Bringing AI into the Classroom: Designing Smart Personal Assistants as Learning Tutors

Rainer Winkler
University of St Gallen, rainer.winkler@unisg.ch

Julian Roos
University of St Gallen, julian.roos@student.unisg.ch

Bringing AI into the Classroom: Designing Smart Personal Assistants as Learning Tutors

Short Paper

Rainer Winkler
University of St. Gallen
Müller-Friedberg-Str. 8,
9000 St. Gallen
rainer.winkler@unisg.ch

Julian Roos
University of St. Gallen
Müller-Friedberg-Str. 8,
9000 St. Gallen
julian.roos@student.unisg.ch

Abstract

Despite a growing body of research about the design and use of Smart Personal Assistants like Amazon's Alexa, little is known about their ability to be learning tutors. New emerging ecosystems empower educators to develop their own agents without deep technical knowledge. The objective of our study is to find and validate a general set of design principles that educators can use to design their own agents. Using a design science research approach, we present requirements from students and educators as well as IS and education theory. Next, we formulate design principles and evaluate them with a focus group before we instantiate our artifact in an everyday learning environment. The findings of this short paper indicate that the design principles and corresponding artifact are able to significantly improve learning outcomes. The completed work aims to develop a nascent design theory for designing Smart Personal Assistants as learning tutors.

Keywords: smart personal assistant, intelligent tutoring system, design science research, quasi field experiment, education

Introduction

New developments, such as ever-growing classrooms at high schools, large-scale lectures at universities with more than 100 learners per lecturer, and massive open online courses (MOOCs) with more than 1,000 learners on average, are increasingly entering the educational landscape (Oeste et al. 2015). For example, in 2018, over 900 universities around the world had announced 2,000 new MOOC courses and there is an upwards trend (ICEF 2019). These learning settings make it possible to reach a wider audience with education. However, individual support from educators is hardly possible due to financial and organizational restrictions (Lehmann et al. 2016). According to constructivist learning theories, individual support is crucial for an increased learning success (Glasersfeld 1987). This raises the question of how educators are able to offer individual support in large-scale learning environments.

Information systems have often tried to address this question. Especially the research stream of intelligent tutoring systems (ITS) has been extensively investigating the value of ITS to improve students' learning outcomes over the last thirty or forty years (Nwana 1990). An ITS is a computer system that aims to provide immediate and customized instruction or feedback to learners (Nwana 1990). A review on the effectiveness of ITS by VanLehn et al. (2011) states that the effect size of ITS is 0.76, nearly as effective as human tutoring with 0.79. However, until now, existing ITS have still found too little entry into learning environments (Nye et al. 2018). One reason for that might be that developing and introducing ITS in learning environments often requires a lot of technological knowledge and a lot of time (Elaine 2015). Moreover, ITS applications are often complex to transfer from one context to the other (e.g., different learning goals, tasks, etc., Woolf et al. 2008). Furthermore, ITS are often standalone software systems designed for devices that students do not use every day (Sottilare and Proctor 2012).

New emerging Smart Personal Assistants (SPAs) such as Amazon's Alexa, Google's Assistant, Apple's Siri, and others have the potential to form a new category of ITS, thereby filling these gaps. SPAs are software agents designed to support users in doing several daily activities by engaging with them via voice (Pais et al. 2015). SPAs are running on 'SPA-enabled devices" (endpoints) such as Apple's iPhone, iPad and Mac, Amazon's Echo, Google's Home, etc. The main functionality, the "brain" of a SPA, is typically housed as a cloud service that processes voice data (converting voice-to-text, performing linguistic context analysis, and providing answers to questions, Chung et al. 2017). The popularity of SPAs has been steadily growing over the past few years (eMarketer, 2018). In the United States, 59 percent of respondents from the ages of 18-24 stated that they were heavy (at least once per day) users of smart personal assistants (Statista 2019a). SPA providers such as Google, Amazon, and IBM offer large eco-systems that allow users to create their own skills without much technological know-how and time, thereby increasing SPA providers' own business value. Consequently, the use of SPAs as ITS has three major advantages. First, it allows educators to develop SPAs for their own learning environments without being dependent on software designers or having to invest a high amount of time. Second, it allows educators to convert the use of SPAs to different contexts (e.g., different learning goals, contents, tasks) in an efficient way. Third, SPAs are becoming daily companions of students in their private and school life, being integrated in devices students use every day (smartphone, tablet, etc.). Until now, little is known about how educators can design and integrate these new kind of dialogue-based information systems in their learning environments (Hobert and Wolff 2019). Most publications in the field focus on specific implementations of ITS and miss to provide transferable insights. Hence, the objective of this study is to find and validate a general set of design principles that researchers and practitioners can use to design effective SPAs as learning tutors. We thereby follow the calls of Hobert and Wolff (2019) and Kim and Baylor (2016) stating that there is a need for generalized design knowledge for SPAs in education. Consequently, we address the following research questions:

RQ1: How can Smart Personal Assistants be designed as learning tutors in order to improve students' learning outcomes?

