See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons

Marcia Nißen [a,d,*], Driton Selimi [b,d], Antje Janssen [c,e], Davinia Rodríguez Cardona [c], Michael H. Breitner [c], Tobias Kowatsch [a,b,d], Florian von Wangenheim [a,d]

[a] *ETH Zurich, Department of Management, Technology, and Economics (D-MTEC), Chair of Technology Marketing, Weinbergstrasse 56/58, 8092, Zürich, Switzerland*
[b] *University of St.Gallen, Institute of Technology Management (ITEM-HSG), Chair of Operations Management, Dufourstrasse 40a, 9000, St.Gallen, Switzerland*
[c] *Leibniz Universität Hannover, Faculty of Economics and Management, Information Systems Institute, Königsworther Platz 1, 30167, Hannover, Germany*
[d] *Centre for Digital Health Interventions (CDHI), ETH Zurich & University of St. Gallen, Switzerland*
[e] *bhn Dienstleistungs GmbH & Co. KG, Hans-Lenze-Straße 1, 31855, Aerzen, Germany*

ARTICLE INFO

ABSTRACT

Users interact with chatbots for various purposes and motivations – and for different periods of time. However, since chatbots are considered social actors and given that time is an essential component of social interactions, the question arises as to how chatbots need to be designed depending on whether they aim to help individuals achieve short-, medium- or long-term goals. Following a taxonomy development approach, we compile 22 empirically and conceptually grounded design dimensions contingent on chatbots' temporal profiles. Based upon the classification and analysis of 120 chatbots therein, we abstract three time-dependent chatbot design archetypes: Ad-hoc Supporters, Temporary Assistants, and Persistent Companions. While the taxonomy serves as a blueprint for chatbot researchers and designers developing and evaluating chatbots in general, our archetypes also offer practitioners and academics alike a shared understanding and naming convention to study and design chatbots with different temporal profiles.

1. Introduction

Chatbots are on the rise to simplify and humanize the access to digital services (Go & Sundar, 2019). Designed to engage individuals in human-like conversations (Araujo, 2018; Scarpellini & Lim, 2020) they accompany people in their everyday lives as "digital coaches" (Fleisch, Franz, & Herrmann, 2021) or "virtual assistants" (Youn & Jin, 2021) providing users with individual, "high-touch" support (Wünderlich, von Wangenheim, & Bitner, 2012) on their smartphones or other devices, anywhere and anytime (Skjuve & Brandtzæg, 2018).

Users' primary motivations to engage with conversational agents[1] are manifold, which is reflected in the variety of conversational agents available ranging from the popular general-purpose voice assistants Siri (Apple, Inc.) or Alexa (Amazon, Inc.) to domain-specific text-based chatbots like the mental health chatbot Woebot or the scheduling assistants Amy and Andrew (x.ai, Inc.). Another basic difference concerns whether users intend to interact with a chatbot only once – for instance, with an e-service chatbot helping users to find a specific product (Chung, Ko, Joung, & Kim, 2020) or for multiple, continuous interactions over longer periods of time, such as with a healthcare chatbot supporting patients to manage a chronic disease (Kowatsch et al., 2018). Thus, designing chatbots fundamentally hinges on their "temporal profile" which encompasses the prospective time horizon of using the chatbot as well as the duration and frequency of individual interactions throughout the entire user-chatbot relationship (Baraka, Alves-Oliveira, & Ribeiro, 2020).

Generally, in human-computer interaction research, temporal aspects have been recognized as a crucial design factor for the

[1] In our research, we focus on text-based conversational agents, which are often also called "chatbots". We will use "conversational agents" when we refer to all types of conversational agents on the reality-virtuality-continuum (De Keyser, Köcher, Alkire (née Nasr), Verbeeck, & Kandampully, 2019) and "chatbots" when we refer to "text-based CAs" only.

contextualization and operationalization of interactive systems (Benford, Giannachi, Koleva, & Rodden, 2009, pp. 709–718; Hildebrandt, Dix, & Meyer, 2004, pp. 1737–1738) and for their impact on user journeys and experiences (Karahasanović, Hollebeek, Chasanidou, & Gurau, 2019, pp. 376–390).

However, despite diverse chatbot characteristics that have previously been investigated with regards to consequential design implications, for example, whether chatbots serve general or domain-specific purposes (Gnewuch, Morana, & Maedche, 2017) or whether chatbots are intended to engage in dyadic one-to-one or in multiparty interactions (Seering, Luria, Kaufman, & Hammer, 2019), there is a scarcity of empirical research on design differences contingent on chatbots' temporal profiles.

Therefore, the goal of the current work is to determine whether and how chatbots' different temporal profiles affect design considerations. The following two research questions guide our work: "Which design elements allow us to distinguish chatbots depending on whether they are aimed to help individuals to achieve short-, medium- or long-term goals?" (**RQ1**) and "How does a chatbot's temporal profile affect its design?" (**RQ2**).

Drawing attention to temporal aspects in user-chatbot relationships and examining the contingency of design choices on a chatbot's temporal profile will be of significant benefit, both, to practitioners using chatbots and managers in the early stages of chatbot development to guide the decision on which design elements need to be tailored to the time horizon of the user-chatbot relationship, and which design elements are neglectable or cumbersome.

This work also raises novel theory-related questions for human-computer interaction and (computer-mediated) communication researchers: If chatbots are used for one-time-only conversations, users will likely seek to get something done quickly via the chatbot, which makes the chatbot a mere "communication medium" (Zhao, 2006, p. 402). In contrast, if chatbots are used to achieve a specific personal long-term goal, users will rather be committed to undergo longer personal learning or development processes together *with* the chatbots which emphasizes the notion of chatbots as "social actors" (Reeves & Nass, 1996).

In the following, we review human-computer interaction and computer-mediated communication literature on time-dependent design aspects and provide a preliminary definition of short-, medium- and long-term chatbots. To answer our first research question, we then develop a taxonomy of time-dependent design aspects of chatbots (Section 3). The taxonomy development process follows the widely used taxonomy development procedure suggested by Nickerson, Varshney, and Muntermann (2013) and consists of two conceptual-to-empirical and five empirical-to-conceptual iterations. In total, we classify 37 chatbots described in research articles and 83 chatbots in the real world. To answer our second research question, we analyze and compare design characteristics of the classified chatbots systematically regarding their temporal profile (Section 4.1). Based on this analysis, we propose three chatbot archetypes (i.e., Ad-hoc Supporters, Temporary Advisors, and Persistent Companions), which allow researchers and practitioners to account for time-dependent aspects in chatbot design and thus provide common ground for further work (Section 4.2). Finally, we discuss broader implications of chatbots' temporal profiles and outline limitations and recommendations for further research (Section 5).

2. Conceptual background

The focal point of our research are time-dependent design aspects of domain-specific text-based conversational agents; here referred to as "chatbots". For the purpose of our analysis, we define chatbots as "software-based systems designed to interact with humans via text-based natural language" (Feine, Adam, Benke, Maedche, & Benlian, 2020, p. 127) that mimic common human-human conversations (Araujo, 2018) within the boundaries of a specific domain-knowledge

(Gnewuch et al., 2017).

Depending on their time horizon configuration, domain-specific chatbots can be characterized as short-term, medium-term, long-term, or life-long chatbots (Baraka et al., 2020). Chatbots designed for helping individuals to achieve short-term goals are defined by a single or very few occasional short interaction(s), while the latter – that is, chatbots designed for supporting individuals in achieving medium-to long-term, or even life-long goals – are comprised of multiple (interdependent) interactions over a certain period (Baraka et al., 2020, p. 29). Typical examples for short-term relationships are chatbots offering brief ad-hoc services such as customer support (e.g., Jaeger-LeCoultre) or self-diagnosis healthcare chatbots such as Babylon or Gyant, whereas typical medium- and long-term examples are chatbots for monitoring chronic conditions (e.g., Woebot) or learning processes (e.g., Duolingo).

Communication researchers have examined the role of time in social interactions in face-to-face (Werner, Altman, & Brown, 1992) and computer-mediated communication (Hesse, Werner, & Altman, 1988), for example, in reference to how relationship building processes in groups or between intimate partners develop through identifiable steps or stages (Hesse et al., 1988), or the impact of chronemic cues on perceived sender's intimacy (Walther & Tidwell, 1995). Construing chatbots as social actors (Ho, Hancock, & Miner, 2018) that can act on their own as "novel, human-created communication entities, playing their own social role" (Hoorn, 2018, p. 1) implies similar relationship-building processes between users and chatbots under the Media Equation Theory umbrella (Reeves & Nass, 1996). However, to account for such relationship processes in short or longitudinal user-chatbot relationships, chatbots likely need to be equipped with specific features or design elements to meet user expectations.

Consequently, some research groups have dedicated their work to understanding longitudinal relationship-building processes with chatbots (e.g., Bickmore & Picard, 2005). In human-robot interaction research, different temporal profiles of robots have long been acknowledged to be a major design characteristic that has crucial implications on their interactions with users (Baraka et al., 2020; Shibata, 2004; Yanco & Drury, 2004). A robot's temporal profile can be characterized by the following time-dependent dimensions: the *time horizon* as the total period during which the user engages with a robot, the *duration* of (individual) interaction(s), and the *frequency* in the case of multiple interactions (Baraka et al., 2020). The fourth dimension in human-robot interaction research concerns *synchronicity* which describes whether a (remotely controlled) robot responds immediately (synchronously) or delayed (asynchronously) when it is located in a more distant place. Considering robots as chatbots at the physical extreme of the "reality-virtuality continuum" (De Keyser, Köcher, Alkire, Verbeeck, & Kandampully, 2019) allows transferring some of those insights to virtual chatbots as well. While synchronicity seems less applicable to messenger-based chatbots which are virtually available around the clock, timespan, duration, and frequency are relevant dimensions to distinguish chatbots with different temporal profiles.

Conversely, while one scientific study of 103 domain-specific real-world chatbots found that the vast majority (84%) of them were developed for short-term purposes (Janssen, Passlick, Rodríguez Cardona, & Breitner, 2020), it is particularly difficult to find research on chatbots with a short-term temporal profile: A SCOPUS search on August 27, 2020, for academic material on the search string ("long*term" AND ("chatbot*" OR "conversational agent*" OR "relational agent*")) revealed 178 documents of which 89 (50.0%) actually dealt with chatbots that are built to foster and maintain long-term relations with users. The same search string but looking for "short*term" instead yielded 99 results of which only seven (7.1%) covered chatbots that are developed to support short-term goals. One reason might be limited research on short-term chatbots. Another might be that is more difficult to identify relevant academic literature on the topic of chatbots developed for short-term purposes because it is not called or tagged as such, making it cumbersome for researchers and practitioners alike to compare and derive

design and evaluation guidelines when developing chatbots for short-term goals.

In line with the two research questions, this work's objective is thus twofold: First, to identify all design elements contingent on the temporal dimension of user-chatbot relationships and to develop a comprehensive design taxonomy that allows us to characterize user-chatbot relationships with different time horizons (RQ1), and, second, to quantitatively assess differences between chatbots for either short-, medium-, or long-term purposes and to illustrate typical design configurations by identifying three chatbot archetypes (RQ2).

3. Methodology

To answer our two research questions, we applied a mixed-methods research approach (Creswell & Clark, 2011) and combined (a) qualitative methods to develop (part 1) and evaluate (part 2) the resulting "Design Taxonomy for Chatbots with Different Temporal Profiles" (RQ1) and (b) quantitative methods to identify differences in the occurrence of design characteristics in chatbots with different temporal profiles and to develop and differentiate three distinct chatbot archetypes aimed at helping users to achieve short-, medium-, or long-term goals (RQ2). The entire research procedure is illustrated in Fig. 1. The applied methodologies in each part and iteration are summarized in Table 1 and the following sections.

