The Anatomy of User Experience with Conversational Agents: A Taxonomy and Propositions of Service Clues

Completed Research Paper

Naim Zierau
University of St.Gallen
Müller-Friedberg-Str. 8,
CH-9000 St. Gallen
naim.zierau@unisg.ch

Thiemo Wambsganss
University of St.Gallen
Müller-Friedberg-Str. 8,
CH-9000 St. Gallen
thiemo.wambsganss@unisg.ch

Andreas Janson
University of St.Gallen
Müller-Friedberg-Str.8,
CH-9000 St. Gallen
andreas.janson@unisg.ch

Sofia Schöbel
University of Kassel
Pfannkuchenstr. 1,
DE-34121 Kassel
sofia.schoebel@uni-kassel.de

Jan Marco Leimeister
University of St.Gallen / University of Kassel
Müller-Friedberg-Str.8, / Pfannkuchenstr. 1,
CH-9000 St. Gallen / DE-34121 Kassel
janmarco.leimeister@unisg.ch

Abstract

Conversational agents (CAs) represent a paradigm shift in regards to how humans use information systems. Although CAs have recently attracted considerable research interest, there is still limited shared knowledge about the distinctive characteristics of CAs from a user experience-based perspective. To address this gap, we conducted a systematic literature review to identify CA characteristics from existing research. Building on classifications from service experience theory, we develop a taxonomy that classifies CA characteristics into three major categories (i.e. functional, mechanic, humanic clues). Subsequently, we evaluate the usefulness of the taxonomy by interviewing six domain experts. Based on this categorization and the reviewed literature, we derive three propositions that link these categories to specific user experience dimensions. Our results support researchers and practitioners by providing deeper insights into service design with CAs and support them in systematizing and synthesizing research on the effects of specific CA characteristics from a user experience-based perspective.

Keywords: Conversational Agent, User Experience, Service Clues, Taxonomy

Introduction

Spurred by technological advances in Artificial Intelligence (AI) especially in the area of Natural Language Processing (NLP), companies increasingly strive to tap into the potential of Conversational Agents (CA) to automate service encounters (Maedche et al. 2019). AI-based CAs, such as Amazon's Alexa or Apple's Siri, assist by engaging with users via natural language (Pfeuffer et al. 2019). These agents are gradually evolving to become the dominant service interface between providers and users (McLean and Osei-Frimpong 2019). On the one hand, they enable to significantly improve service efficiency through intelligent automation (Larivière et al. 2017; Rahwan et al. 2019). On the other hand, they promise to improve service quality by enabling personalization, around the clock availability, and immediate response times (De Keyser et al. 2019; Xu et al. 2017). However, despite technological advances that pay into the above-mentioned capabilities, the interaction of many users with these agents, have yielded mixed results indicating high failure rates (Fuckner et al., 2014). Moreover, market research has shown that users are generally skeptical about the new technology (Krogue 2017). In sum, usage of CAs for different service encounters is becoming increasingly omnipresent, but realizing their potential for service efficiency and service quality still represents a major challenge based on poor user experience.

CAs represent a novel form of information systems (IS) that can be distinguished from a socio-technical perspective by their high degree of interaction and intelligence (Maedche et al. 2019). These capabilities may fundamentally affect user experience and raise several theoretical and design-related questions, most prominently revolving around an emergent conversation-based interaction paradigm (e.g. Clark et al. 2019) and the increasingly autonomous character of AI-based technologies (e.g. Rahwan et al. 2019). Thus, in recent years a multitude of research emerged in different disciplines, most prominently in the *IS* and *Human-Computer-Interaction* (HCI) domain that investigated the effect of different design elements and configurations unique to these agents on various user perceptions such as trust (Nordheim et al. 2019) or social presence (Feine et al, 2019). The growing number of such studies, however, highlight also the necessity to better understand the anatomy of the user experience with these agents from a holistic perspective (Følstad and Brandtzaeg 2017). In this regard, two research needs become especially evident. First, there is a lack of shared knowledge about the design elements and configurations of CAs from an experience-based perspective. Most studies that research about and that design CAs focus on single or few design elements or configurations and their effect on selected user perceptions, leading to a fragmented literature base, and sometimes contradictory research results (Følstad and Bae 2020). However, such an integrative view is needed to fully capture CA aspects from a holistic user-experience view. Second, without a consistent knowledge base on the design elements of CAs, it becomes difficult to interpret and predict how users react to them and how the overall experience can be improved. These shortcomings could be addressed by an integrated analysis to aggregate CA design elements to meaningful categories from a user experience-based perspective, which would increase our understanding of user behaviors and thus enable more effective theorizing on the new phenomena (e.g. Janssen et al. 2020).

Although initial classifications on CAs have emerged during the past years (Diederich and Brendel 2019; Feine et al. 2019; Janssen et al. 2020), the CA research domain is dispersed in a multitude of contextual and theoretical perspectives resulting in a pressing shortage of integrative perspectives (Følstad and Brandtzaeg 2017). For instance, most scientific publications including classification frameworks, focus on specific CA characteristics particularly in regards to their technical or social capabilities, without providing a holistic view (Følstad et al. 2019; Gwenuch et al. 2017). The few integrative classifications focus on structural representations of CAs (e.g. Janssen et al., 2020), which allow the derivation of system archetypes, but do not allow to draw specific design conclusions, e.g., on how to improve user experience with a given CA archetype (Følstad and Bae 2020). In this regard, service research offers a promising perspective, since it allows us to classify design elements into functional, mechanic, and humanic dimensions, which are differently evaluated by users (Berry et al. 2006). Consequently, a classification of CA design elements taking this perspective would enable researchers to more effectively theorize on how different design decisions impact the user experience with a CA. Hence, this paper focuses on the following research question (RQ):

***RQ**: How can dimensions and characteristics of Conversational Agents be classified from a service experience-based perspective and how can they be linked to specific user experiences?*

To achieve this goal, we develop a taxonomy of design elements for CAs facilitating different types of service encounters based on the scientific literature on CA design. We follow the rigorous taxonomy development framework as outlined by Nickerson et al. (2013). Based on five iterations we classify and organize CA dimensions and characteristics embedded in 107 publications. We evaluate and revise our taxonomy regarding both, structure and content, according to recommendations provided by six experts from research and practice familiar with designing and working with CAs. In a second step, we derive three propositions on the influence of service clues on user experience with CAs based on the resulting taxonomy. Lastly, we discuss the results and derive implications for future research.

Theoretical Background and Related Work

User Experience with AI-Based Conversational Agents

CAs are AI-based computer programs that assist users by primarily interacting with them via natural language (Maedche et al. 2019). The conversational character of CAs enables new and potentially more convenient and personal ways to access content and services. In specific, CAs promise to be fast and cost-effective solutions in the form of 24/7 electronic channels to support users (Hopkins and Silverman 2016). Hence, CAs are now deployed in a wide range of application areas such as health (Laumer et al. 2019), education (Wambsganss et al. 2020), and customer service (Qiu and Benbasat 2009). However, contrary to industry expectations, users' adoption of CAs has been relatively low (e.g. Nordheim et al. 2019). Observers note that one reason might be that the development of CAs was initially based more on technology push than on market pull. Consequently, customer wishes and needs were not sufficiently addressed (Coniam 2014). Recently, this has led to a substantial increase in the body of research about user perceptions and preferences regarding CAs and usage of those agents.

In the general research literature on CAs, it is well established that user experience is one of the key drivers of system adoption (Følstad and Bae 2020). Hence, both researchers and practitioners have for some time not only addressed outcomes directly linked to productivity but also focused on outcomes related to users' desires for immersive and engaging activities (Nordheim et al. 2019). Following the international standard for the human-centered design of interactive systems, user experience can be understood as a *"person's perceptions and responses resulting from the use and/or anticipated use of a product, system, or service"* (ISO9241-210 2010, p. 3). Thereby, user experience is considered to be dynamic and context-dependent, which implies that user experience is contingent on the usage context incorporating factors such as time, place, and purpose (Law et al. 2009). To account for this complexity, different research perspectives emerged that aim to explain user experience, which can be distinguished into design- and model-based perspectives. While the design-based perspective focuses on affordances that are the representation of a CA to users (Janssen et al. 2020), a model-based perspective focuses on established experience-related constructs to facilitate comparison and generalization (Hornbæk and Hertzum 2017). Thereby, the proposed attributes of user experience may typically be grouped as ***pragmatic*** or productivity-oriented (i.e., ease of use, usefulness) on the one hand and ***hedonic*** or engagement-oriented (i.e., likeability, trust) on the other (Hassenzahl et al. 2008). In the case of CAs, there is a lack of integrated knowledge regarding what users see as particularly good or particularly poor user experience (Pfeuffer et al. 2019; Zierau et al. 2020); specifically, there is a lack of such knowledge that is grounded in existing theory on user experience (Hornbæk and Hertzum 2017). These issues indicate that there are some inconsistencies about the classification of CAs concerning user experience (Law et al. 2009). To overcome this inconsistency, we propose to identify and classify design elements according to their effects on pragmatic and hedonic experience outcomes.