Our design science research project is grounded on a constructivist view of learning. In specific, we use the ICAP-framework proposed by Chi and Wylie (2014) as our kernel theory. To answer our research questions, we present two cycles of a design science research (DSR) approach as described by Hevner (2007). In our *first cycle*, we derive requirements for SPAs as learning tutors from literature and conduct interviews with educators and students. Based on that, we formulate an initial set of design principles for SPAs as learning tutors. In our *second cycle*, we refine the design principles and build our first instantiation of a SPA, and evaluate it with the help of a quasi field experiment in an everyday learning environment. The *third cycle* will be subject of future research. The goal of the completed research project is to create a set of design principles resulting in a nascent design theory allowing educators to create SPAs for their own learning environments. The remainder of this paper is structured as follows. In the following section, we define and classify SPAs and will present the ICAP-Framework as our kernel theory. The ICAP-Framework defines cognitive engagement activities on the basis of students' overt behaviors and proposes that engagement behaviors can be categorized and differentiated into one of four modes: **I**nteractive, **C**onstructive, **A**ctive, and **P**assive (Chi and Wylie 2014). Next, we explain the research methodology, will discuss the results so far and close with next steps and the expected contribution.

Theoretical Background

In this section, we first define and describe Smart Personal Assistants as Learning Tutors. Next, we elaborate on the ICAP-Framework as our kernel theory for the development of the design principles.

Smart Personal Assistants as Learning Tutors

SPAs are software agents that can automate and ease many of the daily tasks of their users by engaging with them via voice (Pais et al. 2015). SPA providers offer rich ecosystems with intuitive interfaces that allow their users to create and share their own skills, thereby increasing SPA providers' own business value. Compared to traditional user assistance systems, SPAs can be characterized as offering a rather high degree of interaction and intelligence (Pais et al. 2015). They are able to react to users' utterances and can proactively guide users through a complex task. For example, Fast et al. (2018) introduced a SPA supporting users in conducting data science tasks (e.g., a predictive modeling task). In the field of education, using computer tutoring systems is not new. Researchers started to develop intelligent tutoring systems (ITS) as early as the 1970s in order to help students with learning (Nwana 1990). Despite its proven effect on learning success, the distribution of todays' ITS is still very limited because it takes a lot of technological know-how and development time (generally 200 hours of development time for one hour of teaching/instructions), and, thus, many researchers describe the ITS development as notoriously costly (Sottilare et al. 2016). In contrast to that, SPA providers such as Amazon are offering easy-to-use toolkits with a lot of blueprints and tutorials allowing the user to build their own skills for SPAs with very basic technological know-how and low time effort (Amazon 2019).

ICAP – Framework as Kernel Theory

The ICAP framework proposed by Chi and Wylie (2014) is based on a constructivist view of learning. It explains the process of effective learning by classifying observable student behaviours into four modes: **I**nteractive, **C**onstructive, **A**ctive, **P**assive and predicts that these modes will be ordered by effectiveness: interactive > constructive > active> passive. Educators have long recognized that although students can learn from receiving information passively, they learn much better actively. Learning actively requires students to engage cognitively and meaningfully with the tasks they are doing. They really think about their learning material in depth rather than just passively receiving it (King, 1993). Each mode of the ICAP framework corresponds to different types of behaviours and knowledge-change processes predicting different learning outcomes. For example, when watching a video, students can watch it without doing anything else (passive behaviour). Students can also manipulate the tape by pausing, playing, fast-forwarding and rewinding (active). A constructive behavior would be self-explaining the concepts in the video (constructive). The most effective interactive student behavior would be if students discuss the content and its justifications with a peer or tutor (interactive). The ICAP-Framework serves as a basis for the development of the design principles. It explains the mechanism why SPAs might be able to increase students' learning outcomes by changing their knowledge processes when interacting with SPAs compared to non-interactive learning materials. Properly designed SPAs as learning tutors have the capabilities to bring students from a passive to an interactive learning mode (Chi and Wylie 2014).

Research Methodology

We rely on Hevner's (2007) three cycle design science research approach to structure the research process (see Figure 1).

