

Markus, A., Baumann, M., Pfister, J., Carolus, A., Hotho, A., & Wienrich, C. (2025). The impact of persuasion literacy training on anthropomorphization, privacy, and self-determined interactions with intelligent voice assistants. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 19(5), Article 9.
<https://doi.org/10.5817/CP2025-5-9>

The Impact of Persuasion Literacy Training on Anthropomorphization, Privacy, and Self-Determined Interactions With Intelligent Voice Assistants

André Markus¹, Maximilian Baumann², Jan Pfister³, Astrid Carolus², Andreas Hotho³, & Carolin Wienrich¹

¹ Professorship Psychology of Intelligent Interactive Systems, Institute Human-Computer-Media, Julius-Maximilians-University, Wuerzburg, Germany

² Chair of Media Psychology, Institute Human-Computer-Media, Julius-Maximilians-University, Wuerzburg, Germany

³ Chair of Data Science (Informatic X), Julius-Maximilians-University, Wuerzburg, Germany

Abstract

Intelligent voice assistants (IVAs) are widely used, but they also risk influencing users without their knowledge. The anthropomorphic nature of IVAs can tempt users to disclose more personal information and trust IVAs more than intended. Therefore, targeted training is necessary to sensitize users and protect them from the persuasive potential of IVAs. This study developed and evaluated two online training modules designed to promote persuasion literacy among student samples. The modules cover the anthropomorphic characteristics of IVAs, their persuasive power, and potential countermeasures. The results indicate that the training improves understanding of persuasion by IVAs, raises awareness of their persuasive potential, and promotes recognition of persuasion attempts. Additionally, the training modules reduce anthropomorphic perceptions of IVAs and reinforce psychological aspects to ensure privacy and self-determined interaction. The training modules provide an innovative approach to enhancing the competent use of IVAs. As a result, they represent an exciting approach for research and educational institutions to further explore and promote persuasion literacy in the context of IVAs and other AI-related applications. Therefore, this work contributes to enhancing the accessibility of digital educational programs.

Keywords: voice assistant; persuasion; smart speaker; anthropomorphism; digital interaction literacy; e-learning

Editorial Record

First submission received:
November 4, 2024

Revisions received:
July 3, 2025
September 3, 2025

Accepted for publication:
September 3, 2025

Editor in charge:
Alexander P. Schouten

Introduction

Intelligent voice assistants (IVAs), such as Amazon Alexa and Google Assistant, are voice-based AI systems that users utilize to obtain daily information, including weather, traffic, and news. However, they also open up opportunities for companies to deliver persuasive messages (Duggal, 2020; Google, 2024). AI-based persuasion involves generating, supplementing, or modifying a message by communicative AI entities to influence individuals' attitudes, behaviors, and intentions in a targeted manner without their awareness (Dehnert & Mongeau, 2022). In particular, IVAs can influence users at any time and from any location due to their constant presence, which

increases awareness of their persuasive strategies and potential importance (Buijzen et al., 2010; Carolus, Augustin, et al., 2023). One persuasive strategy is “anthropomorphism”, in which technologies are endowed with human-like characteristics to elicit social responses when interacting with users (Guthrie, 1993; Reeves & Nass, 1996). Studies within the “computer are social actors” (CASA) paradigm show that even weak social cues from technologies are sufficient to elicit unconscious social responses from users. IVAs are perceived as particularly human due to their use of language, dialogue capabilities, and human-like names (e.g., Alexa, Siri; Couper et al., 2001; Nass & Brave, 2005; Reeves & Nass, 1996). Studies confirm the persuasive potential of IVAs due to their social and human-like perception (Rzepka et al., 2020; Wienrich et al., 2021). Therefore, the need for specific AI skills to enable people to interact with IVAs self-determinedly and to withdraw from persuasive influences is becoming increasingly important (Carolus, Augustin, et al., 2023; Carolus, Koch, et al., 2023). Previous studies have investigated the effects of educational interventions to counter persuasion in the context of product placement, sponsored influencer posts, and personalized advertising (De Jans et al., 2017; Desimpelaere et al., 2024; Hudders et al., 2016; Lou et al., 2020). However, there is a lack of studies on interventions in the context of IVAs that focus on training approaches that empower users to interact with IVAs self-determinedly and prepare them for their persuasive potential. In addition, many intervention studies in this area rely on commercial content, often only accessible behind a paywall (Eagle, 2007). This work aims to fill this gap by developing and evaluating accessible, non-commercial training interventions that empower young adults, a key IVA user group characterized by high engagement and familiarity (Wardini, 2024), to critically engage with IVAs and prepare for the technologies' persuasive potential and impact. The study addresses two research questions: 1) How can such training approaches be effectively designed, and 2) How does training affect the parameters of competent AI interactions (e.g., persuasion resistance, anthropomorphic perception, privacy-friendly behavior)?

