Please quote as: Tolzin, A.; Körner, A.; Dickhaut, E.; Janson, A.; Rummer, R.; Leimeister, J. M. (2023). Designing Pedagogical Conversational Agents for Achieving Common Ground. International Conference on Design Science Research (DESRIST). Pretoria, South Africa.

Designing Pedagogical Conversational Agents for Achieving Common Ground

Antonia Tolzin[1][0009-0005-4110-4361], Anita Körner[1] [0000-0003-3761-2118], Ernestine Dickhaut[1] [0000-0002-1946-4038], Andreas Janson[2] [0000-0003-3149-0340], Ralf Rummer[1] [0000-0003-3761-2118] and Jan Marco Leimeister[1,2] [0000-0002-1990-2894]

[1] University of Kassel, Kassel, Germany
{antonia.tolzin, anita.koerner, ernestine.dickhaut, rummer, leimeister}@uni-kassel.de
[2] University of St. Gallen, St. Gallen, Switzerland
{andreas.janson,janmarco.leimeister}@unisg.ch

Abstract. As educational organizations face difficulties in providing personalized learning material or individual learning support., pedagogical conversational agents (PCAs) promise individualized learning for students. However, the problem of conversational breakdowns of PCAs and consequently poor learning outcomes still exist. Hence, effective and grounded communication between learners and PCAs is fundamental to improving learning processes and outcomes. As understanding each other and the conversational grounding is crucial for conversations between humans and PCAs, we propose common ground theory as a foundation for designing a PCA. Conducting a design science research project, we propose theory-motivated design principles and instantiate them in a PCA. We evaluate the utility of the artifact with an experimental study in higher education to inform the subsequent design iterations. We contribute design knowledge on conversational agents in learning settings, enabling researchers and practitioners to develop PCAs based on common ground research in education and providing avenues for future research. Thereby, we can secure further understanding of learning processes based on grounding communication.

Keywords: common ground, conversational agent, design science research, education

1 Introduction

Software-based dialogue systems, known as conversational agents (CAs), are common in our everyday lives as they enable communication between humans and computers and aim to simulate human conversations [1]. These dialogue systems show "humanlike behavior" and interact with users through natural language [2]. In the domain of higher education, pedagogical conversational agents (PCAs) can make it easier for learners to study independently at any given place or time. Digital education can be extended to a large audience but in consequence the interaction between learner and educator decreases and it is challenging to replace this interaction. Here, PCAs

may mitigate this problem by providing natural interaction to support learners similar to educators [3, 4]. New digital tools, like PCAs, are increasing didactical possibilities and can improve learning processes and outcomes [5, 6].

Apparently, the perception of CAs in general is shifting from tools to teammates [7]. Therefore, collaboration between humans and CAs is gaining more importance, indicating the need for investigating their conversational aspects for successful interactions. As conversational interactions are fragile, especially in learning situations that mimic teacher-student interactions, they can easily fail after misunderstandings during the dialogue and impact the effectiveness between learners and PCAs [8, 9]. The most common reason for conversational breakdowns is a collapse of the natural language processing and the interpretation by the PCA [10]. PCAs sometimes provide none, wrong, or incomprehensible responses which leads to discomfort, annoyance, and questioning of the capabilities of the PCA, which can end in usage discontinuance and disruption of learning processes and outcomes [11, 12]. Thus, improving the communication between learners and PCAs and a common ground about the learning material is an important goal for researchers and practitioners [13].

Dialogue understanding is an inherently interactive process and understanding each other and the anticipated conversational grounding is a key element for language-based interactions between humans and computers [14]. The interactive process to achieve dialogue understanding is called conversational grounding and describes the coordinative process of dialogue partners to establish a shared understanding also known as common ground [15], which is fundamental for fruitful communication and successful learning. Following, we define common ground as a shared understanding resulting from a coordination process between conversational partners [16]. Nevertheless, identifying the needs and capabilities in human-computer interaction (HCI) and developing presumptions about what the PCA can do and understand is a great challenge [17]. Conversely, it is difficult for programmers and system designers to guess, how the human part of the dialogue will act [15, 16, 18]. While there is ample discussion and research on the *why* of grounding and shared understanding, there is little literature and a lack of research about the *how* [14, 19]. With successful CA interaction through effective conversations based on grounding in the domain of higher education, learning outcomes could be increased [14]. Consequently, this study aims to contribute to improving our understanding of common ground in human-agent interaction and to design components to increase common ground in educational contexts by specifically exploring the following research question:

RQ: How can we design a PCA that builds common ground and improves learning outcomes of students?

The goal of our research is to present a theory-driven design approach to provide a set of design principles to achieve common ground when utilizing a PCA for students in higher education. Hence, we implemented and evaluated a PCA for higher education. To achieve this research goal, we follow the design science approach [20, 21]. Therefore, our research provides a theory of design and action as an improvement for known problems [22]. We designed a prototypical PCA prototype and evaluated the first instantiation in a fully randomized field experiment.