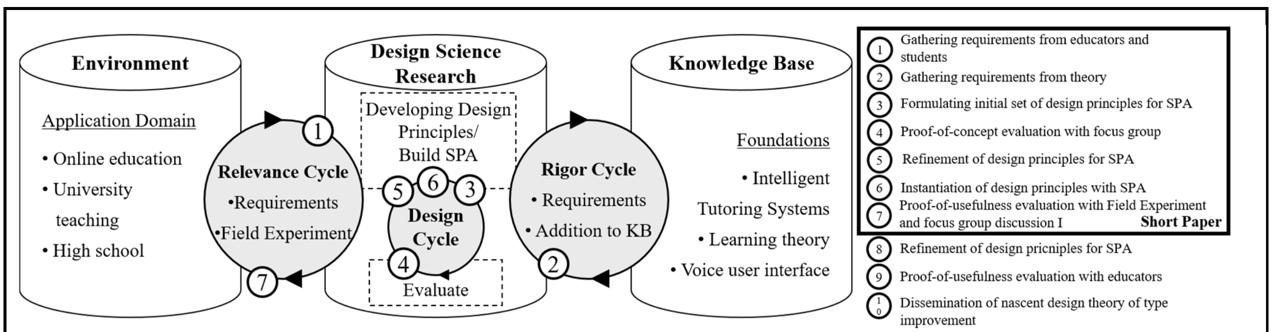

Figure 1. Our DSR Process (Hevner 2007)

In step 1, we initiate the *relevance cycle* by gathering requirements from educator and student interviews. In step 2, we initiate the *rigor cycle* where we gather requirements from three different theoretical perspectives. In step 3, we initiate the *design cycle* by formulating the initial set of design principles for SPAs as learning tutors. In step 4, we conduct a first proof-of-concept evaluation with the help of a student focus group discussion. The objective of this evaluation is to make sure that we capture the most important requirements and logically translate them into design principles (Sonnenberg and Vom Brocke 2012). In step 5 and 6, we refine the design principles and create our first instantiation of design principles. In step 7, we conduct a quasi field experiment and enrich the quantitative data with another focus group discussion. The objective of this evaluation is to prove the usefulness of our design principles and our corresponding artifact. Step 8-10 will be out of the scope of this short paper and will be the subject of the next months.

Designing Smart Personal Assistants as Learning Tutors

Step 1: Gathering Requirements from Students and Educators

Our application domain involves students as well as educators from different levels and teaching formats (Online education, university teaching, high school). In a first step, we wanted to gather a basic understanding of the current problems educators and students face in their everyday life and what they think of using SPAs to enrich existing learning environments. Hence, we conducted twenty interviews with students from a swiss high school and a swiss business university and four interviews with educators from the same school and university. We used our network and social media to recruit the interview partners. Table 1 shows the characteristics of the interview partners. All the interviews were semi-structured and lasted approx. 30-minutes each. We used a paper-based mock-up for showing interviewees our initial idea of SPAs. The paper-based mock-up showed a possible interaction between a student and the SPA when working on a problem task. The problem task dealt with the problem of the rising traffic jam in a Swiss city. Table 1 shows the characteristics of students and educators.

Table 1. Sample Characteristics

Interviewees	Characteristics
Students	**Age:** 21.35, **level of education:** 5 x high school, 10 x bachelor, 5 x master, **nationality:** 8 x Swiss, 11 x German, 1 x Italian, **gender:** 8 women, 12 men
Educators	**Teaching experience:** 5 to 25 years, **type of school:** 1 x high school teacher, 2 x bachelor lecturer, 1 x master lecturer, **experience with different learning environments**: 1 x online, 1 x mass lecture, 2 x small classes

We created the interview questions in an internal workshop of the research team. In the first part of the student and educator interviews, we asked: What are problems you encounter with in a course? Can you explain it based on concrete examples? In the second part, we showed them a mock-up-based SPA-student learning dialogue on a piece of paper. Based on that, we asked them about possible requirements for using this new technology. We transcribed the interviews and analyzed them using the method of user stories proposed by Cohn (2004). We identified user stories in the interviews, coded, clustered and finally translated the user stories into requirements for our SPA application (see Table 2).

Table 2. User-Stories and Requirements from Students and Educators

User Stories (Students)	Requirements from Practice (RP)
USS1: As a *student*, I want the SPA to give me immediate and detailed feedback on my responses.	**RP1:** The SPA should actively process and evaluate the students' responses.
USS2: As a *student*, I want the SPA to reply as fast as a human in a conversation.	**RP2:** The SPA platform should provide a fast hardware and software architecture.
USS3: As a *student*, I want to be able to use the SPA at any time, at any place.	**RP3:** The SPA should be available anytime.
USS4: As a student, I want to access the SPA via voice similar to smartphone assistants (e.g., Google's Assistant or Apple's Siri).	**RP4:** The SPA should include voice in- and output.
User Stories (Educators)	
USE1: As a*n educator*, I want to design SPAs by myself without needing a lot of technological knowledge.	**RP5:** The SPA platform should allow educators to develop SPAs on their own.
USE2: As a*n educator*, I want that the SPA can be used on different devices (e.g., smartphones, tablets, laptops).	**RP6:** The SPA should be accessible via different devices.
USE3: As a*n educator*, I want to use the SPA for different learning goals and tasks.	**RP7:** The SPA should be adaptable to different tasks in an efficient and uncomplicated way.
USE4: As an *educator*, I want to spend little or no money on developing SPAs for my courses.	**RP8:** The development costs of SPAs should be low.

