Proactive and Reactive Help from Intelligent Agents in Identity-Relevant Tasks

Marc Goutier
TU Darmstadt
goutier@ise.tu-darmstadt.de

Christopher Diebel
TU Darmstadt
diebel@ise.tu-darmstadt.de

Martin Adam
TU Darmstadt
adam@ise.tu-darmstadt.de

Alexander Benlian
TU Darmstadt
benlian@ise.tu-darmstadt.de

Abstract

Enabled with artificial intelligence (AI), intelligent agents in information systems have developed from passive tools that only help in return to user prompts (i.e., reactive help) to intelligent agents that can help without requiring user requests (i.e., proactive help). Yet, it is unclear how users react to these different types of help and whether the task creates or reinforces the users' identity (i.e., identity-relevance). Against this backdrop, we drew on self-affirmation and identity theory and conducted a vignette-based online experiment (n = 135). Our results show that proactive (vs. reactive) help decreases users' *willingness to accept help because of* users' *higher perceived self-threat (i.e., threat to their self-image). Identity-relevance of the task moderates this effect – high (vs. low) identity-relevance causes a greater increase in self-threat through proactive (vs. reactive) help. Our study contributes to a better understanding of help from intelligent agents and their implications for effective human-AI collaboration.*

Keywords: Intelligent Agents, Proactive and Reactive Help, Self-Affirmation Theory, Identity-Relevant Tasks.

1. Introduction

Recent research on information systems (IS) suggests that modern intelligent agents (IAs) are becoming increasingly autonomous, requiring less user input in human-AI collaboration, in contrast to former IAs (Spiekermann et al., 2022). While previously users prompted help and IAs offered help in response to user requests–*reactive* help. Recent advances in artificial intelligence enable IAs to act proactively, try to recognize the need for help of users in advance, and offer help without a user prompt–*proactive* help (Cila, 2022). Amazon Alexa is a prominent display of this shift in the behavior of IAs. The AI-based and cloud-enabled virtual assistant has a built-in "latent goal discovery" functionality (Amazon, 2021), which allows Alexa to predict the underlying need for help based on the users' initial question or current conversation. By matching the conversation with pre-trained patterns, context, and the likelihood within all Alexa conversations, Alexa can offer *proactive* help by addressing the assumed underlying need the user has not mentioned explicitly, claiming to provide better support for the user.

While *reactive* help by IAs in human-AI collaboration is broadly considered beneficial and accepted by users, *proactive* help from IAs can be necessary to achieve higher performance than *reactive* help (Baird & Maruping, 2021; Kraus et al., 2021). Users tend to request help at a lower frequency than they should to increase their overall performance, even when the help is objectively beneficial (Fuegener et al., 2021, 2022). This behavior is grounded in a lack of meta-knowledge, the phenomenon in which users cannot assess their capabilities for a specific task correctly, which causes them to overestimate their skills and thus ask for less help than necessary (Fuegener et al., 2022). However, *proactive* help from IAs may overcome this hesitation of asking for help by not requiring a prompt from the user to offer help. With *proactive* help, IAs can offer more help compared to the help users would ask for. Consequently, users may increase their overall performance, assuming the help is objectively useful.

Despite these clear distinctions in the *type of help* from IAs, it is surprising to find how little attention the research community has paid to investigate the effects of *proactive* help compared to *reactive* help from IAs in human-AI collaboration. We identified three important research gaps. First, multiple studies examine forms of *reactive* help by IAs (Fuegener et al., 2021, 2022; Kraus et al., 2021). However, there is limited IS research on comparing and differentiating users' reactions to *proactive* versus *reactive* help from IAs. Second, beneficial help by IAs is often treated as

URI: https://hdl.handle.net/10125/106423
978-0-9981331-7-1

universal, favorable, and welcomed by users (Li & Karahanna, 2015), ignoring cases where beneficial help can cause *self-threat* (i.e., help is seen as a threat to their self-image) (Steele, 1988) when offered *proactively*, leading to a negative effect on users' *willingness to accept this help* (Craig et al., 2019; Strich et al., 2021). Third, studies on help have focused on tasks in the work context in general (Strich et al., 2021), ignoring the importance of the specific task for users to create and reinforce their identity (i.e., the users' *identity-relevance* of the task) (Gendolla, 1998). This is especially important because IAs are becoming increasingly competent, taking over more sophisticated tasks with potentially high *identity-relevance* for the user, e.g., medical diagnostic as a task with potentially high *identity-relevance* for a medical doctor. Taken together, we ask the following research questions:

RQ1: *To what extent does proactive (vs. reactive) help from IAs create self-threat in users and thus influence users' willingness to accept that help?*

RQ2: *How does the identity-relevance of a task influence this effect?*

To answer these research questions, we draw on self-affirmation theory (SAT) and identity theory (Sherman & Cohen, 2006; Steele, 1988), and we conducted a vignette-based online experiment (Atzmüller & Steiner, 2010) with 135 users. We show that *proactive* (vs. *reactive*) help creates feelings of *self-threat* in users, decreasing their *willingness to accept help*. In addition, the effect of *proactive* help triggering *self-threat* can be moderated by the *identity-relevance* of the task – high (vs. low) *identity-relevance* of the task leads to a greater increase in *self-threat* from *proactive* (vs. *reactive*) help.

We make three contributions to IS research on how users react and respond to IAs in human-AI collaboration. First, we provide empirical insights into human-AI collaboration by comparing users' reactions to proactive and reactive help from IAs. Second, we extend the current understanding of help from IAs in human-AI collaborations by unpacking the reasons why proactive help lowers users' willingness to accept help. Third, we contribute to SAT in human-AI collaboration by linking self-threat to the specific task, distinguishing between high and low identity-relevance. In addition, we derive viable suggestions for designers of IAs on how to implement proactive and reactive help in the context of task identity relevance.