2 Theoretical Background

2.1 Pedagogical Conversational Agents

CAs include all types of software that allows people to have a conversation with a computer and have a long history, with memorable representatives like ELIZA, ALICE, Claude, and HeX. They are part of the educational domain as PCAs since the early 1970s [23]. CAs can include voice as an interaction channel [24], for example Amazon´s Alexa, and typically make use of natural language interfaces and machine learning techniques, which allow them to take on tasks more successfully, assisting the users [25]. In contrast, text-based CAs are based on a set of established rules or flow to react to specific queries posed by users. As interactions between learners and PCAs are usually textually mediated [26], we focus on text-based PCAs in this study.

PCAs show great potential to transform education by individualizing and personalizing learning processes, supporting educators, giving insights into learners´ behavior, and engaging learners [3]. This is why we can observe an increasing interest in PCA research in education [27]. PCAs can be seen as a type of intelligent tutoring systems, where they interact with learners through natural language conversations [27, 28]. Through PCAs it is possible to provide learning support to all learners in a personalized way, which is a crucial part of the individualization of learning processes. PCAs can help to keep up motivation during the learning process throughout the interaction, they can improve meta-cognitive skills and help the learner to structure their knowledge actively [29]. In large online classes and where personalized support from educators to learners is not possible, PCAs can facilitate learning, for example by promptly providing students with rehearsal questions [30], assignments [31], course content [32], and study resources [33].

The design of PCAs includes considerations of cognitive, social, emotional and pedagogical elements [34]. Kuhail et al. [3] discovered principles used to design chatbots in education. Most chatbots included personalized learning, some included experimental learning, social dialogue, and collaborative learning, and only few studies included affective learning, learning by teaching, and scaffolding. They can be used in a broad application area because they can act in different human roles like teaching agents, peer agents, teachable agents, and motivational agents [35, 3].

2.2 Common Ground as a Kernel Theory

It is well investigated that senders adapt to the supposed needs and capabilities of the recipient in a dialogue [36]. That is why people speak differently to children, friends, foreigners, and colleagues. Dialogue understanding is an interactive process, aiming to resolve misunderstandings and building shared understanding. Especially, in learning contexts, this shared understanding is crucial to gain procedural and factual knowledge [37]. The process can be improved through additional signals of mutual understanding to individual turns [36]. Thus, we introduce common ground as a kernel theory for our design science research endeavor. We define common ground as a shared understanding resulting from a coordination process between conversational

partners [16]. Common Ground contains the background knowledge on which the communication planning of the conversational partners builds on and can be divided into the global knowledge (all knowledge about the conversational partner and their knowledge requirements) and the situational knowledge (knowledge about the mutual perception conditions and the communication protocol).

Grounding processes lead to a coordination of the background knowledge of the conversational partners [36, 16]. Identifying the needs and capabilities in the context of HCI and developing presumptions about what the PCA can do and understand is a great challenge for most people. The dialogue must therefore be appropriate for the conversational setting including conversational partners, prior dialogue, task context, and lexical context. It is crucial to acknowledge that the recipient has understood the sender [38] through different verbal and non-verbal information (e.g., saying "yes" or "okay", nodding) [39]. Verbal grounding does not bring new information or arguments to the conversation. It is more like a semantic mechanism to check that both dialogue partners received and understood the sender's contribution.

Dialogue partners have different ways to provide evidence of their understanding, which could be explicit acknowledgements (words), display of what has been understood, continued attention, and continuing with next steps [40]. Responses can get interconnected and contingent on what has been said previously by mutually grounding the conversational partner's input [41]. Successful grounding results in a shared context, guided comprehension, instant feedback, and enhanced processes in conveying intent [42]. In essence, we expect that the incorporation of common ground as a kernel theory for the design of a PCA is an important scaffold in the dialogical interaction with a learner. In the following, we outline this design process in detail.

3 Design of a PCA based on Common Ground

3.1 Research Method and Context

For the design of the PCA with the consideration of common ground, we draw on a theory-driven design approach [43, 44] following Peffers et al. [21] on engineering our IT artifact. Therefore, we base our subsequent design decisions on our phenomena of interest and the related theory concept (problem identification and motivation phase). We focus on common ground as an ancillary phenomenon and learning outcomes as the focal phenomenon. We claim that the successful usage of natural language processing in human-computer communication depends on a design that reflects human conversational grounding processes so that each contribution to the ongoing dialogue can be appropriately grounded by both the learner and the PCA (objectives of a solution). Grounding measures are associated with achieving common ground in a learning scenario when utilizing a PCA. Hence, we want to design a PCA that improves common ground and ultimately learning outcomes (design and development phase). For the further design of these grounding elements, we derive requirements from theory and practice [45]. In a first step, we derived requirements from common ground mechanisms and theories found in the literature (T). In a sec-

ond step, we derived requirements from learners. For this purpose, we presented a mockup of a chatbot learning tutor in a mobile application to 33 students and asked them for user requirements that also relate to the use and utility for learning with PCAs [46] (demonstration phase). The answers and requirements given by the students were clustered and requirements (L) were derived. Afterwards, we addressed all requirements by design principles that should influence the common ground with a PCA and linked phenomena like [47] learning outcomes. Ultimately, we present our implementation for the first design iteration that was subject for evaluation (see Section 4) in Western-European lecture (evaluation phase).