Usage Risk Through IVA Persuasion

IVAs exert a significant persuasive influence on users through various psychological and design mechanisms. The following sections highlight key effects and their implications.

Perceived Credibility and Trust Through Design Cues

Research shows that information delivered via IVAs is perceived as more credible than that presented through graphical interfaces, even without source citation (Gaiser & Utz, 2023). A key design factor contributing to this effect is anthropomorphization: for example, when an IVA is given a human name, users tend to trust its recommendations more and comply with its suggestions, which may lead to uncritical acceptance of advice or commercial offers (Voorveld & Araujo, 2020; Zhao, 2006). Additionally, when IVAs disclose personal details early in an interaction, users reciprocate by sharing more information and placing greater trust in the assistant (Harmsen et al., 2023; Saffarizadeh et al., 2023). This reciprocal dynamic increases the risk of users inadvertently disclosing sensitive personal data in everyday scenarios, such as health or financial discussions, often without a full awareness of associated privacy risks.

Human-Like Voices and the Risk of Unconscious Disclosure

Human-like voices in IVAs can increase the perceived social presence and naturalness of the interaction, which tends to lead users to reveal sensitive information unconsciously (Ischen et al., 2020). In practice, users might share private details simply because the voice “feels human”, creating a social bond similar to talking with a real person. Advances in AI-generated voices further amplify this effect, increasing credibility and persuasive power (Dai et al., 2023). Anthropomorphizing an IVA may also foster emotional bonds that users are often unaware of, which can enhance user engagement but also be exploited for commercial gain, bypassing rational decision-making processes (Singh, 2022; Zhang & Rau, 2023). For instance, a user might develop trust or even affection toward an IVA, which could be used to promote products or services more effectively.

Context Dependency and Psychological Framing

The persuasive effects of IVAs depend heavily on the interaction context. Users tend to prefer synthetic voices for functional tasks, associating them with competence (Im et al., 2023). Conversely, more human-sounding voices are perceived as more appropriate for emotional or social interactions, which enhances the sense of connection (Cho & Sundar, 2022). Moreover, the social role attributed to an IVA influences user responses: when framed as a health expert, IVAs increase users' willingness to disclose sensitive data (Wienrich et al., 2021). Similarly, IVAs

perceived as companions or friends encourage more frequent and personal interactions (Markus et al., 2024; Rhee & Choi, 2020; Wienrich et al., 2023). These findings illustrate how subtle psychological framing can guide user behavior without their explicit awareness, raising concerns about potential manipulation.

Commercial Exploitation and Privacy Concerns

In e-commerce, IVAs can make shopping feel seamless and effortless, which increases the likelihood of impulsive and unreflective purchases (Haas & Keller, 2021; Lim et al., 2023; Rzepka et al., 2020). Voice shopping is growing rapidly and may soon surpass traditional online shopping via PCs and laptops (McLean & Osei-Frimpong, 2019; Tassiello et al., 2021). This ease of use can cause users to impulsively buy recommended products without fully considering price or necessity, especially when IVAs leverage emotional or contextual cues to nudge decisions. Particularly concerning are initiatives like Amazon's patents to monitor users' emotional and physical states (e.g., exhaustion, coughing) for personalizing recommendations (Jin & Wang, 2018). Such practices raise serious issues regarding surveillance, data exploitation, and manipulative marketing. This monitoring could translate into scenarios where targeted ads or product offers are presented precisely when users are most vulnerable, such as during illness or fatigue.

In sum, IVAs employ diverse persuasive techniques that significantly influence user attitudes and behaviors, often fostering trust and facilitating interaction. However, these effects are ethically ambiguous, posing risks to privacy, autonomy, and informed consent. A critical perspective is essential for identifying, mitigating, and regulating these risks while harnessing the benefits of IVA technology.