Service Clues to Capture User-Experience with Conversational Agents

In service research, the concept of service clues has been formative as a framework for analyzing user experiences bridging design-and model-based perspectives on user experience. A service clue is everything that a user can perceive or sense (or recognize by its absence) in the service experience (Berry et al. 2006). For instance, the graphical representation of a CA potentially communicates a clue to the user. Accordingly, user experiences in the field of services can be designed through the purposeful arrangement of different elements (i.e. clues) of a service context (Gupta and Varjic 2000; Zomerdijk and Voss 2010). Thereby, literature differentiates between ***functional***, ***mechanic***, and ***humanic*** clues that make up the service experience (Berry et al. 2006). *Functional* clues provide indications on the technical quality of the service. They relate more to a kind of hygiene factor, which is assumed to be *primarily evaluated on a cognitive level* and may lead to dissatisfaction if user expectations are not met (i.e. corresponding mainly to pragmatic value in the user experience). *Mechanic* clues additionally affect the user on an *emotional level and cause delight* (i.e. corresponding mainly to hedonic value in the user experience). They represent the sensory presentation of the service, incorporating all the static characteristics users can perceive with their senses.

Finally, *humanic* clues emerge from the *behavior of the service provider* and are mostly evaluated on the emotional level. They represent features that make a service experience unique to a user possibly leading to engagement (Berry et al. 2006). Accordingly, the user experience with a CA can be characterized as a set of diverse stimuli that can be orchestrated by the service provider, potentially addressing both calculative and emotional experience dimensions (van Doorn et al. 2010) or differently said pragmatic and hedonic value contributions. These stimuli represent any clues that users perceive when interacting with a CA. Accordingly, the user experience may be engineered by managing those clues strategically (Ding et al. 2010). Thus, a concise and rigorously organized structure of service clues for CAs that can be linked to specific user experiences may support researchers and practitioners in understanding and orchestrating the user experience with CAs. However, current literature lacks an *integrated approach to structure service clues of CAs according to their effect on user experience*. Hence, we are proposing to classify clues embedded in interactions with CAs from a service perspective (i.e., functional, mechanic, and humanic clues) linking them to pragmatic and hedonic value contributions.[1] In this sense, we relate our approach to existing CA classifications in the next section, to delineate our theoretical contribution more specifically, especially when taking user experiences with CAs into account.

Existing CA classifications and their relation to user experience

In the literature on CAs, different kinds of CAs classifications already exist. One type of classifications organizes basic design elements of CAs that correspond to functional clues in the user experience to identify different types of service encounters with CAs representing the deployment scenario (Følstad et al. 2019; Gwenuch et al. 2017). For example, Følstad et al. (2019) thus identify different CA archetypes. Another type of classifications focuses on specific classes of CA characteristics that mainly relate to mechanical and humanic clues in the service experience. For instance, Feine et al. (2019) structure characteristics of CAs that relate to their ability to transmit social clues, while Knote et al. (2020) focus on intelligence-related capabilities and functional affordances of CAs. Another type of classifications employs an integrative lens from a design-based perspective and accordingly characterizes CAs for different domains. For instance, Bittner et al. (2019) focus on CAs used in collaborative work. Finally, Janssen et al. (2020) develop a CA taxonomy that integrates different design elements incorporating the use context as well as their interactive and intelligent capabilities. While these classifications provide valuable insights for the design and understanding of service encounters with CAs, we identified theoretical and practical relevant gaps from a user-experience-based perspective as presented in Table 1.

Table 1. Gaps of existing CA classifications with respect to characterizing user experience

Gaps	Description	Resulting Barrier	Examples
Lack of integration from a model-based perspective	Existing taxonomies do not aggregate CA dimensions and characteristics from a perceptual perspective.	• No insights are given on how to design CAs in accordance with user-based processes.	• Diederich and Brendel (2019) • Janssen et al. (2020)
Exhaustiveness of clues forming the user experience	Existing taxonomies are not exhaustive beyond functional and conversational-related dimensions.	• Important clues embedded in the overall service context that might be important for the overall user experience are not covered.	• Gwenuch et al. (2017) • Følstad et al. (2019) • Feine et al. (2019)
Lack of conceptual clarity with regard to the underlying perceptual theoretical basis	Existing taxonomies do not bridge design- and model-based perspectives on CAs.	• An integrated perspective on user experience both spanning pragmatic and hedonic value is missing. • The strategic application of service clues is impaired.	• Bittner et al. (2019) • Janssen et al. (2020)

The first gap relates to a *lack of integrative perspectives* of different types of CA dimensions and characteristics from a user perspective (e.g. Diederich and Brendel 2019; Janssen et al. 2020). This makes it more difficult to apply these taxonomies for both researchers and practitioners, as these taxonomies focus on the representation of the technical system (i.e. design elements), which is useful to generate system archetypes. However, they do not allow sufficient insight into user-based processes. Another gap is related to the *exhaustiveness of identified CA dimensions and characteristics*. While the discussed classifications

[1] In the course of the paper, we refer to user-experience in the domain of services and service experience with CAs as synonyms.

cover some functionality-related (e.g., Følstad et al. 2019; Gwenuch et al. 2017) and in detail conversational-related characteristics (e.g., Feine et al. 2019), those elements that are not directly grounded in the technical system are not sufficiently addressed. However, the service context encompasses different functional, mechanic, and humanic cues (Berry et al. 2006) that can be instrumental to the user experience as it has shown that the adoption of CAs is related to more than the design of the technical system (Banavar 2016). Finally, existing classifications *do not bridge design and model-based perspectives* as most authors focus on design-based conceptualizations (Feine et al. 2019; Janssen et al. 2020). Thus, it is not clear how specific design elements (i.e. service clues) are linked to perceptual outcomes, which impairs our ability to strategically design for specific user experiences and, moreover, to theorize on the perceptual nature of different CA characteristics.

Overall, these gaps mainly arise because of the isolated perspective often taken when classifying dimensions and characteristics of CAs and the lack of conceptual clarity regarding studying their effect from a user experience-based perspective. To overcome these gaps, we suggest 1) to develop a taxonomy of service clues for CAs based on the empirical literature on CAs and 2) to derive propositions based on our taxonomy to illustrate the influence of service clues on user experience, to further strengthen the studies theoretical contributions and implications for future research.

Methodology

As shown in the previous section, a fundamental problem in CA research in regards to user experience is the classification of CAs into clear and meaningful categories. To systematically classify objects of interest we can refer to a taxonomy (Nickerson et al. 2013). Classifications are useful to researchers and practitioners as they enable to structure novel and complex domains, which is especially valuable for young and emerging research fields such as research involving CAs. Classifications highlight the interrelationships between different elements of a phenomenon transparently and coherently, as well as indicate their respective theoretical basis (Bailey 1994; Schöbel et al. 2020). Thus, taxonomies not only have a prescriptive value but rather can also serve as a relevant input for the advancement of theoretical knowledge as in our case a conceptualization to study the effect of CA design elements on user experience. Hence, we follow a rigorous taxonomy development process resulting in four distinct phases (Table 2.):

Table 2. Overview of the four consecutive research phases

	Phase 1: Database Creation	Phase 2: Taxonomy Development	Phase 3: Taxonomy Evaluation	Phase 4: Taxonomy Application
Steps	• Search for relevant papers in IS and HCI literature • Analyze and synthesize literature in regards to user experience clues	• Define meta characteristic • Run taxonomy development iterations until ending conditions are met	• Evaluate dimensions and characteristics with experts based on quality criteria	• Analyze links between dimensions and user perceptions
Method	Literature Review (vom Brocke et al. 2015), Qualitative Coding	Taxonomy Development (Nickerson et al. 2013)	Expert Evaluation (Szopinski et al. 2019)	Analysis of literature (Jeyaraj et al. 2006)
Source	CA literature	Existing Classifications, Database on CA service clues	Semi-structured interviews with experts	CA literature (identified in phase 1)
Results	Database with 107 articles on service clues	Taxonomy of service clues of CAs	Evaluated taxonomy of service clues for CAs	Insights on user experience with CAs