3.1. Part 1: Taxonomy development

To answer our first research question ("Which design elements allow us to distinguish chatbots depending on whether they are aimed to help individuals to achieve short-, medium-, or long-term goals?"), we develop a taxonomy of design elements to classify and differentiate the design of chatbots taking into account their temporal profile.

Taxonomies are well renowned in information systems and human-computer interaction research (Nickerson et al., 2013) as they allow the development of design principles that can inform the design of future artifacts (e.g., chatbots) based on the empirical analysis of structural patterns in existing artifacts (Williams, Chatterjee, & Rossi, 2008). A taxonomy consists of a number of dimensions, each of which has a subset of at least two characteristics. Every object that is classified based on the taxonomy must have exactly one characteristic of each dimension, neither more nor less (Nickerson et al., 2013).

Following the established taxonomy development method proposed by Nickerson et al. (2013), our taxonomy builds on existing conceptual design frameworks (Iterations 1 and 4) and extends them based on empirical observations of chatbots described in scientific articles (Iteration 2) and deployed in practice (Iterations 3 and 7). The *conceptual-to-empirical* iterations ensure that the taxonomy builds on and extends the latest knowledge discussed in the scientific literature, in particular, the taxonomy of design elements for domain-specific chatbots proposed by Janssen, Passlick, Rodríguez Cardona, and Breitner (2020); the *empirical-to-conceptual iterations* and the analysis and classification of the design of short-, medium-, and long-term chatbots introduced in scientific articles and practice increase rigor, relevance, and generalizability of the taxonomy.

Table 1 provides an overview of the applied research approaches, methodologies, and the sample of analyzed chatbots tested in each particular iteration.

Before conducting the first taxonomy development iteration, Nickerson et al.'s (2013) taxonomy development procedure requires the definition of (a) a *purpose* of the taxonomy, and the determination of (b) a *meta-characteristic* as "the most comprehensive characteristic that will serve as the basis for the choice of [all other] characteristics in the taxonomy [that are based] on the purpose of the taxonomy and in turn based on the users and their expected use of the taxonomy" (Nickerson et al., 2013, p. 343) and (c) *ending conditions*, that determine when the taxonomy development is completed.

Specifically, the (a) *purpose* of our taxonomy is to provide a framework of design guidelines for chatbots that support individuals in their short-, medium-, and long-term goals. Therefore, the (b) *meta-characteristics* of our taxonomy are all design elements that have a visible or experiential impact on the user-chatbot interaction.

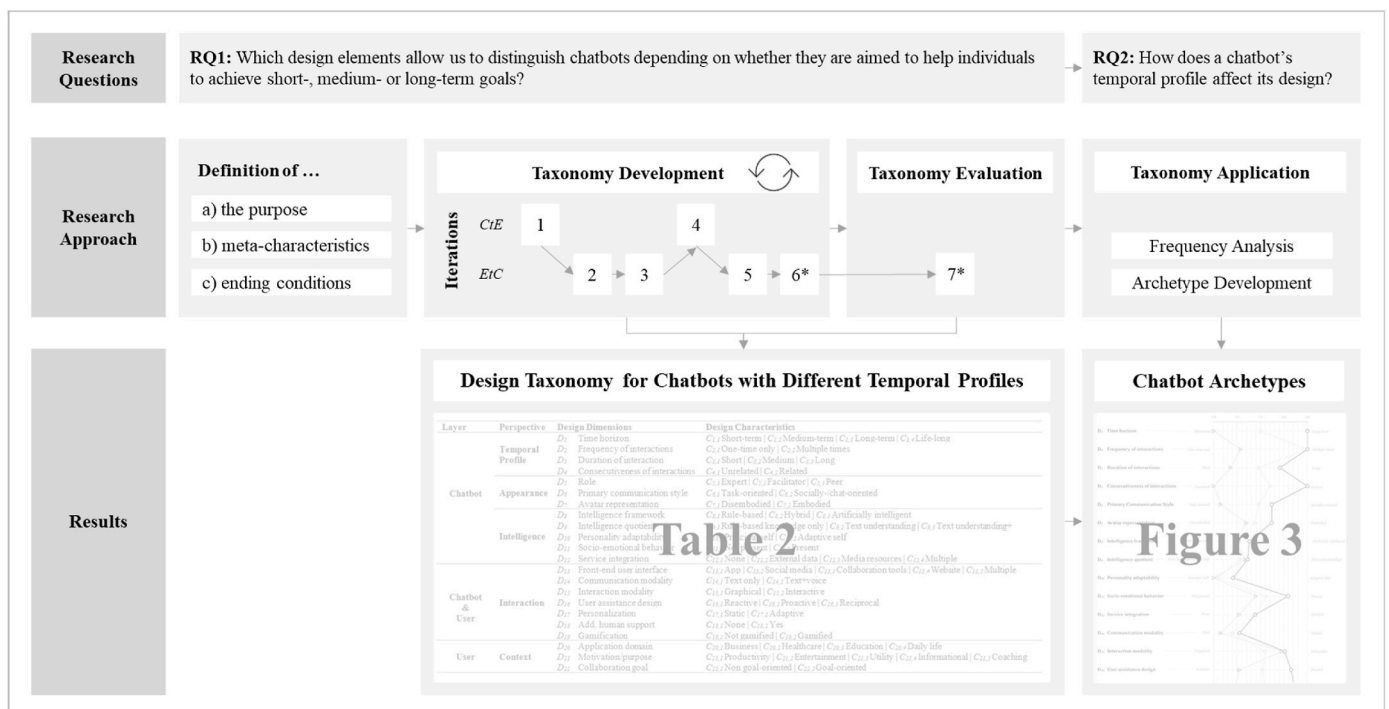Regarding the determination of (c) *ending conditions*, we adopted all

Fig. 1. Research Procedure. Note. CtE = Conceptual-to-Empirical; EtC = Empirical-to-Conceptual, * = all ending conditions fulfilled; Results only display thumbnails of Table 2 and Fig. 3.

Table 1
Research approach and procedure.

	Development						Evaluation	Application
Iteration	1	2	3	4	5	6	7	
Review phase **Approach**	*Nov-Dec 2019* *Conceptual-to-empirical*	*Nov-Dec 2019* *Empirical-to-conceptual*	*Jan-Feb 2020* *Empirical-to-conceptual*	*Apr-May 2020* *Conceptual-to-empirical*	*Jun-Jul 2020* *Empirical-to-conceptual*	*Jun-Jul 2020* *Empirical-to-conceptual*	*Sept-Nov 2020* *Empirical-to-conceptual*	*Jan-Feb 2021* **-**
Data selection approach	Narrative literature review *(cf.* Knote, Janson, Eigenbord, & Söllner, 2018*)*	Systematic literature review *(cf.* Vom Brocke et al., 2015*)*	Systematic search for real-world chatbots in online directories *(cf.* Seering et al., 2019*)*	Search engine alert	*(Re-classification of sample 1 & 2)*	*(Re-classification of sample 1 & 2)*	Secondary analysis of chatbots from the sample identified and analyzed by Janssen, Passlick, et al., 2020	Sample 1, 2, 3
Data basis *Conceptual-to-empirical*	Published chatbot design frameworks (n = 12)	–	–	Published taxonomy of design elements by Janssen, Passlick, et al. (2020)	–	–	–	–
Empirical-to-conceptual	–	**Sample 1:** Chatbot case studies described in scientific articles *(cf.* Table A3*, Web Appendix)*	-	-	**Sample 1**	**Sample 1**	-	**Sample 1**
	–	–	**Sample 2:** Available chatbots in the real-world *(cf.* Table A4*, Web Appendix)*	–	**Sample 2**	**Sample 2**	-	**Sample 2**
	–	–	–	–	–	–	**Sample 3:** Still available real-world chatbots from Janssen, Passlick, et al. (2020) sample *(cf.* Table A5*, Web Appendix)*	**Sample 3**
# Classified Chatbots	-	37	41	-	78	78	42	120

objective and subjective conditions suggested by Nickerson et al. (2013). Throughout the entire taxonomy development process, it was constantly discussed and checked in each iteration whether all identified design dimensions and characteristics fulfilled all ending conditions (cf. Table A7 in the Web Appendix).

In total, the development of the taxonomy, required two conceptual-to-empirical and four empirical-to-conceptual development iterations before all ending conditions were fulfilled (cf. Web Appendix Table A7). Summarized insights into each iteration of the taxonomy development are further outlined in the following subsections. Fig. 2 visualizes all changes on the design dimension level across iterations throughout the taxonomy development process.

3.1.1. Iteration 1 - Conceptual-to-empirical: Identification of chatbot design elements

In the first *conceptual-to-empirical* iteration, which we conducted between November and December 2019, we reviewed published chatbot design classifications and frameworks based on a narrative literature review approach. Since classification schemes and naming conventions for chatbots are fragmented along different thematic axes (e.g., multiple vs. single-user chatbots), use cases (e.g., healthcare vs. shopping chatbots), and across multiple research disciplines (e.g., information systems vs. human-computer interaction), a narrative literature review proves useful and efficient for establishing an overview of the latest developments in a condensed format (Vom Brocke et al., 2015) and to derive an initial set of design dimensions before we set out to run the first empirical-to-conceptual iteration.

The initial set of conceptually identified design dimensions is visible in Fig. 2 (this manuscript) and Table A1 (cf. Web Appendix). The detailed description of the research procedure, the used databases, and search strings as well as the initial set of design dimensions and characteristics are described in depth in the Web Appendix.

3.1.2. Iteration 2 - Empirical-to-conceptual: Classification of proof-of-concept chatbots described in scientific articles

For the second iteration, we chose an *empirical-to-conceptual* approach to complement and test our initial set of design dimensions based on published chatbot design and development case studies described in scientific articles (Cooper, 1988; Knote et al., 2018). Concentrating first on chatbots described in scientific articles (cf. Table A3, Web Appendix) allowed us to understand which design dimensions chatbot developers and researchers focused on when developing different types of chatbots. Thereby, in this iteration, we could classify all sampled chatbots described in the studies with regards to the set of design dimensions already identified and could simultaneously look for new design dimensions that the articles' authors explicitly mentioned or discussed, and which we had not identified based on the review of conceptual design frameworks in Iteration 1.

To obtain a comprehensive set of 37 scientific articles that either focused on conceptualizing or developing chatbots in parts or as a whole, we followed a systematic literature review approach (cf. Figure A1 in the Web Appendix) including using search string in

Table 2
Design taxonomy for chatbots with different temporal profiles.

