Towards Empowering Educators to Create their own Smart Personal Assistants

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

Matthias Söllner
University of Kassel
Henschelstr. 4,
34127 Kassel
soellner@uni-kassel.de

Abstract

Despite a growing body of research about the design and use of Smart Personal Assistants such as Amazon's Alexa or Google's Assistant, little is known about their ability to help educators offering individual support in large-scale learning environments. Smart Personal Assistant ecosystems empower educators to develop their own agents without deep technological knowledge. The objective of this paper is to design and validate a method that helps educators to create Smart Personal Assistants for their learning environments. Using a design science research approach, we first gather requirements from students and educators as well as from information systems and education theory. Next, we create an alpha version of our method and evaluate it with a focus group before we instantiate our artifact in an everyday learning environment. The findings indicate that our method is able to empower educators to design Smart Personal Assistants that significantly improve students' learning success.

1. Introduction

New developments, such as ever-growing classrooms at high schools, large-scale lectures at universities with more than 100 students per lecturer, and massive open online courses (MOOCs) with more than 1,000 students on average, are increasingly entering the educational landscape [25]. For example, in 2018, over 900 universities around the world had announced 2,000 new MOOC courses and this trend is likely to continue [30]. These learning environments make it possible to offer top courses to a wider audience of students. However, this positive development also brings new challenges. In these learning scenarios, educators are hardly able to offer individual support to their students due to financial and organizational restrictions [20]. In contrast to this, predominant constructivistic learning theories tell us that individual support is crucial for learning success [12]. For example, Vygotsky's social development theory argues that social interaction plays a fundamental role in the process of cognitive development [36]. This means that effective learning processes occur when educators individually interact with their students.

Information technology (IT) has often tried to address the problem of offering individual support. Especially the research stream of intelligent tutoring systems (ITS) has been extensively investigating the value of IT to improve students' learning outcomes over the last thirty or forty years [23]. A review on the effectiveness of ITS by VanLehn et al. [32] states that the effect size of ITS is 0.76, nearly as effective as human tutoring with 0.79. Despite its proven usefulness, until now, existing ITS have still found too little entry into learning environments [24]. One reason might be that developing and introducing ITS in learning environments often requires a lot of technological knowledge and a lot of time [10]. Moreover, ITS applications are often complex to transfer from one context to the other (e.g., different learning goals, tasks, etc., [37]). Furthermore, ITS are often standalone software systems designed for certain devices, which hinders students to use them on their everyday devices (smartphones, tablet, PC, [29]). 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 natural language [26]. 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 hosted as a cloud service that processes voice data (converting voice-to-text, performing linguistic context analysis, and providing answers to questions, [7]). The popularity of SPAs has been steadily growing over the past few years [2].

Globally, the market of SPAs is predicted to increase from 3.0 billion U.S. dollars in 2017 to 15.7 billion U.S. dollars in 2021 across different sectors [34]. 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 [31]. SPA providers such as Google, Amazon, and IBM offer large ecosystems 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 new emerging SPAs in learning environments has four major advantages. First, it allows educators to develop SPAs for their own learning environments without being dependent on the technological knowledge of software designers or having to invest a high amount of time. Second, SPA ecosystems offer most on-the-edge technology necessary to bring students into an interactive learning mode that is hardly possible with existing systems. Third, it allows educators to easily convert the use of SPAs to different contexts (e.g., different learning goals, contents, tasks). And finally, SPAs are becoming daily companions of students in their private and school life, being integrated in devices they use everyday (smartphone, tablet, etc.). Hence, students are more motivated to use SPAs compared to other standalone systems. Until now, past research in the field of ITS mainly focused on design knowledge for software designers [4]. Moreover, recent publications regardings SPAs in learning environments focused on specific implementations and miss to provide transferable insights and in-depth evaluations [3]. For example, Arend [3] developed a specific SPA with the help of Apple's Siri in a secondary school class for a specific homework assignment without comparing the learning success with a control group. Little is known about how educators can design effective SPAs for their own learning environments [15]. Hence, the objective of this study is to create a method that allows researchers and educators to design effective SPAs as learning tutors to address the problem of individual support in large-scale learning environments. We thereby follow the calls of Hobert and Wolff [15] and Kim and Baylor [17] stating that there is a need for generalized design knowledge for SPAs in education. Consequently, we address the following research questions:

RQ1: How can educators design Smart Personal Assistants in order to improve students' learning outcomes?