Phase 1: Database Creation

To identify relevant literature as the basis for the systematic and stepwise development of a taxonomy, we conducted a Systematic Literature Review (SLR) following Webster & Watson (2002) and vom Brocke et al. (2015). The overall scope of the conducted SLR can be defined along the dimensions of *process*, *source*, *coverage*, and *techniques* of the SLR (vom Brocke et al. 2015). To establish the basis for the taxonomy development and conceptualization, we used a *comprehensive set of techniques* (i.e., keyword search, backward search, and forward search). To reach a high level of reproducibility and transparency of our research, we describe in this section the single methodical steps that we undertook:

Selection of search string: To identify a wide range of literature on CAs, the search string is chosen to be rather broad. Based on recent literature reviews (Winkler and Söllner 2018; Zierau et al. 2020), we identified different keywords researchers used to describe CAs. This resulted in the following search string: *("conversational agent"* OR "*chat bot*" OR "*chatbot*" OR "*dialogue system*" OR "*smart personal assistant*" OR "*smart assistant*" OR "*intelligent agent*" OR "*intelligent assistant"*). In the SLR we used all variations of the keywords – singular, plural, hyphenated, or not hyphenated.

Selection of outlets: As our goal is to identify representative literature samples of different research perspectives on design elements of CAs from a user perspective, our search covers multiple journals and conference proceedings. For the selection of outlets, we identified two broad areas for deriving design elements of CAs – *Information Systems* (IS) and *Human-Computer Interaction* (HCI) – as they cover a substantial share of literature on CAs. To safeguard the relevance of our results, we discussed our selection of journals and conference proceedings with two senior researchers from the field of interest, who were not involved in the writing process of the papers. In sum, we selected 20 journals and proceedings for our keyword search (see Table 3).

Table 3. Overview of searched journals and conference proceedings

Field	**Outlets** *(Hits / Relevant Publications in Brackets / Forward & Backward)*
IS	ACM Transactions on Information Systems *(16/0/0)*, Decision Sciences *(6/0/0)*, Decision Support Systems *(39/5/2)*, European Journal of Information Systems *(6/0/0)*, Information Systems Journal *(2/0/1)*, Information Systems Research *(6/0/0)*, Journal of Information Technology *(1/0/0)*, Journal of Management Information Systems *(21/1/1)*, Journal of Strategic Information Systems *(0/0/0)*, Journal of the Association for Information Systems *(1/0/0)*, Management Information Systems Quarterly *(0/0/0)*, Proceedings of the International Conference on Information Systems *(14/2/3) Proceedings of the European Conference on Information Systems (3/0/3)*
HCI	ACM Transactions on Computer-Human Interaction *(17/6/0)*, International Journal on Human-Computer Studies *(43/6/1)*, Journal of Computer-Mediated Communication *(6/0/0)*, Journal of the ACM *(1/0/0)*, User-Modelling and User-Adapted Interaction *(17/2/2)*, Human Computer Interaction (12/3/0) Proceedings of the Conference on Human Factors in Computing Systems *(172/51/7), Additional Forward & Backward (0/0/9)*

Selection of papers: Searching in the title, abstract, and keywords of the papers, the outlet-based search reveals 383 hits. This number still contains literature not relevant to this paper. In an initial screening process, the identified papers are analyzed based on their abstracts. We only included papers that referred to any type of CAs and which provide information on service clues as a central focus concept and unit of analysis of the papers. This resulted in 76 papers. Finally, the forward and backward search was carried out. Through screening the references and applying forward searches using *GoogleScholar*, 31 articles were added to the list, resulting in the final number of 107 papers.

Analysis of papers: The 107 relevant papers are analyzed from a concept-centric perspective based on an abductive approach. Thereby, to aggregate the insights from identified studies, we developed a list of master codes and master code descriptions representing service clues of CAs and user experience outcomes. Moreover, we identified service clues based on the information given on the nature of the service (i.e. task) provided by the CA. This process was iterative and required multiple rounds of coding of the identified papers by different researchers. Thereby, the iterative process started by two of the researchers to independently code a subset of 20 randomly chosen articles. For each of the 20 studies, we listed each service clue often represented as the independent variable and user experience outcome represented as the dependent variable as named by the author(s), which together form our initial list of "author variables" and "author variable" descriptions. Subsequently, we met to discuss how to combine variables across studies, which resulted in a list of "master variables" and "master variable descriptions". In some cases, both of these lists are identical, while other variables required more consideration. Those that were not identical were discussed by both coders till both agreed on a variable. Next, we re-examined the initial subset set of 20 articles and mapped author variables to our master variables. During the next iterations, two researchers coded independently the rest of the articles. Thereby, we coded service clues and user experience outcomes, and also mapped these variables to the growing list of master variables and descriptions. Afterward, these researchers met independently to discuss their findings. In case, the findings differed a third researcher was involved to discuss the differences. Thus, in each iteration, we added new master variables and descriptions until all papers were coded.

Phase 2: Taxonomy Development

We aimed to integrate and classify service clues from identified publications from a user experience-based perspective. Therefore, we decided to develop a taxonomy of service clues for CAs that influence user experience to provide a systematic representation of existent scientific knowledge and to integrate these findings with existing service experience literature to arrive at a sound conceptual model. We decided to base our taxonomy development on Nickerson et al. (2013) since it is the most prominent and widely used approach in the field of IS. Moreover, it offers a systematic and step-by-step method for developing taxonomies while ensuring the completeness of the identified dimensions and characteristics of an object.

We started by defining a meta-characteristic, which reflects the purpose of the taxonomy and determines the selection of dimensions and characteristics. Ultimately, we aimed to theorize on the feature-related perceptual mechanisms of CA design. Thus, we choose *the concept of service clues* as our meta-characteristic, which refers to the distinctive technical and situational features that can be perceived by the user during the service interaction with a CA and that together frame the structure of the user experience (Berry et al. 2006). To account for the complex nature of CAs, we subdivided the taxonomy dimensions into subclasses of established service clues as introduced in the theoretical background adapted to the context of CAs – *functional, mechanic,* and *humanic* service clues (i.e. design elements).

Next, we determined the subjective and objective ending conditions (EC) according to Nickerson et al. (2013) that determine the termination of the taxonomy development process:

A) At least one service clue is classified under every characteristic of every dimension.
B) No new dimension or characteristic has been added in the last iteration.
C) Every dimension and every characteristic in its dimension are unique and do not repeat.
D) Every known service clue is classified in the taxonomy.

The iterative taxonomy building process either starts with an *empirical-to-conceptual* or a *conceptual-to-empirical* approach. In subsequent iterations, these approaches can be interchanged. A *conceptual-to-empirical* approach involves the examination of empirical cases in regards to, how they fit with the initial conceptualization, while an *empirical-to-conceptual* approach involves starting with empirical data clusters before conceptualizing the nature of each cluster (Nickerson et al. 2013). Figure 1 shows how the taxonomy evolved over the process.