Layer	Perspective	Design Dimensions	Design Characteristics
Chatbot	**Temporal Profile**	D_1 Time horizon	$C_{1,1}$ Short-term $\mid$ $C_{1,2}$ Medium-term $\mid$ $C_{1,3}$ Long-term $\mid$ $C_{1,4}$ Life-long
		D_2 Frequency of interactions	$C_{2,1}$ One-time only $\mid$ $C_{2,2}$ Multiple times
		D_3 Duration of interaction	$C_{3,1}$ Short $\mid$ $C_{3,2}$ Medium $\mid$ $C_{3,3}$ Long
		D_4 Consecutiveness of interactions	$C_{4,1}$ Unrelated $\mid$ $C_{4,2}$ Related
	Appearance	D_5 Role	$C_{5,1}$ Expert $\mid$ $C_{5,2}$ Facilitator $\mid$ $C_{5,3}$ Peer
		D_6 Primary communication style	$C_{6,1}$ Task-oriented $\mid$ $C_{6,2}$ Socially-/chat-oriented
		D_7 Avatar representation	$C_{7,1}$ Disembodied $\mid$ $C_{7,2}$ Embodied
	Intelligence	D_8 Intelligence framework	$C_{8,1}$ Rule-based $\mid$ $C_{8,2}$ Hybrid $\mid$ $C_{8,3}$ Artificially intelligent
		D_9 Intelligence quotient	$C_{9,1}$ Rule-based knowledge only $\mid$ $C_{9,2}$ Text understanding $\mid$ $C_{9,3}$ Text understanding+
		D_{10} Personality adaptability	$C_{10,1}$ Principal self $\mid$ $C_{10,2}$ Adaptive self
		D_{11} Socio-emotional behavior	$C_{11,1}$ Not present $\mid$ $C_{11,2}$ Present
		D_{12} Service integration	$C_{12,1}$ None $\mid$ $C_{12,2}$ External data $\mid$ $C_{12,3}$ Media resources $\mid$ $C_{12,4}$ Multiple
Chatbot & User	**Interaction**	D_{13} Front-end user interface	$C_{13,1}$ App $\mid$ $C_{13,2}$ Social media $\mid$ $C_{13,3}$ Collaboration tools $\mid$ $C_{13,4}$ Website $\mid$ $C_{13,5}$ Multiple
		D_{14} Communication modality	$C_{14,1}$ Text only $\mid$ $C_{14,2}$ Text + voice
		D_{15} Interaction modality	$C_{15,1}$ Graphical $\mid$ $C_{15,2}$ Interactive
		D_{16} User assistance design	$C_{16,1}$ Reactive $\mid$ $C_{16,2}$ Proactive $\mid$ $C_{16,3}$ Reciprocal
		D_{17} Personalization	$C_{17,1}$ Static $\mid$ $C_{17,2}$ Adaptive
		D_{18} Add. human support	$C_{18,1}$ None $\mid$ $C_{18,2}$ Yes
		D_{19} Gamification	$C_{19,1}$ Not gamified $\mid$ $C_{19,2}$ Gamified
User	**Context**	D_{20} Application domain	$C_{20,1}$ Business $\mid$ $C_{20,2}$ Healthcare $\mid$ $C_{20,3}$ Education $\mid$ $C_{20,4}$ Daily life
		D_{21} Motivation/purpose	$C_{21,1}$ Productivity $\mid$ $C_{21,2}$ Entertainment $\mid$ $C_{21,3}$ Utility $\mid$ $C_{21,4}$ Informational $\mid$ $C_{21,5}$ Coaching
		D_{22} Collaboration goal	$C_{22,1}$ Non goal-oriented $\mid$ $C_{22,2}$ Goal-oriented

scientific databases, forward, backward, and similarity search, which "takes a structured approach to identifying, evaluating, and synthesizing research" (Vom Brocke et al., 2015, p. 9). We concentrated the search on chatbots in application domains characterized by processes that show a progressive evolution over time and where we expected to find examples of short-, medium- and long-term chatbots (i.e., healthcare, education, and business). All parameters of the systematic literature review, search strings, the screening, review and coding procedure, as well as the results of this iteration and changes in the taxonomy are described in detail in the Web Appendix.

3.1.3. Iteration 3 - Empirical-to-conceptual: Classification of real-world chatbots

In this iteration, we chose the empirical-to-conceptual path again. To ensure the relevance of our taxonomy, we aimed at triangulating the sample of short-, medium-, and long-term chatbots described in the latest scientific articles with state-of-the-art examples of actually available chatbots in the real-world and systematically sampled a set of chatbots from online chatbot directories such as botlist.co or thereisaboutforthat.com and from curated chatbot platforms and magazines (cf. Table A4, Web Appendix). As there are still no standardized procedures that determine how to sample or analyze chatbots "in the wild" (Seering et al., 2019), we describe our systematic approach (i.e., sampling strategy, data selection, coding, and classification procedures) in full detail in the Web Appendix.

3.1.4. Iteration 4 - Conceptual-to-empirical: Refinement of the taxonomy

The publication of a "Taxonomy of Design Elements for Domain-specific Chatbots" by Janssen, Passlick, et al. (2020) on April 6, 2020, allowed us to challenge and further refine our taxonomy in another conceptual-to-empirical iteration. Therefore, we compared both taxonomic structures, all design dimensions, and design characteristics hitherto and identified that we had eleven dimensions in common that were identical or very similar in meaning, four dimensions that had not been included in the aforementioned taxonomy, and five which we had not listed in ours yet. Interweaving and complementing the taxonomies promised a more comprehensive understanding of differences between chatbots with different temporal profiles since it synthesized design dimensions identified based on the analysis of chatbots from different application domains. The merging process is visible in Fig. 2 and described in the Web Appendix in full detail as well.

3.1.5. Iteration 5 - Empirical-to-conceptual: Re-classification of all chatbots based on refined taxonomy

To test and assess the new structure of the taxonomy again, we proceeded to carry out another empirical iteration path and coded our two chatbot samples from Iteration 2 and 3 based on the taxonomic structure and terminologies retrieved in Iteration 4. Whenever one of the design dimensions could not be assessed based on the stored data, we revisited the chatbots and updated the chat logs accordingly. As a result of continued discussions during this iteration, we substantially reordered the structure of the design dimensions as visible in Fig. 2 (cf. full methodological details in the Web Appendix).

3.1.6. Iteration 6 - Empirical-to-conceptual: Re-classification of chatbots to meet ending conditions

Due to the addition of one new dimension and merging three design characteristics into one in Iteration 5 (cf. Fig. 2), the ending conditions were still not fulfilled in the last iteration (cf. Figure A7 in the Web Appendix), which rendered another empirical-to-conceptual iteration imperative. In this iteration, we classified all chatbots again with a focus on the newly added dimension and the changes in the discussed design characteristics. After this iteration, all ending conditions were fulfilled and the taxonomy development process was complete.

3.2. Part 2: Taxonomy evaluation

To confirm that the taxonomy could be applied to other application domains and by individuals not involved in the development of the taxonomy, we conducted an *evaluation iteration* based on a new sample of real-world chatbots (cf. Table A5 in the Web Appendix). In line with the taxonomy evaluation framework by Szopinski, Schoormann, and Kundisch (2019) this evaluation iteration was characterized by the following directions:

Regarding the *subject of evaluation* (the 'who'), we involved two additional researchers (C & D) with chatbot domain and taxonomy method expertise who had not been involved in the taxonomy development process before. Before starting with the actual evaluation phase, these two researchers provided feedback concerning the interpretation of dimensions and characteristics defined in the taxonomy codebook provided by researchers A & B (cf. Table A2 in the Web Appendix) which led to a refinement of the definition of the *frequency of interactions* design dimension.

Regarding the *method of evaluation* (the 'how'), we followed the
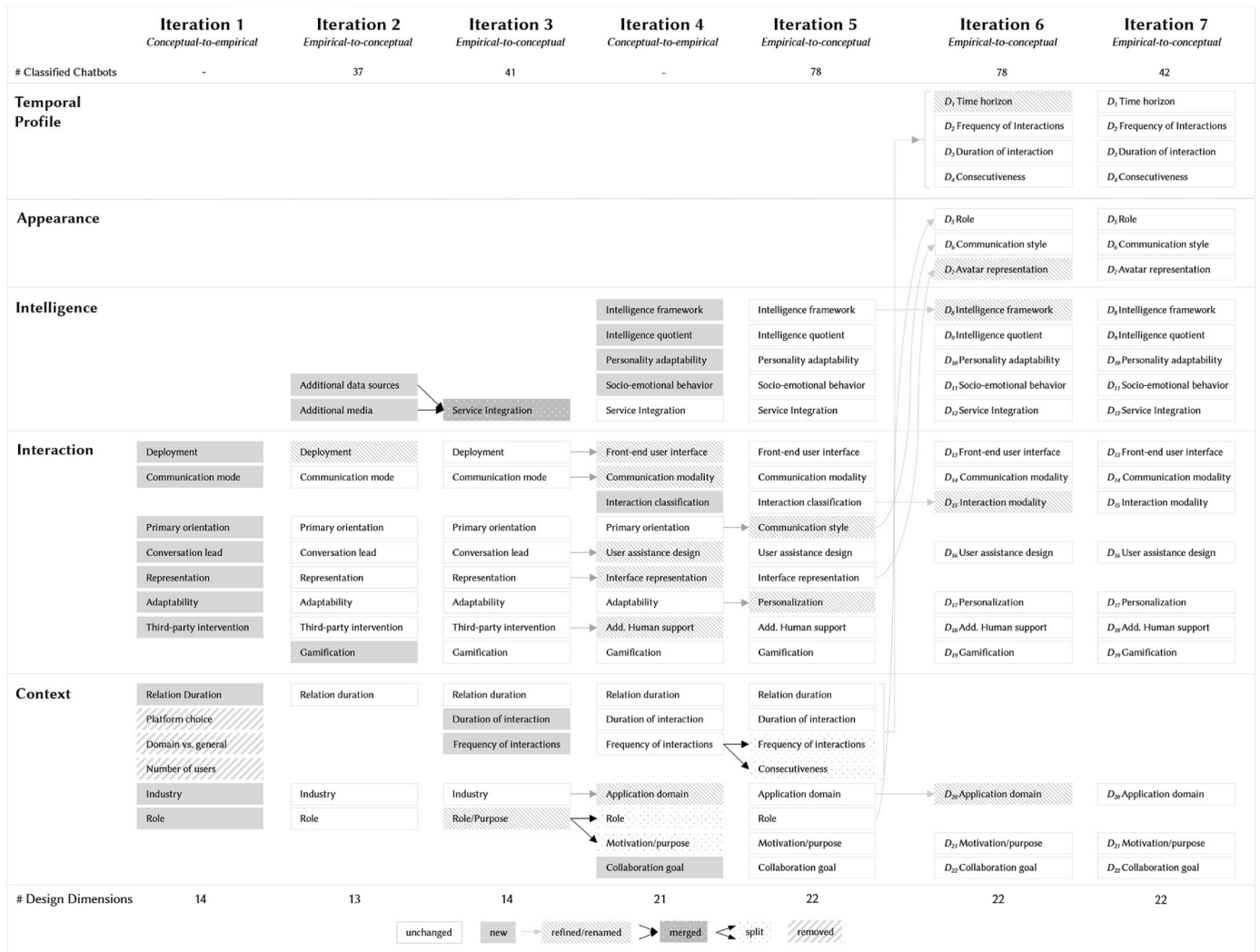
Design Dimension Development Across Iterations

Fig. 2. Design dimension development across iterations.

illustrative scenario technique for which the researchers C & D applied the present taxonomy to a new set of real-world objects based on the last version of the design dimensions' and characteristics' definitions: "Applying a present taxonomy to real-world objects allows researchers to evaluate their […] usefulness for classifying, differentiating, and comparing objects as well as to evaluate their robustness, utility, efficacy, stability, and completeness" (Szopinski et al., 2019, p. 11).

Regarding the *object of evaluation* (the 'what'), we re-applied the 103 real-world chatbots identified and classified by Janssen, Passlick, et al. (2020). An analysis of this sample had revealed that the chatbots were completely disjunctive from our sample as Janssen, Passlick, et al. (2020) had applied a different sampling strategy that had focused on sourcing chatbots from the chatbot directories "chatbots.org" and "botlist.org" in May 2019 while we had pursued a purposive sampling strategy to identify the most popular or renowned short- and long-term real-world chatbots per application domain. Additionally, the analysis of the aforementioned sample had revealed that the sample provided a large number of chatbots that could be attributed to the application domains *Business*, which was less dominant in our sample, and *Daily Life*, which we, therefore, added as a new design characteristic to the taxonomy as well.