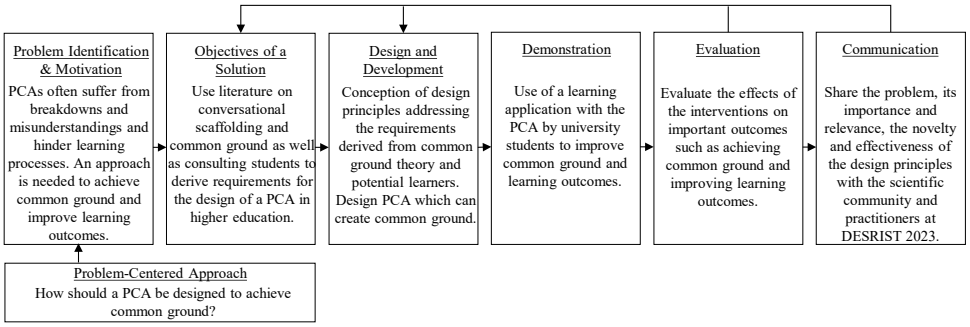

Fig. 1. Research Approach for PCA Development. Adapted from [21].

3.2 Requirements From Common Ground Theory and Practice

In this section, we will describe how we gathered the preliminary requirements from literature and practice to address the first two parts of the objectives of the solution phase of the design science approach of Peffers et al. [21]. The problem formulation (step one), described in the introduction, serves as the foundation for the derivation of the requirements. First, we derived requirements to design our PCA from common ground theory and literature.

Examining more closely the grounding processes, two different ways have been suggested as the mechanism responsible for grounding: Alignment [48] and complementarity [16, 49]. Alignment theory emphasizes increasing alignment and similarity between dialogue partners over time [48], which can be achieved through mimicking different language characteristics of each other (e.g., on phonetic, lexical, or syntactic levels) [14]. Studies proved that alignment could enhance human-computer interaction [50, 51] and surprisingly humans align more to computers than they align to other humans. In contrast, complementarity emphasizes the complementarity between dialogue partners. This theory assumes, that the mutual understanding is grounded in dissimilar contributions, which supportively relate to each other and therefore create a greater whole. Complementarity can be distinguished into two subtypes, interpersonal synergy and audience design. Rothwell [52] showed a consistent advantage of the two complementarity theories, considering that syntax seems to be essential and indicates evidence for the alignment model.

Table 1. Requirements from Theory and Practice.

Requirements from Common Ground Theory (T) and Learners (L)	**Common Ground Mechanisms**
T1) The PCA should be transparent concerning its purpose, goals, intentions, and abilities.	Social Features, Joint Action, Mental Model
T2) The PCA should elaborate, "close-the-loop" constantly regarding the shared goal and knowledge.	Interpersonal Synergy, Knowledge Base, Joint Action
T3) The PCA should check the understanding of communication and the learning material constantly.	Audience Design, Mental Model
T4) The PCA should integrate signals of mutual understanding by acknowledging that the learner has understood the sender and task.	Alignment, Social Features
T5) The PCA should make the background knowledge transparent and motivate the learner to do so.	Knowledge Base
L1) The PCA should indicate if a misunderstanding/breakdown is appearing.	Practice & Audience Design, Alignment, Social Features
L2) The PCA should inform the learner about what it can do (and what not) and what to expect from it.	Practice & Inoculation Theory, Mental Model

Furthermore, in a prior systematic review we identified five mechanisms to achieve common ground in the interaction with CAs [53]. Embodiment (1) describes the presentation of an identifiable conversational counterpart [54–56]. Learners are more willing to put effort into establishing common ground if the interaction resembles a human-to-human interaction. Social Features (2) are needed to show authenticity: the CA has to be transparent concerning its purpose, abilities, and knowledge, must be able to learn from experiences and remember the communication protocol [57–59]. Joint Action (3) describes that in a collective task the goals and intentions of the dialogue partners need to be salient. Common Ground is what both dialogue partners know in a transparent way [60, 61]. The Knowledge Base (4) refers to the knowledge organisation of the CA: To establish common ground, it is crucial that all dialogue partners suppose the others have access to the same information [62]. Lastly, it is desirable that learners develop a Mental Model of the CA (5) [63, 64]. This influences learners' expectations of the CA and whether the CA interaction is effective and therefore whether it is worth putting effort into establishing common ground. Hence, a shared mental model is a sign of common ground. Through these mechanisms, common ground can be achieved to improve learning outcomes.