Countermeasure: Persuasion Literacy

Previous studies have investigated the effects of educational training aimed at countering persuasion in various contexts, such as product placement, sponsored influencer posts, advergames, and personalized advertising (De Jans et al., 2017; Desimpelaere et al., 2024; Hudders et al., 2016; Lou et al., 2020). However, despite their high persuasive potential, IVAs have received little attention in this regard. Carolus, Augustin, et al. (2023) address this gap by introducing the concept of "persuasion literacy"—a competence that enables users to recognize and manage persuasive attempts by IVAs, thereby facilitating more self-determined interactions. Persuasion literacy is embedded within the broader framework of digital interaction literacy, which encompasses a range of skills essential for competent IVA use, including understanding technical capabilities, evaluating associated risks, and preventing misunderstandings (Carolus, Augustin, et al., 2023).

Persuasion literacy is conceptualized as a developing, multidimensional competence that consists of two distinct but interrelated sub-competencies: persuasion knowledge and persuasion understanding (Carolus, Augustin, et al., 2023; Carolus, Koch, et al., 2023). These sub-competencies differ in their depth and function. **Persuasion knowledge** refers to a basic awareness that a message is designed to influence attitudes or behavior. It involves recognizing persuasive intent without necessarily understanding the underlying mechanisms (Kunkel, 2010; Lapierre, 2015). In contrast, **persuasion understanding** involves a more advanced comprehension of how and why persuasive messages work. It includes knowledge about the dynamics, timing, techniques, and contextual factors that shape persuasive communication (Friestad & Wright, 1994). Taken together, these dimensions reflect the broader concept of persuasion literacy as an umbrella term, encompassing both the ability to recognize persuasion (knowledge) and the capacity to analyze and respond to it (understanding) critically. While persuasion knowledge enables detection, persuasion understanding facilitates resistance through deeper insight (Campbell & Kirmani, 2000; Friestad & Wright, 1994). Together, they form the foundation for a critical and reflective engagement with persuasive attempts by IVAs. While persuasion literacy is commonly viewed as a multidimensional competence encompassing persuasion knowledge and understanding (Carolus, Augustin, et al., 2023; Carolus, Koch, et al., 2023), its theoretical conceptualization and operationalization are still evolving in the literature. Nevertheless, building on this framework and prior empirical findings, targeted training to enhance persuasion knowledge and persuasion understanding can serve as an effective countermeasure against persuasive influences. Against this background, we assume:

H1: Conducting persuasion literacy training increases a) persuasion understanding, b) persuasion knowledge, and c) persuasion literacy in the context of IVAs.

Persuasion Literacy and Anthropomorphism

Anthropomorphic cues, defined as features that make non-human agents appear human-like, can increase susceptibility to persuasion (Epley et al., 2007). Such cues in messages or technologies often lead to positively biased product evaluations favoring the communicator or manufacturer (Koo et al., 2019; Touré-Tillery & McGill, 2015; Zhou et al., 2019). To ensure competent and objective interaction, users should be empowered to recognize, differentiate, and critically assess these cues, enabling them to regulate their perception of such influences effectively. The anthropomorphic characteristics of IVAs contribute to their persuasive power (Guthrie, 1993). When people anthropomorphize technologies, they apply their understanding of the concept of 'human' to infer that these technologies possess human-like traits, emotions, or intentions (Epley et al., 2007). The more successful the transfer and application of this concept, for example, through similarities (e.g., language ability and names of IVAs), the stronger and more likely the anthropomorphization (Morewedge et al., 2007). Conversely, the tendency to anthropomorphize decreases with increasing knowledge about non-human beings (Epley et al., 2007). Therefore, increasing knowledge about IVAs related to their overly anthropomorphic influences and design features could help reduce anthropomorphic perceptions. Thus, it is hypothesized:

H2: Conducting persuasion literacy training reduces anthropomorphic perceptions of IVAs.