Step 2: Gathering Requirements from Theory

We initiate the *rigor cycle* by gathering requirements from theory. We conducted a systematic literature review following established methodical approaches from Cooper (1988) and vom Brocke et al. (2015). Based on that, we (1) defined the review scope, (2) conceptualized the topic, (3) searched the literature, and (4) analyzed the findings regarding requirements. Regarding step 1 (define the review scope), we primarily focused our literature review on research outcomes that show successfully implemented SPA design features. Furthermore, our goal is to identify requirements on a conceptual level with a focus on an espousal of position and a representative coverage (Cooper 1988). Regarding step 2 (conceptualization of the topic), we identified three perspectives useful for deriving requirements: a technical, educational and voice user interface perspective. We used these three perspectives, because learning with a tutor is a very complex phenomenon being investigated through different lenses by psychologists, educationists, computer scientists and various others (Lajoie and Azevedo 2006). Regarding step 3 (literature search), we conducted a keyword search in databases to identify relevant publications. We started with all eight journals of the AIS Senior Scholars' Basket of Journals as well as the three most prestigious journals in educational research and then opened the range to other peer-reviewed journals and conferences to cover all the three different perspectives. We selected the following databases: "AIS Electronic Library", "ACM Digital Library", "IEEE Xplore Digital Library", "Science Direct", "EBSCOhost Business Source Complete" and "ERIC". Additionally, we used the search terms "smart personal assistant", "pedagogical conversational agent", "voice Assistant", "artificial intelligence teaching assistant", "conversational agent" and "design". In total, we obtained 478 articles. We defined criteria for inclusion and exclusion and reviewed titles and abstracts of all search results in a first step. We only included papers that explicitly tested the effectiveness of some kind of design features of a SPA in order to make sure that we derive successfully tested requirements. Based on that, we selected 36 papers. Regarding step 4 (literature analysis), we clustered similar requirements resulting in five requirement clusters. Table 3 depicts the perspectives, an exemplary paper, and the requirements from theory (RT).

Table 3. Requirements from Theory

Perspectives	Exemplary Paper	Requirements from Theory (RT)
Educational	**ICAP-Framework** (Chi and Wylie, 2014)	**RT1:** The SPA should include a proactive and reactive interaction model, where the SPA and the learner both make a contribution.
	Scaffolding (Kim and Hannafin 2011)	**RT2:** The SPA should provide different kinds of scaffolds during task completion.
Technical	**Personification** (Woolf et al. 2008)	**RT3:** The SPA should include personification attributes.
	Error handling (Guo and Goh 2015)	**RT4:** The SPA should resolve errors and dysfunctions so that the user can continue.
Voice User Interface	**Transparent Systems** (Xu et al. 2014)	**RT5:** The SPA should help the users to build an internal model of the functionality of the system.
	Engaging Response Structure (Pearl 2016)	**RT6:** The SPA should use an engaging response structure different to text interfaces.

Step 3: Formulating Design Principles based on Requirements

Based on our identified requirements from theory and practice, we formulated an initial set of design principles. Figure 2 shows an overview of requirements, design principles and design decisions. The rectangles in dark grey represent the requirements and design principles which we have added in cycle 2 (after the first evaluation). *Interactivity* (**DP1**) is a fundamental principle of every computer tutor. According to constructivist learning theories like the ICAP framework (Chi & Wylie, 2014), interactivity is the main determinant for learning success. Thus, the SPA tutor should try to trigger students' interactive learning behavior by using strategies to elicit constructivist responses from the student. *Scaffolding for complex tasks* (**DP2**) is the second main principle coming from educational research. Scaffolding is an effective standard tutoring method. It nudges students to acquire skills on their own and is applicable in almost every learning scenario. Therefore, a SPA tutor must provide different kinds of scaffolds. *Humanness* (**DP3**) relates to how the SPA is represented. For example, Woolf et al. (2008) revealed that personified computer artifacts show a higher user satisfaction and more intense usage. As a consequence, the SPA tutor should be designed in such a way that it appears as human as possible.