RQ2: Do students using Smart Personal Assistants perform better in final exam-related assignments than those learners who do not use them?

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 [6] as our kernel theory. To answer our research questions, we present three cycles of a design science research (DSR) approach as described by Hevner [14]. In our *first cycle*, we derive requirements for our Smart Personal Assistant for Education Method (SPAEM) from literature and conduct qualitative interviews with educators and students. Based on that, we formulate the first version of SPAEM and evaluate it with the help of a focus group. In our *second cycle*, we refine the method and use it to build our first instantiation of a SPA, and evaluate its usefulness with the help of a quasi field experiment in an everyday learning environment. In the *third cycle*, we explain SPAEM to a selection of educators to proof its use for practitioners. The method should help educators to create SPAs that help them to bring students into an interactive learning mode, a behavior that is regarded as gold standard according to the ICAP-framework. 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. Next, we explain the research methodology and will elaborate on the development of SPAEM in detail. Finally, we discuss the results and end with limitations and a future outlook.

2. Theoretical Background

2.1 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 natural language [26]. 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 [26]. SPAs can be divided into two types: (1) built-in SPAs that use multi-purpose devices and (2) stand-alone SPAs that use dedicated devices. Examples of built-in SPAs include Siri (for Apple products) and Cortana (for Windows-based PCs). Examples of stand-alone SPAs include Alexa (that uses Echo, Echo Dot, and Tab dedicated devices) and Google Assistant (that uses Google Home dedicated devices, [7]). SPAs are able to react to users' utterances and can proactively guide users through a complex task. For example, Fast et al. [11] introduced a SPA into their learning environment to support users in conducting data science tasks (e.g., a predictive modeling task).

In the field of education, using computer tutor systems is not new. Researchers started to develop intelligent tutoring systems (ITS) as early as the 1970s in order to offer individual support to students [23]. Despite its proven effect on learning success, the distribution of todays' ITS is still very limited because it requires a lot of technological know-how and takes a lot of 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 [4]. In contrast to that, SPA providers such as Amazon are offering easy-to-use toolkits with a lot of blueprints and tutorials allowing users to build their own skills for SPAs with very basic technological know-how and low time effort [22].

2.2 ICAP – Framework as Kernel Theory

The ICAP framework proposed by Chi and Wylie [6] 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 and **p**assive. It 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 [18]. 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 to self-explain the concepts in the video (constructive). The most effective student behavior would be if students discuss the content and its justifications with a peer or tutor (interactive). Properly designed SPAs as learning tutors have the capabilities to build up dialogs with students, bringing them from a passive to an interactive learning mode [6]. Based on that, the implementation of SPAs in learning environments save a lot of human resources and are a promising solution to enrich current (online) learning environments.

3. Research Methodology

We follow a design science research (DSR) approach to answer our research questions. In specific, we rely on Hevner's three cycle view to structure the research process (see Figure 1). This approach is well suited for our project, because of its three-cycle view. The relevance cycle connect the application domain of the research project with our design science activitities. The rigor cycle makes sure that the design science activities consider the existing knowledge base of scientific foundations, experience, expertise. The central design cycle iterates between the core activities of building and evaluating our design artifact [14].

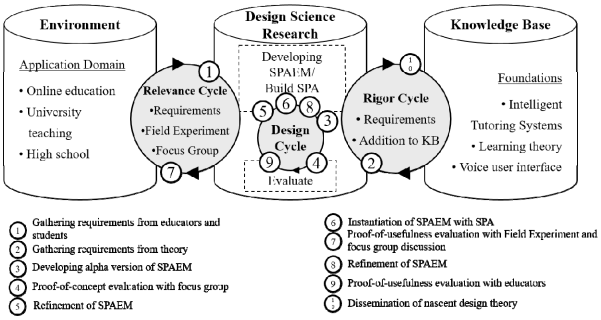


Figure 1. Our DSR process [14]

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 theory using three different perspectives. In step 3, we initiate the *design cycle* by developing the alpha version of SPAEM. In step 4, we conduct a first proof-of-concept evaluation with the help of a student focus group discussion. This evaluation helps to test our first design hypothesis: *SPAEM is logically derived from the requirements [28].* In step 5 and 6, we refine SPAEM and instantiate it by creating and introducing a SPA for a real learning environment. 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 SPAEM by testing our second design hypothesis: *SPAs created with SPAEM are better than existing learning aids in practice.* In step 8 and 9, we further refine SPAEM and explain it to educators. With that, we test our third design hypothesis: *SPAEM is useful for practitioners.* Last but not least, we discuss our results and disseminate a nascent design theory.

