ICIS 2025 Proceedings

International Conference on Information Systems (ICIS)

December 2025

# Real-Time Guidance Design: Principles for AI-Augmented Proactive Guidance Systems Supporting Human Dialogue in Customer Service

Marc Christopher Grau
*University of St. Gallen*, marcchristopher.grau@unisg.ch

Ivo Blohm
*University of St. Gallen*, ivo.blohm@unisg.ch

# Real-Time Guidance Design: Principles for AI-Augmented Proactive Guidance Systems Supporting Human Dialogue in Customer Service

*Short Paper*

**Marc Christopher Grau**
University of St. Gallen
Müller-Friedbergstrasse 8, 9000 St. Gallen, Switzerland
marcchristopher.grau@unisg.ch

**Ivo Blohm**
University of St. Gallen
Müller-Friedbergstrasse 8, 9000 St. Gallen, Switzerland
ivo.blohm@unisg.ch

## Abstract

*Live customer-service calls place heavy demands on agents' limited attentional resources: they must listen, speak, retrieve information, and follow prescribed procedures simultaneously. Yet most AI support tools remain reactive, relying on user-initiated queries and thus amplifying the very cognitive load they are meant to relieve. To address this issue, we examined AI-augmented proactive guidance embedded directly in service agents' interfaces. Using a design-science research methodology, we carried out three iterative development and evaluation cycles. These cycles yielded three design principles, rooted in theory and practice, and guided our transition from a low-fidelity prototype to a fully instantiated web-application artefact. Findings from controlled evaluations in the form of semi-structured interviews with customer-service practitioners validate our nascent design theory. Alongside the design principles, we provide a transferable system architecture for practical deployment and outline avenues for further evaluation.*

**Keywords:** Artifical Intelligence, AI, AI Guidance, Proactive, UI

## Introduction

Real-time customer service interactions place immense cognitive demands on service employees (SEs). Agents must simultaneously monitor customer dialogue, assess context, follow service processes, and plan solutions, all in a highly synchronous environment. This multitasking burdens strains working memory and attentional resources (Poser, Wiethof, Banerjee, et al., 2022). Conversational Artificial Intelligence (AI) solutions aimed at automating customer interactions have not eliminated these challenges. In fact, while chatbots can handle routine queries, they often increase SEs' cognitive load by introducing new tasks like AI-to-human handoffs (Benner et al., 2021). For example, when a virtual agent transfers a customer to a human because it is unable to resolve the request, the SE must rapidly grasp the situation and the context that the virtual agent handled. Hence, AI assistance should not end at the transfer to the SE but can also assist during the hand-off procedure and augment the SE during the conversation with customer (Brynjolfsson et al., 2025). However, when AI assist SEs (e.g. providing on-demand information) during the human-human conversation, the support is typically reactive – the SE must query the AI or search a knowledge base for help while handling the conversation with the customer in parallel (Poser, Wiethof, & Bittner, 2022). Such reactive interaction paradigms force SEs to interrupt the conversation flow and switch systems, taxing their limited attention resources. Thus, current support tools often require the SE's invocation and handling, adding to the very workload they intend to reduce. Therefore, we posit that augmenting the SE during the human-human conversation requires a rethinking.

Emerging agentic information systems can anticipate user needs without explicit user requests (Baird & Maruping, 2021). This opens an opportunity to reimagine AI support for human-human service interactions through proactive guidance. Rather than waiting for the employee to ask for help, a proactive AI-augmented guidance system should continuously monitor the conversation and infer how it can help by offering timely suggestions or context-aware information. Nonetheless, how human augmentation with Generative AI should be designed requires careful considerations as it can either reduce or intensify cognitive load (Shao et al., 2024). Prior research suggests need anticipation can enhance performance (Bar-Or & Meyer, 2019) and efficiency. For instance, proactive support has been shown to improve decision accuracy and speed in various contexts (Maniktala et al., 2022; Reinhard et al., 2024). However, designing for proactivity is delicate. If done poorly, it may overwhelm users or reduce their satisfaction (Diebel et al., 2025). An already old but prominent failed example would be Microsoft's well-known Clippy, that fell short in actually assisting the user (Baym et al., 2019). While the content of assistance is obviously crucial, the design of the user interface and timing of invocation plays a central role in adoption and efficient usage. For instance, research on proactive guidance from GitHub Copilot indicate that slight changes to the interface, such as showing a whole block of code instead of the next piece, positively affects user productivity but also imposes higher cognitive load. Further, with the suggestion providing a starting point, guidance eased solution development for stuck users (Vaithilingam et al., 2022). Thus, an interface intended for proactive guidance demands for careful design considerations.