2. Theoretical Background

2.1. Intelligent Agents and Willingness to Accept Help

IAs, such as automated, sales, or recommendation agents (Adam et al., 2021; Kuehl et al., 2022; Maedche et al., 2019), have become indispensable in industry and private households. This trend comes with an increasing interest in understanding human-AI collaborations (Adam et al., 2022; Feine et al., 2019). IAs can help users increase their task performance, e.g., in the form of recommendations (Maedche et al., 2016). The term *help* thereby refers to the various possibilities in how IAs can support users in performing tasks, ranging from IAs providing advice, recommendations, or guidance that allow users to make better decisions (Berger et al., 2021; Morana et al., 2020; Morana et al., 2017), to IAs autonomously executing tasks after users' delegation of decision and execution rights to the IAs (Berente et al., 2021; Pinski et al., 2023; Schneider et al., 2020). In this study, we differentiate between two types of help from IAs, namely *proactive* and *reactive* help, introduced by Baird and Maruping (2021), and their different effects on users' *willingness to accept help*: *Reactive* help, where help is user-invoked, in response to user demand, and *proactive* help, where help is offered without being asked for (i.e., help is invoked intelligently or automatically by IAs) (Gregor & Benbasat, 1999; Lee et al., 2019; Morana et al., 2017).

In regards to IAs, *reactive* help is the most established and studied form of help. For instance, while previous research on interface design features such as the representation from IAs, e.g., anthropomorphism or personalization, is mainly grounded in a *reactive* context where the user has to interact first and ask for help (Komiak & Benbasat, 2006; Qiu & Benbasat, 2009), *proactive* acting IAs has received little attention yet. However, *proactive* help from IAs is an emerging phenomenon that is gaining increasing attention in recent IA literature (Kraus et al., 2021; Wenninger et al., 2022) but also in IAs' development and application (Amazon, 2021; GitHub, 2021). *Proactive* help can be more efficient and lead to higher performance than *reactive* help since the user can be offered help when they might not have asked for it (Kraus et al., 2021). However, receiving *proactive* help can also cause negative reactions in users, such as feeling threatened in competence, leading to a lower *willingness to accept help* (Harari et al., 2021). In the following, we draw on self-affirmation theory and identity theory as a theoretical lens to explain how users may react to help.

2.2. Self-Affirmation Theory and Self-Threat

SAT is first popularized by Steele (1988) and is motivated by humans' need to maintain their self-image (Sherman & Cohen, 2006; Steele, 1988). According to SAT, the overarching goal of individuals is to preserve their own self-integrity as "adaptively and morally adequate, that is, as competent, good, coherent, unitary, stable, capable of free choice, capable of controlling important outcomes, and so on" (Steele, 1988, p. 262). The self-affirmation process will be activated by information threatening their self-image and persist until their self-view is restored (Steele, 1988). When individuals' self-image is threatened – e.g., by being questioned about being competent in their jobs – they will experience *self-threat*.

Following SAT, individuals can address *self-threat* in three ways to restore their self-view: First, experienced *self-threat* can be reduced directly through defensive mechanisms such as avoiding or bypassing the threat. Second, humans accommodate the threat by changing their behavior or mindset. Third, *self-threat* can be addressed through self-affirmation with self-integrity activity that is not necessarily linked directly to the threat itself. Within this paper, we focus on the first and most prominent way. Users can decline the help offered by IAs (i.e., show a lower *willingness to accept help)*, reducing the self-threat (Craig et al., 2019; Harari et al., 2021).

2.3. Identity Theory and Identity-Relevance

According to Stets and Burke (2000), identity theory attributes humans with the capability to see themselves as objects. This capability enables humans to name, classify, or categorize themselves in relation to social categories, norms, or other classifications, a process called identification. This process of identification leads to an identity. Therefore, identity is seen as the sum of all self-beliefs of a human, which means that identity comprises all beliefs about who the individual is (Craig et al., 2019).

Humans need to create and reinforce their identity, which they do through actions and behaviors. These actions must be well balanced with their self-beliefs (Ashforth, 2016). A task, e.g., an activity at work, can be an action that is used to create and reinforce identity. From the individual's perspective, some tasks have high importance in the identity-creation process, i.e., these tasks have a high *identity-relevance*. Other tasks can have a limited or no meaning for the individual's identity, i.e., these tasks have a low *identity-relevance* compared to other tasks (Gendolla, 1998).

3. Hypotheses Development

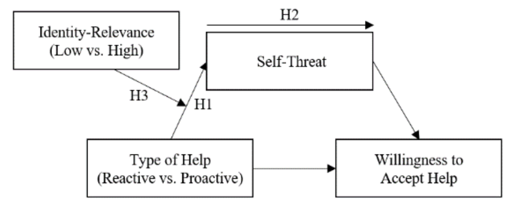

Figure 1. Research Model

This section provides an overview of our research model (Figure 1). According to SAT, humans are intrinsically compelled to retain a positive self-image. Humans want to maintain their self-image to be morally adequate, e.g., to be good, coherent, unitary, stable, and competent. Events and information that endanger their self-image will lead to *self-threat* (Sherman & Cohen, 2006; Steele, 1988).