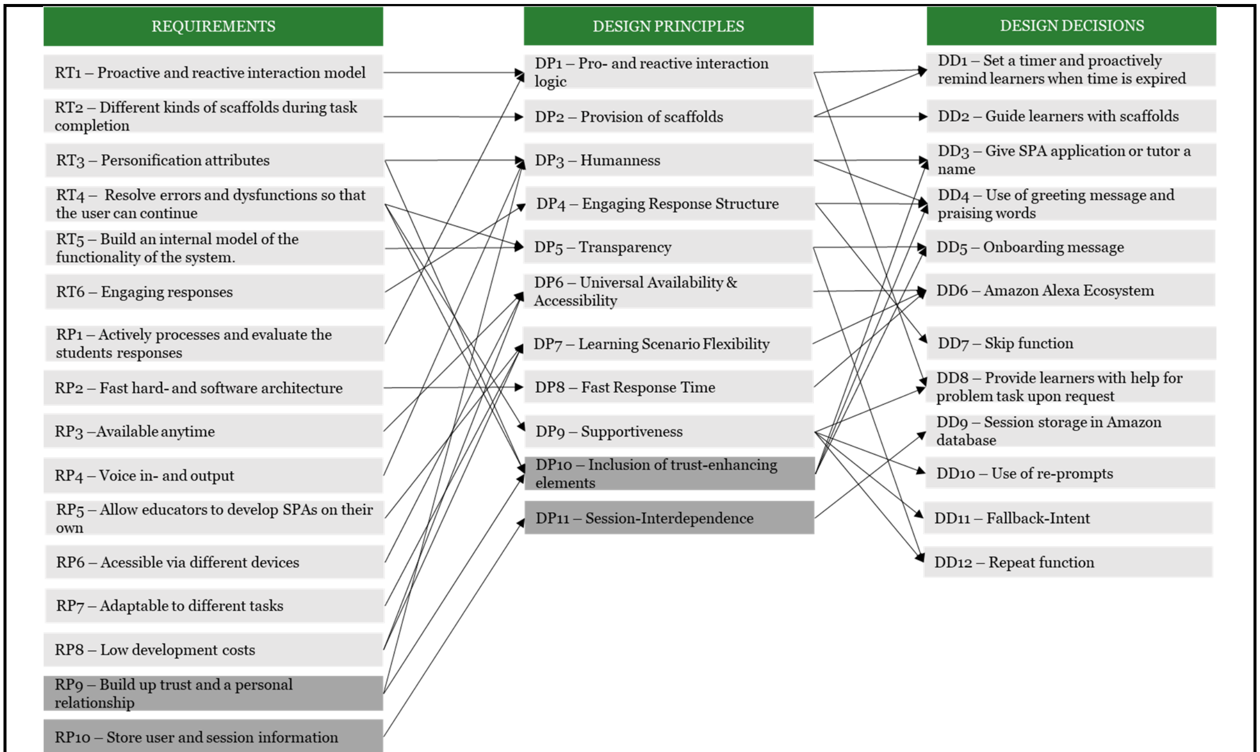

Figure 2. Requirements, Design Principles, and Design Decisions

Engagement **(DP4)** is the fourth design principle. Looking at voice interaction strategies it becomes clear that a major goal and challenge is to keep users engaged during the conversation (Pearl 2016). The fifth DP is *transparency* **(DP5)**. According to Xu et al. (2014), the central problem of purely voice-controlled systems is that they are invisible. To avoid dysfunctions and user frustration, the SPA tutor must incorporate different strategies to give the user orientation. The sixth DP is *availability and accessibility* (**DP6**). Today, students want to access applications on their everyday devices at any time, at any place. Our interviews with students and educators clearly confirmed that a SPA tutor should be accessible across devices and also available outside the class room. The seventh DP is *learning scenario flexibility* (**DP7**). A key reason why ITS have not spread larger are the high costs and time efforts for changing learning contexts. By enabling educators to create their own SPAs, the costs and effort can be significantly reduced. The eighth DP is *response time* **(DP8)**. The conversational capabilities of the SPA tutor are an essential characteristic. During the interviews, we found that it is very frustrating for users if they expect a human-like conversation, but the conversation feels clumsy because the system takes too long to respond. The ninth DP is *supportiveness* **(DP9)**. As Pearl (2016) pointed out, SPAs are far from being able to respond to every user input. Together with the problem of voice user interfaces not being self-evident, it is likely that dysfunction occur during usage. Therefore, it is important to offer support when errors happen.