Early approaches in customer service have explored AI in a supporting role. For example, Reinhard et al. (2024) introduce a hybrid intelligence agent that aids new employees during onboarding, and proactivity has been discussed as a design dimension in human-AI service systems (Wiethof et al., 2022). Yet, how to systematically design for proactive AI guidance in live customer-service human-human dialogues remains largely unexplored (Wenninger et al., 2022). Approaches have been successfully applied in other domains (Chen et al., 2025), but research on interface design for proactive guidance in the customer service domain remains scarce. To contribute design research in this area, we draw on selective attention theory (Broadbent, 1958) with a focus on Wickens' model of multiple attentional resources (Wickens, 2021; Wickens & Carswell, 2021) – to ensure interface design accounts for the SEs' limited attentional resources. Designs must account for different channels (visual, auditory, etc.) and avoid overloading any single channel with excessive information. Prior work on adaptive user interfaces demonstrates the value of interfaces that adjust to users' attentional load in real-time (Toreini et al., 2022). In line with these insights, we envision a system that provides proactive guidance dynamically during the conversation, integrated directly into the SE's workflow.

*RQ: How can proactive AI-augmented guidance systems be designed to assist customer service employees during human-human dialogue while attending to the SE's limited attentional resources?*

To address this question, we used a Design Science Research (DSR) approach to iteratively develop and evaluate an interface that proactively guides SEs based on conversation context. After three design cycles, we instantiated our nascent design theory in a working web application. We divided the complete evaluation into formative and summative evaluation episodes as proposed in the FEDS process by Venable et al. (2016) following the Technical Risk & Efficacy strategy. In this research-in-progress we present the design theory, the instantiated web application and the evaluation episodes. In the following, we review the theoretical background, describe our methodology, and present the nascent design theory. Finally, we show our preliminary results and discuss the next steps of our future research and intended contributions.

## Related Work and Theoretical Background

### *Proactive Guidance in Human-AI Collaboration*

Traditional decision support in service contexts relies on reactive user-initiated help, which imposes cognitive overhead. Users not only have to recognize the need for help but also formulate queries or commands to request it (Bansal et al., 2019). By contrast, proactive guidance – where the system infers users' needs from ongoing interaction and offers unprompted, context-relevant suggestions – shifts initiation from the human to the AI. In fast-paced service interactions (e.g., a customer requests status updates on an order and a SE must quickly navigate multiple information systems to find relevant information), this can divert attention from the customer. Research shows that people often fail to decide optimally when to seek AI assistance, resulting in suboptimal outcomes (Bansal et al., 2021). Proactive guidance systems ease this decision by taking the lead. They infer when assistance might be needed and display corresponding suggestions without an explicit request (e.g., the system infers the request from the dialogue and fetches the information from subsystems and shows them in the SE's interface). This approach can relieve working memory and has been shown to improve performance. In an experimental study in a programming context proactive assistance reveals significant benefits in productivity and user experience (Chen et al., 2025), though users may react negatively if the system's help is seen as intrusive or undermining autonomy (Diebel et al., 2025). Prior work in human-AI collaboration suggests that AI agents which intervene at appropriate moments can significantly boost efficiency (Fügener et al., 2022), but can also lead to detrimental effects when perceived as interrupting (Kim et al., 2024). Despite these advantages, most conversational AI research has focused on the benefits of reactive interaction paradigms (Lin et al., 2024), or proactive behavior in AI-driven dialogues (Deng et al., 2023). Our work contributes design knowledge for developing proactive guidance by AI in a human-to-human dialogue and how the context-aware assistance should be provided keeping interruptions to a minimum.