4. Designing Smart Personal Assistant for Education Method

4.1 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). The main goal of our method is that mainstream educators, after a short training period in the method and its description, are able to create a SPA for their own learning environment. In a first step, we want 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 high schools and universities. Moreover, we wanted to gather a basic understanding of educators' needs when developing a SPA for their learning environments. Thus, we conduct four interviews with educators. Table 1 shows the characteristics of the interview partners. All the interviews are semi-structured and lasted approx. 30-minutes each. We used a paper-based mock-up of a SPA dialogue for students and a mock-up of SPAEM for educators.

Table 1. Sample characteristics

Interviewees	Characteristics
Students (N=20)	**Average Age:** 21.35, **level of education:** 5 x high school, 10 x bachelor, 5 x master, **nationality:** 8 x Swiss, 11 x Germany, 1 x Italy, **gender:** 8 women, 12 men
Educators (N=4)	**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 course, 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 in regard to individual support? Can you explain it based on concrete examples? In the second part, we showed students a mock-up based SPA on a piece of paper demonstrating a student-SPA learning dialogue. In contrast, we showed educators a first draft of a SPA development method. Based on that, we asked students about possible requirements for using SPAs and educators for requirements of SPAEM itself. We transcribed the interviews and analyzed them using the method of user stories proposed by Cohn [8]. We identified user stories in the interviews, coded, clustered and finally translated the user stories into requirements for SPAEM. In the interviews, students have often complained about a lack of individual support during the learning phases. Educators have often complained that they would need too much time to deploy a technology solution for their learning environment. See Table 2 for more detailed user-stories and requirements.

Table 2. User-Stories and requirements from students and educators

User Stories Students (USS)	Requirements from Practice (RP)
USS1: As a *student*, I want the SPA to give me immediate and detailed feedback on my responses.	**RP1:** SPAEM should help educators to include an evaluation mechanism for students' utterances.
USS2: As a *student*, I want the SPA to reply as fast as a human in a conversation.	**RP2:** SPAEM should help educators in choosing a SPA platform with 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:** SPAEM should help educators to build SPAs that are available at any time, at any place.
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:** SPAEM should help educators to build voice-based SPAs
User Stories Educators (USE)	
USE1: As an *educator*, I want to first analyze for which learning goals, contents and methods I want to use the SPA.	**RP5:** SPAEM should include a step where educators can analyze the intended use scenario.
USE2: As a*n educator*, I want to design SPAs by myself without needing a lot of technological knowledge.	**RP6:** SPAEM should help educators to choose SPA ecosystems with easy-to-use tutorials and blueprints that allow educators to develop SPAs on their own.
USE3: As a*n educator*, I want that the SPA can be used on different devices (e.g., smartphones, tablets, laptops).	**RP7:** SPAEM should help educators to create SPAs that are accessible via different devices.
USE4: As a*n educator*, I first want to test the SPA by myself before I use it in my classes.	**RP8:** SPAEM should include a step where educators can iteratively test and adapt their initial version of a SPA.

4.2 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 [9] and vom Brocke et al. [35]. 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 SPAs in learning environments. Furthermore, our goal is to identify requirements on a conceptual level with a focus on an espousal of position and a representative coverage [9]. Regarding step 2 (conceptualization of the topic), we identified three perspectives useful for deriving requirements for SPAEM: 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 [19]. 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 "method". 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 systematically displayed the development process of the SPAs and similar computer tutors in order to make sure that we derive proper requirements for SPAEM. Based on that, we selected 36 papers. Regarding step 4 (literature analysis), we scanned the 36 papers regarding necessary steps to include in developing SPAs and clustered similar requirements resulting in four requirement clusters. Table 3 depicts the perspectives, an exemplary paper, and the requirement clusters (RT).

Table 3. Requirements from theory

Perspectives	Exemplary Paper	Requirements from Theory (RT)
Educational	**ICAP-Framework** [6]	**RT1:** SPAEM should help educators to include a proactive and reactive interaction logic, where the SPA and the learner both make a contribution.
	Scaffolding [16]	**RT2:** SPAEM should help educators to include different kinds of scaffolds when creating the interaction model.
Techn-ological	**Error handling** [13]	**RT3:** SPAEM should help educators to include error handling mechanisms to resolve errors and dysfunctions so that the user can continue.
Voice User Interface	**Engaging Response Structure** [27]	**RT4:** SPAEM should help educators to use an engaging response structure that is different to visual interfaces.