Based upon the above-described theoretical tenets and the requirements derivation with learners, we identified seven requirements (T1-T5 and L1-2 in Table 1) to establish common ground with a PCA. The first requirement from theory **(T1)** refers to the transparent goals (see Joint Action, [53]), purpose, intention, and abilities of the PCA as these are typical social interaction elements and support the development of an appropriate mental model. The second requirement **(T2)** deals with the mechanisms of interpersonal synergy [52] and joint action [53] and proposes that common ground is based on complementary actions. A PCA must contain a constant elabora-

tion and references to the dialogue partners' contributions. During joint actions, common ground can be achieved by sending relevant information to the dialog partner, verifying what each partner knows, establishing or negotiating shared meaning, requesting information or repairing insufficiencies in shared knowledge. Moreover, it has been shown, that good teams "close the loop" more often than bad teams [65]. The third requirement **(T3)** refers to the mechanisms of audience design [52] and mental model of the PCA [53], emphasizing that the learner uses theory of mind and perceives that their dialogue partner possibly does not share their perspective and knowledge. The sender develops assumptions of their partner's perspective and can adjust their contribution. To achieve an internal representation of the dialogue partner's perspective, constant checking of each other's perspective is fundamental. The next requirement **(T4)** deals with the interactive alignment mechanism [48] [52]and social features [53], in which shared understanding results in imitation of the partner's language characteristics as a signal of mutual understanding by acknowledging this understanding, which is a key element of conversation. The last requirement **(T5)** is based on the mechanism of the knowledge base [53]. To achieve common ground, the PCA needs to adapt to the learner's level of knowledge and the level of common ground between them. Therefore, transparency of the partner's background knowledge is needed.

In addition to the theory requirements, we address the second part of the objectives of a solution phase of the design science approach by deriving requirements from Learners (L). In a large-scale lecture, we identified the needs, wishes, and expectations of 33 learners considering the PCA. The clustered answers and considerations of the students resulted in two requirements (L1 and L2 in Table 1). Students required information about the possible arising of a misunderstanding **(L1)**, which can be linked to the mechanisms of audience design, alignment and social features as constant checking of the understanding and indicating if something is going wrong is an essential part of successful human communication. The second requirement **(L2)** can be seen as a prevention strategy for possible conversational breakdowns as inoculation messages can prevent breakdowns [12]. By informing the learner about its abilities, the PCA can foster the learner's development of an appropriate mental model of the PCA.

3.3 Design Principles to Achieve Common Ground

In this section, we will describe how we derived the design principles (DPs) addressing the design and development phase of the design science approach. The goal is to develop DPs that consider common ground for PCA interactions as a key driver of learning processes. In Table 2 we describe four DPs derived from requirements from theory and learners and describe how these DPs were implemented in the PCA prototype. Figure 2 shows how the different DPs are instantiated in the evaluated PCA prototype. We argue that a PCA that instantiates our DPs improves common ground between learners and the PCA, therefore resulting in increased learning outcomes. To evaluate our design, we developed a PCA based on the derived DPs.

Table 2. Design Principles for Pedagogical Conversational Agents.

Title	Design Principle	Req.
DP 1) Active Communication Protocol	Provide a PCA that makes its knowledge transparent and is grounding the learner's contributions in order to interconnect conversational responses, so they become contingent on what had been said previously.	T3, T4, T5
DP 2) Introduction	Provide a PCA that initial informs the learner about what it can do and what to expect from it in order to make its purpose, (shared) goals, and abilities transparent and show authenticity.	L2, T1, T2
DP 3) Student Question Generation	Provide a PCA that instructs students to practice question generation to make students' knowledge transparent and show the boundaries of common ground.	L1, T2, T3, T5
DP 4) Re-check	Provide a PCA that includes agreements, questions, and rewordings to avoid misunderstandings, breakdowns, and make the knowledge of the dialogue partners salient.	L1, T2, T3, T5

The first design principle **(DP1)** considers an active communication protocol and the repetition of personalized information ("what was said previously") so that responses become contingent and the knowledge of the PCA is more transparent. This design principle was implemented through different answers of the PCA depending on the type of information (question, explanation, definition, example) required from the student and given before (see Figure 2). The second design principle **(DP2)** addresses the need for detailed instruction by the PCA with information about what it can do (and what not) and what learners can expect from it. In the introduction part of the conversation the abilities of the PCA are stated and a shared goal of the interaction is built. The PCA prototype started the conversation with an introduction of itself, in which its abilities (You can ask me questions and determine what information I give to you; My ability to understand natural language is unfortunately limited) and goal (equal understanding of what aspects of theories increase or decrease falsifiability) became transparent. The third design principle **(DP3)** refers to the importance of student question generation which has positive effects on learning [66]. Moreover, student question generation makes students' knowledge transparent and shows the boundaries of the common ground. At the end of every learning unit, the PCA prototype encouraged the students to generate a question before the next learning unit started. The last design principle Re-check **(DP4)** relates to the establishment of a common ground as a result of agreements, questions, and rewordings. This prevents misunderstandings and makes the knowledge of the dialogue partners salient. This DP was implemented through repeated questions about the deeper understanding of the learning material and joint agreements about the further course of the conversation and the given information.