### *Selective Attention and Cognitive Resources*

We ground our design in theories of selective attention to ensure the system supports rather than hinders the SE. Classic attention research established that humans have limited processing capacity and must filter information, as the mind is not capable of attending to all inputs at once (Broadbent, 1958; Kahneman, 1973). Kahneman's (1973) capacity model conceptualizes attention as a resource that can be allocated but is finite. When tasks compete for the same attentional resources, performance deteriorates. Wickens' (1991) multiple resource theory refines this by proposing separate pools of resources for different modalities and stages of processing. Tasks that draw on the same modality or cognitive function interfere with each other more than tasks that use distinct modalities (Wickens, 2021). In a live customer call, an SE's visual channel (e.g., scanning screens) and auditory channel (listening to the customer), are both heavily engaged leaving little spare capacity. For instance, SEs must continuously keep track of the dialogue, process visual and auditory cues, and make rapid decisions. Thus, any extra action such as a switching to a different information system or adding extra clicks in information fetching adds to the overall load. If a support system forces frequent visual scanning or multitasking, it risks overload. Selective attention describes how we cope with this conflict, deploying filters to admit only the most relevant signals and suppress extraneous ones (Wickens, 2021). In customer service calls, these filters become critical. As SEs listen and scan multiple screens, they must continuously decide which visual or auditory inputs to prioritize. Design should therefore help users select the important information and suppress distractions. Research shows that not only too many simultaneous stimuli or too much irrelevant detail will significantly impair performance, but shifting attention from one task to another incurs its own switching costs (Chun et al., 2011). Every time an SE distracts away from the current task, they pay a cognitive toll. At the same time, working memory, the short-term buffer holding few chunks of information, cannot reliably store or manipulate the multiple context elements an SE would need if they must manually retrieve or remember past details (Baddeley, 2007). Designing an interface for these constraints means intelligently pacing information flow while avoiding unnecessary interruptions in the interface leading to decreased performance (Ohly & Bastin, 2023). The system should infer what the agent needs next and present that insight directly, rather than disturbing the user with irrelevant information or demanding separate visual searches or clicks. Prior studies on dashboard design demonstrate that adapting what is shown based on the user's current task load can reduce cognitive overload, reduce attention shifts and improve task efficiency (Toreini et al., 2022). In our context, this means a proactive system should deliver guidance in a way that minimizes unnecessary attentional shifts. Relevant information should be pushed to the SE in situ, formatted to be quickly digestible, and updated automatically so that the agent does not have to manually retrieve it or remember it.

## Research Approach

We followed a DSR methodology (Hevner et al., 2004) to iteratively develop and refine our solution and abstract design knowledge. Our project comprised three design cycles, following the general stages of problem awareness, suggestion, development, and evaluation in each cycle. In the first two cycles, the emphasis was on understanding the problem and user needs. The third cycle focused on consolidating design knowledge and building the final artifact. We adopted the process model of Kuechler and Vaishnavi (2012), which encourages cyclical refinement and learning. For the evaluation of the cycles we followed the

Technical Risk & Efficacy strategy suggested by Venable et al. (2016). In all cycles we conduct artificial formative evaluations of the individually suggested design principles to test our design hypotheses. Finally, we conclude the evaluation cycles with a summative evaluation moving from artificial to naturalistic evaluation to assess the artefacts' effectiveness. In the first two cycles, we conducted 16 semi-structured interviews with stakeholders from a corporate customer service center (including frontline SEs, business analysts, and team leaders) and domain experts outside the company. These interviews provided rich qualitative insights into current pain points and expectations for our system we intend to build. We analyzed interview transcripts using open coding and thematic analysis to derive user requirements following the guidelines by Saldaña (2009). The knowledge gained informed each subsequent design iteration. The final evaluation in design cycle 3 is two-fold. First, we conduct a formative evaluation via an eye-tracking study to empirically validate the impact of the individual design principles on SEs' attentional resources. Second, we conduct a naturalistic, summative evaluation deploying the artefact in an organizational setting with real users testing the solution in a real context. Below we briefly outline the design cycles and the final evaluation, a complete overview of the cycles is illustrated in Figure 1.