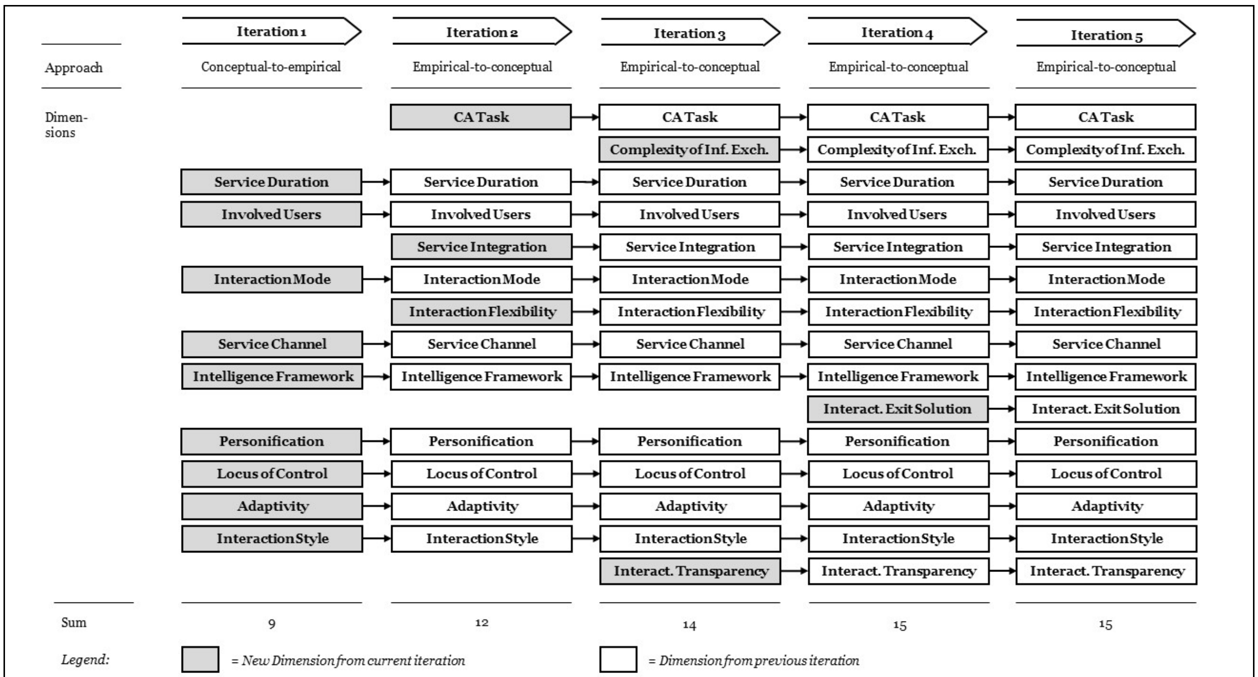

Figure 1. Taxonomy development process based on Nickerson et al. 2013

For the first iteration, we chose the conceptual-to-empirical approach, to add on and integrate prior classifications of CAs from a user experience-based perspective. Therefore, we established the first dimensions based on prior characterizations. During this iteration, we added nine dimensions: service duration (Følstad et al. 2019; Janssen et al. 2020), involved humans (Janssen et al. 2020), interaction mode (Knote et al. 2020), service channel (Janssen et al. 2020), intelligence framework (Knote et al. 2020), personification (Feine et al. 2019; Janssen et al. 2020), locus of control (Følstad et al. 2019), adaptivity (Knote et al. 2020) and interaction style (Feine et al. 2019). In the subsequent four empirical-to-conceptual iterations, we inductively adjusted the latest status of our taxonomy by classifying service clues representing the coded dependent variables embedded in convenience samples of 25 publications. Thus, in the second cycle, we added CA task, service integration, and interaction flexibility to the taxonomy. In the next cycle, we identified with complexity of information exchange and interaction transparency two additional dimensions. In the fourth iteration, only one dimension was added – interaction exit solution. In the final iteration, where we looked at identified service clues from the remaining 32 publications, we could not identify new dimensions. Thus, we classified all service clues identified in the 107 publications from our systematic literature review from phase 1 in five iterations until all ECs were met.

Phase 3: Taxonomy Evaluation

The iterative methodology development process comes to an end when all ECs are met (Nickerson et al. 2013). Our taxonomy presents the diversity of service clues of CAs that are relevant to the user experience. We found a set of service clues diverse enough to demonstrate their commonalities and differences and comprehensive enough to display all found design characteristics of CAs in our systematic literature review. However, to ensure the quality of a taxonomy, it is recommended to assess it against the following five criteria: *conciseness, robustness, comprehensibility, extendibility*, and *explanatory power* (Nickerson et al. 2013). Hence, to evaluate the taxonomy we conducted, in line with the taxonomy evaluation suggestions from Szopinski et al. (2019), semi-structured interviews with six experts that either had expertise in CA research, CA development in practice, or taxonomy development. Table 4. (Appendix A) provides information on each of the interviewees. We conducted the interviews via skype or phone between September and October 2019. The shortest interview lasted 22 minutes, while the longest 54 minutes. The interview guideline consisted of 17 open questions which were based on the five evaluation criteria. Therefore, the final version of our taxonomy and definitions of each dimension were sent to the interviewees via email one week before the interviews. For the preparation of the interviews, the interviewees were asked to make comments and note the potentials for revision and improvement. Subsequently, we will present the evaluation results for each criterion:

- *Conciseness* refers generally to the number of dimensions, which should cover the phenomenon sufficiently, while at the same time should not overwhelm the reader. Most experts think that the dimensions were well chosen, except for one expert that demanded small changes, which we followed on. In general, the subdivision in functional, mechanic, and humanic clues was evaluated particularly positively.
- *Robustness* means that based on the dimensions and characteristics it can be differentiated between objects of interest. The experts considered these elements to be sufficiently disjunct.
- *Comprehensiveness* describes the ability of a taxonomy to classify all objects of a phenomenon of interest. The experts agreed that the taxonomy is comprehensive with regard to the state of the art. However, they stressed that based on ongoing rapid developments in the area of AI, new dimensions and characteristics may need to be added in the future.
- *Extendibility* refers to the ability of a taxonomy to include new dimensions and new characteristics. As noted in the previous point, the taxonomy illustrates the state-of-the-art and is not finitely complete, but the experts agreed that the taxonomy is easily extendible based on the categorization into functional, mechanic, and humanic clues.
- *Explanatory power* refers to the ability of a taxonomy to highlight the interrelationships between different elements and characteristics transparently and thus to uncover previously unknown aspects of a phenomenon. Our taxonomy was mainly developed based on an inductive approach, which enables a clearer understanding of design elements of CAs from an overall user experience-based perspective. The experts agree that the taxonomy describes the design elements of CAs well from a user interaction point of view. In particular, they think that the taxonomy is suitable to be integrated with theoretical insights from previous literature to shed light on experience evaluation.

Phase 4: Taxonomy Application

The objective of the fourth phase was to apply the taxonomy to analyze the identified literature sample in regards to the effects of service clue dimensions on user experience. The objective of this phase is to highlight links between the dimensions and specific user perceptions based on the reported empirical data. Thus, concurrent with the creation of a list of service clues and user perceptions, we also coded the empirical relationships found between a dimension and a dependent variable representing a user perception within each study from the database. Thereby, adapting an approach of Jeyaraj et al. (2006), we assigned the value "+1" to the relationship between a service clue (i.e., independent variable) and a specific user perception (i.e., dependent variable): "*+1*". We used $p < 0.10$ as the cut-off requirements for a significant positive or negative relationship. Thus, by counting the number of significant results we can highlight linkages between service clue dimensions and user perceptions that can be distinguished into pragmatic and hedonic value dimensions.

Taxonomy of Service Clues for Conversational Agents

In the following section, we present our consolidated version of the taxonomy after conducting five iterations and the revision based on the feedback from the expert interviews. All presented design elements are paramount for defining the service experience according to the reviewed literature. Followingly, we will introduce the different dimensions and their characteristics as presented in Table 4.

Table 4. Taxonomy of service clues for Conversational Agents

	Dimensions	Characteristics		
Functional	CA Task	Provide Information	Collect Information	Two-Way Information Exchange
	Complexity of Information Exchange	Structured Information	Primitive Unstructured Information	Compound Unstructured Information
	Service Duration	Short-Term	Long-Term	
	Involved Users in Service Interaction	Single	Several	
	Service Integration	Stand-Alone	Integrated	
Mechanic	Interaction Mode	Text-Based	Voice-Based	Multi-Modal
	Interaction Flexibility	Solely Predefined	Partly Predefined	Free Text / Free Speech
	Service Channel	Application	Social Media	Website
	Intelligence Framework	Rule-based	Hybrid Learning	Self-Learning
	Interaction Exit Solution	Manual	Active Handling	Human Agent
Humanic	Personification	Disembodied (Anonym)	Embodied (Personal)	
	Locus of Control	User-Driven	CA-Driven	
	Adaptivity	Static CA Behavior	Customized CA Behavior	
	Interaction Style	Transactional	Relational	
	Interaction Transparency	None	Viewing Rights	Editing Rights

Overall, the service experience with CAs can be categorized into *functional, mechanic,* and *humanic dimensions*. Functional clues of CAs capture the structural characteristics of the service (i.e., service creation) such as the type and complexity of the service task. Mechanic clues of CAs capture characteristics related to the technical specification of the user interface (i.e., service delivery), such as the mode of interaction and service channel. Humanic clues of CAs capture characteristics that are related to the interaction with the user from a behavioral perspective (i.e., service interaction) such as the style of communication and the adaptivity of the CA (Berry et al. 2006; Huang and Rust 2020). The scaffolding division into these three perspectives aims to increase the usability of the taxonomy in regard to theorizing on the effect of these dimensions on pragmatic and hedonic value experiences. To that end, we strive for a precise and unambiguous description of the different classifications, to allow for a robust categorization of identified service clues.