Adapted to the context of help from IAs, we hypothesize that *proactive* (vs. *reactive*) help can threaten the users' self-image. Unsolicited help by definition, such as *proactive* help, is not prompted by users. Therefore, users might perceive that they have, consciously or unconsciously, signaled the need for help to others, e.g., signaling that they cannot accomplish the task with the expected performance by others (Ilgen & Davis, 2000). Having the perception that others (IAs which offer help as well as potentially involved humans) see themselves in a state of needing help causes *self-threat*. In addition, IAs are capable of *proactively* helping users in their core activities at work. This kind of help at work can threaten users, giving users the impression that IAs can take over their activities at work and eventually replace the user (Strich et al., 2021). The fear of being replaced at work can lead to *self-threat* as well. Since *proactive* (vs. *reactive*) help is usually unexpected, the *self-threatening* effect can be enhanced (Ashford et al., 2003). As a consequence, we state the following hypothesis:

H1: *Proactive (vs. reactive) help from IAs increases users' self-threat.*

In SAT, a *self-threat* is considered to jeopardize the positive self-image. As a consequence, humans need to respond adequately to reaffirm their self-image. One possible response to reaffirm their self-image is to react defensively, e.g., avoiding or bypassing the threat (Steele, 1988). Recent studies on collaboration between IAs and humans have shown that IAs can threaten the user, e.g., challenging their competence at work by offering *proactive* help (Seeber et al., 2020), leading to defensive mechanisms by users to protect their self-image. While some users

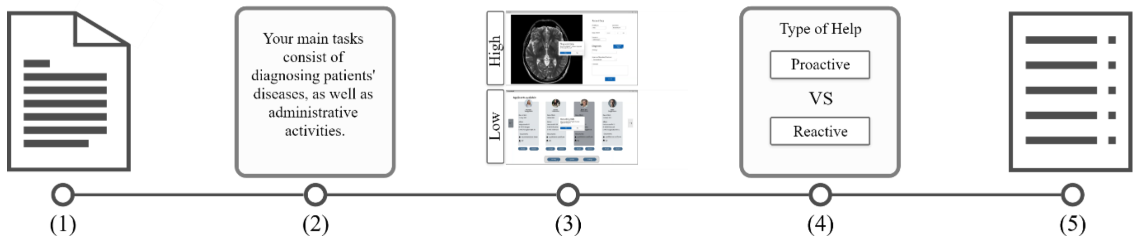

Figure 2. Five major steps of our experimental procedure

simply refuse to work with IAs (Strich et al., 2021), defensive mechanisms by users can even surpass pure refusal and turn into aggressive resistance (Lapointe & Rivard, 2005) against the IAs, e.g., containing the production capability of the IA (Puntoni et al., 2021) or user data manipulation by entering false data to force the IA to come to another decision (Strich et al., 2021). Therefore, we hypothesize that humans will respond defensively in return to *self-threat*, triggered by *proactive* help and will be reflected by declining help from IAs, because it is perceived as the option with the best "relative effectiveness-to-cost ratio" (Steele, 1988, p. 293). Accordingly, we derive our following hypothesis:

H2: *Self-threat mediates the effect of proactive (vs. reactive) help on the users' willingness to accept help, so an increase in self-threat leads to a decrease in users' willingness to accept help.*

According to identity theory, identity can be defined as the sum of all self-beliefs (beliefs about "who am I") (Craig et al., 2019). Humans can view themselves as an object and to categorize, classify, or name themselves. By this self-categorizing or self-identifying process in relation to other social categories or norms, an identity is created (Stets & Burke, 2000). Subsequently, humans use actions and behavior to create and reinforce their identity (Ashforth, 2016). At work, humans can use tasks for identity-creation, i.e., identity-relevant tasks (Gendolla, 1998). *Self-threat* occurs when humans believe that they will be viewed as "socially undesirable, incompetent, […] or irresponsible or immoral" (Leary & Baumeister, 2000, p. 17, p. 17) because of the actions or behaviors they perform. Applied to the work context, having the perception of signaling to coworkers a lack of competence or status with how one performs a task can cause *self-threat*. (Harari et al., 2021; Kakkar et al., 2019; Nadler & Fisher, 1986; Pettit et al., 2010). In addition, *self-threat* is influenced when the threatening domain (e.g., competence, status, etc.) is personally important or part of the human's identity (Sherman & Cohen, 2006). Thus, *self-threat* may be increased even more when the task is used for identity creation or identity-reinforcement, i.e., the task has a high *identity-relevance* for the user (Gendolla, 1998). When the task that signals a lack of competence or status is used for identity-creation, i.e., is *identity-relevant*, the identity is challenged, which further increases *self-threat*. However, if the task itself is not part of the identity-creation, i.e., not identity-relevant, the identity is not challenged, and the threat to the self-image is limited (Sherman & Cohen, 2006). To conclude, we state our third hypothesis:

H3: *High (vs. low) identity-relevance of the task leads to a greater increase in self-threat from proactive (vs. reactive) help, therefore identity-relevance moderates the effect of type of help on self-threat and consequently on the users' willingness to accept help.*

4. Research Methodology

4.1. Experimental Design and Manipulations

We conducted a vignette-based online experiment with 2x2 (type of help: *reactive* vs. *proactive*; *identity-relevance*: high vs. low) full-factorial design and between-subject treatments. Figure 2 depicts the five major steps of our experiment. Following established standards for this method (Aguinis & Bradley, 2014; Atzmüller & Steiner, 2010), after introducing the users to the experiment (*Step 1*), users were told to express their feelings after they stepped into the shoes of a fictional medical doctor (framed as gender-neutral, pronouns: they/them/their) and experienced a simulated human-AI collaboration with an IA (*Step 2*). For our experiment, we decided to focus on the context of radiology in the medical sector because IAs are increasingly essential to validate medical diagnoses (Dilsizian & Siegel, 2014). In practice, an increasing number of companies rely on using IAs for medical purposes. For example, Aignostics is a German startup that uses AI-based algorithms to support medical research and diagnostics (Aignostics, 2022). This trend is also reflected in the growing interest in recent literature to explore IA in human-AI collaboration as a decision-support tool for medical diagnosis (Jussupow et al., 2021; Wang et al., 2021).