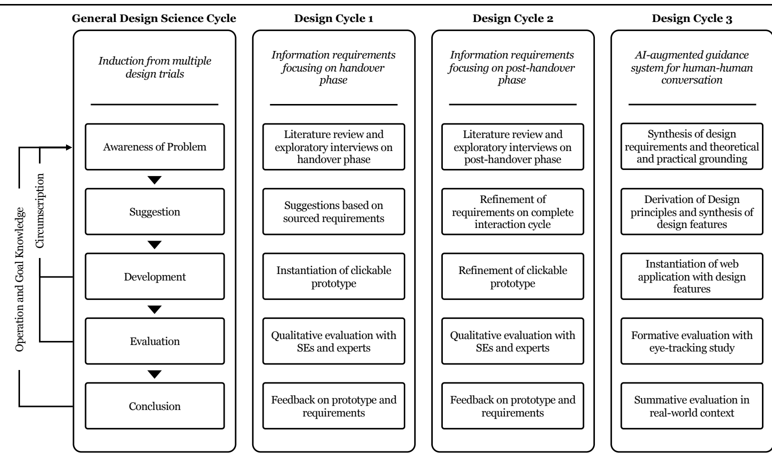


**Figure 1. UI of the Web Application**

**Design Cycle 1 – Problem Exploration & Initial Prototype:** The first cycle focused on the handover phase where a customer conversation transitions from an AI chatbot to a human agent. We reviewed literature on AI-human handovers and interviewed 12 participants to capture requirements for this scenario. From these insights, we identified initial design requirements and developed a low-fidelity clickable prototype. This prototype was essentially a rough interface model addressing attentional challenges in handover phases. We then gathered feedback through follow-up interviews, which confirmed usability issues, information requirements and additional needs, guiding improvements.

**Design Cycle 2 – Extending Scope:** In the second cycle, we broadened the scope to the post-handover phase, i.e. the remainder of the live human-customer interaction after the initial transfer. We refined requirements and added features to the prototype to support SEs throughout the conversation, not just at the moment of takeover. Again, we consulted theory and conducted 4 more interviews. Practitioners validated the new features and provided input on remaining gaps, leading to further refinements of both requirements and the interface.

**Design Cycle 3 – Consolidation & High-Fidelity Artifact:** In the final cycle, we synthesized the knowledge from cycles 1 and 2. The various specific requirements were abstracted into a set of meta-requirements (MR) that generalize the needs of SEs in real-time support contexts to improve attentional resource allocation. From these, we derived design principles (DP) to guide the solution at a broader level. We then mapped the principles to concrete design features and implemented a designed instantiation. This artifact constitutes a proactive AI-augmented guidance system attending to SEs limited attention resources. The development was iterative, incorporating feedback from domain experts along the way. The concluded design cycles ensured the artifact and nascent design theory is grounded in both theoretical and practical insights.

**Final Formative Evaluation:** In designing the eye-tracking study, we closely follow the proven approach of Toreini et al. (2022). We conduct a between-subjects controlled laboratory experiment with 100 participants. We use students for the lab experiments as they can be considered novice SEs and we achieve both an adequate sample size and comparable levels of knowledge. Each student receives 15 USD as a financial incentive for participating in the experiment. Participants are asked to perform an artificial customer service task. The task is defined as an information request concerning a service offering and fee structure overview of a credit card. In structuring the task, we closely follow the call center guideline provided by one of our practice partners. We choose this task because of its simplicity and common knowledge of the product. Still, before starting the experiment we provide participants a brief introduction to the product. Participants are separated into two experimental conditions. The treatment group performs the task with the instantiated design principles on, the control group performs the task with the instantiated design principles off. Participants will be able to access the same information in the same UI to ensure internal validity. We measure attentional resource allocation and cognitive load using the Tobii Glasses Pro 3 and tracking software incorporated into the web application. The task is conducted at the authors' university's behavioral lab. We measure attentional resource allocation by first defining areas of interest (AOI). Following Cheung et al. (2017) we measure attentional resource allocation by comparing fixation duration and count. Second, we track gaze transitions (Rayner, 1998) assessing the frequency and patterns of transitions between AOI. Pupil dilation informs us about the user's cognitive load (Beatty & Lucero-Wagoner, 2000). Following the experiment, we conduct a short semi-structured interview with each participant to qualitatively assess the perception of the interface. Based on the tracking of attentional resource allocation and cognitive load, we can quantitatively assess each instantiated design principles resource allocation depending on what kind of activity participants are currently doing. Combined with the qualitative evaluation we can infer their individual utility to the overall solution.

**Final Summative Evaluation:** To increase ecological validity, we conduct a summative evaluation in a naturalistic setting. We recruit 25 SEs from practice partners and let them conduct the same experiment as outlined above. SEs have different levels of experience which we consider a moderator in the experimental setting. Nonetheless, to account for experience levels and task complexity differences, we introduce a second more complex task of providing information for an investment fund. In constructing the task, we follow the same guideline from a practice partner. Unlike the formative evaluation, we focus on the evaluation in a real-world organizational setting with the key property of evaluation being the artefact's effectiveness. To this end, we let SEs conduct the tasks in their organizational setting. Ex-post we evaluate the effectiveness of the solution via semi-structured interviews.