Functional Dimension

The functional design elements capture the structural characteristics of the service facilitated by the CA. They determine high-level requirements that should be met to guarantee basic service functionality.

According to our analysis, the functional dimension can be distinguished into the following five clue dimensions using 12 characteristics to describe them: First, according to the CA task, *service encounters* with CAs can be distinguished into *unidirectional* or *bidirectional* interactions. Unidirectional CAs characterize service encounters where the agent either provides (Cha et al. 2019) or collects information (Elkins and Derrick 2013). Bidirectional CAs are designed to facilitate interactions, where both parties exchange information with each other representing a more complex service interaction (Bickmore and Mauer 2006). Second, *complexity of information exchange* relates to the requirement to dismantle and process natural language of different complexity levels. The simplest form is the processing of collected or manually entered data *(structured information)*, followed by simple natural language commands, such as "turn light on" (*primitive structured information*), followed by compound natural language commands, such as "turn light on every day at 6 am except for weekends" (Knote et al. 2020). Third, the interaction type refers to the length and depth of an interaction required to facilitate the service on a timeline. In this context, services can be differentiated that either facilitate a transactional or *short-term* encounter (e.g., conducting a survey) (Kim et al. 2019) or a relational service over a *longer* period (e.g., health behavior change) (Ren et al. 2014). Fourth, the *involved humans in service interaction* dimension described whether one or more humans are involved in the service interaction process. For instance, there are several agents primarily in social media environments that interact with several humans at the same time (Bittner et al. 2019). Finally, the service integration refers to the i*nterconnectedness* of the service offered. Thereby, services can be distinguished that are completely facilitated by the CA or where the service is integrated into an experience covering several service interfaces (Cho 2019).

Mechanic Dimension

The mechanic dimension relates to the technical specification of the user interface of the CA. The key design challenge at this abstraction level is to enable the facilitation of the service functionality. We could identify 19 characteristics of mechanic features, which describe the user interface. These characteristics can be integrated into five dimensions: First, the *interaction mode* refers to the primary way(s) a user communicates with a CA and vice-versa. Thereby the interaction is either primarily *text-based* or *voice-based* (Cho 2019). However, there are also CAs that include *multi-modal communication* covering both chat and voice, and in some cases, even other forms of communication such as haptics (Pfeuffer et al. 2019). Second, the dimension *interaction flexibility* relates to the way the user interacts with the CAs. Thereby, in some cases, the user only chooses between predefined answers, where in other cases the user is free in his answers. Some CAs use a mixed form depended on the type of interaction (Iovine et al. 2020). Third, the *service channel* indicates the respective service platform, which the CA has been integrated into. Thereby, services with CAs are primarily offered via a *stand-alone application* or are integrated within *social media applications* or *websites* (Janssen et al. 2020). Third, the *intelligence framework* depicts the underlying cognitive system design delimiting the technical principles, under which a CA communicates, processes information and/or selects an action or response (Diederich and Brendel 2019; Knote et al. 2020). Rule-based chat intents are usually less adaptive to new user content, e.g., colloquial questions, whereas adaptively trained chat intents based on a hybrid-or self-learning framework result in more flexibility in the interaction with the bot (Kontogiorgos et al. 2019). Finally, to deal with breaks in the interaction flow and to overcome misunderstanding between a CA and a user, designers can choose from several conversation exit solutions (Jain et al. 2018). Some CAs actively offer alternatives (e.g., repetition of question), or they offer to talk to a humane live-agent. Moreover, mixed forms can be observed (Ashktorab et al. 2019).

Humanic Dimension

The *humanic* dimension emerges from clues that are related to the behavior of the CA that enhance the interaction between the user and the CA. In this regard, we derive eleven characteristics within five design dimensions. First, the *interface personification* illustrates the extent to which a CA incorporates visual or physical anthropomorphic or personification features in the form of static, animated, or reactive avatars (Nunamaker et al. 2011). Second, *locus of control* determines which actor guides the service interaction either being the CA or the user (Følstad et al. 2019). Third, *adaptivity* refers to the system's ability to take into account usage and context data and adapt accordingly. Thus, some CAs can adjust their interaction behavior to different users in the same context (Liao et al. 2018). Therefore, a CA can either be characterized to show *static behavior* if the system's behavior and capabilities remain the same throughout use or *adaptive behavior* if its behavior is customized according to context and personal use data (Hess et al.

2018). Fourth, the *communication style* distinguishes between *transactional* and *relational* forms of communication. While transactional conversation refers to a purely task-oriented form of interaction, relational CAs use language that is aimed at forming a relationship with the user beyond fulfilling a task. Finally, *interaction transparency* describes the degree of transparency provided about the interaction by the user. Some systems provide no option to review conversations (Jain et al. 2018), while other systems allow viewing or even editing the interaction log (Saffarizadeh et al. 2017).

Service Clues of Conversational Agents and their Relation to User Experience

Having our taxonomy in mind, we want to highlight how the agents' dimensions and characteristics that we have identified, are linked to user experience and the different types of evoked experience perceptions. Aligned with the results of our literature review, based on which we identified the dimensions and characteristics of our taxonomy, we screened each paper to identify, how these CA characteristics were liked to user experience measurements. Thus, for each dimension of our taxonomy, we counted the number of significant empirical results identified in the SLR on various user perceptions. A summary can be seen in Table 5. The functional dimension is not included in this table, since we found very few empirical findings that directly link them to user experience outcomes. This stems from the fact that the functional dimensions were mostly added based on the conceptual-to-empirical cycle and based on service clues identified from information on the nature of the service provided by the CA. They can be understood as contextual variables that describe the nature of the service provided by the CA. Thus, they may not be perceived in isolation and could be mainly evaluated implicitly (Berry et al. 2006). Moreover, the results of our analysis show that mechanic clues seem to be instrumental for pragmatic value contributions, while humanic clues seem to be more associated with hedonic value contributions. In sum, we coded 258 relationships between service clue dimensions and user experiences.

Table 5. User perceptions for mechanic and humanic clues based on analyzed studies

Clue Type	Dimension	Pragmatic (*)	Perceptions	Hedonic (*)	Perceptions
Mechanic	Interaction Mode	27	*Ease of Use, Usefulness, Satisfaction*	15	*Trust, Social Presence, Rapport*
	Interaction Flexibility	13	*Ease of Use, Satisfaction*	2	*Social Presence*
	Service Channel	4	*Satisfaction, Ease of Use*	2	*Trust, Rapport*
	Intelligence Framework	5	*Satisfaction*	4	*Trust, Rapport*
	Interaction Exit Solution	1	*Usefulness*	0	
Humanic	Embodiment	12	*Satisfaction, Ease of Use*	49	*Likeability, Trust, Social Presence, Engagement*
	Locus of Control	9	*Satisfaction, Ease of Use*	0	
	Adaptivity	10	*Satisfaction, Ease of Use*	35	*Social Presence, Involvement, Closeness*
	Interaction Style	13	*Satisfaction, Ease of Use*	51	*Likeability, Rapport, Social Presence. Closeness, Involvement*
	Interaction Transparency	1	*Ease of Use*	5	*Trust*

*: *Number of empirical found relationships in the SLR that were significant ($p < 0.10$)*

Towards a Model for Understanding Service Clues of CAs

Having the relationship of the taxonomies dimensions to hedonic and pragmatic values in mind, we next want to discuss our results by referring to different propositions. In sum, we can observe that functional clues mostly refer to structural variables that define the essential service task to be delivered by the CA and that provide context to how mechanic and human clues are evaluated by users. Thereby, mechanic clues lead in a first instance to pragmatic value experiences as they add to the effective and efficient delivery of the service task, while humanic clues primarily refer to hedonic value experiences as they are centered around a delightful service interaction process. In the following, we will theorize on each of these links based on the notion of service clues (Berry et al. 2006) and user experience (Følstad and Bae 2020) as well as an illustrative discussion of the empirical results of the identified papers and based on this derive several propositions. Figure 3. presents an overview of these propositions.