Between September and November 2020, researchers C & D revisited all 103 chatbots. If a chatbot was no longer available via the original

URL, the chatbot's (or company's) name was used to search for the chatbot via Google Search. Eventually, only 42 chatbots were still accessible. All other chatbots were either no longer detectable on the websites (e.g., SOFIA (TRAVEL)), the websites were offline (e.g., SOA SEKS CHECK) or the chatbot did not answer anymore (e.g., IFRS ROOKIES). Some chatbots also had been (re-)replaced with live chats with human agents in the meantime (e.g., AXA).

To make sure that all possible design dimensions could be assessed for the remaining 42 chatbots, researchers C & D followed the updated semi-structured conversation guidelines (cf. details on Iteration 3 and Table A6 in the Web Appendix) to engage in conversations with the chatbots. Similar to Iteration 3, chat logs, screenshots, and personal notes were stored in an independent database which was later merged with the previous database as the analysis of the different chatbot archetypes was performed based on the full sample across all iterations.

To check the extent to which the classification of the two evaluation researchers matched with those of previous iterations, all 42 chatbots were also classified by researcher A again who had also classified all other chatbots in the previous iterations. From this, inter-coder reliability was calculated for the entire taxonomy as well as for each dimension and each inter-coder combination (C & D, C & A, and D & A). All inter-coder reliabilities were above 90% and, thus, considered satisfactory (Kassarjian, 1977). The largest variation appeared in the
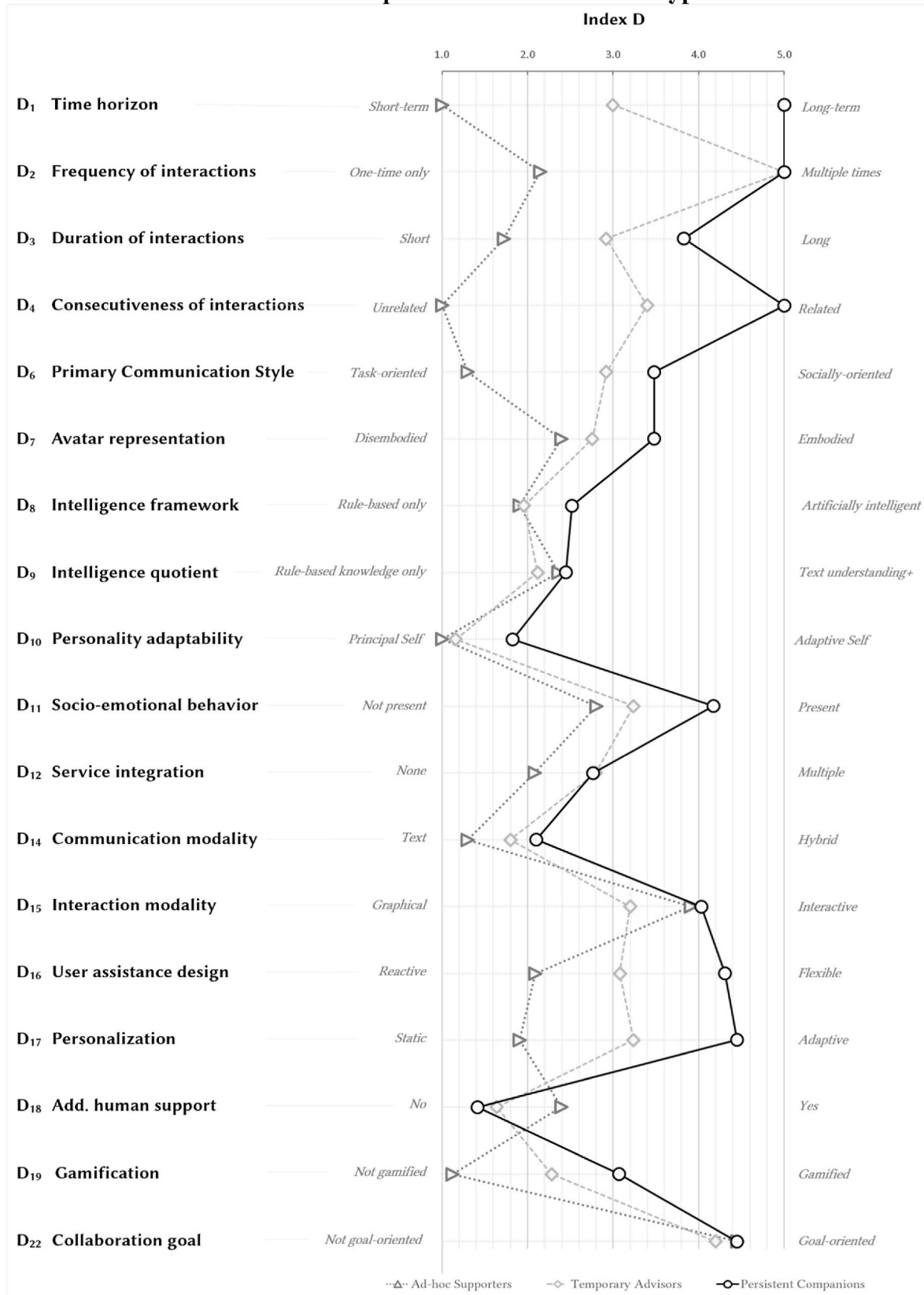
Fig. 3. Comparison of chatbot archetypes.

dimensions D_{21} motivation for chatbot use and D_{12} service integration, which led to a refinement of their definitions. Overall, the evaluation participants declared the taxonomy useful, complete, and comprehensible. Since no characteristics or dimensions were deleted, added, or split in the evaluation iteration, all ending conditions were fulfilled and the final taxonomy could be confirmed (cf. Table A7, Web Appendix).

3.3. Part 3: Taxonomy application

To answer our second research question (i.e., "How does a chatbot's temporal profile affect its design?") we (i) analyzed and evaluated the distribution of design characteristics per design dimension and per temporal profile in all 120 chatbots that we had sampled for the taxonomy development (i.e., all 37 chatbots sampled from scientific articles in Iteration 2, all 41 real-world chatbots sampled in Iteration 3, and all 42 chatbots sampled to evaluate the taxonomy in Iteration 7) and (ii) developed an index to abstract three time-dependent chatbot archetypes to better understand differences in the design configuration of short-, medium- and long-term chatbots.

3.3.1. Frequency analysis of chatbots' time-dependent design characteristics

Since each chatbot was classified by exactly one design characteristic per design dimension (Nickerson et al., 2013), resulting in 2640 codes (22 design dimension codes * 120 chatbots), frequency analysis is suitable to be applied. Frequency analysis is a "relevant brick to bridge the gap between qualitative and quantitative methods (mixed-methods research)" and can be described "as a process that breaks down complex behaviors into smaller units [by counting] their occurrences" (Rack, Zahn, & Mateescu, 2018, chap. 14, p. 278). We counted occurrences of design characteristics per design dimension and per chatbot temporal profile. Lastly, tests of independence were conducted to detect statistically significant differences in the distribution of design characteristics between short-, medium-, and long-term chatbots: that is χ^2- Tests or Fisher's Exact Tests (FETs) respectively. χ^2-Tests are recommended when all cells have expected frequencies greater than or equal to 5 (Field, 2009, p. 692); FETs are particularly recommended when any expected frequencies are less than 1 (Sauro & Lewis, 2016, p. 79). Results are presented in a comprehensive contingency table (Table 3).

3.3.2. Chatbot archetypes development

Since "reports of descriptive metrics such as frequencies are [...] in most cases not sufficient to fully understand complex aspects [...] on a more general level" (Rack et al., 2018, chap. 14, p. 288), we used the frequency analysis results as a basis for further statistical analysis by calculating an index per design dimension that can be used to compare chatbots with different temporal profiles and to systematically derive time-dependent chatbot archetypes.

We computed this "*Index I_d*" for each of the 17 design dimensions D whose characteristics can be ordered (e.g., the characteristics $C_{3,1}$ short, $C_{3,2}$ medium, and $C_{3,3}$ long of the dimension D_3 duration of interactions). This Index I_d can take a value between 1 and 5 and will be computed as shown in Equation (1), where C_i represents the frequency of the i-th design characteristic C_d and n the number of design characteristics per design dimension D:

$$Index\ I_d = \frac{1}{n} * \sum_{i=1}^{n} C_i * \left(1 + \left(\frac{4}{n-1}\right) * (i-1)\right) \qquad (1)$$

That is, Index I_d is the mean of the factored frequencies of all design characteristics of a design dimension. Index I_d is computed for each design dimension and for each short-, medium-, and long-term chatbot archetype separately. All index values per design dimension and chatbot archetype are plotted in a "design configurator" on semantic differential scales to compare all three archetypes simultaneously (Fig. 3). The elements of the design dimensions D_5 role, D_{13} front-end user interface,

D_{20} application domain, and D_{21} motivation/purpose for chatbot use could not be ordered in a meaningful way and are therefore not represented in the figure.

4. Results

In line with our two research questions, our results are presented in two parts as well: First, the final "Design Taxonomy for Chatbots with Different Temporal Profiles", resulting from the taxonomy development and evaluation procedure is introduced (RQ1) and, second, the results from applying the taxonomy to 120 chatbots to analyze differences between short-, medium- and long-term chatbots (RQ2) are presented.

4.1. Part 1 & 2: Design taxonomy for chatbots with different temporal profiles

The final taxonomy (Table 2) provides chatbot designers and researchers with a framework of design dimensions and characteristics for chatbots with different temporal profiles (RQ1). In the following, we present all design dimensions and design characteristics following the structure of the five overarching design perspectives, which themselves can be differentiated with regards to whether they relate to the (i) chatbot, to the (ii) user-chatbot relationship, or to the (iii) user alone (Baraka et al., 2020, p. 3). The perspectives *temporal profile, appearance,* and *intelligence* relate to the chatbot, the *interaction* perspective reflects the user-chatbot relationship, and *context* relates to the user's circumstances and intentions to engage with the chatbot in the first place.

4.1.1. Temporal profile

The first overarching perspective, a chatbot's temporal profile, can be characterized by the D_1 time horizon of the user-chatbot relationship, the D_2 duration of (individual) interactions, the D_3 frequency, and the D_4 consecutiveness of interactions with the user.

The D_1 time horizon of a user-chatbot relationship can be either $C_{1,1}$ short-, $C_{1,2}$ medium-, $C_{1,3}$ long-term, or $C_{1,4}$ life-long (Baraka et al., 2020). Short-term relationships are characterized by only a single or few occasional interactions (e.g., self-diagnosis healthcare chatbots like BABYLON or GYANT). Medium- and long-term relationships always consist of multiple interactions over a certain period (Baraka et al., 2020, p. 29). A typical example for a medium-term chatbot is an educational chatbot that teaches a particular course's defined junk of content (e.g., CODE-MONKEY or BOOKBUDDY) over a defined period (e.g., one school semester). A typical long-term example is a chatbot that monitors a patients' weight-loss progress (e.g., WEIGHTMENTOR) for a sustained period. Life-long relations are different from long-term chatbots as they aim to offer companionship similar to a partner- or friendship that may persist through major changes in a person's life (Baraka et al., 2020, p. 30).

We furthermore included the design dimension D_4 consecutiveness of interactions to capture whether multiple interactions are $C_{4,1}$ unrelated or $C_{4,2}$ related as a chatbot-based service. Unrelated interactions are typical for chatbots that provide style recommendations based on the current product database (e.g., Levis's INDIGO) or chatbots that curate information, for example, about HIV (e.g., SHIHBOT). Related interactions are, for example, typical for a language teaching chatbot such as DUO-LINGO, which tutors multiple sequential units of a topic.

4.1.2. Appearance

The appearance perspective gathers all design dimensions that characterize a chatbot's look and feel. Concretely, a chatbot's appearance can be characterized by its D_5 role (i.e., $C_{5,1}$ expert, $C_{5,2}$ facilitator, or $C_{5,3}$ peer; Bittner, Oeste-Reiß, & Leimeister, 2019), its D_6 primary communication style (i.e., $C_{6,1}$ task- or $C_{6,2}$ socially/chat-oriented; Verhagen, van Nes, Feldberg, & van Dolen, 2014), and its D_7 avatar representation (i.e., disembodied or embodied).