## Preliminary Results

Our iterative design cycles surfaced four key meta-requirements (MR1-MR4) for proactive guidance systems in customer service, which we present below. Building on attention theory, we propose a nascent design theory that addresses limited attentional resources by prescribing design principles for AI-augmented proactive guidance systems and a designed instantiation. In so doing, we bridge the gap between prescriptive knowledge and explanatory knowledge (Baskerville et al., 2018; Gregor & Hevner, 2013), thereby offering both a practical solution and theoretical insights.

**Meta-Requirements:** Based on the literature and our interview findings, we distilled four overarching requirements that proactive guidance systems for SEs should meet. In brief, the system should: *(MR1) Minimize information noise and visual distractions* – present information cleanly and only when needed to avoid cluttering the interface; *(MR2) Minimize concurrent task execution* – reduce the need for the SE to juggle multiple applications or tasks at the same time; *(MR3) Provide proactive guidance in a synchronous fashion* – offer real-time hints or next-steps suggestions during the conversation without waiting to be asked; and *(MR4) Dynamically adapt to SE needs and interaction context* – tailor the content and detail of guidance based on the situation and the individual SE's preferences or level of expertise.

**Design Principles:** To fulfill the above requirements, we formulated three DPs, following the notion of high-level design rules of form and function (Gregor et al., 2020). Each principle is derived from and traceable to one or more meta-requirements. The design principles are:

*DP1: Dynamic hierarchy with progressive information disclosure.* The interface should organize and prioritize information in a dynamic hierarchy, revealing detailed information progressively as needed. This principle directly targets information overload. By not showing all data at once, minimizes noise (MR1) and reduces the need for the SE to switch contexts or collect information across multiple screens (MR2). Our interviews consistently raised the problem of fragmented information across many systems. One SE lamented, *"We have many systems, and it would be easier if we had a system where we could integrate multiple tasks… one would not have to open six tools just to access all the information" (SE)*. This reflects the elevated mental workload caused by constantly switching between disparate tools to gather customer info. Another interviewee observed, *"I think a complete transcription of what was said would be too much, as one would have to read through everything" (Business Analyst)*, emphasizing that filtering information presented to users is crucial. By integrating relevant data into one interface and using progressive disclosure (e.g. expanding details on demand or context driven), DP1 aims to give SEs a coherent view without overwhelming them by providing too much information at once on a single channel. This approach aligns with selective attention and multiple resource theory: reducing unnecessary visual scanning, prevent split-attention (Sweller et al., 2019) and avoiding memory and information load improves efficiency (Wickens, 2021).

*DP2: Real-time suggestions embedded in the workflow*. The system should proactively provide suggestions and guidance within the user's primary workflow interface, in real-time. Instead of requiring the SE to consult a separate source for help, the recommendations (e.g. next-best-action prompts, relevant knowledge snippets) are displayed in context. Participants in our study strongly supported this idea of contextual, on-the-fly assistance: *"When I then take over the customer, I already know what I need to do [as suggested by the system]. I think that can be very helpful" (Team Leader Customer Center)*. With DP2, as the conversation progresses, the SE sees prompts or options on-screen guiding the next steps, without having to pause and search elsewhere. This design principle addresses the need for synchronous guidance at the moment of decision (MR3). It also helps with MR1 and MR2 by preventing disruptive task switching. One interviewee noted that currently *"we have to consider the next-best actions [in another system]… we are constantly being reminded to address these" (SE),* meaning SEs must mentally track process checklists outside the chat interface. Embedding suggestions into the interface itself means the SE does not leave the conversation view, reducing memory strain and multitasking. DP2 provdes support on multiple levels. At a high level, the system maintains an overview of the process status – indicating which stage of the service process the case is in and what steps are upcoming – so the SE always knows the broader context of the call. At a micro level, the system offers concrete response suggestions or guidance for the immediate next turn in the dialogue. By presenting several suggestion options rather than a single answer, the system preserves the SE's autonomy and decision-making power (i.e. they can choose how to respond). This multi-option approach was deliberate, since forcing a single course of action could reduce user satisfaction (Poser, Wiethof, Banerjee, et al., 2022).