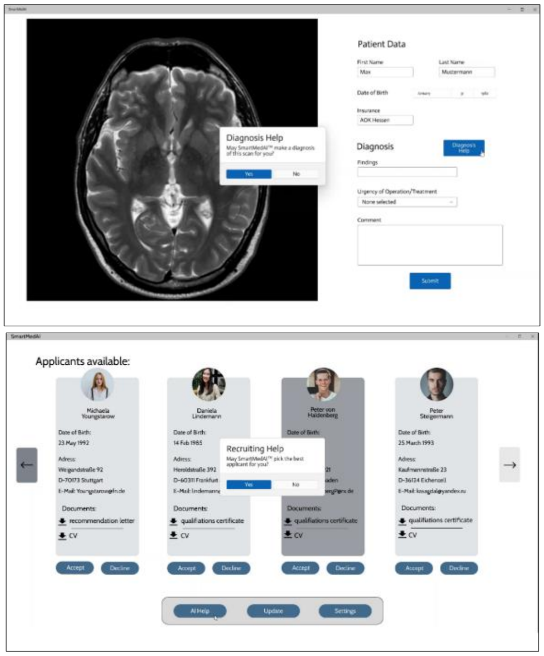

Figure 3. Screenshot of the high identity-relevance manipulation (CT scenario; top) and the low identity-relevance manipulation (hiring scenario; bottom)

The medical doctor was introduced as a traditional medical doctor working in diagnostics. Traditional medical doctors' motives lie in patient care and the doctor-patient relationship (Martin et al., 2021), unlike doctorial managers who consider administrative tasks an essential part of their identity (Cascon-Pereira et al., 2016). To set the boundaries for the medical doctor's identity, we informed the users that their main tasks as a medical doctor consist of diagnosing and analyzing patients' diseases and administrative activities. However, the medical doctor was presented as a person whose passion lies in core tasks of a traditional medical doctor - diagnosing and analyzing patients' diseases. For our *identity-relevance* manipulation, we simulated two different scenarios at work that the medical doctor experiences (*Step 3*). In both scenarios, the medical doctor must complete an equally important task that differs in *identity-relevance* (Figure 3). For our scenario with high *identity-relevance* we choose to simulate the diagnosis process using computed tomography (CT), a well-known practice for diagnosing and analyzing patients' diseases. A software represents the CT scan that the medical doctor has to analyze. In addition, the medical doctor must fill out a form with each relevant patient's information such as the diagnosis and the urgency of treatment.

For our scenario regarding low *identity-relevance*, we chose the hiring process of a new employee–an administrative activity of high importance, especially in the medical sector. Hiring the right people can lead to higher performance, reduced costs, and improved care efficiency (Meyer, 2018). However, this task is not considered of high *identity-relevance* for a traditional medical doctor because it is not directly related to patient care or the doctor-patient relationship, the main identity of a traditional medical doctor (Martin et al., 2021). A software represents each applicants' short profile, including the option to download their applications. The medical doctor has to decide which candidate will be invited to a job interview. Both scenarios have the possibility for the medical doctor to receive help from an IA, which we introduced as SmartMedAI™.

To compare the effects of type of help (*proactive* vs. *reactive*) on users' *willingness to accept help* from IAs, we designed two different human-AI collaborations for each scenario (*Step 4*). In both collaborations, SmartMedAI™ offers help to the medical doctor, either in the form of *reactive* or *proactive* help. Users experienced the collaboration in 30-second videos to increase realism. To make sure the users understood what happened in the video, they also read a textual description of the human-AI collaboration. In the *reactive* help condition, the medical doctor actively requested help which was operationalized in both scenarios by a moving mouse cursor that clicks on an "Ask AI for help" button during the work process. In the *proactive* help condition, SmartMedAI™ *proactively* offered help without any request from the user. The offer happened after the medical doctor had enough time to look closely at the CT scan or job candidates and when he was also not interacting with the mouse and keyboard to minimize the perception of interruption or intrusion. In both conditions, SmartMedAI™ opened a message window and asked the user if help was needed with the task, giving the option to accept or decline the help. After the users saw the videos and read the text, the experiment concluded with a questionnaire to gather data on the variables measured and users' demographic information (*Step 5*).

4.2. Sample Description, Variables Measured, and Manipulation Checks

In total, 155 participants responded to our survey and completed the questionnaire. We removed five participants due to low-effort responses, such as always selecting the same number in our 7-point Likert-type scale (Huang et al., 2015), two participants who reported problems within the experiment (e.g., the video was not working), and 13 participants completing the experiment in over 50 minutes, more than double of the average completion time. This leads to a final data set of 135 participants for our statistical analysis. The descriptive statistics are displayed in Table 1. Since our participants were randomly selected, we did not filter for medical expertise. With a random sample, we ensured that participants could immerse themselves fully in the role during the experiment. By preventing an overrepresentation of participants with medical expertise, we aimed to minimize potential biases that could extend beyond the identity-relevance of the tasks.

We chose *willingness to accept help* as our dependent variable, measured by three items adapted by Harari et al. (2021) (e.g., "*I would allow SmartMedAI^{TM} to help me.*"). We measured *self-threat*, our mediator, with six items adapted from literature (e.g., "*SmartMedAI^{TM} makes me feel less confident that I understand things well enough to get work done.*") (Burke & Reitzes, 1991; Carter & Grover, 2015; Craig et al., 2019; Fleming & Courtney, 1984; Harari et al., 2021; Rosenberg, 1965). In addition, we measured age, gender (*female, male, and diverse*), personal innovativeness (e.g., "*If I heard about a new technology, I would look for ways to experiment with it.*") (Agarwal & Prasad, 1998), and negative affectivity (e.g., "*My feelings are hurt rather easily.*") (Ayyagari et al., 2011) as our control variables which we assumed may have an impact on our dependent variables (Spector & Brannick, 2011). We measured all variables besides demographics within a 7-point Likert-type scale.