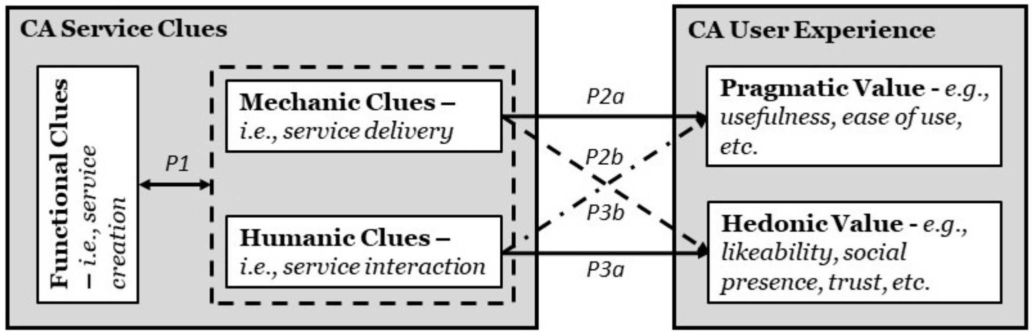

Figure 3. Conceptual model linking service clues and user experience

First, we judge about functional clues of CAs as general, contextual driven clues that are both relevant for mechanic and humanic clues in the same way. Functional clues are concerned with the core of the service facilitated by the agent because they address the problem that the user wants to get solved. They comprise the evaluation of the core service and managing them well is fundamental to meeting user expectations. Ultimately, by combining them, they try to find an answer to the question of how to create a valuable service for the user (Berry et al. 2006). In this regard, the results of our analysis suggest that the CA task needs to be carefully tailored to what is required by a target group. For instance, one study has shown that deploying a bi-directional CA is often misplaced when unidirectional information exchange is sufficient, as interaction complexity and reliability of the interaction seem to be in an anachronistic relationship to each other (Bickmore and Mauer 2006). In a similar way, the complexity of the information exchange may distinctively define the service facilitated by the CA and how a possible representation may look like (Knote et al. 2020). In general, tailoring these characteristics to a specific group of users can thus be seen as a make or break criteria for the user experience as they are related to the fundamental value proposition of the service. Under this light, we can also assume that there are different kinds of configurations of functional clues, depending on what users need and want, that is up to the designer to decide about based on a detailed analysis of the CAs target group and service context. In sum, we posit that:

P1: Functional clues need to be adapted to a specific group of users and to a specific context with the goal of supporting the service creation of CAs to positively influence how users experience mechanic and humanic clues.

Mechanic clues arise from the user interface of the CA and offer the basic technical representation of the intangible service delivered by the CA (Berry et al. 2006). Initially, the user who is considering using a CA cannot directly assess the CAs functionality but can perceive mechanic clues that serve as influential surrogate evidence reassuring the user that the CA can deliver the service effectively (Clark et al. 2019). A reality of service consumption in general and especially with regards to CAs based on their conversational character is that users need to decide if they want to interact with a CA before they were able to fully experience it. Hence, a potentially important role of mechanic clues is to signal pragmatic value especially at the beginning of the service experience. For instance, it was shown that the choice of interaction mode dramatically affects the disclosure behavior of users (Schroeder and Schroeder 2018), which suggests that effective service consumption is depended on interface design. Thus, mechanic clues may function as implicit service promises suggesting to users what the service should be like. In that, they also may give an indication, if the service is pleasurable to use. For example, underlying this claim, it was shown that users will expect a more distinctive experience, with a higher level of personal attention, when they use the CA via a mobile app compared to a CA embedded on a website (Araujo 2018). In sum, the main objective of mechanical clues of CAs is to make service delivery as effective and as efficient as possible, but they can also support delightful service experiences. We argue that:

P2a: A combination of mechanic clues facilitates the effective and efficient service delivery of CAs and thus in a first instance positively influences the users' pragmatic experience with CAs.

P2b: A combination of mechanic clues supports a delightful service experience with CAs and thus in a second instance positively influences the users' hedonic experience with CAs.

Humanic clues are related to the behavioral clues emitted by the CA and characterize service interaction (Berry et al. 2006). Literature has shown the more personal and immersive the CA-user interaction, the more pronounced humanic effects are likely to be, e.g., leveraging social presence perceptions (Cho 2019). Combining human clues offer the chance to cultivate emotional connectivity that can extend respect and esteem to users and, in so doing, exceed their expectations, strengthen their trust and deepen their affinity (Qiu and Benbasat 2009). Humanic clues are typically most important in regard to exceeding user expectations, especially in those contexts where treatment of the user is central to these service experiences (i.e. long-term service encounters) (Kujala et al. 2011). In these cases, excellent mechanic clues rarely overcome poor humanic clues. For instance, the desired communication style can add distinctively to the user experience. While in formal and commercial settings a transactional style is preferred and expected, in casual or long-term service encounters users prefer a CA that acts relationally such that they can build rapport with them. Still, humanic clues can also foster pragmatic experiences, e.g. by using elements such as a transactional conversation style that support efficient service interaction (Xiao et al. 2007). In sum, the main objective of combining humanic clues of CAs is to make the service interaction as delightful as possible, but they can also support efficient and effective service experience. Hence, we posit:

P3a: A combination of humanic clues facilitates the delightful service interaction between users and CAs and thus in a first instance positively influences the users' hedonic experience with CAs.

P3b: A combination of humanic clues supports the effective and efficient service experience with CAs and thus in a second instance positively influences the users' pragmatic experience with CAs.

Contributions, Limitations, and Future Research

From a theoretical perspective, we can make the following contributions. First, we integrate existing literature including classifications on CAs, by developing a new taxonomy, going beyond existing classifications structuring and grouping CA characteristics from a user experience-based perspective. With a common classification of CA characteristics assuming a holistic perspective, we provide a better understanding of what needs to be considered when designing and working with CAs. Thus, by combining the dimensions and characteristics of our taxonomy, we can now better conjecture about which characteristics influence which kind of user experience variables. This provides more room for future research studies and more detailed insights on how to design CAs. Second, we identify and classify new dimensions and characteristics beyond the technical system that are part of a service context and that play an important role in forming user perceptions. Thus, we combine a service-related view with the components and characteristics of a CA on different dimensions, which contributes to a holistic perspective on CAs. Third, the different characteristics are categorized into different clue categories that can be linked to pragmatic and/or hedonic value dimensions. Thus, based on our taxonomy that bridges design-and model-based perspectives on user experience, researchers and practitioners can combine and analyze specific CA characteristics and determine the conceptual lens based on which to study their effect on user experience. Given the immense growth of CA-based service encounters and the potential of CAs to improve the service experience, further research on this topic is warranted. Such research will require a solid theoretical understanding. In this paper, we offer such an understanding by presenting our taxonomy and a related conceptual framework linking categories of service clues to user experience dimensions. In that, we expand the knowledge base on factors that are related to how humans experience the use of configurational, interactive, and intelligent IS such as CAs.

From a practical perspective, a common understanding of the user experience with CAs, also gives rise to several insights for user experience designers. Our taxonomy of service clues supports researchers and practitioners in strategically managing and combining clues embedded in the user journey with CAs to facilitate a superior user experience, which is paramount to CA adoption as well as the differentiation from other service offerings. For instance, at a basic level, our taxonomy determines a number of high-level design decisions a service designer has to take when configuring a CA for a specific service task. Moreover, our taxonomy may support the design of CAs for specific experience outcomes based on different types of services. Based on our taxonomy, designers can now identify different kinds of CAs by combining different characteristics of CAs, depending on the context and on the target group, a CA is used for.

Our research has some limitations that provide avenues for future research. First and foremost, we only identified service clues based on scientific publications, since our main objective was to categorize CA characteristics from an experience-based perspective. Therefore, future research may adjust and extend our taxonomy based on an in-depth analysis of real-life use-cases. Second, this categorization is based on a cut-off criterion of $p < 0.1$ for the considered perceptual evidence, which comes at the expense of rigor as we acknowledge. However, to integrate research findings of a rather young and emerging field, that has yet to develop more mature findings, we believe that our criterion is useful and valid for deriving a first conceptual framework that can guide future research in this domain. Third, based on technological advancements more design elements may need to be added in the future. However, based on our evaluation with experts, we believe that our taxonomy resembles a sufficient state-of-the-art tool for analyzing service experiences with CAs. Fourth, despite our best effort to rigorously derive propositions based on a discussion of established theories in the user experience context and the analysis of identified literature, our proposed model of the service experience with CAs needs further substantiation. For instance, this could be done by zooming into each of the taxonomies dimensions and to analyze on a more granular level how they influence individual user experiences. Finally, as part of future research, we also suggest analyzing different patterns of CAs that can be developed and interpreted by using our taxonomy. Thus, by characterizing different types of service archetypes based on a clustering of empirical cases, successful configurational approaches may be identified to get a better understanding of the relationship of service clues to hedonic and pragmatic user experiences.