Persuasion Literacy and Privacy

Internet users are frequently exposed to social engineering attacks, which use persuasive tactics (e.g., eliciting sympathy, creating fake profiles, impersonating employees) to build trust and persuade users to take actions that compromise their privacy (e.g., disclosing sensitive information or bypassing security features; Hage et al., 2020; Hasan et al., 2010). Similarly, IVAs can increase the likelihood of users disclosing personal information if they establish intimacy with users or are perceived as trustworthy experts (Ki et al., 2020; Sah & Peng, 2015; Wienrich et al., 2021). Defending against such attacks requires users to be aware of the potential negative consequences of their actions, which can be challenging because they often lack the necessary awareness and knowledge to protect their privacy (Aïmeur et al., 2019; Masur, 2019). Understanding persuasion tactics helps users resist persuasion attempts more effectively, protects their privacy better, and can reduce trust or self-disclosure in IVAs (Hallam & Zanella, 2017; Mangleburg & Bristol, 2013; Wright et al., 2005; Zhao et al., 2012). Trust is critical when it comes to disclosing or withholding personal information. When we trust someone, we are more likely to share information (Altman, 1975; Mesch, 2012). Conversely, recognizing manipulative or persuasive intentions can damage that trust, leading to greater reluctance to share information, which can benefit privacy (Hallam & Zanella, 2017; Zhao et al., 2012). A deeper understanding of persuasion can also increase perceived control to resist persuasion attempts (Ham et al., 2015). As a result, this enhanced control may translate into greater privacy control, which is understood as an individual's ability to regulate what personal information is shared and under what conditions (Xu et al., 2011). Understanding persuasion processes and tactics can also increase privacy awareness (Schaub et al., 2016), which refers to the recognition of privacy risks and the importance of protecting personal data (Ermakova et al., 2014; Xu et al., 2008). A deeper understanding of how IVAs persuade could increase awareness and interest in protecting one's privacy (Ham, 2017; Lim et al., 2023; Morimoto, 2021; Youn, 2009). Therefore, it is assumed that:

H3: Conducting persuasion literacy training reduces a) trust and b) self-disclosure towards IVAs.

H4: Conducting persuasion literacy training increases a) privacy control and b) privacy awareness concerning IVAs.

Persuasion Literacy and Self-Determined Interaction

As previously stated, digital interaction literacy encompasses the skills necessary for self-determined interaction with IVAs. Self-determined interaction refers to users actively and consciously controlling their engagement by making reflective and informed decisions, rather than reacting automatically or being unknowingly influenced by persuasive tactics (Carolus, Augustin, et al., 2023). This concept centers on enabling users to maintain autonomy through reflection, emotional regulation, and persuasion literacy, key components that collectively support critical and intentional interaction with IVAs. Reflection involves questioning one's usage behavior to individual needs, ethical aspects, and possible risks, and is influenced by existing knowledge concepts (Carolus, Augustin, et al., 2023; Hatlevik, 2012; Kember et al., 2008). The concepts of existing knowledge about persuasion contribute to a more critical questioning and evaluation of persuasive messages (Tutaj & van Reijmersdal, 2012). Emotion-regulating processes, such as indulgence—the deliberate tolerance of one's own negative emotions without

immediate reaction—can enhance the quality of interactions with IVAs (Okada, 2005; Xu & Schwarz, 2009). For example, indulgence helps maintain behavioral control when users experience frustration from misunderstood voice commands (Carolus, Augustin, et al., 2023; Schweitzer et al., 2019). Although indulgence is desirable for IVA interactions, a negative effect of indulgence is assumed for the present persuasion literacy training. When users recognize persuasion or perceive the persuasive intentions as inappropriate, this can lead to negative emotional reactions (e.g., reactance) or defensive attitudes that reduce the risk of being persuaded (Darke & Ritchie, 2007; Fransen, Verlegh, et al., 2015; Friestad & Wright, 1994; Mikołajczak-Degrauwe & Brengman, 2014). From this, it is assumed that:

H5: Conducting persuasion literacy training increases a) reflection and reduces b) indulgence concerning IVAs.

Present Study

Research has shown that persuasion literacy is a necessary AI-related skill to protect against the unnoticed influence of IVAs and to promote self-determined interaction. Therefore, training is needed to promote this ability in the context of IVAs. Studies have shown the potential of digital training to promote skills for competent IVA usage, including functional understanding (Markus et al., 2024). This work aims to transfer this potential to persuasion literacy and investigates how targeted training approaches can be designed to develop and evaluate the first training based on current models of self-determined IVA interaction. This study focuses on young adults, an important and highly engaged user group of IVAs with a high level of technological familiarity, making them a relevant population for investigating persuasion literacy in AI interactions (Wardini, 2024). Based on the Digital Interaction Literacy Model (Carolus, Augustin, et al., 2023), two complementary training modules were developed addressing key dimensions of persuasion literacy: 1) understanding how human-like characteristics of IVAs influence users' perceptions and behaviors, targeting unconscious persuasion effects related to anthropomorphism (Study 1), and 2) detecting specific IVA persuasion tactics to enable conscious identification and defense against persuasive attempts (Study 2). These modules target both psychological influence mechanisms and tactical detection skills, reflecting complementary functional roles within the realm of persuasion literacy. This modular approach allows examination of how improving each dimension individually affects users' competent AI interaction, including anthropomorphism (e.g., perceived sociality), privacy (e.g., trust, awareness), and self-determined interaction (e.g., reflection). Overall, the trainings aim to promote competent IVA use by helping users better assess potential risks, reduce misunderstandings, and improve everyday interactions with AI.