We calculated Cronbach's alpha and average variance extracted to demonstrate the reliability and validity of our measurement variables. Cronbach's alpha was greater than the threshold value of 0.70 (Nunnally, 1978), indicating internal consistency for all constructs. The average variances extracted (AVEs) were above the recommended threshold of 0.50 (Hair Jr. et al., 2021), while the square roots of the AVEs were greater than correlations between the constructs, providing evidence for discriminant validity (Fornell & Larcker, 1981). To derive the effectiveness of our manipulations *type of help* and *identity-relevance*, we conducted several one-way ANOVAs and t-tests of our manipulation checks. Our results indicate that users perceived the help provided by the IA in the *reactive* help condition as more *reactive* (e.g., "*SmartMedAI^{TM} helped me because I made it clear I wanted its help*.") (Harari et al., 2021) than in the *proactive* condition ($F = 60.92$, $p < 0.001$). Moreover, users in the *proactive* help condition perceived help provided by the IA as more *proactively* (e.g., "*SmartMedAI^{TM} offered help without me asking for help.*") (Harari et al., 2021) than in the *reactive* help condition ($F = 57.63$, $p < 0.001$). In addition, there was no statistical difference in the perceived task importance (e.g., "*How important do you regard hiring qualified personnel for your clinic/doing CT scan analysis for detecting disease?*") of both *identity-relevance* conditions ($T = 1.20$, $p = 0.23$), but they were perceived as different identity-relevant (e.g., "*You identify yourself as a doctor, because you are good at conducting administrative activities/good at diagnosing diseases*") ($T = 21.33$, $p < 0.001$). Therefore, we can assume that the manipulations were successful.

Table 1. Descriptive Statistics

	Mean	**StD**
Willingness to Accept Help	5.14	1.44
Self-Threat	2.90	1.20
Age[1]	2.41	0.93
Gender (Female)	0.30	0.46
Negative Affectivity	3.60	1.35
Personal Innovativeness	4.85	1.28

[1] "younger than 18 years old" = 1; "19 to 29 years old" = 2; "30 to 40 years old" = 3; "41 to 50 years old" = 4; "51 to 60 years old" = 5 and "61 or older" = 6.

4.3. Hypotheses Testing

We conducted a hierarchical multiple linear regression analysis with four stages to test our hypotheses (Table 2). We performed linear regressions on the mediator *self-threat* and the dependent variable *willingness to accept help*. For all stages, we included the control variables age, gender (female = 1, other (male, diverse) = 0), negative affectivity, and personal innovativeness. We coded our independent variables type of help (reactive help = 0; proactive help = 1) and identity-relevance (high identity-relevance = 0; low identity-relevance = 1) as binary. In **support of H1,** our results in stage 1 display a significant effect of *type of help* on *self-threat* ($\beta = 0.43$, $p < 0.05$). Thus, *proactive* (vs. *reactive*) help from IAs leads to significantly more *self-threat* ($M = 2.93$ vs. $M = 2.49$). In addition, we find in stage 4 a significant mediation effect of *self-threat* on *willingness to accept help* ($\beta = -0.26$, $p < 0.05$). This provides **support for H2**; the increase in *self-threat* from *proactive* (vs. *reactive*) help from IAs decreases users' *willingness to accept help* ($M = 4.78$ vs. $M = 5.48$, total effect = -0.65, direct effect = -0.54, indirect effect = -0.11). Lastly, we analyzed the moderation effect of *identity-relevance*. Our results in stage 2 demonstrate **initial support for H3**. We find a significant interaction between *type of help* and *identity-relevance* in relation to *self-threat*.

To increase the robustness of our moderated mediation analysis, we conducted a bootstrap analysis with 5,000 bootstrap samples and confidence intervals of 95% using PROCESS model 7 (Hayes, 2022). In

Table 2. Linear regressions on self-threat and willingness to accept help

	Self-Threat		Willingness to Accept Help	
	Stage 1	Stage 2	Stage 3	Stage 4
Intercept	2.63^{***} (0.60)	2.43^{***} (0.62)	6.12^{***} (0.73)	6.88^{***} (0.76)
Manipulations				
Type of Help	0.43^{*} (0.20)	0.93^{**} (0.28)	-0.65^{**} (0.24)	-0.54^{*} (0.24)
Identity-Relevance	-	0.24 (0.27)	-	-
Mediation				
Self-Threat	-	-	-	-0.26^{*} (0.10)
Moderation				
Type of Help x Identity-Relevance	-	-0.99^{**} (0.38)	-	-
Controls				
Age	0.02 (0.11)	0.02 (0.11)	-0.34^{**} (0.13)	-0.34^{**} (0.13)
Gender	-0.08 (0.24)	-0.03 (0.23)	0.26 (0.29)	0.24 (0.28)
Negative Affectivity	0.21^{**} (0.08)	0.20^{*} (0.08)	-0.23^{*} (0.10)	-0.18^{T} (0.10)
Personal Innovativeness	-0.18^{*} (0.08)	-0.17^{*} (0.08)	0.17^{T} (0.10)	0.13 (0.10)
R^2	0.13	0.18	0.12	0.19

Note: n = 135; $^{T}p < 0.10$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$; () = standard error

support of H3, our results lend support to the relationship that identity-relevance interacts with *type of help* in the form of a moderated mediation (Index of moderated mediation = 0.26, confidence interval [0.01, 0.59]). High (vs. low) *identity-relevance* of the task leads to a greater increase in *self-threat* from *proactive* (vs. *reactive*) help. Specifically, we find that the effect of *type of help* on *willingness to accept help* via *self-threat* is significant when the IA provides help for a task with high *identity-relevance* (*M* = 3.34 vs. *M* = 2.41, total effect = -0.78, direct effect = -0.54, indirect effect = -0.24, confidence interval [-0.56, -0.01]) but not when the IA provides help for a task with low *identity-relevance* (*M* = 2.49 vs. *M* = 2.55, total effect = -0.53, direct effect = -0.54, indirect effect = 0.01, confidence interval [-0.16, 0.15]).