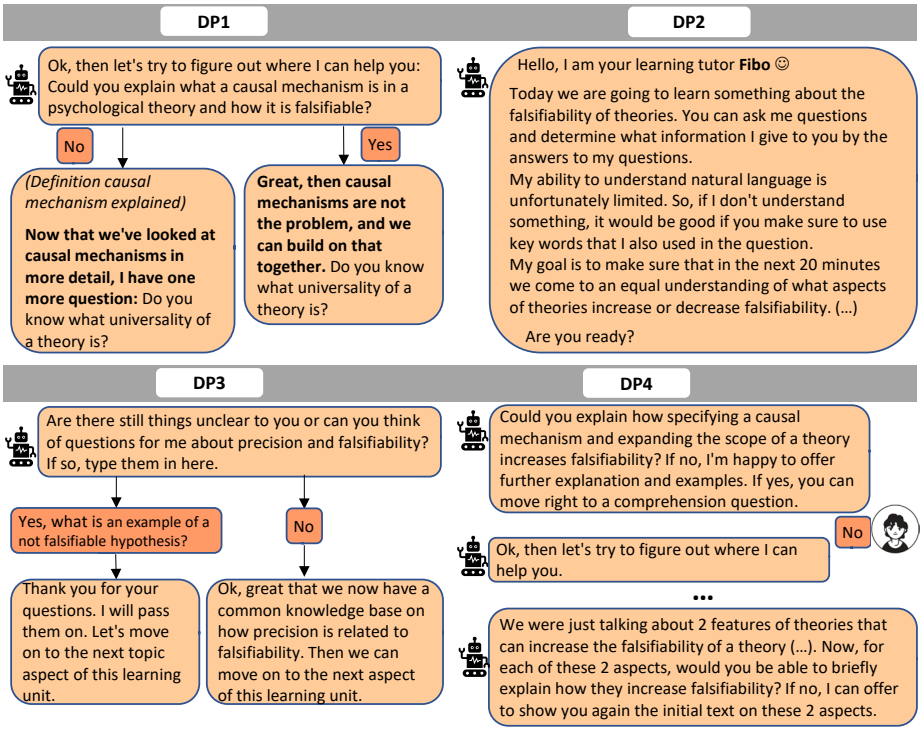

Fig. 2. Exemplary Conversations of the PCA based on the Design Principles.

4 Demonstration and Evaluation of the PCA

4.1 Demonstration

We demonstrate the prototypical instantiation in a fully randomized field experiment. In our evaluation, we followed the DSR evaluation frameworks proposed by [67] and adopt a formative ex-post evaluation approach [67, 68]. This approach will allow us to iteratively refine and improve the design of our method in the early stages of its development. We conducted an online experiment in a lecture at a Western European university. The students who participated majored in psychology, were enrolled in the course "General Psychology I", and were undergraduates (freshmen). The learning material (falsifiability of theories) presented by the PCA was part of the curriculum for this course. 71 students started the study. From these, we collected 54 valid data sets in total, of which only 33 participants completed all learning performance questions. Therefore, evaluations are based on 54 participants, and learning performance on 33 participants. The sample consisted of 24 female students and seven male students (and two who did not want to answer this question) with an average of 21 years.

The experiment consisted of three main parts: 1) a pre-test, 2) the interaction with the PCA, and 3) a post-test. The pre- and post-tests were identical for all partici-

pants. The pre-test took place before the use of the PCA and captures the student knowledge level of the learning material. The post-test was conducted at the end of the implementation phase to investigate common ground establishment and learning success. During the interaction with the PCA, the experimental group interacted with a PCA based on the four DPs presented to achieve common ground, while the control group interacted with a traditional PCA without grounding elements. To avoid any confounds, all learning material was the same across groups. The PCA instantiation was built with Google DialogFlow. Moreover, we kept the interaction style transactional and thus, language simple and design functional. Our aim was not to design a humanic CA interaction in order to investigate our hypotheses without potential confounds by the interaction design.

43 students were randomly assigned to one of two groups. To examine learning outcomes, learners' information retention and transfer ability levels were measured as this are typical, established learning outcome measures [69]. Common Ground was measured using a modification of the Five Factor Perceived Shared Mental Model Scale (5-PSMMS) [74], as there are no widely accepted measures of common ground, either among humans or between humans and PCAs, and the mental model is part of common ground [22]. Besides our key dependent variables, we assessed satisfaction with the PCA, subjective complexity of the learning material [70], judgments of learning [71], personal innovativeness [72], trust in PCA [73], and previous experience with PCAs.

4.2 Results

Previous experiences with PCAs, personal innovativeness and pre-test test performance did not differ between the participants using the common ground PCA and the control PCA, indicating that the two groups of participants were similar concerning their previous knowledge and relevant experiences. When evaluating their experience with the PCA, satisfaction and judgments of learning did not differ significantly between the two groups.

However, participants who used the common ground (vs. control) PCA spent more time with the PCA, $t(51) = 2.11$, $p = .02$, $d = 0.58$. Additionally, participants who used the common ground (vs. control) PCA evaluated the common ground between themselves and the PCA as higher, $t(50) = 1.73$, $p = .045$, $d = 0.46$. The common ground scale consisted of four subscales concerning shared representations concerning communication, task performance, mental models, and content of learning materials. Participants interacting with the common ground (vs. control) PCA especially judged their and the PCA's understanding of the content of the learning materials, such as central terms and concepts, to be more similar, $t(51) = 2.53$, $p = .007$, $d = 0.69$. However, this did not lead to increased trust in the PCA or greater liking of the PCA for the common ground compared to the control version. Concerning the learning success, there was a non-significant tendency towards better performance for participants with the common ground compared to the control PCA, $t(27) = 1.33$, $p = .097$, $d = 0.47$, indicating that with a larger sample size, the common ground PCA might facilitate learning success compared to a control PCA.