*DP3: Context-aware content personalized to the user and interaction*. The system's guidance should be tailored both to the ongoing conversation context and to the individual SE's needs or preferences. While users desire help, they warned against being flooded with too much information. An interviewee expressed concern, *"My concern with AI-generated suggestions is always whether the voicebot has accurately understood the keywords it processes. That's why I prefer to quickly review what was recorded from the customer and their response" (SE)*. Closely related another SE acknowledged, "An advantage would be if the voice function could transcribe the customer's conversation live [to adjust system output based on it].", indicating context-awareness, modality switching and system content adaption is a crucial component. Combined with insights from selective attention and multiple resource theory, these insights led to DP3, which stresses adaptive content presentation. The system should intelligently decide how much information to show (MR4). For example, provide a concise summary of the last customer exchanges rather than the full transcripts. Content presentation should also account for the SE's expertise level. Novice SEs may benefit from more detailed guidance, whereas experts might find that unnecessary or intrusive (Silver, 1991). Therefore, DP3 calls for flexibility and personalization. In practice, this could involve user-controlled settings (SEs can choose to hide or show certain components or select mode of guidance) and algorithm-driven adjustments (the system detects if a conversation is complex or if the user is struggling and responds by adjusting the level of guidance). This principle is rooted in the notion of selective attention. To conserve the SE's limited attentional capacity, the interface should filter and prioritize information in real-time. Only the most relevant data to the current query or problem should be foregrounded, with secondary details available on demand. Together, our three DPs translate the abstract MRs into practical design guidance, which we implemented in a high-fidelity protoype shown in Figure 2.

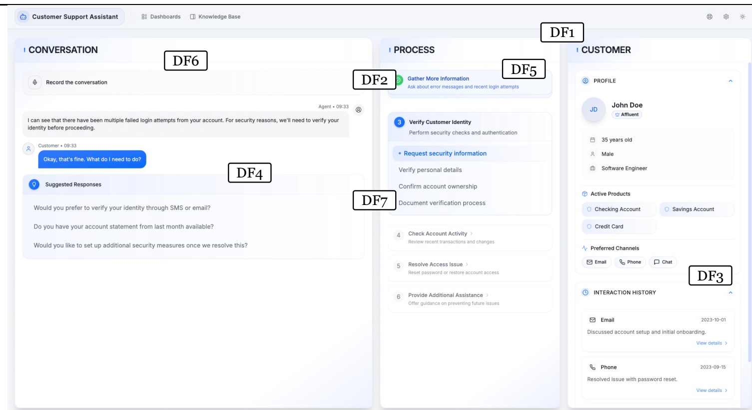


**Figure 2. Design Cycles of the Project**

We built the proactive guidance system as a Next.js web application whose backend uses LLMs for language understanding and generation, with Azure Cognitive Services for speech-to-text. A Retrieval-Augmented Generation pipeline fetches relevant documents (e.g., process guidelines, past cases) and integrates them into context-specific AI-generated suggestions. The interface mirrors the SE's chat workflow. The component "Conversation" shows the most recent messages of the live dialogue and proactive suggestions (DF4) directly below recent customer inputs. An adjacent Sidebar presents stable customer data and interaction history. It uses AI-generated summaries (DF3) to provide context without overloading the user. The interface includes collapsible components, such as the optional process tracking, to enforce progressive disclosure (DF1) while reducing visual clutter. The "Process" component (DF5) displayed alongside the conversation indicates major workflow steps (e.g., "Identify issue → Verify customer → Provide solution") and automatically updates as the call progresses, acting as a real-time checklist. The system's design (DF2 and DF6) ensures that guidance considers the full dialogue context, incorporates external knowledge bases (DF7) improving the relevance and coherence of recommended next actions.

## Next Steps, Future Research, and Intended Contributions

This study intends to make three contributions, which we will discuss along with the next steps we plan to take to shape this ongoing research from level 1 and 2 to a level 3 DSR contribution as defined by (Gregor & Hevner, 2013). *First*, we present a nascent design theory on the design of proactive AI-augmented guidance systems which accounts for limited attentional resources. The design theory has been developed and evaluated over multiple evaluation rounds grounded in practice and theory. We are adding to the growing design knowledge and discourse of attentive information systems (Toreini et al., 2022) by extending the scope in using components based on generative AI. Thereby, we address that current reactive interaction paradigms with GenAI might not yet fully capture the potential of context-aware assistance of SEs. With the complementary research of the eye-tracking study and summative evaluation, the next steps for strengthening this contribution are already in place. Furthermore, we intend to apply the nascent design theory in different contexts to further evolve and validate it to a point it can be called a design theory (Gregor & Hevner, 2013). *Second*, we provide an instantiated artefact that is adaptable to different use cases in human-human dialogues requiring minimal customization. We plan to open-source the codebase for future developments and applications in so far unstudied contexts. Thereby, we contribute to practice by providing a publicly available system architecture to build upon and disseminate the so far collected design knowledge. *Third*, we contribute to an emerging stream of research on generative user interfaces, using AI capabilities to create dynamic user interfaces (Cao et al., 2025). We focus on proactivity as crucial but delicate aspect of design. In future research, we plan to use the insights gathered from our evaluation to validate the downstream task effectiveness of using AI capabilities within interfaces.