5. Discussion

5.1. Contributions to Research

Our study provides three valuable contributions to IS research on how users react and respond to IAs in human-AI collaboration. First, we demonstrate that users' *willingness to accept help* significantly differs between *proactive* and *reactive* help. We investigated users' responses to *proactive* and *reactive* help from IAs at work. In recent literature on how users react and respond to IAs in human-AI collaborations, help was mainly investigated as *reactive* help or without distinguishing between the *type of help* (Fuegener et al., 2022; Kraus et al., 2021), while the insights, derived from IAs research on help, have been generalized and applied to all types of help. Our study shows that user reactions to *proactive* help can deviate from the positive reaction to *reactive* help from IAs. *Proactive* help is seen way more negatively than reactive help, leading to a higher rate of declining help from IAs. This finding stresses the need for a clear distinction between the types of help in research.

Second, we increase our understanding of what drives the differences in users' *willingness to accept* for different types of *help*. We investigated *self-threat* as an underlying explanation of why user reactions differ between *proactive* and *reactive* help. Recent research has focused on the beneficial aspects of help from IAs, where help is favorable and welcomed by users since users benefit from help objectively, e.g., higher performance or satisfaction (Li & Karahanna, 2015). However, research on human-AI collaborations has so far neglected cases where help from IAs or other intelligent technology systems was declined or pushed back (Craig et al., 2019; Strich et al., 2021). Our research shows that *self-threat* is an explanatory mechanism for a negative reaction to help. More specifically, we demonstrate that *proactive* (vs. *reactive*) help faces more negative user reactions due to higher feelings of *self-threat*, leading to lower *willingness to accept help*. In addition, having revealed *self-threat* as a mechanism that explains negative user reactions, our research also provides a useful intervention point to investigate how to lower *self-threat* and consequently reduce or eliminate negative reactions toward *proactive* help.

Finally, we shed light on the nature of the task as a critical boundary condition that shapes the effect of proactive (vs. reactive) help on willingness to accept help via self-threat. Previous research focused on help from IAs in the workplace from a general perspective

(Strich et al., 2021). Our study demonstrates that it is important to consider at the specific task for which IAs provide help. We differentiate between tasks based on their identity-relevance, with high identity-relevant tasks being essential for creating and reconfirming the users' identity, and low identity-relevant tasks having a limited effect on users' identity-creation. We show that high (vs. low) identity-relevance of the task leads to a greater increase (vs. decrease) of self-threat from proactive (vs. reactive) help by IAs and, therefore, to a greater decrease (vs. increase) of the users' willingness to accept help. This finding stresses the importance of analyzing help on a task level and not in general, since the users' reactions to help can differ based on the task's nature. In addition, distinguishing between high and low identity-relevance provides a new lens to research to make more accurate predictions and recommendations.

5.2. Implications for Practice

Our study provides insights for research and valuable implications for designers of IAs in organizations. The insights are of considerable importance since evermore IAs are becoming capable of providing help in a *proactive* manner, and former assumptions about users' reactions to IAs - based on *reactive* help - might not be applicable anymore. IAs are becoming more intelligent, e.g., predict that the user could need help, allowing designers to design help from IAs that offer this help without a user prompt (i.e., *proactive*). However, our research suggests that, although the *proactive* help from IAs might be objectively useful, users tend to decline this help. Derived from our study, we suggest that designers shape help from IAs in consideration of the *identity-relevance* of the task. Users performing tasks with a low *identity-relevance* will be profiting from *proactive* help. In contrast and whenever possible, designers of IAs in organizations should implement *reactive* help instead of *proactive* help when it is related to a task with a high *identity-relevance* to avoid a decline of the help or, even worse, face resistance and manipulation toward the IAs (Strich et al., 2021).

However, there may be some cases where *proactive* help is needed because of its significant benefits, e.g., in time-critical situations such as in medicine or security breaches. If this is the case, we suggest designers to reduce *self-threat* since our research revealed high *self-threat* as the driver for negative response to help from IAs. In addition, we revealed *identity-relevance* as a potential factor influencing self-threat generation. Therefore, we recommend designing *proactive* help from IAs in such a fashion that the IA does not interrupt or interfere in the identity-creation process, e.g., by signaling that accepting the help is a one-time event with no consequences for future human-AI collaborations, allowing users to continue to create and confirm their identity with the task in the future.

6. Conclusion, Limitations and Directions for Future Research

In this study, we asked users to what extent *proactive* (vs. *reactive*) help from IAs creates *self-threat* in them, how *self-threat* influences the users' *willingness to accept help*, and if the *identity-relevance* of a task influences this effect. We showed that users are less *willing to accept help* from IAs when the help is *proactive* (vs. *reactive*). Our results hint that *proactive* help creates high *self-threat*, leading to a lower *willingness to accept help*. In addition, we show that if the help is related to a task of high (vs. low) *identity-relevance*, the *self-threat* increases, which decreases users' *willingness to accept help*.

Despite the theoretical and practical contributions based on our findings revealed in this study, we still faced limitations that opens space for further research. First, our study is based on a vignette-based experiment, which mainly focuses on user intentions and perceptions. We would encourage future research to improve generalizability and confirm the robustness of our results by conducting field experiments with actual user behavior, e.g., with traditional medical doctors at their workplaces. It would be very interesting to investigate users' *willingness to accept help* and performance when an IA offers *proactive* (vs. *reactive*) help in an actual performed work-related task. In doing so, this could increase the understanding of human-AI collaboration and thus lead to further implications for design principals to improve the acceptance of help offered by IAs that enhances user performance. Second, our study investigates human-AI collaborations in a work-related context. We recommend testing and validating our findings in contexts beyond. IAs in the form of voice assistants are available to more and more private households. Functions that offer *proactive* help, such as Amazon Alexa's "Latent goal discovery"(Amazon, 2021), could be perceived as more positive and not *self-threatening* in not work-related contexts.