Methods

General Design of Training Modules

The dynamics of social networking sites (SNSs) challenge the privacy management of users. Specifically, users of SNSs are confronted with multiple and invisible audiences, context collapse and a merging between public and private sphere (boyd, 2010). Over the years, much research studied boundary coordination and privacy in the context of SNSs (e.g., boyd & Hargittai, 2010; boyd & Marwick, 2011; Litt, 2013; Marwick & boyd, 2014; Stutzman & Hartzog, 2012; Wisniewski et al., 2012).

The persuasion literacy training modules are based on the Digital Interaction Literacy Model (Carolus, Augustin, et al., 2023). Specific content for the instructional texts was developed through a literature review, including Diederich (2020), Reeves and Nass (1996), and Fogg (1997). Training Module 1, "The Humanity of Technology", explains why people perceive technologies such as IVAs as human or social, and highlights the risks associated with this perception and its use. Training Module 2 "Understanding and Avoiding Persuasion" presents anthropomorphic design principles that can have persuasive influences and introduces user personality traits that are susceptible to anthropomorphism. It also presents coping strategies for recognizing and avoiding persuasion attempts. Further details of each training module can be found in the relevant sections of Study 1 and Study 2.

The training modules were developed using multimedia learning principles (Mayer, 2014) and recommendations from instructional psychology that promote learning, such as the seductive detail effect (Harp & Mayer, 1997), the signaling effect (Mayer & Fiorella, 2014), and the Hamburg comprehensibility concept (Langer et al., 2019). Learning-relevant and decorative images with positive valence were integrated into instructional texts to enhance the learning experience (Schneider et al., 2016; Schneider et al., 2018). The training modules were created using

the e-learning tool Captivate Classic (Adobe Systems, 2019) and had the following structure: 1) welcome screen, 2) presentation of the learning goals, 3) instructional texts, and 4) exercises based on the learning content (e.g., multiple choice, gap-fill texts). Incorrectly solved exercises could be repeated three times. After each exercise, participants viewed a sample solution, regardless of the correctness of their answer, ensuring a standardized level of information across all participants. Excerpts from the training modules are shown in Figure 1, and examples of exercises are shown in Table A1.

Figure 1. Exemplary Excerpts From the Training.



Measures

The following measures were used in both studies to evaluate the training modules. All scales were rated on a five-point Likert scale (1 = *not at all*; 5 = *very much*). Items and sample items of each scale can be found in Tables A2–A5.

To provide a holistic measurement, persuasion literacy was measured and conceptualized using the following three scales (Table A2): To measure participants' **persuasion understanding** in the context of IVAs, two subscales were developed based on the definition of persuasion literacy by Carolus, Augustin, et al. (2023): 1) IVA Human Likeness ($\alpha = .77$; 5 Items) and 2) IVA Persuasion Type ($\alpha = .83$; 5 Items). Tutaj and van Reijmersdal (2012) developed items to measure **persuasion knowledge** (4 items; $\alpha = .83$), conceptualizing it as an understanding of persuasion and sales intentions. We utilized these items and adapted the wording for IVAs ($\alpha = .83$). The **AI Persuasion Literacy** subscale ($\alpha = .75$) of the Meta AI Literacy Scale (Carolus, Koch, et al., 2023) uses three items to measure perceived resistance to, avoidance, and recognition of persuasive influences from AI-based applications. In this work, these items were used and adapted for IVAs.

The Human-Robot Interaction Evaluation Scale by Spatola et al. (2021) was used to measure changes in the **anthropomorphic perception** of IVAs through training ($\alpha = .93$). It consists of the subscales Sociability (items: warm, likable, trustworthy, friendly), Animacy (items: human-like, real, alive, and natural), Agency (items: self-reliant, rational, intentional, intelligent), and Disturbance (creepy, scary, uncanny, weird). Participants rated the extent to which they associated characteristics of the subscales with IVAs

In the context of privacy, changes in privacy-related **trust** in IVAs as a result of the training were measured with three items from Kim et al. (2019; $\alpha = .85$). The willingness to engage in **self-disclosure**—defined as the willingness to disclose personal information—to IVAs was measured using three items developed by Pal et al. (2020; $\alpha = .73$).

The **privacy control** subscale from Xu et al. (2011) was used to measure the effect of training on the feeling of control over one's privacy ($\alpha = .89$; 4 items). Changes in **privacy awareness** were measured using the corresponding subscale from Xu et al. (2008; $\alpha = .87$; 3 items). Sample items for the scales are given in Table A3.