4.3 Step 3: Developing Alpha Version of SPAEM based on Requirements

Based on our identified requirements from practice and theory, we developed our initial version of SPAEM. The main goal of our artifact SPAEM is to help educators to create SPAs for their own learning environments without much technological knowledge and with a short training period in the method and its description. The development of SPAEM follows the definition of a method according to Brinkkemper [5]: *"A method is an approach to perform a systems development project, based on a specific way of thinking, consisting of directions and rules, structured in a systematic way in development activities with corresponding development products."* When designing the method, we tried to bring all the requirements into a sense-making order. We used grey circles for representing development activities and grey rectangles for development products. Figure 2 shows SPAEM and indicates the addressed requirements (small grey circles). The white rectangles show the main questions in the respective steps and the dashed figures show which process steps were added after the first evaluation. In step 1, the educators need to analyze their learning environment and decide for which learning goals, content and method they want to use the SPA. This results in a determination of the use scenario. For example, an educator might want to use the SPA for helping students to repeat English vocabulary as preparation for an exam. This decision has to be made by educators before they decide which platform to choose. In step 2, the educator needs to choose a SPA ecosystem (e.g., Amazon's Alexa) based on criteria such as interaction mode (voice-based and/or text-based), accessibility (smartphone and/or laptop), etc. For example, when the SPA should be text