5 Contributions, Limitations, and Future Work

Common ground is a key element of human dialogue and a necessary capability for PCAs. Overall, this study should help designers to implement CAs, especially in educational contexts. As it is important to address the mechanisms of common ground, focus on domain (e.g., education) specific requirements for PCA design, and derive DPs, this study aims to fill this gap. We provide the first steps of designing a PCA to achieve common ground and improve learning outcomes. As our contribution, we provide a new solution to the problem of conversational breakdowns and unsatisfactory learning processes with PCAs to counteract poor learning outcomes. In terms of design science research, our research can be classified as an improvement according to Gregor and Hevner [22] since we address a known problem with a new solution. We contribute to theory by systematically considering theory-driven requirements to achieve common ground in the design of a PCA. That is how we can secure further understanding of learning processes based on grounding communication. We enable practitioners to design PCAs based on common ground and enhance learning processes. We present a theory-driven design of PCA to achieve common ground. Thus, we demonstrate and evaluate the design according to Peffers et al. [21]. The results provide deeper insights into the different ways how common ground could be successfully implemented in human-agent interaction by presenting design solutions in the form of DPs and one possible instantiation. With our study, we provide a first design iteration that consider common ground mechanisms. Thus, future research should use the evaluation results to re-design the PCA, e.g., aiming at different educational contexts.

A limitation for the examination of common ground is that there are no widely accepted measures of degree of common ground, neither between humans nor between humans and PCAs. In dyadic conversations, explicit acknowledgments (e.g., how often feedback occurs of having perceived, understood, and accepted the other's message) have been assessed [40]. However, explicit acknowledgments might not seem necessary when all conversational partners are sure that common ground has been established. Here, we therefore employed an adapted version of the Five Factor Perceived Shared Mental Model Scale (5-PSMMS) [74]. Shared mental models are one aspect of common ground. As common ground violations can go undetected, however, such a subjective measure can also lead to biased common ground assessments. Therefore, a measure that assesses the conceptual content on which common ground is necessary for the respective conversation would be desirable.

Our PCA-prototype yielded promising results. However, the sample size in our evaluation study was too small to draw firm conclusions. Moreover, the sample was also too small to examine how usage of the PCA influences learning success. A larger study is necessary to gain a detailed understanding of which design features meant to enhance common ground are most successful and most conducive to learning success.

6 Acknowledgments

The results presented in this article were developed in the research project Komp-HI funded by the German Federal Ministry of Education and Research (BMBF, grant 16DHBKI073). The fourth author acknowledges funding from the Basic Research Fund of the University of St. Gallen.

References

1. Bittner EAC, Oeste-Reiß S, Leimeister JM (2019) Where is the Bot in our Team? Toward a Taxonomy of Design Option Combinations for Conversational Agents in Collaborative Work. HICSS
2. Vassallo G, Pilato G, Augello A et al. (2010) Phase Coherence in Conceptual Spaces for Conversational Agents. In: Sheu PC-Y (ed) Semantic computing. IEEE Press; John Wiley & Sons, Piscataway, NJ, Hoboken, N.J., pp 357–371
3. Kuhail MA, Alturki N, Alramlawi S et al. (2022) Interacting with educational chatbots: A systematic review. Educ Inf Technol.
4. Winkler R, Söllner M (2018) Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. Proceedings 2018:15903.
5. Mikic Fonte FA, Burguillo JC, Nistal ML (2012) An intelligent tutoring module controlled by BDI agents for an e-learning platform. Expert Systems with Applications
6. Hayashi Y (2020) Gaze awareness and metacognitive suggestions by a pedagogical conversational agent: an experimental investigation on interventions to support collaborative learning process and performance. Intern J Comput.-Support Collab Learn 15:469–498.
7. Seeber I, Bittner E, Briggs RO et al. (2020) Machines as teammates: A research agenda on AI in team collaboration. Information & Management 57:103174.
8. Benner D, Elshan E, Schöbel S et al. (2021) What do you mean? A Review on Recovery Strategies to Overcome Conversational Breakdowns of Conversational Agents. ICIS.
9. Luger E, Sellen A (2016) "Like Having a Really Bad PA". In: Kaye J (ed) Proceedings of the 2016 CHI. ACM, New York, pp 5286–5297
10. Myers C, Furqan A, Nebolsky J et al. (2018) Patterns for How Users Overcome Obstacles in Voice User Interfaces. In: Mandryk R (ed) 2018 CHI. ACM, NY, pp 1–7
11. Chakrabarti C, Luger GF (2015) Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics. Expert Syst. Appl.:6878–6897
12. Weiler S, Matt C, Hess T (2021) Immunizing with information – Inoculation messages against conversational agents' response failures. Electron Markets.
13. Meredith J (2017) Analysing technological affordances of online interactions using conversation analysis. Journal of Pragmatics:42–55
14. Rothwell CD, Shalin VL, Romigh GD (2021) Comparison of Common Ground Models for Human--Computer Dialogue. ACM Trans Comput.-Hum Interact 28:1–35.
15. Koulouri T, Lauria S, Macredie RD (2016) Do (and Say) as I Say: Linguistic Adaptation in Human–Computer Dialogs. Human–Computer Interaction 31:59–95.
16. Clark HH (1996) Using language, Sixth printing. [ACLS Humanities E-Book]. Cambridge University Press, Cambridge
17. Diederich S, Brendel AB, Morana S et al. (2022) On the Design of and Interaction with Conversational Agents: An Organizing and Assessing Review of Human-Computer Interaction Research. JAIS 23:96–138.