7. Acknowledgements

The authors gratefully acknowledge the funding support by the German Research Foundation (DFG) as part of the project "Helpful/Harmful Aid from Agentic IS" (project number 52002093).

8. References

Adam, M., Roethke, K., & Benlian, A. (2022). Human Versus Automated Sales Agents: Shifting Customer Responses Across Sales Stages. *Information Systems Research, Articles in Advance*, 1-21.

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, *31*(2), 427-445.

Agarwal, R., & Prasad, J. (1998). A conceptual and operational definition of personal innovativeness in the domain of information technology. *Information Systems Research*, *9*(2), 204-215.

Aguinis, H., & Bradley, K. J. (2014). Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organizational Research Methods*, *17*(4), 351-371.

Aignostics. (2022). *Aignostics - A New Era of Pathology*. Retrieved 26.04.2022 from https://aignostics.com/

Amazon. (2021). *Latent Goal Discovery*. Retrieved 9.09.2022 from https://developer.amazon.com/en-US/alexa/alexa-haus/latent-goal-discovery

Ashford, S. J., Blatt, R., & VandeWalle, D. (2003). Reflections on the looking glass: A review of research on feedback-seeking behavior in organizations. *Journal of Management*, *29*(6), 773-799.

Ashforth, B. E. (2016). Distinguished Scholar Invited Essay: Exploring Identity and Identification in Organizations: Time for Some Course Corrections. *Journal of Leadership & Organizational Studies*, *23*(4), 361-373.

Atzmüller, C., & Steiner, P. (2010). Experimental Vignette Studies in Survey Research. *Methodology: European Journal of Research Methods for The Behavioral and Social Sciences*, *6*, 128-138.

Ayyagari, R., Grover, V., & Purvis, R. (2011). Technostress: Technological Antecedents and Implications. *MIS Quarterly*, *35*(4), 831-858.

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.

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS Quarterly*, *45*(3), 1433-1450.

Berger, B., Adam, M., Ruhr, A., & Benlian, A. (2021). Watch Me Improve-Algorithm Aversion and Demonstrating the Ability to Learn. *Business & Information Systems Engineering*, *63*(1), 55-68.

Burke, P. J., & Reitzes, D. C. (1991). An Identity Theory Approach to Commitment. *Social Psychology Quarterly*, *54*(3), 239-251.

Carter, M., & Grover, V. (2015). Me, My Self, and I(T): Conceptualizing Information Technology Identity and Its Implications. *MIS Quarterly*, *39*(4), 931-957.

Cascon-Pereira, R., Chillas, S., & Hallier, J. (2016). Role-meanings as a critical factor in understanding doctor managers' identity work and different role identities. *Social Science & Medicine*, *170*, 18-25.

Cila, N. (2022). Designing Human-Agent Collaborations: Commitment, responsiveness, and support. CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA.

Craig, K., Thatcher, J. B., & Grover, V. (2019). The IT Identity Threat: A Conceptual Definition and Operational Measure. *Journal of Management Information Systems*, *36*(1), 259-288.

Dilsizian, S. E., & Siegel, E. L. (2014). Artificial Intelligence in Medicine and Cardiac Imaging: Harnessing Big Data and Advanced Computing to Provide Personalized Medical Diagnosis and Treatment. *Current Cardiology Reports*, *16*(1).

Feine, J., Gnewuch, U., Morana, S., & Maedche, A. (2019). A Taxonomy of Social Cues for Conversational Agents. *International Journal of Human-Computer Studies*, *132*, 138-161.

Fleming, J. S., & Courtney, B. E. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology*, *46*(2), 404-421.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39-50.

Fuegener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). Will Humans-in-the-Loop Become Borgs? Merits and Pitfalls of Working with AI. *MIS Quarterly*, *45*(3), 1527-1556.

Fuegener, 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.

Gendolla, G. H. E. (1998). Effort as assessed by motivational arousal in identity-relevant tasks. *Basic and Applied Social Psychology*, *20*(2), 111-121.

GitHub. (2021). *GitHub Copilot - Your AI pair programmer*. Retrieved 9.09.2022 from https://github.com/features/copilot

Gregor, S., & Benbasat, I. (1999). Explanations from intelligent systems: Theoretical foundations and implications for practice. *MIS Quarterly*, *23*(4), 497-530.

Hair Jr., J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)* (Vol. Third Edition). Sage Publications.

Harari, D., Parke, M. R., & Marr, J. C. (2021). When Helping Hurts Helpers: Anticipatory versus Reactive Helping, Helper’s Relative Status, and Recipient Self-Threat. *Academy of Management Journal*.

Hayes, A. F. (2022). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Vol. Third Edition). Guilford Press.

Huang, J., Liu, M. Q., & Bowling, N. A. (2015). Insufficient Effort Responding: Examining an Insidious Confound in Survey Data. *Journal of Applied Psychology*, *100*(3).

Ilgen, D. R., & Davis, C. A. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology-an International Review-Psychologie Appliquee-Revue Internationale*, *49*(3), 550-565.

Jussupow, E., Spohrer, K., Heinzl, A., & Gawlitza, J. (2021). Augmenting Medical Diagnosis Decisions? An Investigation into Physicians' Decision-Making Process with Artificial Intelligence. *Information Systems Research*, *32*(3), 713-735.

Kakkar, H., Sivanathan, N., & Pettit, N. C. (2019). The impact of dynamic status changes within competitive rank-ordered hierarchies. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(46), 23011-23020.