Conclusion

In sum, our results provide deeper insights into the user experience with CAs and support researchers in systematizing and synthesizing research on the effects of specific CA characteristics from a user experience-based perspective. We conducted five iterations, one being conceptually based in current CA classification literature and four iterations being empirically grounded on a set of 107 articles on CA dimensions and characteristics that we identified through a systematic literature review in the IS and HCI field. To demonstrate its usefulness, we evaluated the taxonomy with six domain experts. Our taxonomy maps service clues of CAs and their characteristics to a prominent categorization of service clues (i.e. functional, mechanic, and humanic clues). Based on this categorization and the reviewed literature, we derived three propositions that link these categories to specific user experience dimensions (i.e. pragmatic and hedonic perceptions). With our taxonomy and the related propositions, we demonstrate new ways of what to research about when working with CAs. Therefore, researchers and practitioners can use the results of our study not only to derive individual CA designs by referring to different characteristics but also to research more on the relationship between CA characteristics and user-experience outcomes.

Acknowledgments

We thank the Swiss National Science Foundation for funding parts of this research (100013_192718).

References

Araujo, T. 2018. "Living up to the Chatbot Hype: The Influence of Anthropomorphic Design Cues and Communicative Agency Framing on Conversational Agent and Company Perceptions," *Computers in Human Behavior* (85), pp. 183–189.

Ashktorab, Z., Jain, M., Vera Liao, Q., and Weisz, J. D. 2019. "Resilient Chatbots: Repair Strategy Preferences for Conversational Breakdowns," in *2019 Conference on Human Factors in Computing Systems - Proceedings*.

Bailey, K. D. 1994. *Typologies and Taxonomies: An Introduction to Classification Techniques*, Newbury Park, CA: Sage.

Banavar, G. 2016. "Learning to Trust Artificial Intelligence Systems," *IBM Research (September 2016).*

Berry, L. L., Wall, E. A., and Carbone, L. P. 2006. "Service Clues and Customer Assessment of the Service Experience: Lessons from Marketing," *Academy of Management Perspectives*, Academy of Management, pp. 43–57.

Bickmore, T., and Mauer, D. 2006. "Modalities for Building Relationships with Handheld Computer Agents," in *Conference on Human Factors in Computing Systems - Proceedings*, pp. 544–549.

Bittner, E., Oeste-Reiß, S., and Leimeister, J. M. 2019. "Where Is the Bot in Our Team? Toward a Taxonomy of Design Option Combinations for Conversational Agents in Collaborative Work," *Proceedings of the 52nd Hawaii International Conference on System Sciences* (6), pp. 284–293.

vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., and Plattfaut, R. 2015. "Standing on the Shoulders of Giants: Challenges and Recommendations of Literature Search in Information Systems Research,"

Communications of the Association for Information Systems (37:1), pp. 205–224.
Cha, Y., Hong, Y., Jang, J., and Yi, M. Y. 2019. "'Jack-of-All-Trades': A Thematic Analysis of Conversational Agents in Multi-Device Collaboration Contexts," in *2019 Conference on Human Factors in Computing Systems - Proceedings*.
Cho, E. 2019. "Hey Google, Can i Ask You Something in Private? The Effects of Modality and Device in Sensitive Health Information Acquisition from Voice Assistants," in 2019 *Conference on Human Factors in Computing Systems - Proceedings*.
Clark, L., Munteanu, C., Wade, V., Cowan, B. R., Pantidi, N., Cooney, O., Doyle, P., Garaialde, D., Edwards, J., Spillane, B., Gilmartin, E., and Murad, C. 2019. "What Makes a Good Conversation?," in *Proceedings of the 2019 CHI Conference Human Factors in Computing Systems*.
Coniam, D. 2014. "The Linguistic Accuracy of Chatbots: Usability from an ESL Perspective," *Text and Talk* (34:5), pp. 545–567.
Diederich, S., and Brendel, A. B. 2019. *Towards a Taxonomy of Platforms for Conversational Agent Design Digital Nudging View Project Chatbots and Gamification View Project*, (February), pp. 1100–1114.
Ding, D. X., Hu, P. J. H., Verma, R., and Wardell, D. G. 2010. "The Impact of Service System Design and Flow Experience on Customer Satisfaction in Online Financial Services," *Journal of Service Research* (13:1), pp. 96–110.
van Doorn, J., Lemon, K. N., Mittal, V., Nass, S., Pick, D., Pirner, P., and Verhoef, P. C. 2010. "Customer Engagement Behavior: Theoretical Foundations and Research Directions," *Journal of Service Research* (13:3), pp. 253–266.
Elkins, A. C., and Derrick, D. C. 2013. "The Sound of Trust: Voice as a Measurement of Trust During Interactions with Embodied Conversational Agents," *Group Decision and Negotiation* (22:5), pp. 897–913.
Feine, J., Gnewuch, U., Morana, S., and Maedche, A. 2019. "A Taxonomy of Social Cues for Conversational Agents," *International Journal of Human-Computer Studies*.
Følstad, A., and Bae, P. 2020. "Users ' Experiences with Chatbots : Findings from a Questionnaire Study," *Quality and User Experience*, Springer International Publishing, pp. 1–14.
Følstad, A., and Brandtzaeg, P. B. 2017. "Chatbots and the New World of HCI," *Interactions* (24:4), Association for Computing Machinery, pp. 38–42.
Følstad, A., Skjuve, M., and Brandtzaeg, P. 2019. "Different Chatbots for Different Purposes: Towards a Typology of Chatbots to Understand Interaction Design," in *International Conference on Internet Science Proceedings*, pp. 145–156.
Fuckner, M., Barthes, J. P., and Scalabrin, E. E. 2014. "Using a Personal Assistant for Exploiting Service Interfaces," in *Proceedings of the 2014 IEEE 18th International Conference on Computer Supported Cooperative Work in Design*, pp. 89–94.
Gupta, S., and Varjic, M. 2000. "The Contextual and Dialectical Nature of Experiences," in *New Service Development: Creating Memorable Experiences*, J. A. Fitzsimmons and M. J. Fitzsimmons (eds.), Thousand Oaks, CA: Sage, pp. 33–51.
Gwenuch, U., Morana, S., and Maedche, A. 2017. "Towards Designing Cooperative and Social Conversational Agents for Customer Service," in *Thirty Eighth International Conference on Information Systems, South Korea 2017*.
Hassenzahl, M., Schöbel, M., and Trautmann, T. 2008. "How Motivational Orientation Influences the Evaluation and Choice of Hedonic and Pragmatic Interactive Products: The Role of Regulatory Focuse," *Interacting with Computers* (20:4–5), pp. 473–479.
Hess, T., Fuller, M., and Campbell, D. 2018. "Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents," *Journal of the Association for Information Systems* (10:12), pp. 889–919.
Hornbæk, K., and Hertzum, M. 2017. "Technology Acceptance and User Experience: A Review of the Experiential Component in HCI," *ACM Transactions on Computer-Human Interaction* (24:5).
Huang, M. H., and Rust, R. T. 2020. "Engaged to a Robot? The Role of AI in Service," *Journal of Service Research*.
Iovine, A., Narducci, F., and Semeraro, G. 2020. "Conversational Recommender Systems and Natural Language:: A Study through the ConveRSE Framework," *Decision Support Systems* (131), p. 113-250.
ISO9241-210. 2010. *Ergonomics of Human Sys- Tem Interaction—Part 210: Human-Centred Design for Interactive Systems. International Organization for Standardization (ISO), Geneva, Switzerland*.
Jain, M., Kota, R., Kumar, P., and Patel, S. 2018. "Convey: Exploring the Use of a Context View for Chatbots," in *Conference on Human Factors in Computing Systems - Proceedings* (Vol. 2018-April).

Janssen, A., Passlick, J., Cordona, D. R., and Breitner, M. H. 2020. "Virtual Assistance in Any Context: A Taxonomy of Desgin Elements for Domain-Specific Chatbots," *Business & Information Systems Engineering*.