18. Montemayor C (2021) Language and Intelligence. Minds & Machines 31:471–486.
19. Khosrawi-Rad B, Rinn H, Schlimbach R et al. (2022) Conversational Agents in Education – A Systematic Literature Review. ECIS 2022 Research Papers
20. Hevner AR, March ST, Park J et al. (2004) Design Science in Information Systems Research. MIS Quarterly 28:75–105.
21. Peffers K, Tuunanen T, Rothenberger MA et al. (2007) A Design Science Research Methodology for Information Systems Research. JMIS 24:45–77.
22. Gregor S, Hevner AR (2013) Positioning and Presenting Design Science Research for Maximum Impact. MIS Quarterly 37:337–355.
23. Laurillard D (2002) Rethinking university teaching: A conservational framework for the effective use of learning technologies, 2nd ed. Taylor & Francis, London
24. Schmitt A, Zierau N, Janson, A., & Leimeister, J. M. (2021) Voice as a Contemporary Frontier of Interaction Design. ECIS 2021 Research Papers
25. Budiu R (2018) The User Experience of Chatbots. www.nngroup.com/articles/chatbots/
26. Song D, Oh EY, Rice M (2017) Interacting with a conversational agent system for educational purposes in online courses. HSI. IEEE, pp 78–82
27. Hobert S, Meyer von Wolff R (2019) Say Hello to Your New Automated Tutor – A Structured Literature Review on Pedagogical Conversational Agents. WI.
28. Gupta S, Bostrom R (2009) Technology-Mediated Learning: A Comprehensive Theoretical Model. Journal of the Association for Information Systems:686–714
29. Gabriel C, Hahne C, Zimmermann A et al. (2021) The Virtual Tutor: Tasks for conversational agents in Online Collaborative Learning Environments. In: Bui T (ed) HICSS.
30. Sinha S, Basak S, Dey Y et al. (2020) An Educational Chatbot for Answering Queries. In: Mandal (ed) Emerging Technology in Modelling and Graphics, Singapore, pp 55–60
31. Ismail M, Ade-Ibijola A (2019) Lecturer's Apprentice: A Chatbot for Assisting Novice Programmers. IMITEC:1–8.
32. Cunningham-Nelson S, Boles W, Trouton, Luke, & Margerison, Emily (2019) A review of chatbots in education: Practical steps forward. AAEE: Educators Becoming Agents of Change: Innovate, Integrate, Motivate:299–306
33. Mabunda K, Ade-Ibijola A (2019) PathBot: An Intelligent Chatbot for Guiding Visitors and Locating Venues. In: 2019 ISCMI, pp 160–168
34. Gulz A, Haake M, Silvervarg A et al. (2011) Building a Social Conversational Pedagogical Agent. In: Perez-Marin D, Pascual-Nieto I (eds) Conversational agents and natural language interaction: Techniques and effective practices. IGI Global, pp 128–155
35. Kerly A, Hall P, Bull S (2007) Bringing chatbots into education: Towards natural language negotiation of open learner models. Knowledge-Based Systems 20:177–185.
36. Clark HH, Brennan SE (1991) Grounding in communication. In: Resnick LB, Levine JM, Teasley SD (eds) Perspectives on socially shared cognition. APA, pp 127–149
37. Trede F, Higgs J, Rothwell R (2008) Critical Transformative Dialogues: A Research Method Beyond the Fusions of Horizons. Forum Qualitative Sozialforschung: Qualitative Social Research, (2009): Qualitative Research on Intercultural Communication.
38. Jurafsky D, James M (2017) Dialog Systems and Chatbots. Speech and Language Processing
39. Richardson DC, Dale R (2006) Grounding Dialogue: Eye Movements Reveal the Coordination of Attention During Conversation and the Effects of Common Ground. Proceedings of the Annual Meeting of the Cognitive Science Society