Table 3
Distribution of design characteristics per design dimension and temporal profile.

	Total		Short-term		Medium-term		Long-term		χ²-/Fisher's Exact Tests				Cramer's V	
# Chatbots analyzed	120		67		24		29		Test	df	Value	p	V	p
# Codes	2640		1474		528		638							
	n	%	n	%	n	%	n	%						
Temporal Profile														
D₁ Time horizon														
C₁,₁ Short-term	66	55.0	66	100.0	0	0.0	0	0.0	FE		217.87	.000	.816	.000
C₁,₂ Medium-term	25	20.8	0	0.0	25	100.0	0	0.0						
C₁,₃ Long-term	28	23.3	0	0.0	0	0.0	28	96.6						
C₁,₄ Life-long	1	0.8	0	0.0	0	0.0	1	3.4						
D₂ Frequency of interactions														
C₂,₁ One-time only	47	39.2	47	71.2	0	0.0	0	0.0	X²	2	63.21	.000	.726	.000
C₂,₂ Multiple times	73	60.8	19	28.8	25	100.0	29	100.0						
D₃ Duration of interactions														
C₃,₁ Short	51	42.5	44	66.7	5	20.0	2	6.9	FE		50.02	.000	.459	.000
C₃,₂ Medium	49	40.8	20	30.3	16	64.0	13	44.8						
C₃,₃ Long	20	16.7	2	3.0	4	16.0	14	48.3						
D₄ Consecutiveness of interactions														
C₄,₁ Unrelated	76	63.3	66	100.0	10	40.0	0	0.0	X²	2	94.16	.000	.886	.000
C₄,₂ Related	44	36.7	0	0.0	15	60.0	29	100.0						
Appearance														
D₅ Role														
C₅,₁ Expert	26	21.7	19	28.8	4	16.0	3	10.3	FE		25.36	.000	.334	.000
C₅,₂ Facilitator	74	61.7	45	68.2	16	64.0	13	44.8						
C₅,₃ Peer	20	16.7	2	3.0	5	20.0	13	44.8						
D₆ Communication style														
C₆,₁ Task-oriented	85	70.8	61	92.4	13	52.0	11	37.9	X²	2	34.38	.000	.535	.000
C₆,₂ Socially-/chat-oriented	35	29.2	5	7.6	12	48.0	18	62.1						
D₇ Avatar representation														
C₇,₁ Disembodied	68	56.7	43	65.2	14	56.0	11	37.9	X²	2	6.09	.046	.225	.046
C₇,₂ Embodied	52	43.3	23	34.8	11	44.0	18	62.1						
Intelligence														
D₈ Intelligence framework														
C₈,₁ Rule-based	59	49.2	36	54.5	13	52.0	10	34.5	FE		8.37	.049	.219	.023
C₈,₂ Hybrid	58	48.3	30	45.5	12	48.0	16	55.2						
C₈,₃ Artificially intelligent	3	2.5	0	0.0	0	0.0	3	10.3						
D₉ Intelligence quotient														
C₉,₁ Rule-based knowledge only	44	36.7	21	31.8	12	48.0	11	37.9	FE		8.94	.040	.199	.043
C₉,₂ Text understanding	72	60.0	45	68.2	12	48.0	15	51.7						
C₉,₃ Text understanding+	4	3.3	0	0.0	1	4.0	3	10.3						
D₁₀ Personality adaptability														
C₁₀,₁ Principal self	113	94.2	66	100.0	24	96.0	23	79.3	FE		13.18	.001	.364	.001
C₁₀,₂ Adaptive self	7	5.8	0	0.0	1	4.0	6	20.7						
D₁₁ Socio-emotional behavior														
C₁₁,₁ Not present	52	43.3	35	53.0	11	44.0	6	20.7	X²	2	8.96	.010	.274	.010
C₁₁,₂ Present	67	55.8	30	45.5	14	56.0	23	79.3						
D₁₂ Service integration														
C₁₂,₁ None	38	31.7	23	34.8	7	28.0	8	27.6	FE		7.84	.243	.178	.274
C₁₂,₂ External data	44	36.7	31	47.0	6	24.0	7	24.1						
C₁₂,₃ Media resources	25	20.8	9	13.6	7	28.0	9	31.0						
C₁₂,₄ Multiple	13	10.8	3	4.5	5	20.0	5	17.2						
Interaction														
D₁₃ Front-end user interface														
C₁₃,₁ Application	23	19.2	6	9.1	4	16.0	13	44.8	FE		39.86	.000	.410	.000
C₁₃,₂ Social media	34	28.3	15	22.7	9	36.0	10	34.5						
C₁₃,₃ Collaboration tools	11	9.2	4	6.1	5	20.0	2	6.9						
C₁₃,₄ Website	45	37.5	39	59.1	4	16.0	2	6.9						
C₁₃,₅ Various	7	5.8	2	3.0	3	12.0	2	6.9						
D₁₄ Communication Modality														
C₁₄,₁ Text	102	85.0	61	92.4	20	80.0	21	72.4	FE		7.01	.023	.241	.035
C₁₄,₂ Text + voice	18	15.0	5	7.6	5	20.0	8	27.6						
D₁₅ Interaction Modality														
C₁₅,₁ Graphical	35	29.2	18	27.3	10	40.0	7	24.1	X²	2	2.27	.329	.138	.329
C₁₅,₂ Interactive	84	70.0	48	72.7	14	56.0	22	75.9						
D₁₆ User assistance design														
C₁₆,₁ Reactive	47	39.2	39	59.1	8	32.0	0	0.0	X²	4	36.63	.000	.391	.000
C₁₆,₂ Proactive	36	30.0	18	27.3	8	32.0	10	34.5						
C₁₆,₃ Reciprocal	37	30.8	9	13.6	9	36.0	19	65.5						
D₁₇ Personalization														
C₁₇,₁ Static	66	55.0	51	77.3	11	44.0	4	13.8	X²	2	34.35	.000	.535	.000
C₁₇,₂ Adaptive	54	45.0	15	22.7	14	56.0	25	86.2						
D₁₈ Add. Human support														
C₁₈,₁ No	90	75.0	43	65.2	21	84.0	26	89.7	X²	2	7.82	.023	.255	.020
C₁₈,₂ Yes	30	25.0	23	34.8	4	16.0	3	10.3						
D₁₉ Gamification														
C₁₉,₁ No	95	79.2	64	97.0	17	68.0	14	48.3	X²	2	31.35	.000	.511	.000
C₁₉,₂ Yes	25	20.8	2	3.0	8	32.0	15	51.7						

Table 3 (*continued*)

		Total		Short-term		Medium-term		Long-term		χ^2-/Fisher's Exact Tests				Cramer's V	
# Chatbots analyzed		120		67		24		29		Test	df	Value	p	V	p
# Codes		2640		1474		528		638							
		n	%	n	%	n	%	n	%						
Context															
D_{20} Application Domain	$C_{20,1}$ Business	45	37.5	40	60.6	5	20.0	0	0.0	FE		50.21	**.000**	.441	**.000**
	$C_{20,2}$ Education	24	20.0	6	9.1	10	40.0	8	27.6						
	$C_{20,3}$ Healthcare	36	30.0	13	19.7	5	20.0	18	62.1						
	$C_{20,4}$ Daily Life	15	12.5	7	10.6	5	20.0	3	10.3						
D_{21} Motivation/purpose	$C_{21,1}$ Productivity	9	7.5	9	13.6	0	0.0	0	0.0	FE		72.51	**.000**	.531	**.000**
	$C_{21,2}$ Entertainment	6	5.0	3	4.5	2	8.0	1	3.4						
	$C_{21,3}$ Utility	47	39.2	35	53.0	7	28.0	5	17.2						
	$C_{21,4}$ Informational	25	20.8	19	28.8	5	20.0	1	3.4						
	$C_{21,5}$ Coaching	32	26.7	0	0.0	10	40.0	22	75.9						
D_{22} Collaboration goal	$C_{22,1}$ Not goal-oriented	18	15.0	9	13.6	5	20.0	4	13.8	FE		0.75	.738	.072	.783
	$C_{22,2}$ Goal-oriented	102	85.0	57	86.4	20	80.0	25	86.2						

Note. FE = Fisher's Exact Test, significant values at **p < .05 are in boldface.**

4.1.3. Intelligence

In contrast to the appearance perspective, the intelligence perspective entails all design dimensions that are characteristics of a chatbot's inner working mechanisms on which its functionalities are based. These include its D_8 intelligence framework, its D_9 intelligence quotient, and its capabilities to D_{10} adapt its personality, to adequately and autonomously react D_{11} socio-emotionally to user sentiments, and D_{12} to integrate and process information from further services and sources such as $C_{12,2}$ additional external data or $C_{12,3}$ media resources. The differentiation of service integrations into $C_{12,3}$ additional media or $C_{12,2}$ external data highlights whether a chatbot is capable of broadcasting media such as video and pictures to users (e.g., Sephora Kik Bot) or of integrating and processing external data, for example, from a product database (e.g., 1-800-Flowers) or from users' devices (e.g., Lark). Furthermore, we attenuated whether D_{11} socio-emotional behaviors were $C_{11,2}$ present or $C_{11,1}$ not and classified spontaneous empathic reactions as present (e.g., "that's great to hear" in reaction to a user who had indicated that she had slept well; e.g., Youper, Brook).

4.1.4. Interaction

The interaction perspective comprises all design dimensions that are related to the interactions between the user and the chatbot. This includes the D_{13} front-end user interface for which a chatbot has been developed, which allows a user to access (or not) a certain chatbot, for example, via $C_{13,2}$ social media platform messengers such as Facebook, via stand-alone $C_{13,1}$ apps, on $C_{13,4}$ websites, via $C_{13,3}$ communication and collaboration messenger platforms such as kik or Whatsapp, or $C_{13,5}$ combinations of these interfaces.

D_{14} Communication modality (i.e., $C_{14,1}$ text, $C_{14,2}$ speech or $C_{14,3}$ hybrid), D_{15} interaction modality (i.e., $C_{15,1}$ graphical via quick response buttons only or $C_{15,2}$ interactive which allows for free text inputs) and D_{16} user assistance design define the mode of operation of the user-chatbot relationship. D_{16} User assistance denotes the "locus of control" (Følstad et al., 2019) and indicates who is in charge of the conversational flow, meaning whether the chatbot only reacts to user inputs ($C_{16,1}$ reactive), whether it steers the conversation ($C_{16,2}$ proactively) or whether it is capable to alternate ($C_{16,3}$ reciprocal).

D_{17} Personalization refers to a chatbot's capability to adapt a conversation based on previous interactions and inputs from a user. D_{18} Additional human support refers to the possibility of human interventions that complement or accede the user-chatbot interaction (Kowatsch et al., 2017). D_{19} Gamification specifies whether gamification elements such as quizzes are $C_{19,2}$ present or $C_{19,2}$ not.

4.1.5. Context

The context perspective gathers all design dimensions that refer to users' initial motives to engage with a chatbot. This is reflected in the D_{20} application domain, whether a user seeks a D_{22} goal-oriented collaboration (or not) and a user's primary D_{21} motivation to engage with a chatbot in the first place.