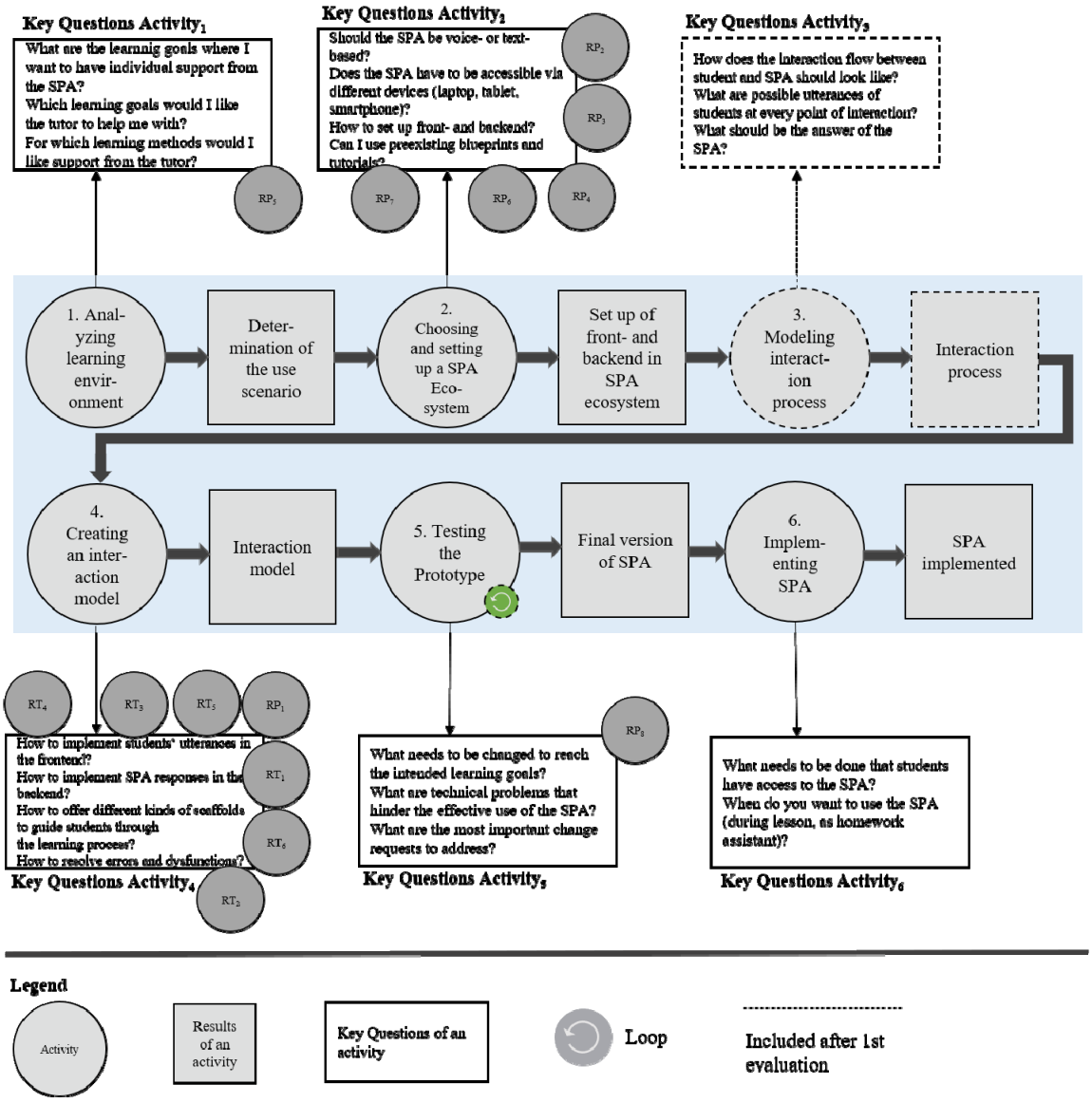


Figure 2. Initial and second version of SPAEM

and voice-based, Amazon's Alexa SPA ecosystem might be suitable. Next, the educators need to set up the front- and backend on their preferred SPA ecosystem. SPA providers, such as Amazon's Alexa or Google's Assistant, offer step-by-step video tutorials on how to set up the front- and backend and the connection between those two. Educators should watch these tutorials and try to imitate every step until the platform is ready. They can also use preexisting blueprints or tutorials for that step (e.g., Alexa's Quiz Game Skill Template [1]). For setting up the SPA ecosystem, almost no technological knowledge is required. The setting up of the ecosystem serves as a basis for the later content of the SPA. In step 3, the educators model the desired interaction process between the student and the SPA and add possible students' utterances and SPA responses at each point of interaction. For modeling the interaction process, educators can use prototyping tools or simply a piece of paper where they write down which statements the SPA should make and where they try to anticipate what the students will say. For example, educators specify which kind of vocabulary the SPA should ask the students and which kind of student' answers might be wrong or right.

In the next step, the educators integrate the interaction process into the front- and backend. In the front-end, they add possible students' intents and in the back-end they set the rules for the intents. For example, they can state that when a student says X, the SPA should say

whether the answer is right or wrong. In step 5, after creating the interaction model, the educator starts to test the initial prototype internally or externally with a few students. In step 6, after some loops of adapting and testing, the SPA can be be implemented in the learning environment.

4.4 Step 4 and 5: Proof-of-Concept Evaluation and Refinement of SPAEM

According to Sonnenberg and vom Brocke [28], 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. Our first evaluation concentrates on the constituents of the artifact. Focus group discussions are considered a suitable method for evaluating designs within DSR projects [28, 33]. The goal of the focus group discussion was to check if the translation of the requirements into the method make sense. Thereby, we tested our first design hypothesis: *SPAEM is logically derived from the requirements.* Moreover, we wanted to gather further change requests for improving SPAEM. 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 management 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 explained students our derived requirements and the corresponding SPAEM. We asked students to identify possible, logical breaks and to collect further change requests. The students suggested to add the loop sign to the prototype testing phase (step 5) to show that this step is iterative. Moreover, they mentioned that it is important to model the interaction process (step 3) first before technically creating the interaction model. Last but not leasts, they mentioned that step 2 "choosing and setting up a SPA-Ecosystem" is crititcal and should be described carefully so that the method can really achieve a benefit for the educator.

4.5 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 SPAEM (see Figure 2) and designing our first instantiation. We decided to instantiate SPAEM by creating a SPA for a real learning environment. We chose a real learning environment, because we wanted to proof that SPAEM and the resulting SPA is able to increase students' learning success. We were able to test the SPA in a second-grade class at a Swiss vocational business school. In specific, we used SPAEM to create a SPA that helped students to conduct their homework assignments over a period of 5 weeks and compared it with a control class that used only paper-based learning aids. The experiment class consists of 22 students and the control class of 23 students. The relevant subject was business and law.

In step 1 (analyzing learning environment), together with the educator we analyzed the learning goals, method and the content of the relevant subject. The main goal of the SPA was to help students with their homework assignments in order to increase their problem-solving skills in the relevant subject. In step 2 (choosing and setting up a SPA ecosystem), 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. Amazon's Alexa is accessible via different devices (smartphone, tablet, laptop) and offers voice and text interaction modes. We used the preexisting Template "Alexa Quiz Game Skill" on GitHub and followed the step-by-step tutorial.