Jeyaraj, A., Rottman, J. W., and Lacity, M. C. 2006. "A Review of the Predictors, Linkages, and Biases in IT Innovation Adoption Research," *Journal of Information Technology* (21:1), pp. 1–23.

De Keyser, A., Köcher, S., Alkire (née Nasr), L., Verbeeck, C., and Kandampully, J. 2019. "Frontline Service Technology Infusion: Conceptual Archetypes and Future Research Directions," *Journal of Service Management* (30:1), Emerald Group Publishing Ltd., pp. 156–183.

Kim, S., Lee, J., and Gweon, G. 2019. "Comparing Data from Chatbot and Web Surveys Effects of Platform and Conversational Style on Survey Response Quality," in *Conference on Human Factors in Computing Systems - Proceedings*.

Knote, R., Janson, A., Söllner, M., and Leimeister, J. M. 2020. "Value Co-Creation in Smart Services: A Functional Affordances Perspective on Smart Personal Assistants," *Journal of the Association for Information Systems (JAIS)*.

Kontogiorgos, D., Pereira, A., Andersson, O., Koivisto, M., Rabal, E. G., Vartiainen, V., and Gustafson, J. 2019. "The Effects of Anthropomorphism and Non-Verbal Social Behaviour in Virtual Assistants," in *IVA 2019 - Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, pp. 133–140.

Krogue, K. 2017. "Artificial Intelligence Is Here to Stay, but Constumer Trust Is a Must for AI in Business," *Forbes, (https://www.forbes.com/sites/kenkrogue/2017/09/11/artificial-intelligence-is-here-to-stay-but-consumer-trust-is-a-must-for-ai-in-business/#2cc10a2b776e), (accessed:09-03-2020)*.

Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E., and Sinnelä, A. 2011. "UX Curve: A Method for Evaluating Long-Term User Experience," *Interacting with Computers* (23:5), pp. 473–483.

Larivière, B., Bowen, D., Andreassen, T. W., Kunz, W., Sirianni, N. J., Voss, C., Wünderlich, N. V., and De Keyser, A. 2017. "'Service Encounter 2.0': An Investigation into the Roles of Technology, Employees and Customers," *Journal of Business Research* (79), Elsevier Inc., pp. 238–246.

Laumer, S., Maier, C., and Gubler, F. T. 2019. "Chatbot Acceptance in Healthcare: Explaining User Adoption of Conversational Agents for Disease Diagnosis," in *Proceedings of the 27th European Conference on Information Systems*.

Law, E. L. C., Roto, V., Hassenzahl, M., Vermeeren, A. P. O. S., and Kort, J. 2009. "Understanding, Scoping and Defining User Experience: A Survey Approach," in *Conference on Human Factors in Computing Systems - Proceedings* (April), pp. 719–728.

Liao, Q. V., Hussain, M. M. U., Chandar, P., Davis, M., Khazaen, Y., Crasso, M. P., Wang, D., Muller, M., Shami, N. S., and Geyer, W. 2018. "All Work and No Play? Conversations with a Question-and-Answer Chatbot in the Wild," in *Conference on Human Factors in Computing Systems - Proceedings* (Vol. 2018-April).

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., and Söllner, M. 2019. "AI-Based Digital Assistants," *Business & Information Systems Engineering* (61:4), pp. 535–544.

McLean, G., and Osei-Frimpong, K. 2019. "Hey Alexa … Examine the Variables Influencing the Use of Artificial Intelligent In-Home Voice Assistants," *Computers in Human Behavior* (99:April), pp. 28–37.

Nickerson, R. C., Varshney, U., and Muntermann, J. 2013. "A Method for Taxonomy Development and Its Application in Information Systems," *European Journal of Information Systems* (22:3), pp. 336–359.

Nordheim, C. B., Følstad, A., and Bjørkli, C. A. 2019. "An Initial Model of Trust in Chatbots for Customer Service—Findings from a Questionnaire Study," *Interacting with Computers* (31:3), pp. 317–335.

Nunamaker, J. F., Derrick, D. C., Elkins, A. C., Burgoon, J. K., and Patton, M. W. 2011. "Embodied Conversational Agent-Based Kiosk for Automated Interviewing," *Journal of Management Information Systems* (28:1), pp. 17–48.

Pfeuffer, N., Benlian, A., Gimpel, H., and Hinz, O. 2019. "Anthropomorphic Information Systems," *Business & Information Systems Engineering* (61:4), pp. 523–533.

Qiu, L., and Benbasat, I. 2009. "Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems," *Journal of Management Information Systems* (25:4), pp. 145–182.

Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., Crandall, J. W., Christakis, N. A., Couzin, I. D., Jackson, M. O., Jennings, N. R., Kamar, E., Kloumann, I. M., Larochelle, H., Lazer, D., McElreath, R., Mislove, A., Parkes, D. C., Pentland, A. 'Sandy,' Roberts, M. E., Shariff, A., Tenenbaum, J. B., and Wellman, M. 2019. "Machine Behaviour," *Nature* (568:7753), Springer US, pp. 477–486.

Ren, J., Jack, B., Schulman, D., and Bickmore, T. 2014. "Supporting Longitudinal Change in Many Health Behaviors," in *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1657–1662.

Saffarizadeh, K., Boodraj, M., and Alashoor, T. M. 2017. "Conversational Assistants: Investigating Privacy Concerns, Trust, and Self-Disclosure," in *Proceedings of the 38th International Conference on Information Systems*.

Schöbel, S., Janson, A., and Söllner, M. 2020. "Caputuring the Complexity of Gamification Elements: A Holistic Approach for Analysing Existing and Deriving Novel Gamification Designs," *European Journal of Information Systems (EJIS)*, Forthcoming.

Schroeder, J., and Schroeder, M. 2018. "Trusting in Machines: How Mode of Interaction Affects Willingness to Share Personal Information with Machines," in *Proceedings of the 51st Hawaii International Conference on System Sciences* (Vol. 9).

Szopinski, D., Schoormann, T., and Kundisch, D. 2019. "BECAUSE YOUR TAXONOMY IS WORTH IT: TOWARDS A FRAMEWORK FOR TAXONOMY EVALUATION," in *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Wambsganss, T., Winkler, R., Schmid, P., and Söllner, M. 2020. "Unleashing the Potential of Conversational Agents for Course Evaluations: Empirical Insights from a Comparison with Web Surveys," *Twenty-Eighth European Conference on Information Systems (ECIS2020)* (May).

Webster, J., and Watson, R. T. 2002. "Analyzing the Past to Prepare for the Future: Writing a Literature Review," *MIS Quarterly - Management Information Systems* (26:2), pp. 13–23.

Winkler, R., and Söllner, M. 2018. "Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis," *Academy of Management Proceedings*.

Xiao, J., Stasko, J., and Catrambone, R. 2007. "The Role of Choice and Customization on Users' Interaction with Embodied Conversational Agents: Effects on Perception and Performance," in *Conference on Human Factors in Computing Systems - Proceedings*.

Xu, A., Liu, Z., Guo, Y., Sinha, V., and Akkiraju, R. 2017. "A New Chatbot for Customer Service on Social Media," in *Conference on Human Factors in Computing Systems - Proceedings* (Vol. 2017-May).

Zierau, N., Engel, C., Söllner, M., and Leimeister, J. M. 2020. "Trust in Smart Personal Assistants: A Systematic Literature Review and Development of a Research Agenda," in *WI2020 Zentrale Tracks*, pp. 99–114.

Zomerdijk, L. G., and Voss, C. A. 2010. "Service Design for Experience-Centric Services," *Journal of Service Research* (13:1), pp. 67–82.

Appendix

Table 6. Expert Panel for Taxonomy Evaluation

No.	Function	Organization	Expertise in
1	Researcher	International Business School	*Taxonomy Development* – Developed for analytic-based service
2	Researcher		*CA Research* – Conducted experimental and design-oriented research with CAs in customer service context
3	Researcher	University	*CA Research* – Conducted experimental and design-oriented research with CAs for learning support
4	Researcher		*CA Research* – Conducted experimental and design-oriented research with CAs in customer service context
5	System Developer	IT Consultancy	*CA Practice* – Builds CAs for different use-cases in various industries
6	IT Strategy Consultant	Insurance Company	*CA Practice* – Conducts requirements analyses and proofs of concepts for CAs in the insurance industry