## Acknowledgements

This work was supported by InnoSuisse and is part of the project 104.954 IP-SBM.

## References


Baddeley, A. (2007). *Working memory, thought, and action*. Oxford University Press.

Baird, A., & Maruping, L. M. (2021). The Next Generation of Research on Is Use: A Theoretical Framework of Delegation to and from Agentic Is Artifacts. *MIS Quarterly*, *45*(1), 315–341. https://doi.org/10.25300/MISQ/2021/15882

Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., & Horvitz, E. (2019). Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance. *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, *7*, 2–11. https://doi.org/10.1609/hcomp.v7i1.5285

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M. T., & Weld, D. (2021). Does the Whole Exceed its Parts? The Effect of AI Explanations on Complementary Team Performance. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3411764.3445717

Bar-Or, S., & Meyer, J. (2019). What Is Good Help? Responses to Solicited and Unsolicited Assistance. *International Journal of Human–Computer Interaction*, *35*(2), 131–139. https://doi.org/10.1080/10447318.2018.1437866

Baskerville, R., Baiyere, A., Gregor, S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems*, *19*(5).

Baym, N., Shifman, L., Persaud, C., & Wagman, K. (2019). INTELLIGENT FAILURES: CLIPPY MEMES AND THE LIMITS OF DIGITAL ASSISTANTS. *AoIR Selected Papers of Internet Research*. https://doi.org/10.5210/spir.v2019i0.10923

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In *Handbook of psychophysiology, 2nd ed* (pp. 142–162). Cambridge University Press.

Benner, D., Elshan, E., Schöbel, S., & Janson, A. (2021). What do you mean? A Review on Recovery Strategies to Overcome Conversational Breakdowns of Conversational Agents. *ICIS 2021 Proceedings*.

Broadbent, D. E. (1958). *Perception and communication*. Pergamon Press.

Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at Work*. *The Quarterly Journal of Economics*, *140*(2), 889–942. https://doi.org/10.1093/qje/qjae044

Cao, Y., Jiang, P., & Xia, H. (2025). Generative and Malleable User Interfaces with Generative and Evolving Task-Driven Data Model. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–20. https://doi.org/10.1145/3706598.3713285

Chen, V., Zhu, A., Zhao, S., Mozannar, H., Sontag, D., & Talwalkar, A. (2025). Need Help? Designing Proactive AI Assistants for Programming. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–18. https://doi.org/10.1145/3706598.3714002

Cheung, M., Hong, W., & Thong, J. (2017). Effects of Animation on Attentional Resources of Online Consumers. *Journal of the Association for Information Systems*, *18*(8). https://doi.org/10.17705/1jais.00464

Chun, M. M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, *62*, 73–101. https://doi.org/10.1146/annurev.psych.093008.100427

Deng, Y., Lei, W., Lam, W., & Chua, T.-S. (2023). A survey on proactive dialogue systems: Problems, methods, and prospects. *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*, 6583–6591. https://doi.org/10.24963/ijcai.2023/738

Diebel, C., Goutier, M., Adam, M., & Benlian, A. (2025). When AI-Based Agents Are Proactive: Implications for Competence and System Satisfaction in Human–AI Collaboration. *Business & Information Systems Engineering*. https://doi.org/10.1007/s12599-024-00918-y

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, *33*(2), 678–696. https://doi.org/10.1287/isre.2021.1079

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Q.*, *37*(2), 337–356. https://doi.org/10.25300/MISQ/2013/37.2.01

Gregor, S., Kruse, L. C., & Seidel, S. (2020). Research Perspectives: The Anatomy of a Design Principle. *Journal of the Association for Information Systems*, *21*(6). https://doi.org/10.17705/1jais.00649

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Q.*, *28*(1), 75–105.