Step 4 and 5: Proof-of-Concept Evaluation and Refinement of Design Principles

According to Sonnenberg and vom Brocke (2012), it is important to direct the foci of evaluations on two aspects: (1) the constituents of the artifact and the design decisions take as well as on (2) the evaluation of the usefulness of the artifact. The first evaluation concentrates on the constituents of the artifact. Focus group discussions are considered a suitable method for evaluating designs within DSR projects (Sonnenberg and Vom Brocke 2012; Venable et al. 2012). The goal of the focus group discussion was to check the validity of requirements and also the translation to the resulting design principles. Moreover, we wanted to gather further requirements and design principles for our next cycle. The focus group discussion lasted 60 minutes with one of the researchers as facilitators. The participants were six master and four undergraduate students with an average age of 23.5 (m=3, w=7) and were collected from a Swiss business university (different students than in step 1). The focus group discussion was structured as follows. In the first part, one of the researchers introduced the goal of the group discussion. In the second part, we showed students our derived requirements and design principles and asked them to check for weaknesses regarding the formulation of design principles. In the third part, we used a small Wizard-of-Oz Simulation where one of the researchers (not the facilitator) was located in a different room using an Alexa Echo Dot and its intercom feature to talk to the students. We also placed an Alexa Echo Dot in the room of the students so that the researcher was able to guide them through a short 10-minute learning task. After the simulation, we asked learners to collect further requirements and design principles. We added the additional requirements and design principles to the overview in Figure 2.

Step 6: Instantiating of a SPA in an Everyday Learning Environment

Based on the insights from the focus group discussion, we started the next DSR cycle by refining our design principles, designing our first instantiation of a SPA and evaluating the artifact with the help of a quasi field experiment. We decided to use Amazon's Alexa platform as this platform offers state-of-the-art speech recognition and natural language processing as well as easy-to-use toolkits. The use of the Amazon Alexa ecosystem itself is a design decision (**DD6**) that instantiates several DPs (see Figure 2). Amazon Alexa is accessible and available for different devices **(DP6)**, has a very powerful voice recognition system in the backend **(DP8)**, and provides a lot of blueprints and tutorials **(DP7)**. The learners will be guided through the problem-exercises through scaffolds **(DP2)** based on problem-solving steps provided by Kim & Hannafin (2011) **(DD2)**. After each instruction, the SPA tutor gives the learners a certain amount of time to work on the task (**DD1**). To ensure interactivity **(DP1)** and supportiveness **(DP9)**, the students are able to skip to the next step **(DD7)**. Moreover, the SPA proactively asks the learners if they are finished with the task or if they need more time or help. If they choose help, the SPA tutor will give them hints on how to solve the task and remembers their mistakes (**DD1, DD8, DD9**). To further strengthen the supportiveness, we included re-prompts (**DD10**) as well as fallback-intents (**DD11**). To make the voice interaction model transparent (**DP5**), we decided to have an extensive onboarding message after the launch of the SPA application that informs learners about the capabilities of the tutor (**DD5**). Moreover, we integrated conversation markers such as greetings and praises **(DD4)** to make the tutor more engaging (**DP4**). To make sure that the SPA recognizes a user when they continue with a task **(DP9)**, we connected the SPA application to an Amazon database **(DP11)**. To make the SPA application appear more personified and trust-enhancing (**DP3** and **DP 10**), we gave the tutor the name "Sarah" and included an onboarding message (**DD3, DD5**). Furthermore, we included conversation markers **(DD4)** to make the tutor more engaging (**DP4**). Another problem of voice user interface is that it is difficult to make the system transparent for new users but efficient to use for experienced users. To not lose the engagement of both user groups **(DP4)**, we connect the SPA application to an Amazon database, where session information about the user is saved (**DD9**). Hence, we can recognize if the user is new or experienced as well as continue the session of a user (**DP11**). Finally, to further increase transparency **(DP5)** and supportiveness **(DP9)**, we included a task repeat function **(DD12)**. Figure 3 presents an exemplary SPA interaction with a student in the experiment class and the corresponding design decisions.