Figure 3. Excerpt of an exemplary dialog between a student and the SPA.

First, we built the front-end by changing the already existing intents. Next, we connected the front-end with the back-end and used the provided code from the

template and adjusted the text according to our needs. For example, we changed the onboarding text and added information about the homework assignments. In step 3 (modelling interaction process), we modelled student and SPA dialogues and collected possible students' utterances and SPA responses for four homework assignments. An exemplary interaction process for homework assignment 2 can be seen in Figure 3 on the previous page. In step 4, we implemented the interaction process into Amazon's Alexa platform by creating intents (possible students' utterances) and the corresponding SPA utterances. In step 5, we tested the prototype internally and with the teacher of the second-grade school class and noted some change requests. Next, we tested the prototype with a couple of students and further refined the SPA. Last but not least, we introduced the SPA to the second-grade class and installed the SPA on students' devices.

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

The second evaluation concentrates on the usefulness of the artifact. We thereby test our second design hypothesis: *SPAs created with SPAEM are better than existing learning aids in practice.* Overall, 45 students participated in the quasi field experiment. Additionally, the experiment class participated in a focus group discussion afterwards to receive more insights into the usefulness of SPAs. 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 ensures that the two groups are equal regarding gender, age, pre-knowledge, pre-experience with SPAs, and personal innovativeness ($p < 0.05$). The experiment class (22 learners) used our developed SPA as learning tutor for their four 30-minute homework assignments. The control class used paper-based scaffolds as a baseline. Both learning aids had the same scope of information. 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 (Amazon's Alexa Echo Dot Device) in the experiment class and installed the Alexa software on their everyday devices (smartphone, tablet, laptop). Next, we asked them to use Alexa for conducting their homework 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, $p = 1.88e\text{-}06$, confidence intervals = 22.081 and 38.227, $N = 45$). Cohen's d is 1.70 (confidence intervals = 1.002 and 2.401). Since this effect is considered to be high, we can conclude, that our artifact proved to add value to this real learning environment.

Regarding the focus group data, several students mentioned that receiving challenging questions from the SPA helped them to think of the next solution steps. Some other learners perceived it as more entertaining compared to "business as usual" paper-based learning materials. Some students mentioned that it would be great if the SPA can remember students' knowledge gaps and then concentrates on helping them filling these gaps.

4.7 Step 8: Proof-of-Usefulness Evaluation with Educators

After proofing the usefulness of SPAEM in a real-life learning environment, there remained one design hypothesis to be tested: *SPAEM is useful for practitioners.* We tested our last design hypothesis with the help of semi-structured interviews with a selection of educators. We chose four educators from different school levels (2 x high school, 2 x university, teaching experience: 2 to 10 years, 2 educators with experience in online learning environments). We chose different educators regarding school level and experience with learning environments to identify particularities in using SPAEM for their purposes. The interviews were structured as follows. In the first part, we shortly explained SPAEM to the educators. The educators had no experience with SPAs before. In a next step, we asked them to explain every step in their own words and to note if something is unclear. In the second part, we asked them if they could imagine to use SPAEM and corresponding SPAs in their own learning environment. We analyzed the interviews with a qualitative content analysis by Mayring [21]. All of the educators were able to explain the most important sub steps of SPAEM. They agreed that SPAs can be helpful in offering individual support to their students. However, three out of four interviewees mentioned that step 2 "choosing and setting up a SPA-ecosystem" might be critical. Two educators emphasized that they wished to have a more fine-grained step-by-step guide in form of screenshots or videos. All of the interviewees mentioned that they could imagine to use SPAEM to create a SPA for their own learning environment. Furthermore, one educator suggested to collect created SPAs in a database and to share it with other educators so that they do not have to create SPAs from scratch.

5. Discussion

In this paper, we developed a method that help educators to create SPAs for their own learning environments. New emerging SPAs such as Amazon's Alexa and Google's Assistant promise to be a helpful enrichment for learning environments by offering individual support that otherwise would not be possible due to financial and organizational restrictions. SPAEM should help educators to address the problem of individual support in large-scale learning environments. We proposed three design hypotheses to test whether SPAEM is able to fulfill the overarching goal. Within our design hypothesis 1 *(SPAEM is logically derived from the requirements)*, we evaluated if we correctly derive parts of our method according to students' and educators' requirements. In design hypothesis 2 **(***SPAs created with SPAEM are better than existing learning aids in practice),* we created an actual SPA with the help of SPAEM and proofed that it was able to increase students' learning outcome. In design hypothesis 3 (*SPAEM is useful for practitioners),* we explained SPAEM to educators and asked them to check whether this method helps them to create their own SPAs.