Komiak, S. Y. X., & Benbasat, I. (2006). The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly*, *30*(4), 941-960.

Kraus, M., Wagner, N., Callejas, Z., & Minker, W. (2021). The Role of Trust in Proactive Conversational Assistants. *IEEE Access*, *9*, 112821-112836.

Kuehl, N., Schemmer, M., Goutier, M., & Satzger, G. (2022). Artificial intelligence and machine learning. *Electronic Markets*, *32*(4), 2235-2244.

Lapointe, L., & Rivard, S. (2005). A multilevel model of resistance to information technology implementation. *MIS Quarterly*, *29*(3), 461-491.

Leary, M. R., & Baumeister, R. F. (2000). The nature and function of self-esteem: Sociometer theory. *Advances in Experimental Social Psychology*, *32*, 1-62.

Lee, H. W., Bradburn, J., Johnson, R. E., Lin, S.-H. J., & Chang, C.-H. D. (2019). The Benefits of Receiving Gratitude for Helpers: A Daily Investigation of Proactive and Reactive Helping at Work. *Journal of Applied Psychology*, *104*(2), 197-213.

Li, S. S., & Karahanna, E. (2015). Online Recommendation Systems in a B2C E-Commerce Context: A Review and Future Directions. *Journal of the Association for Information Systems*, *16*(2), 72-107.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., & Sollner, M. (2019). AI-Based Digital Assistants Opportunities, Threats, and Research Perspectives. *Business & Information Systems Engineering*, *61*(4), 535-544.

Maedche, A., Morana, S., Schacht, S., Werth, D., & Krumeich, J. (2016). Advanced User Assistance Systems. *Business & Information Systems Engineering*, *58*(5), 367-370.

Martin, G., Bushfield, S., Siebert, S., & Howieson, B. (2021). Changing Logics in Healthcare and Their Effects on the Identity Motives and Identity Work of Doctors. *Organization Studies*, *42*(9), 1477-1499.

Meyer, M. A. (2018). Qualifications and Skills Required for Performance Improvement: A Content Analysis of Job Postings. *Journal of Healthcare Management*, *63*(6), 427-444.

Morana, S., Pfeiffer, J., & Adam, M. T. P. (2020). User Assistance for Intelligent Systems. *Business & Information Systems Engineering*, *62*(3), 189-192.

Morana, S., Schacht, S., Scherp, A., & Maedche, A. (2017). A review of the nature and effects of guidance design features. *Decision Support Systems*, *97*, 31-42.

Nadler, A., & Fisher, J., D. (1986). The Role of Threat to Self-Esteem and Perceived Control in Recipient Reaction to Help: Theory Development and Empirical Validation. *Advances in Experimental Social Psychology*, *19*, 81-122.

Nunnally, J. C. (1978). An Overview of Psychological Measurement. In B. B. Wolman (Ed.), *Clinical Diagnosis of Mental Disorders: A Handbook* (pp. 97-146). Springer.

Pettit, N. C., Yong, K., & Spataro, S. E. (2010). Holding your place: Reactions to the prospect of status gains and losses. *Journal of Experimental Social Psychology*, *46*(2), 396-401.

Pinski, M., Adam, M., & Benlian, A. (2023). AI Knowledge: Improving AI Delegation through Human Enablement. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, Article 25.

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and Artificial Intelligence: An Experiential Perspective. *Journal of Marketing*, *85*(1), 131-151.

Qiu, L. Y., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, *25*(4), 145-181.

Rosenberg, M. (1965). Rosenberg self-esteem scale (RSES). *APA PsycTests*.

Schneider, D., Klumpe, J., Adam, M., & Benlian, A. (2020). Nudging users into digital service solutions. *Electronic Markets*, *30*(4), 863-881.

Seeber, I., Waizenegger, L., Seidel, S., Morana, S., Benbasat, I., & Lowry, P. B. (2020). Collaborating with technology-based autonomous agents Issues and research opportunities. *Internet Research*, *30*(1), 1-18.

Sherman, D. K., & Cohen, G. L. (2006). The psychology of self-defense: Self-affirmation theory. *Advances in Experimental Social Psychology, Vol 38*, *38*, 183-242.

Spector, P. E., & Brannick, M. T. (2011). Methodological Urban Legends: The Misuse of Statistical Control Variables. *Organizational Research Methods*, *14*(2), 287-305.

Spiekermann, S., Krasnova, H., Hinz, O., Baumann, A., Benlian, A., Gimpel, H., Heimbach, I., Koster, A., Maedche, A., Niehaves, B., Risius, M., & Trenz, M. (2022). Values and Ethics in Information Systems A State-of-the-Art Analysis and Avenues for Future Research. *Business & Information Systems Engineering*, *64*(2), 247-264.

Steele, C. M. (1988). The Psychology of Self-Affirmation: Sustaining the Integrity of the Self. *Advances in Experimental Social Psychology*, *21*, 261-302.

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly*, *63*(3), 224-237.

Strich, F., Mayer, A.-S., & Fiedler, M. (2021). What Do I Do in a World of Artificial Intelligence? Investigating the Impact of Substitutive Decision-Making AI Systems on Employees' Professional Role Identity. *Journal of the Association for Information Systems*, *22*(2), 304-324.

Wang, D., Wang, L., Zhang, Z., Wang, D., Zhu, H., Gao, Y., Fan, X., & Tian, F. (2021). Brilliant AI Doctor in Rural Clinics: Challenges in AI-Powered Clinical Decision Support System Deployment. CHI Conference on Human Factors in Computing Systems, Yokohama, Japan.

Wenninger, A., Rau, D., & Roglinger, M. (2022). Improving customer satisfaction in proactive service design A Kano model approach. *Electronic Markets*, *32*(3), 1399-1418.