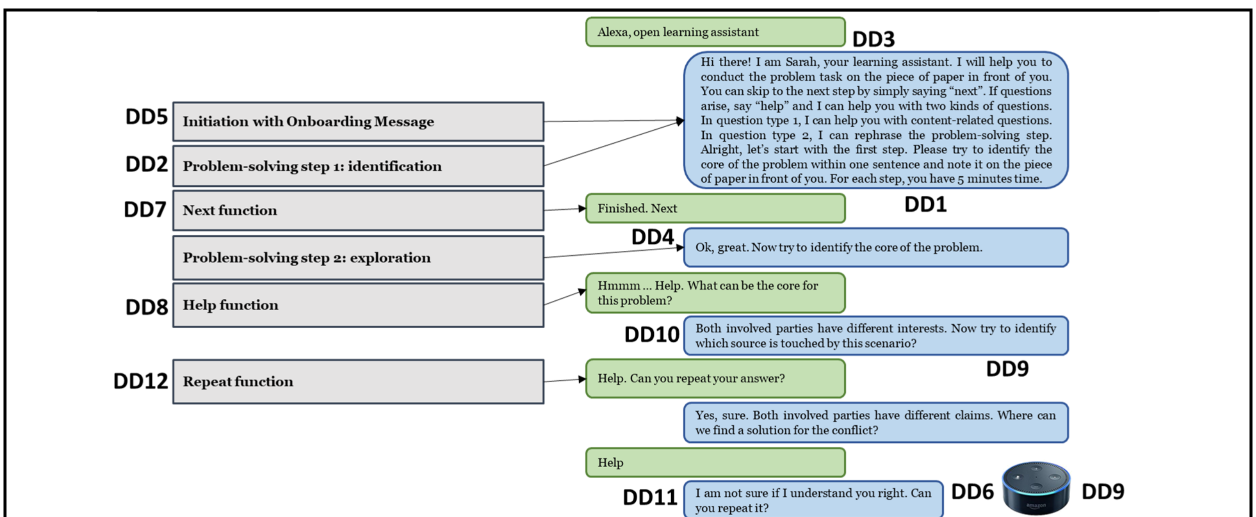

Figure 3. Exemplary SPA-student interaction and corresponding design decisions.

Step 7: Proof-of-Usefulness Evaluation in an Everyday Learning Environment

This second evaluation concentrates on the usefulness of the artifact with the help of a quasi field experiment in a business vocational school in Switzerland over the period of five weeks. Overall, 45 students participated in the quasi field experiment. The experiment class participated in a focus group discussion afterwards to receive more insights into the usefulness of our artifact. The average age of the experiment class was 17.4 with 12 males and 10 females. The average age of the control class was 17.2 with 12 males and 11 females. The results of an ANOVA test ensure that the two groups are equal regarding gender, age, pre-knowledge, pre-experience with SPAs, and personal innovativeness. The experiment and the control class received four paper-based homework assignments. The experiment class (22 learners) used our developed SPA as learning tutor for their four 30-minute homework assignments (see Figure 3). The control class used paper-based scaffolds as a baseline. The paper-based scaffolds were separate scripts, which contained hints for each of the four homework assignments. Both learning aids were information equivalent (all SPA hints were also available on the paper-based scaffolds). The classroom teacher administered a 30-minute, 3-subtask pre-knowledge test to all learners one week before the experiment period starts. Both classes had the same teacher using the same teaching methods. At the beginning of week 1, we introduced the SPA devices (Alexa Echo Dot Device) in the experiment class and installed the Alexa software on their everyday devices (smartphone, tablet, laptop). Next, they used Alexa for conducting their assignments. In week 5, they had to conduct a 30-minute post-test (similar to pre-test) and post-survey. After another week passed by, week 6, we conducted a 45-minute focus group discussion with the whole experiment class. The ANCOVA (with pretest as covariate) indicates that there is a highly significant relation between SPA usage and learning outcomes (($F (2, 42) = 30.573$, r^2 adjusted $= 0.42$, $p = 1.88e{-}06$, confidence intervals = 22.081 and 38.227, $N = 45$). Cohen's d is 1.70. Since this effect is considered to be high, we can conclude, that our artifact proved to add value to everyday learning environments in this setting.

Next Steps and Expected Contribution

In our next steps, we will start with DSR cycle 3 by analyzing the focus group discussion from our field experiment. Next, we will use our refined design principles to instruct a couple of educators from different education levels to create a SPA for their own learning environments. Cycle 2 proved that our artifact is useful for increasing learning success of students. Cycle 3 aims to evaluate the usefulness of our artifact for educators. After this evaluation, we document the gained design knowledge in the form of a nascent design theory (Gregor and Jones 2007). Thus, we contribute to the field of computer tutoring by providing prescriptive knowledge on how educators can design new forms of ITS (SPAs) to enrich their learning environments and improve students' learning outcomes. As SPAs will increasingly enter students' everyday lives, it is important to understand how SPAs can be used in education. Regarding practical implications, we exemplarily show educators how to build SPAs for their own learning environments.

References

Amazon 2019. *New Amazon Alexa Skill Kit 2.0.* https://developer.amazon.com/blogs/alexa/post/1a4e8b01-663d-4680-8efd-c28e96e31655/now-available-version-2-of-the-ask-software-development-kit-for-java. Accessed 20 November 2018.

Chi, M. T. H., and Wylie, R. 2014. "The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes," *Educational Psychologist* (49:4), pp. 219–243.

Chung, H., Iorga, M., Voas, J., and Lee, S. 2017. "Alexa, Can I Trust You?" *Computer* (50:9), pp. 100–104.

Cohn, M. 2004. *User stories applied*: *For agile software development:* Addison-Wesley Professional.

Cooper, H. M. 1988. "Organizing knowledge syntheses: A taxonomy of literature reviews," *Knowledge, Technology & Policy* (1:1), pp. 104–126.

Elaine, K. 2015. "Design of a domain-independent, interactive, dialogue-based tutor for use within the GIFT framework," in *Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym3),* Orlando, USA, p. 161.