40. Dillenbourg P, Traum D (2006) Sharing Solutions: Persistence and Grounding in Multimodal Collaborative Problem Solving. Journal of the Learning Sciences 15:121–151.
41. Sundar SS, Qian X, Bellur S (2010) Designing Interactivity in Media Interfaces: A Communications Perspective. CHI 2010: Perspectives on Design:2247–2256
42. Brennan SE (1998) The grounding problem in conversations with and through computers. Social and Cognitive Approaches to Interpersonal Communication. Susan R. Fussell and Roger J. Kreuz (Eds.):201–225
43. Briggs RO (2006) On theory-driven design and deployment of collaboration systems. International Journal of Human-Computer Studies:573–582
44. Gehlert A, Schermann M, Pohl K et al. (2009) Towards a Research Method for Theorydriven Design Research. Wirtschaftinformatik Proceedings 2009
45. Dickhaut E, Janson A, Söllner M et al. (2023) Lawfulness by design – development and evaluation of lawful design patterns to consider legal requirements. EJIS:1–28.
46. Brenner W, Karagiannis D, Kolbe L et al. (2014) User, Use & Utility Research. Bus Inf Syst Eng 6:55–61.
47. Serban IV, Sankar C, Germain M et al. (2017) A Deep Reinforcement Learning Chatbot. arXiv
48. Pickering MJ, Garrod S (2004) Toward a mechanistic psychology of dialogue. Behav Brain Sci 27:169-90; discussion 190-226.
49. Fusaroli R, Rączaszek-Leonardi J, Tylén K (2014) Dialog as interpersonal synergy. New Ideas in Psychology 32:147–157.
50. Branigan HP, Pickering MJ, Pearson J et al. (2010) Linguistic alignment between people and computers. Journal of Pragmatics 42:2355–2368.
51. Cowan BR, Branigan HP, Obregón M et al. (2015) Voice anthropomorphism, interlocutor modelling and alignment effects on syntactic choices in human−computer dialogue. International Journal of Human-Computer Studies 83:27–42.
52. Rothwell CD, Shalin VL, Romigh GD (2021) Comparison of Common Ground Models for Human--Computer Dialogue. ACM Trans Comput.-Hum Interact 28:1–35.
53. Tolzin A, Janson A (2023) Mechanisms of Common Ground in Human-Agent Interaction: A Systematic Review of Conversational Agent Research. HICSS.
54. Corti K, Gillespie A (2016) Co-constructing intersubjectivity with artificial conversational agents: People are more likely to initiate repairs of misunderstandings with agents represented as human. Computers in Human Behavior 58:431–442.
55. Pustejovsky J, Krishnaswamy N (2021) Embodied Human Computer Interaction. Künstl Intell 35:307–327. https://doi.org/10.1007/s13218-021-00727-5#auth-James-Pustejovsky
56. Maybin J, Mercer N, & Stierer B (1992) 'Scaffolding' Learning in the classroom. In: Norman K (ed) Thinking Voices: Thework of the National Oracy Project. Hodder Arnold
57. Clark L, Pantidi N, Cooney O et al. (2019) What Makes a Good Conversation? In: Brewster S, Fitzpatrick G, Cox A et al. (eds) 2019 CHI. ACM, New York, NY, USA, pp 1–12
58. Neururer M, Schlögl S, Brinkschulte L et al. (2018) Perceptions on Authenticity in Chat Bots. MTI 2:60.
59. Cassell J (2003) Negotiated Collusion: Modeling Social Language and its Relationship Effects in Intelligent Agents. User Modeling and User-Adapted Interaction 13:89–132.
60. Bernard D, Arnold A (2019) Cognitive interaction with virtual assistants: From philosophical foundations to illustrative examples in aeronautics. Comput. Ind. 107:33–49.

61. Pinhanez CS, Candello H, Pichiliani MC et al. (2018) Different but Equal: Comparing User Collaboration with Digital Personal Assistants vs. Teams of Expert Agents
62. Blache P (2017) Dialogue management in task-oriented dialogue systems. In: Chaminade T, Interaction ASIGoC-H (eds) Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents. ACM, , pp 4–8
63. Frijns HA, Schürer O, Koeszegi ST (2021) Communication Models in Human–Robot Interaction: An Asymmetric MODel of ALterity in Human–Robot Interaction (AMODAL-HRI). Int J of Soc Robotics.
64. Kiesler S (2005) Fostering common ground in human-robot interaction. In: Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on. IEEE, pp 729–734
65. Fischer U, McDonnell L, Orasanu J (2007) Linguistic correlates of team performance: toward a tool for monitoring team functioning during space missions. Aviat Space Environ Med 78:B86-95
66. Fu-Yun Yu (2012) Learner-centered pedagogy + adaptable and scaffolded learning space design—online student question-generation. International Conference on Computers in Education 2012
67. Venable J, Pries-Heje J, Baskerville R (2016) FEDS: a Framework for Evaluation in Design Science Research. European Journal of Information Systems 25:77–89.
68. Mettler T, Eurich M, Winter R (2014) On the Use of Experiments in Design Science Research: A Proposition of an Evaluation Framework. CAIS 34.
69. Soderstrom NC, Bjork RA (2015) Learning versus performance: an integrative review. Perspect Psychol Sci 10:176–199.
70. Gupta S, Bostrom R (2013) Research Note —An Investigation of the Appropriation of Technology-Mediated Training Methods Incorporating Enactive and Collaborative Learning. Information Systems Research 24:454–469.
71. Rhodes MG (2016) Judgments of learning: Methods, data, and theory. In: Dunlosky J, Tauber SK (eds) The Oxford handbook of metamemory. Oxford University, New York
72. Agarwal R, Prasad J (1998) A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. Information Systems Research
73. Schmitt A, Wambsganss T, Janson A (2022) Designing for Conversational System Trustworthiness: The Impact of Model Transparency on Trust and Task Performance. ECIS 2022 Research Papers
74. van Rensburg JJ, Santos CM, Jong SB de et al. (2021) The Five-Factor Perceived Shared Mental Model Scale: A Consolidation of Items Across the Contemporary Literature. Front Psychol 12:784200.