4.2. Part 3: Frequency analysis and archetype development

4.2.1. Distribution of design characteristics per temporal profile

Comparing the distribution of design characteristics per design dimension, χ^2-Tests or FETs respectively revealed significant differences between short-, medium-, and long-term chatbots (RQ2) for 19 out of 22 design dimensions (cf. Table 3). There were no significant differences in the distribution of design characteristics per design dimension between the different temporal profiles for D_{12} service integration, $FET = 7.84$, $p = .243$, $V = 0.178$, D_{15} interaction modality, $\chi^2(120, 2) = 2.27$, $p = .329$, $V = 0.138$, and D_{22} collaboration goal, $FET = 0.75$, $p = .738$, $V = 0.072$.

4.2.1.1. Temporal profile.
In total, 66 (55.0%) of all 120 chatbots in our sample were classified as short-term, 25 (20.8%) as medium-term, and

29 (24.2%) as long-term chatbots. The only chatbot in our sample that could potentially be classified as a life-long chatbot (i.e., REPLIKA) was here classified as long-term as it still is not mature enough to answer the claim of life-long companionship.

Regarding the D_2 frequency of interactions, most short-term chatbots (71.2%) provided one-time-only interactions. Still, a proportion of 28.8% was classified as offering multiple occasions for interactions. For example, the chatbot GYANT is a symptom-checking and medical screening chatbot that can be consulted multiple times; however, each interaction with the chatbot starts as an independent conversation. Contrarily, medium- and long-term chatbots are exclusively (100%) characterized by multiple interactions, $\chi^2(120, 2) = 63.21$, $p < .001$, $V = 0.726$.

Correspondingly, Fisher's Exact tests show that the chatbots' temporal profile is indicative of the D_3 duration of interactions, $FET = 50.02$, $p < .001$, $V = 0.459$, and the D_4 consecutiveness of interactions, $\chi^2(120, 2) = 94.16$, $p < .001$, $V = 0.886$: As short-term chatbots only provide one-time only interactions, interactions are never related (0%). Interestingly, only 60.0% of all interactions with medium-term chatbots are $C_{4,2}$ related, but 100% of all long-term chatbots. Furthermore, 66.7% of all short-term chatbots are characterized by $C_{3,1}$ short interaction, but only 20.0% of all medium-term and 6.9% of all long-term chatbots. Contrarily, 80% (93.1%) of all medium-term (long-term) chatbots engage in $C_{3,2}$ medium to $C_{3,3}$ long conversations.

4.2.1.2. Appearance. Fisher's Exact tests show that the design dimensions D_5 role and D_6 primary communication style depend on a chatbot's temporal profile: While the majority of chatbots are equipped with a $C_{5,2}$ facilitator role (61.7% of all short-term, 64.0% of all medium-term and 44.8% of all long-term chatbots), 44.8% of all long-term are personified as $C_{5,3}$ peer characters, compared to 20.0% of all medium-term and only 3.0% of all short-term chatbots. Eventually, 21.7% of all short-term chatbots exhibit an $C_{5,1}$ expert role compared to 16.0% of all medium-term and 10.3% of all long-term chatbots, $FET = 25.36$, $p < .001$, $V = 0.334$.

With regards to their communication style, 70.8% of all chatbots (i. e., 92.4% of all short-term, 52.0% of all medium-term but only 37.9% of all long-term chatbots) employ a $C_{6,1}$ task-oriented communication style. Short-term chatbots are least likely to engage in $C_{6,2}$ socially-oriented chat (7.6%) compared to medium-term (48.0%) or long-term (62.1%) chatbots, $\chi^2(120, 2) = 34.38$, $p < .001$, $V = 0.535$.

Long-term chatbots are slightly more likely to be personified by an $C_{7,2}$ embodied avatar than medium- or short-term chatbots (62.1% of all long-, 44.0% of all medium-, and 34.8% of all short-term chatbots, $\chi^2(120, 2) = 6.09$, $p = .046$, $V = 0.046$).

4.2.1.3. Intelligence. Concerning the intelligence layer, D_8 intelligence framework, D_9 intelligence quotient, D_{10} personality adaptability, and D_{11} socio-emotional behavior are significantly associated with the different temporal profiles of chatbots. While the majority of chatbots either relied on purely $C_{8,1}$ rule-based text generation (49.2%) or on a $C_{8,2}$ hybrid solution generally following a rule-based conversational path but integrating some natural language processing capabilities to learn from the conversation (48.3%), only three (10.3% of all) long-term chatbots (i.e., REPLIKA, BRAINBOT, VIRTUAL IMAGINARY INTERLOCUTOR) could be classified as purely $C_{8,3}$ artificially intelligent chatbots, $FET = 8.37$, $p = .049$, $V = 0.219$. The same distribution is reflected in chatbots' D_9 intelligence quotient that differed significantly across temporal profiles, $FET = 8.94$, $p = .040$, $V = 0.199$: Overall, 36.7% of all chatbots could be classified as possessing $C_{9,1}$ rule-based knowledge only, 60.0% possessed some more or less basic $C_{9,2}$ text understanding capabilities and one medium- and three long-term chatbots (3.3% of all chatbots) could process information from other sources than text, for example, from images of an injury (i.e., MBOT).

While only seven (5.8%) out of all 120 chatbots are capable of

adapting their personality to the user, all of them were either medium-term (1) or long-term (6) chatbots, $FET = 13.18$, $p = .001$, $V = 0.364$. Accordingly, D_{11} socio-emotional behaviors were only present in 45.5% of all short-term chatbots and in 56.0% of all medium-term, but in 79.3% of all long-term chatbots, $FET = 8.96$, $p = .010$, $V = 0.274$.

4.2.1.4. Interaction. In the interaction layer, Fisher's Exact tests show significantly different distributions of design characteristics with regards to the temporal profile of the chatbots for six of seven design dimensions: D_{13} front-end user interface, D_{14} communication modality, D_{16} user assistance design, D_{17} personalization, D_{18} additional human support, and D_{19} gamification.

A majority of all chatbots in the total sample are deployed either as pop-up windows on $C_{13,4}$ websites (37.5%) or as artificial contacts in $C_{13,2}$ social media messengers (28.3%). There were significant differences between chatbots' temporal profiles, $FET = 39.86$, $p < .001$, $V = 0.410$. While 59.1% of all short-term chatbots were accessible via $C_{13,4}$ websites and only 9.1% via standalone $C_{13,1}$ applications, 44.8% of all long-term chatbots were only accessible via a standalone $C_{13,1}$ application, and two (6.9%) only on websites (i.e., MONDLY and KIM).

While we only included text-based chatbots in our sample, some of them (15.0%) also allowed for voice input or offered voice output ($C_{14,2}$ text + voice). Differences between chatbots with different temporal profiles were significant, $FET = 7.01$, $p = .023$, $V = 0.241$: 7.6% of all short-term, 20.0% of all medium-term, and 27.6% of all long-term chatbots could not only communicate via text messages but also via voice.

With regards to the user assistance design, 39.2% of all chatbots could be described as $C_{16,1}$ reactive, 30.0% as $C_{16,2}$ proactive, and 30.8% as $C_{16,3}$ reciprocal (i.e., they react to user input but also steer the conversation themselves). A χ^2-test shows that differences between the temporal profiles are significant, $\chi^2(120, 2) = 36.63$, $p < .001$, $V = 0.391$: While the majority (59.1%) of short-term chatbots are reactive and only 13.6% reciprocal, the majority (65.5%) of long-term chatbots are flexible and ten (34.5%) chatbots are proactive. Contrarily, medium-term chatbots are almost evenly distributed: 32.0% are reactive, 32.0% proactive and 36.0% reciprocal.

Similarly, chatbots are different with regards to their capability to adapt their content based on what they (already) know about the user and thus to personalize the conversation, $\chi^2(120, 2) = 34.35$, $p < .001$, $V = 0.535$: While 86.2% of all long-term chatbots could be characterized as $C_{17,2}$ adaptive, only 22.47% of all short-term chatbots personalize the conversation. Again, medium-term chatbots were more evenly distributed: 44.0% can be characterized as $C_{17,1}$ static, 56.0% as $C_{17,2}$ adaptive.

Furthermore, 25.0% of all chatbots offered the possibility to connect the user with a human (live) agent with significant differences between the temporal profiles, $\chi^2(120, 2) = 7.82$, $p = .023$, $V = 0.255$: While 23 (34.8%) of all short-term chatbots were wired to a human agent, this was only true for four (16.0% of all) medium-term (i.e., @DAWEBOT, DROPSHIPPING ASSISTANT, STINA, and ANA COPA AIRLINES) and three (10.3% of all) long-term chatbots (i.e., BROOK, WYSA, and ANNA/LUKAS).

Chatbots were significantly different with regards to the integration of D_{19} gamification elements, $\chi^2(120, 2) = 31.35$, $p < .001$, $V = 0.511$: While 51.7% of all long-term chatbots were $C_{19,2}$ gamified, only 3.0% of all short-term and 32.0% of all medium-term chatbots were $C_{19,1}$ gamified.

4.2.1.5. Context. Regarding their D_{20} application domain, 45 (37.5%) of all 120 chatbots were classified as $C_{20,1}$ business chatbots, 36 (30.0%) as $C_{20,2}$ healthcare chatbots, 24 (20%) as $C_{20,3}$ education chatbots, and 15 (12.5%) as $C_{20,4}$ daily life chatbots. Fisher's Exact tests revealed significant differences between the distribution of temporal profiles across application domains, $FET = 50.21$, $p < .001$, $V = 0.441$: While none of the $C_{20,1}$ business chatbots was classified as long-term, 50.0% of all $C_{20,2}$ healthcare chatbots, 33.3% of all $C_{20,3}$ education chatbots, and

20.0% of all $C_{20,4}$ daily life chatbots covered long-term purpose; on the other hand, 88.9% of all $C_{20,1}$ business chatbots were classified as short-term chatbots, but only 46.7% of all $C_{20,4}$ daily life chatbots, 36.1% of all $C_{20,2}$ healthcare chatbots, and 25.0% of all $C_{20,3}$ education chatbots. Overall, the distributions of short- and long-term chatbots in business and healthcare in our total sample are largely in line with previous research (Tudor Car et al., 2020); for daily life and educational chatbots, we could not find similar analyses.

While the presence of a D_{22} collaboration goal was not associated with the different temporal profiles, chatbots' D_{21} usage motivation/purpose differed significantly across the temporal profiles, $FET = 72.51$, $p < .001$, $V = 0.531$: $C_{21,5}$ coaching and supervision was the dominant motivation for using medium- or long-term chatbots (75.9% of all long-term, 40.0% of all medium-term and none of the short-term chatbots were classified as such); $C_{21,3}$ utility is the dominant motivation for using short-term chatbots (53.0%), followed by $C_{21,4}$ informational (28.8%).

4.2.2. Time-dependent chatbot archetypes

Comparing the calculated indices per design dimension and temporal profile revealed a consistent pattern of differences between short-, medium-, and long-term chatbots (RQ2). Since the design characteristics were always ordered hierarchically from none/low/less to high/more the design of long-term chatbots can be characterized as more advanced, complex, and sophisticated than medium- or short-term chatbots in all design dimensions except for the availability of D_{18} additional human support which was more frequently present in short- and medium-term chatbots.

The results are depicted in Fig. 3. Since the design characteristics are mutually exclusive, the visualization in Fig. 3 allows emphasizing fundamental design differences between the different temporal profiles, while accounting for the fact that the archetypes ought not to be understood as separate, dichotomous entities that cannot share common characteristics or tendencies. The differences are subsequently conceptualized into three time-dependent archetypes.