Kahneman, D. (1973). *Attention and Effort*. Prentice-Hall.

Kim, M., Lee, J., Koh, Y., Lee, C., Lee, U., & Kim, A. (2024). Interrupting for Microlearning: Understanding Perceptions and Interruptibility of Proactive Conversational Microlearning Services. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–21. https://doi.org/10.1145/3613904.3642778

Kuechler, W., & Vaishnavi, V. (2012). A Framework for Theory Development in Design Science Research: Multiple Perspectives. *Journal of the Association for Information Systems*, *13*(6). https://doi.org/10.17705/1jais.00300

Lin, X., Wang, X., Shao, B., & Taylor, J. (2024). How Chatbots Augment Human Intelligence in Customer Services: A Mixed-Methods Study. *Journal of Management Information Systems*, *41*(4), 1016–1041. https://doi.org/10.1080/07421222.2024.2415773

Maniktala, M., Chi, M., & Barnes, T. (2022). Enhancing a student productivity model for adaptive problem-solving assistance. *User Modeling and User-Adapted Interaction*, *33*(1), 159–188. https://doi.org/10.1007/s11257-022-09338-7

Ohly, S., & Bastin, L. (2023). Effects of task interruptions caused by notifications from communication applications on strain and performance. *Journal of Occupational Health*, *65*(1), e12408. https://doi.org/10.1002/1348-9585.12408

Poser, M., Wiethof, C., Banerjee, D., Shankar Subramanian, V., Paucar, R., & Bittner, E. A. C. (2022). Let's Team Up with AI! Toward a Hybrid Intelligence System for Online Customer Service. *17th International Conference on Design Science Research in Information Systems and Technology*, 142–153. https://doi.org/10.1007/978-3-031-06516-3_11

Poser, M., Wiethof, C., & Bittner, E. (2022). Integration of AI into Customer Service: A Taxonomy to Inform Design Decisions. *ECIS 2022 Research Papers*.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Reinhard, P., Neis, N., Kolb, L., Wischer, D., Li, M. M., Winkelmann, A., Teuteberg, F., Lechner, U., & Leimeister, J. M. (2024). Augmenting Frontline Service Employee Onboarding via Hybrid Intelligence: Examining the Effects of Different Degrees of Human-GenAI Interaction. *Design Science Research for a Resilient Future*, 384–397. https://doi.org/10.1007/978-3-031-61175-9_26

Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage Publications Ltd.

Shao, Y., Huang, C., Song, Y., Wang, M., Song, Y. H., & Shao, R. (2024). Using Augmentation-Based AI Tool at Work: A Daily Investigation of Learning-Based Benefit and Challenge. *Journal of Management*, 01492063241266503. https://doi.org/10.1177/01492063241266503

Silver, M. (1991). Decisional Guidance for Computer-Based Decision Support. *Management Information Systems Quarterly*, *15*(1).

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*, *31*(2), 261–292. https://doi.org/10.1007/s10648-019-09465-5

Toreini, P., Langner, M., Maedche, A., Morana, S., & Vogel, T. (2022). Designing Attentive Information Dashboards. *Journal of the Association for Information Systems*, *23*(2), 521–552. https://doi.org/10.17705/1jais.00732

Vaithilingam, P., Zhang, T., & Glassman, E. L. (2022). Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models. *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3491101.3519665

Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: A Framework for Evaluation in Design Science Research. *European Journal of Information Systems*, *25*(1), 77–89. https://doi.org/10.1057/ejis.2014.36

Wenninger, A., Rau, D., & Röglinger, M. (2022). Improving customer satisfaction in proactive service design. *Electronic Markets*, *32*(3), 1399–1418. https://doi.org/10.1007/s12525-022-00565-9

Wickens, C. (2021). Attention: Theory, Principles, Models and Applications. *International Journal of Human–Computer Interaction*, *37*(5), 403–417. https://doi.org/10.1080/10447318.2021.1874741

Wickens, C., & Carswell, C. M. (2021). Information processing. In *Handbook of Human Factors and Ergonomics, 5th ed* (pp. 111–149). John Wiley & Sons, Inc.

Wickens, C. D. (1991). Processing resources and attention. In *Multiple Task Performance*. CRC Press.

Wiethof, C., Poser, M., & Bittner, E. (2022). Design and Evaluation of an Employee-Facing Conversational Agent in Online Customer Service. *PACIS 2022 Proceedings*.