Fast, E., Chen, B., Mendelsohn, J., Bassen, J., and Bernstein, M. 2018. "Iris: A Conversational Agent for Complex Tasks," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Montreal, Canada: ACM, p. 473.

Glasersfeld, E. v. 1987. "Constructivism," *The concise Corsini encyclopedia of psychology and behavioral science* (6), pp. 19–21.
Gregor, S., and Jones, D. 2007. "The anatomy of a design theory," *Journal of the Association for Information Systems* (8:5), p. 312.
Guo, Y. R., and Goh, D. H.-L. 2015. "Affect in Embodied Pedagogical Agents: Meta-Analytic Review," *Journal of Educational Computing Research* (53:1), pp. 124–149.
Hevner, A. R. 2007. "A three cycle view of design science research," *Scandinavian journal of information systems* (19:2), p. 4.
Hobert, S., and Wolff, R. M. von 2019. "Say Hello to Your New Automated Tutor–A Structured Literature Review on Pedagogical Conversational Agents,"
ICEF 2019. *The year in MOOCs: Increase revenue and more degrees*. http://monitor.icef.com/2019/01/the-year-in-moocs-increased-revenue-and-more-degrees/.
Kim, M. C., and Hannafin, M. J. 2011. "Scaffolding 6th graders' problem solving in technology-enhanced science classrooms: a qualitative case study," *Instructional Science* (39:3), pp. 255–282.
Kim, Y., and Baylor, A. L. 2016. "based design of pedagogical agent roles: A review, progress, and recommendations," *International Journal of Artificial Intelligence in Education* (26:1), pp. 160–169.
Lajoie, S. P., and Azevedo, R. 2006. "Teaching and learning in technology-rich environments," *08058493*.
Lehmann, K., Söllner, M., and Leimeister J. M. 2016. "Design and Evaluation of an IT-based Peer Assessment to Increase Learner Performance in Large-Scale Lectures," in *International Conference on Information Systems (ICIS),* Dublin, Ireland.
Nwana, H. 1990. "Intelligent tutoring systems: an overview," *Artificial Intelligence Review* (4:4).
Nye, B. D., Pavlik, P. I., Windsor, A., Olney, A. M., Hajeer, M., and Hu, X. 2018. "SKOPE-IT (Shareable Knowledge Objects as Portable Intelligent Tutors): overlaying natural language tutoring on an adaptive learning system for mathematics," *International journal of STEM education* (5:1), p. 12.
Oeste, S., Lehmann, K., Janson, A., Söllner, M., and Leimeister, J. M. 2015. "Redesigning University Large Scale Lectures: How To Activate The Learner," in *75th Academy of Management Annual Meeting*, Vancouver, British Columbia.
Pais, S., Casal, J., Ponciano, R., and & Lourenço, S. 2015. "Unsupervised assistive and adaptive intelligent agent in smart environment," in *Intelligent Environments and System Conferences ICIES,* Paris, France.
Pearl, C. 2016. *Designing Voice User Interfaces: Principles of Conversational Experiences:* " O'Reilly Media, Inc.".
Sonnenberg, C., and Vom Brocke, J. (eds.) 2012. *Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research,* Springer.
Sottilare, R. A., Graesser, A. C., Hu, X., Olney, A., Nye, B., and Sinatra, A. M. 2016. *Design Recommendations for Intelligent Tutoring Systems: Volume 4-Domain Modeling:* US Army Research Laboratory.
Sottilare, R. A., and Proctor, M. 2012. "Passively Classifying Student Mood and Performance within Intelligent Tutors," *Educational Technology & Society* (15:2), pp. 101–114.
Statista 2019a. *US Voice Technology Usage by Age*. https://www.statista.com/statistics/879907/us-voice-technology-usage-age/. Accessed 1 May 2019.
Statista 2019b. *Virtual Digital Assistant Market Worldwide*. https://www.statista.com/statistics/589079/worldwide-virtual-digital-assistants-consumer-market/. Accessed 1 May 2019.
Vanlehn, K. 2011. "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems," *Educational Psychologist* (46:4), pp. 197–221.
Venable, J., Pries-Heje, J., and Baskerville, R. 2012. "A Comprehensive Framework for Evaluation in Design Science Research," *Design Science Research in Information Systems. Advances in Theory and Practice* (7286:1), pp. 423–438.
Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., and Cleven, A. 2015. "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research," *CAIS* (37), p. 9.
Woolf, B. P., Aïmeur, E., Nkambou, R., and Lajoie, S. (eds.) 2008. *Intelligent Tutoring Systems*, Berlin, Heidelberg: Springer Berlin Heidelberg.
Xu, J. D., Benbasat, I., and Cenfetelli, R. T. 2014. "The Nature and Consequences of Trade-off Transparancy in the Context of Recommendation Agents," *MIS quarterly* (38:2).