4.2.2.1. Ad-hoc Supporters.
Short-term chatbots designed for short, isolated, one-time-only interactions are denominated as *Ad-hoc Supporters*. With regards to their level of development, they are often based on less advanced and less complex technologies. Furthermore, they are usually not designed to offer services beyond the services that the website on which they are typically implemented already provides. Thus, Ad-hoc Supporters are generally not intended to replace but to complement a company's existing communication channels. Ad-hoc Supporters are furthermore strongly task-oriented, which is not only reflected in their primary communication style but also in the fact that they neither (need to) possess the ability to adapt their personality nor other aspects of the conversation to prior interactions with the user. Furthermore, to ensure that the underlying problems for which users primarily approach them (mostly informational and utilitarian reasons) are solved efficiently, they tend to connect users more quickly with a human agent than chatbots with other temporal profiles.

4.2.2.2. Persistent Companions.
Long-term chatbots, which are designed for longer, interdependent, and perpetual interactions, are denominated as *Persistent Companions*. To meet expectations that come along with long-term oriented relationships, they appear more advanced and flexible along multiple dimensions: They are designed in a way that allows users to steer a conversation in a certain direction, but they are also capable of proposing new directions proactively, for example, when a conversation is stuck. Overall, Persistent Companions appear to possess characteristics that support relationship-building processes with users: Their socially-oriented communication style often allows for social talk and chitchat besides the primary objectives of the conversation. Furthermore, they are intended to adapt their personality in the course

of the relationship-building process and to personalize conversations based on what they have learned about a user's profile in prior interactions. To further increase the variability of the interaction and to account for people's primary motivation to engage with Persistent Companions (i.e., entertainment and coaching), Persistent Companions are also more likely to integrate gamification elements.

To endow Persistent Companions with the necessary technical capabilities, they tend to be developed as stand-alone applications, which allows implementing a variety of functionalities beyond the features that platform-dependent specifications dictate. Thereby, they can integrate additional services and embed information from external databases or process information retrieved from media elements. Overall, Persistent Companions are likely aimed at (partly) replacing or superseding existing offerings.

4.2.2.3. Temporary Advisors.
Chatbots for medium-term relationships are here denominated as *Temporary Advisors*. While they meander between short- and long-term chatbots with regards to their design characteristics, their temporal profile is more comparable to long-term than short-term chatbots as they are rather designed for multiple, (partly) associated medium-length interactions. However, similar to short-term chatbots, they are less likely to adapt their personality and rely on less advanced technologies. More similar to long-term chatbots again, they refrain from providing instant contact to a human agent and are instead more likely to integrate additional services, and features (such as processing of external data and gamification elements) to help solve a user's cause to approach the chatbot in the first place.

5. General discussion

Since chatbots are becoming increasingly prevalent across all industries, managers' success increasingly depends on their ability to adapt a chatbot's design to the conditions it is developed for, which includes, in particular, for how long users intend to interact with the chatbot.

In order to answer our first research question (**RQ1:** "Which design elements allow us to distinguish chatbots depending on whether they are aimed to help individuals to achieve short-, medium- or long-term goals?"), following an established taxonomy development method, we developed a design taxonomy to characterize user-chatbot relationships with different time horizons comprised of 22 design dimensions and 61 design characteristics organized into five overarching perspectives that are visible or experiential within the user-chatbot relationship. In order to answer our second research question (**RQ2:** "How does a chatbot's temporal profile affect its design?"), we classified and analyzed 120 chatbots with regards to their distribution of design characteristics per design dimension. Frequency analysis revealed significant differences between the chatbots depending on whether they are aimed to help individuals to achieve short-, medium- or long-term goals for 19 out of 22 design dimensions. Further analyses allowed us to derive, distinguish, and conceptualize three distinct time-dependent chatbot archetypes (i.e., Ad-hoc Supporters, Temporary Advisors, and Persistent Companions) that allow scientists and practitioners alike to understand, study, and take into account design particularities inherent in the time horizon of the user-chatbot relationship.

Overall, our findings offer important theoretical contributions beyond existing classifications and typologies of chatbots to human-computer interaction (e.g., Baraka et al., 2020) and conversational agent design research (e.g., Araujo, 2018; Diederich, Brendel, & Kolbe, 2019; Scarpellini & Lim, 2020), and actionable managerial implications, which we outline in the remaining sections.

5.1. Theoretical contribution

Even though researchers from different fields acknowledged the

importance of temporal design aspects for users' engagement (Baraka et al., 2020; Hildebrandt et al., 2004, pp. 1737–1738; Karahasanović et al., 2019, pp. 376–390) and their individual "trajectories of interaction" (Benford et al., 2009, p. 109), time-dependent design aspects of chatbots have not been investigated systematically so far. Thereto, our research offers four main theoretical contributions:

First, so far, research that took into account a chatbot's temporal profile as a determining design factor had been "one-sided" and predominantly focused on investigating design factors that drive users' engagement with long-term chatbots (Bickmore & Picard, 2005; Hobert & Berens, 2020) – without ever questioning the transferability of possibly successful design factors to chatbots that were developed for short-term relationships. Thus, similar to prior research that has provided classifications and frameworks for various specific foci such as healthcare chatbots (Laranjo et al., 2018), business-to-business chatbots (Janssen, Rodríguez Cardona, & Breitner, 2020), or collaborative team chatbots (Bittner et al., 2019; Seering et al., 2019), this work takes up on prior calls for research to consider time as an important factor in user-chatbot relationships and provides a holistic perspective on chatbots' different temporal profiles. The development and provision of a comprehensive taxonomy of time-dependent design elements for chatbots with different temporal profiles (RQ1) enables researchers and practitioners to compare fundamental design differences between chatbots for short-, medium-, or long-term purposes.

Second, by providing insights about the impact of a chatbot's temporal profile on its design (cf. RQ2), the derivation and differentiation of three time-dependent chatbot archetypes allows researchers to extend the conceptual understanding of chatbots' social roles in user-chatbot encounters (Scarpellini & Lim, 2020). Apparently, chatbots supporting short-term goals rather aim at assisting and complementing human actors and work as receptionists that connect users with a human agent to make sure customer problems are solved quickly. Their task-oriented communication style reflects their aspiration for efficiency and the minimization of "cost, effort, and time allocated to the interaction" (Verhagen et al., 2014, p. 534), fulfilling a social role similar to a supportive "assistant". Quiet contrarily, chatbots that are developed to accompany users over longer periods are often developed in greater depth and with more complexity to be ultimately capable of working independently from any human agent (De Keyser et al., 2019). The comparatively more pronounced manifestations of characteristics that anthropomorphize and personalize the interaction, such as the integration of socio-emotional behaviors, the socially-oriented communication style, the inclusion of gamification elements, and the adaptation of the chatbot's character in the course of the user-chatbot relationship, are all indicators that – for chatbots that want to help individuals achieve long-term goals – "greater emphasis is put on the feeling of solving a problem together, being more [responsive] to personal needs and enhancing social contagion" (Verhagen et al., 2014, p. 535). While some of the aforementioned design manifestations likely also depend on other factors, for instance, the functional purpose for which a chatbot has been developed (Scarpellini & Lim, 2020), our analysis offers evidence that the temporal profile is a decisive factor as well and strengthens the notion of chatbots' dual role as a communication medium and, in their role as social actors, as communication partners.

Third, with the development of a chatbot taxonomy, we not only contribute to the chatbot research community with new knowledge but also methodologically to the taxonomy development field. We not only used the ending conditions as suggested by Nickerson et al. (2013) to evaluate the taxonomy's comprehensiveness (Iteration 6) but also asked two previously uninvolved researchers in iteration 7 to classify a new set of chatbots using the developed taxonomy. Furthermore, one of the researchers involved in the development process hitherto classified this new set as well which allowed us to show that external researchers can apply the taxonomy correctly and that the taxonomy is applicable to a new dataset (cf. Table A5, Web Appendix). Moreover, analogous to the use of interview guides in other qualitative research, our research is the

first to classify real-world chatbots from a temporal-based perspective through chat logs obtained from standardized, semi-structured dialogue guides (cf. Table A6, Web Appendix), which should become a standard for the analysis of real-world chatbots to ensure comparability.

Lastly, we demonstrate that a taxonomy can be used to test differences with regards to a specific, predefined design characteristic (i.e., temporal profile) instead of using cluster analyses to discover latent archetypes within a dataset as in other taxonomy-based research (e.g., Diederich et al., 2019) and suggest a novel visual representation of the design configurations for the differentiated archetypes (Fig. 3). In this vein, our approach serves as a guideline for future research that strives to recognize structures with a focus on a certain superordinate feature.

5.2. Managerial implications

Even though our analysis is based on 120 real-life chatbots and, thus, based on past design choices of practitioners and chatbot developers, a systematic understanding of time-dependent aspects in the design of chatbots was missing. Therefore, the present research offers actionable guidelines and a salient framework that can guide practitioners from the first day in designing, developing, and implementing a chatbot with a specific time horizon.

First, our definitions of three time-dependent chatbot archetypes offer practitioners an explicit representation of the time horizon as a determining factor for a chatbot's design. Having a common understanding and definition of a chatbot's temporal profile and being aware of this factor helps to prevent communication problems within companies between product managers and developers.

Second, the taxonomy of time-dependent design elements together with the conceptualization of three chatbot archetypes lay a solid foundation for streamlining the design process of structures and architectures of domain-specific chatbots when the intended temporal profile of the user-chatbot relationship is clear, which in turn reduces designers' efforts, cost, and time to develop and implement new chatbot-based services.

Third, while each of the design dimensions entails challenges and opportunities, the taxonomy gives designers the flexibility to add and combine dimensions to prototype and tailor the chatbot development quickly to any desired target group or use case while taking into account boundary conditions and restrictions (e.g., available budget or development expertise).

5.3. Limitations, further research, and concluding remarks

As with any research, this work has some limitations, which offer opportunities for future research directions (RDs).

While having thoroughly followed an established taxonomy development procedure protocol (Nickerson et al., 2013), the limitations of this study mainly stem from the subjective choices inherent in any qualitative research approach. This subjectivity, for example, may underlie to a certain extent in the construction of our sample(s). While the final taxonomy relies on the classification of 120 chatbots from three different samples and sampling strategies which ensures wide coverage of available chatbots, further analyzes of archetypes could investigate boundary conditions that relate to interactions with situational design factors (RD1) other than the temporal profile, for example, across different domains (Feine et al., 2020) or purposes (Scarpellini & Lim, 2020). Notwithstanding, we applied a systematic empirical evaluation process to analyze the final sample under the same structure and attributes, assuring homogeneity in quality and data format. Likewise, we maintained a consistent unit of analysis throughout our research relying on the same aforementioned sample to develop the taxonomy and to analyze differences between the three identified temporal profiles.

Furthermore, while our research offers insights into a time-aware design of chatbot-based services, we have purposefully limited the scope of our study to domain-specific, text-based chatbots. Yet, it needs

to be investigated, to which extent our taxonomy can be used to inform the design of other types of conversational agents, such as general-knowledge and/or voice-based conversational agents (RD2). Voice assistants, such as ALEXA, CORTANA, or SIRI often accompany users over longer periods (Knote, Janson, Söllner, & Leimeister, 2019), yet, at the same time, they can usually be equipped with so-called "skills" that support individuals' short-term goals. Such skills are often provided by third-party developers, and it needs to be well understood how they can be best integrated into an existing relationship with the voice assistant (RD3). Similarly, many chatbots that are overall aimed to help individuals achieve a specific long-term goal (e.g., losing weight), are simultaneously also designed to help individuals achieve short-term goals (e.g., to reach a certain number of daily steps). Therefore, differentiating subtypes of long-term chatbots likely renders different design choices necessary (RD4).

Although our research does not attempt to assess the success (e.g., user engagement, satisfaction, retention) of the current state-of-the-art configurations of the different identified temporal chatbot archetypes, the technologies that are enabling and driving chatbots' capabilities are advancing quickly. Since "taxonomies are not static but change over time as new objects that may or may not fit into an existing taxonomy are developed or identified" (Nickerson et al., 2013, p. 355), these issues can be addressed in further research projects by re-iterating the taxonomy development procedure from time-to-time (RD5) and by connecting the identified design elements to specific success factors (RD6) to avoid upcoming gaps between theory and practice.