In the remainder of this section we describe our contributions to research and practice, as well as our study's limitations and directions for future research.

Our work makes several contributions to research. First, it contributes to research regarding the ICAP-Framework by showing that educators can create IT-based systems that are able to bring students into an interactive learning behavior considered as gold standard. When students interact with SPAs, they dive deeper into the learning material which results in an increased learning outcome. Second, it contributes to Intelligent Tutoring System research by proposing a new method that helps educators to use the potential of new emerging SPAs as learning tutors. Until now, to the best of our knowledge, no such method for educators exist. Existing methods regarding the design of ITS addressed software designers. Consequently, there was a gap between the technological know-how of software designers and the pedagogical know-how of educators. As a result, ITS were not widespread enough in learning environments. Third, this method helps researcher to create new emerging SPAs for learning environments in order to further investigate their value compared to other kinds of learning aids.

Our work also has several implications for practice. First of all, our proposed method helps educators to offer individual support to students without increasing costs or time resources. Since there is a significant increase of online and distributed learning environments, SPAs as a new type of intelligent tutoring systems are more relevant than ever before. Second, our method helps educators to transfer SPAs from one context to the other without being dependent on the technological know-how of software designers.

5.3 Limitations and Future Outlook

A number of limitations have to be considered with respect to our study. First, the list of requirements are derived from specific theoretical perspectives and from a certain selection of students and educators. It is likely that the requirements would be different if we have used different perspectives and interview partners. However, we tried to include the most important theoretical perspectives and select a representative sample of interview partners. Second, we proofed the value of SPAEM and the corresponding SPA in a very narrow context (second-grade vocational business school). Moreover, this learning environment cannot really be considered as a large-scale environment. It might be that the experiment have let to different results in a different learning environment. Nevertheless, to proof the usefulness of SPAs in real learning environments, our setting seemed to be suitable. Furthermore, we were able to identify a large effect size which let us assume that SPAs have a positive effect on learning success. Last but not least, we used Amazon's Alexa platform to create our SPA. Other platforms might have let to different results.

For future research, it would be very interesting to explain our method to educators and to let them create a real SPA for their own learning environment. Especially the field of online and distributed learning seems to be a promising area for this new kind of technology.

References

[1] https://github.com/alexa/skill-sample-java-quiz-game/blob/master/instructions/1-voice-user-interface.md, accessed 8-26-2019.

[2] https://www.emarketer.com/Article/Alexa-Say-What-Voice-Enabled-Speaker-Usage-Grow-Nearly-130-This-Year/1015812, accessed 11-23-2018.

[3] Arend, B., "Investigating Siri as a virtual assistant in a learning context", in Proceedings of 12th annual International Technology, Education and Development Conference, Proceedings of 12th annual International Technology, Education and Development Conference 2018, Seville, Spain. 2018.

[4] Barbhuiya, R., K.M. Karim, and J. Suraiya, "Design specifications for a generic Intelligent Tutoring System", in Proceedings of the International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE), Proceedings of the International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE). 2011. The Steering Committee of The World Congress in Computer

