scale and distance learning scenarios. The findings from the evaluation will be summarized as a nascent design theory (Gregor and Hevner, 2013) for student-centered Q&A systems. Overall, we aim to contribute to different large-scale and distance-learning scenarios to enable students to learn self-regulated effectively independent of an instructor, time, and location.
Reference


A Student-Centered Q&A System

the Association for Information Systems, Forthcomin.

Twenty-Ninth European Conference on Information Systems (ECIS 2021), Marrakesh, Morocco.