In conclusion, the present paper demonstrates that the relationship duration is a central factor in the design of chatbots and offers new directions for investigating nuances of engaging time-dependent chatbot design. In this spirit, this work strives to serve as a foundation for further researchers undertaking design-related research projects that ultimately enable the optimization of the development of chatbots.

Credit author statement

Marcia Nißen: Conceptualization, Data Curation, Methodology, Formal Analysis, Project Administration, Writing – Original Draft, Writing - Review & Editing. **Driton Selimi**: Data Curation, Conceptualization, Formal Analysis, Visualizations, Writing – Review & Editing. **Antje Janssen & Davinia Rodríguez Cardona:** Data Curation, Validation, Writing – Review & Editing. **Michael H. Breitner, Tobias Kowatsch, Florian von Wangenheim**: Supervision, Writing – Review & Editing.

Financial disclosure

Declaration of competing interest

Appendix A. Supplementary data

Supplementary data to this article can be found online in a Web Appendix at https://doi.org/10.1016/j.chb.2021.107043.

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*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Baraka, K., Alves-Oliveira, P., & Ribeiro, T. (2020). An extended framework for characterizing social robots. *Human-Robot Interaction, 12*, 21–64. https://doi.org/10.1007/978-3-030-42307-0_2

Benford, S., Giannachi, G., Koleva, B., & Rodden, T. (2009). From interaction to trajectories: Designing coherent journeys through user experiences. *Proceedings of the conference on human factors in computing systems.* CHI '09. https://doi.org/10.1145/1518701.1518812

Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction, 12*(2), 293–327. https://doi.org/10.1145/1067860.1067867

Bittner, E., Oeste-Reiß, S., & 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 52th Hawaii international conference on system sciences (HICSS).* https://doi.org/10.24251/HICSS.2019.035

Chung, M., Ko, E., Joung, H., & Kim, S. J. (2020). Chatbot e-service and customer satisfaction regarding luxury brands. *Journal of Business Research, 117*, 587–595. https://doi.org/10.1016/j.jbusres.2018.10.004

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society, 1*(1), 104–126. https://doi.org/10.1007/BF03177550

Creswell, J. W., & Clark, V. L. P. (2011). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

De Keyser, A., Köcher, S., Alkire, L., Verbeeck, C., & Kandampully, J. (2019). Frontline service technology infusion: Conceptual archetypes and future research directions. *Journal of Service Management, 30*(1), 156–183. https://doi.org/10.1108/JOSM-03-2018-0082

Diederich, S., Brendel, A. B., & Kolbe, L. M. (2019). Towards a taxonomy of platforms for conversational agent design. *Proceedings of the International Conference on Wirtschaftsinformatik (WI '19)*.

Feine, J., Adam, M., Benke, I., Maedche, A., & Benlian, A. (2020). Exploring design principles for enterprise chatbots: An analytic hierarchy process study. Proceedings of the 15th international conference on design science research in information systems and technology (DESRIST 2020). In S. Hofmann, et al. (Eds.), *Lecture notes in computer science: Vol. 12388. Designing for digital transformation. Co-creating services with citizens and industry* (pp. 126–141). https://doi.org/10.1007/978-3-030-64823-7_13

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage Publications.

Fleisch, E., Franz, C., & Herrmann, A. (2021). *The digital pill: What everyone should know about the future of our healthcare system.* Bingley, UK: Emerald Publishing.

Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. *Proceedings of the 38th international conference on information systems (ICIS 2017)*.

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior, 97*, 304–316. https://doi.org/10.1016/j.chb.2019.01.020

Hesse, B. W., Werner, C. M., & Altman, I. (1988). Temporal aspects of computer-mediated communication. *Computers in Human Behavior, 4*, 147–165. https://doi.org/10.1016/0747-5632(88)90023-4

Hildebrandt, M., Dix, A., & Meyer, H. A. (2004). Time design. *Proceedings of the conference on human factors in computing systems (CHI '04 - extended abstracts).* https://doi.org/10.1145/985921.986208

Hobert, S., & Berens, F. (2020). Small talk conversations and the long-term use of chatbots in educational settings – experiences from a field study. In A. Følstad, et al. (Eds.), *Vol. 11970. Chatbot research and design (CONVERSATIONS 2019), lecture notes in computer science* (pp. 260–272). https://doi.org/10.1007/978-3-030-39540-7_18

Ho, A., Hancock, J. T., & Miner, A. S. (2018). Psychological, relational, and emotional effects of self-disclosure after conversations with a chatbot. *Journal of Communication, 68*(4), 712–733. https://doi.org/10.1093/joc/jqy026

Hoorn, J. F. (2018). *Theory of robot communication: I. The medium is the communication partner. arXiv preprint. arXiv:1812.04408*.

Janssen, A., Passlick, J., Rodríguez Cardona, D., & Breitner, M. H. (2020). Virtual assistance in any context. *Business & Information Systems Engineering, 62*(3), 211–225. https://doi.org/10.1007/s12599-020-00644-1

Janssen, A., Rodríguez Cardona, D., & Breitner, M. H. (2020). More than FAQ! Chatbot taxonomy for business-to-business customer services. In A. Følstad, et al. (Eds.), *Lecture notes in computer science: Vol. 12604. Chatbot research and design (CONVERSATIONS 2020)* (pp. 175–189). https://doi.org/10.1007/978-3-030-68288-0_12

Karahasanović, A., Hollebeek, L. D., Chasanidou, D., & Gurau, C. (2019). Temporality of customer engagement in service innovation: A theoretical model. *Handbook of research on customer engagement.* Edward Elgar Publishing. https://doi.org/10.4337/9781788114899.00025

Kassarjian, H. D. (1977). Content analysis in consumer research. *Journal of Consumer Research, 4*(1), 8–18.

Knote, R., Janson, A., Eigenbord, L., & Söllner, M. (2018). The what and how of smart personal assistants: Principles and application domains for is research. *Proceedings of the International Conference on Wirtschaftsinformatik (WI '18)*.

Knote, R., Janson, A., Söllner, M., & Leimeister, J. M. (2019). Classifying smart personal assistants: An empirical cluster analysis. *Proceedings of the 52th Hawaii international conference on system sciences (HICSS)*. https://doi.org/10.24251/HICSS.2019.245

Kowatsch, T., Nißen, M., Rüegger, D., Shih, C.-H. I., Volland, D., Filler, A., et al. (2017). Text-based healthcare chatbots supporting patient and health professional teams: Preliminary results of a randomized controlled trial on childhood obesity. In *Proceedings of the persuasive embodied agents for behavior change (PEACH 2017) workshop, co-located with the 17th international conference on intelligent virtual agents (IVA '17)*.

Kowatsch, T., Nißen, M., Rüegger, D., Stieger, M., Flückiger, C., Allemand, M., et al. (2018). The impact of interpersonal closeness cues in text-based healthcare chatbots on attachment bond and the desire to continue interacting: An experimental design. In *Proceedings of the 26th European conference on information systems (ECIS '18)*.

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., et al. (2018). Conversational agents in healthcare: A systematic review. *Journal of the American Medical Informatics Association, 1*(25), 1248–1258. https://doi.org/10.1093/jamia/ocy072

Nickerson, R. C., Varshney, U., & Muntermann, J. (2013). A method for taxonomy development and its application in information systems. *European Journal of Information Systems, 22*(3), 336–359. https://doi.org/10.1057/ejis.2012.26

Rack, O., Zahn, C., & Mateescu, M. (2018). Coding and counting: Frequency analysis for group interaction research. *The cambridge handbook of group interaction analysis*. Cambridge University Press. https://doi.org/10.1017/9781316286302.015

Reeves, B., & Nass, C. (1996). *The media equation: How pepole treat computers, television, and new media like real people and places*. Cambridge, UK: Cambridge University Press.

Sauro, J., & Lewis, J. R. (2016). Chapter 5 - is there a statistical difference between designs? In J. Sauro, & J. R. Lewis (Eds.), *Quantifying the user experience: Practical statistics for user research* (2nd ed., pp. 61–102). Cambridge: Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-802308-2.00005-9.

Scarpellini, I., & Lim, Y. (2020). Role-based design of conversational agents: Approach and tools. In Stephanidis, et al. (Eds.), *HCI international 2020 - late breaking posters. Communications in computer and information science* (Vol. 1293, pp. 366–375). Cham: Springer. https://doi.org/10.1007/978-3-030-60700-5_47.

Seering, J., Luria, M., Kaufman, G., & Hammer, J. (2019). Beyond dyadic interactions: Considering chatbots as community members. In S. Brewster, et al. (Eds.), *Proceedings of the conference on human factors in computing systems (CHI '19)*. https://doi.org/10.1145/3290605.3300680

Shibata, T. (2004). An overview of human interactive robots for psychological enrichment. *Proceedings of the IEEE, 92*(11), 1749–1758. https://doi.org/10.1109/JPROC.2004.835383

Skjuve, M. B., & Brandtzæg, P. B. (2018). Chatbots as a new user interface for providing health information to young people. In Y. Andersson, et al. (Eds.), *Youth and news in a digital media environment: Nordic-Baltic perspectives* (pp. 59–66).

Szopinski, D., Schoormann, T., & 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 '19)*.

Tudor Car, L., Dhinagaran, D. A., Kyaw, B. M., Kowatsch, T., Joty, S., Theng, Y.-L., et al. (2020). Conversational agents in health care: Scoping review and conceptual analysis. *Journal of Medical Internet Research, 22*(8), Article e17158. https://doi.org/10.2196/17158

Verhagen, T., van Nes, J., Feldberg, F., & van Dolen, W. (2014). Virtual customer service agents: Using social presence and personalization to shape online service encounters. *Journal of Computer-Mediated Communication, 19*(3), 529–545. https://doi.org/10.1111/jcc4.12066

Vom Brocke, J., Simons, A., Riemer, K., Niehaves, B., Plattfaut, R., & Cleven, A. (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*(9), 205–224. https://doi.org/10.17705/1CAIS.03709

Walther, J. B., & Tidwell, L. C. (1995). Nonverbal cues in computer-mediated communication, and the effect of chronemics on relational communication. *Journal of Organizational Computing, 5*(4), 355–378. https://doi.org/10.1080/10919399509540258

Werner, C. M., Altman, I., & Brown, B. B. (1992). A transactional approach to interpersonal relations: Physical environment, social context and temporal qualities. *Journal of Social and Personal Relationships, 9*(2), 297–323. https://doi.org/10.1177/0265407592092008

Williams, K., Chatterjee, S., & Rossi, M. (2008). Design of emerging digital services: A taxonomy. *European Journal of Information Systems, 17*(5), 505–517. https://doi.org/10.1057/ejis.2008.38

Wünderlich, N. V., von Wangenheim, F., & Bitner, M. J. (2012). High tech and high touch: A framework for understanding user attitudes and behaviors related to smart interactive services. *Journal of Service Research, 16*(1), 3–20. https://doi.org/10.1177/1094670512448413

Yanco, H. A., & Drury, J. (2004). Classifying human-robot interaction: An updated taxonomy. *Proceedings of the IEEE international conference on systems, man & cybernetics (SMC '04)*. https://doi.org/10.1109/ICSMC.2004.1400763

Youn, S., & Jin, S. V. (2021). "In a.I. We trust?" the effects of parasocial interaction and technopian versus luddite ideological views on chatbot-based customer relationship management in the emerging "feeling economy". *Computers in Human Behavior, 119*, 106721. https://doi.org/10.1016/j.chb.2021.106721

Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media & Society, 8*(3), 401–419. https://doi.org/10.1177/1461444806061951