Science, Computer Engineering and Applied Computing (WorldComp): Las Vegas, USA.
[5] Brinkkemper, S., "Method engineering: engineering of information systems development methods and tools", Information and Software Technology, 38(4), 1996, pp. 275–280.
[6] Chi, M.T.H. and R. Wylie, "The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes", Educational Psychologist, 49(4), 2014, pp. 219–243.
[7] Chung, H., M. Iorga, J. Voas, and S. Lee, "Alexa, Can I Trust You?", Computer, 50(9), 2017, pp. 100–104.
[8] Cohn, M., User stories applied: For agile software development, Addison-Wesley Professional, 2004.
[9] Cooper, H.M., "Organizing knowledge syntheses: A taxonomy of literature reviews", Knowledge, Technology & Policy, 1(1), 1988, pp. 104–126.
[10] Elaine, K., "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), Generalized Intelligent Framework for Tutoring (GIFT) Users Symposium (GIFTSym3), Orlando, USA. 2015.
[11] Fast, E., B. Chen, J. Mendelsohn, J. Bassen, and M. Bernstein, "Iris: A Conversational Agent for Complex Tasks", in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018. ACM: Montreal, Canada.
[12] Glasersfeld, E.v., "Constructivism", The concise Corsini encyclopedia of psychology and behavioral science, 6, 1987, pp. 19–21.
[13] Guo, Y.R. and D.H.-L. Goh, "Affect in Embodied Pedagogical Agents: Meta-Analytic Review", Journal of Educational Computing Research, 53(1), 2015, pp. 124–149.
[14] Hevner, A.R., "A three cycle view of design science research", Scandinavian journal of information systems, 19(2), 2007, p. 4.
[15] Hobert, S. and R.M. von Wolff, "Say Hello to Your New Automated Tutor–A Structured Literature Review on Pedagogical Conversational Agents", 2019.
[16] Kim, M.C. and M.J. Hannafin, "Scaffolding 6th graders' problem solving in technology-enhanced science classrooms: a qualitative case study", Instructional Science, 39(3), 2011, pp. 255–282.
[17] Kim, Y. and A.L. Baylor, "based design of pedagogical agent roles: A review, progress, and recommendations", International Journal of Artificial Intelligence in Education, 26(1), 2016, pp. 160–169.
[18] King, A., "From sage on the stage to guide on the side", College teaching, 41(1), 1993, pp. 30–35.
[19] Lajoie, S.P. and R. Azevedo, "Teaching and learning in technology-rich environments", 08058493, 2006.
[20] Lehmann, K., M. Söllner, and Leimeister J. M., "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. 2016.
[21] Mayring, P., "Qualitative content analysis: theoretical foundation, basic procedures and software solution", 2014.
[22] 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 11-20-2018.
[23] Nwana, H., "Intelligent tutoring systems: an overview", Artificial Intelligence Review, 4(4), 1990.
[24] Nye, B.D., P.I. Pavlik, A. Windsor, A.M. Olney, M. Hajeer, and X. Hu, "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), 2018, p. 12.
[25] Oeste, S., K. Lehmann, A. Janson, M. Söllner, and J.M. Leimeister, "Redesigning University Large Scale Lectures: How To Activate The Learner", in 75th Academy of Management Annual Meeting, 75th Academy of Management Annual Meeting. 2015: Vancouver, British Columbia.
[26] Pais, S., J. Casal, R. Ponciano, and S. & Lourenço, "Unsupervised assistive and adaptive intelligent agent in smart environment", in Intelligent Environments and System Conferences ICIES, Intelligent Environments and System Conferences ICIES, Paris, France. 2015.
[27] Pearl, C., Designing Voice User Interfaces: Principles of Conversational Experiences, " O'Reilly Media, Inc.", 2016.
[28] Sonnenberg, C. and J. Vom Brocke, eds., Evaluations in the science of the artificial–reconsidering the build-evaluate pattern in design science research, Springer, 2012.
[29] Sottilare, R.A. and M. Proctor, "Passively Classifying Student Mood and Performance within Intelligent Tutors", Educational Technology & Society, 15(2), 2012, pp. 101–114.
[30] http://monitor.icef.com/2019/01/the-year-in-moocs-increased-revenue-and-more-degrees/.
[31] https://www.statista.com/statistics/879907/us-voice-technology-usage-age/, accessed 5-1-2019.
[32] Vanlehn, K., "The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems", Educational Psychologist, 46(4), 2011, pp. 197–221.
[33] Venable, J., J. Pries-Heje, and R. Baskerville, "A Comprehensive Framework for Evaluation in Design Science Research", Design Science Research in Information Systems. Advances in Theory and Practice, 7286(1), 2012, pp. 423–438.
[34] https://www.statista.com/statistics/589079/worldwide-virtual-digital-assistants-consumer-market/, accessed 5-1-2019.
[35] Vom Brocke, J., A. Simons, K. Riemer, B. Niehaves, R. Plattfaut, and A. Cleven, "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research", CAIS, 37, 2015, p. 9.
[36] Vygotsky, L.S., Mind in society: The development of higher mental process, Cambridge, MA: Harvard University Press, 1978.
[37] Woolf, B.P., E. Aïmeur, R. Nkambou, and S. Lajoie, eds., Intelligent Tutoring Systems, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.