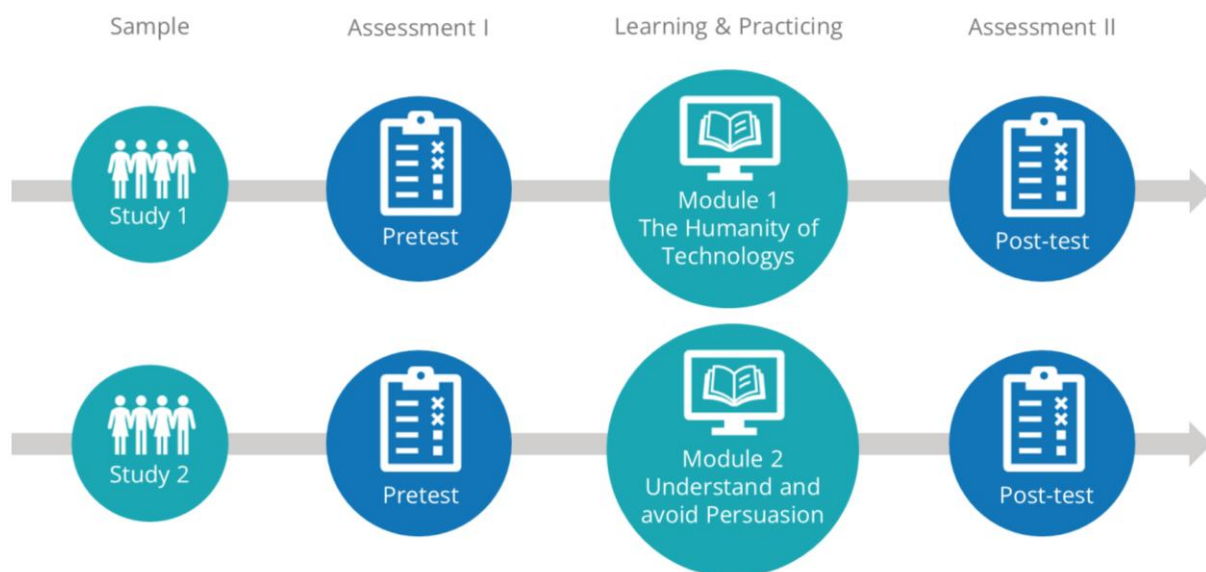
The **Reflection** (8 items; $\alpha = .91$) and **Indulgence** (3 items; $\alpha = .87$) scales developed by Carolus, Augustin, et al. (2023) were used to assess self-determined interaction with IVAs (Table A4). The reflection subscale measures the willingness to reflect on one's use of IVA, its effects, and expectations of the systems. The indulgence subscale measures the willingness to indulge due to negative experiences with IVAs.

Training quality was measured using the Training Evaluation Inventory ($\alpha = .73$), developed by Ritzmann et al. (2020; Table A5). The inventory measures various quality criteria, including Subjective Fun (3 Items), Perceived Usefulness (4 Items), Perceived Difficulty (4 Items), Subjective Knowledge Growth (3 Items), and Attitude towards Training (3 Items). Results served as an assessment of the learning experience and provided insight into possible training optimizations.

Procedure

Figure 2 illustrates the course of the two studies. The training modules were conducted online during April 2024 and lasted approximately 25 minutes. First, participants in the online study consented to data collection and analysis, and they were introduced to IVAs as voice-based AI systems used in everyday applications, such as smartphones and smart speakers. This was followed by scales to measure persuasion literacy, anthropomorphism, privacy, and self-determined interaction (pretest). Then, participants completed the training module. After the training, participants again rated the scales presented in the pretest (post-test). Finally, participants provided demographic details (age and gender), assessed the training quality, and were informed about the purpose of the study.

Figure 2. Illustration of the Experimental Course.



Data Analysis

T-tests for dependent samples were used to analyze changes in the dependent variables between the pre- and post-tests, with a significance level of $\alpha < .05$. Additionally, marginal effects ($p < .10$) were considered to detect subtle changes and trends. Due to the exploratory nature of the training study, the Benjamini-Hochberg correction is used to control for the risk of alpha error and to identify relevant effects without compromising statistical power through overly conservative corrections (Benjamini & Hochberg, 1995; Storey, 2002). The Training Evaluation Inventory (Ritzmann et al., 2020) lacks predefined quality criteria or benchmarks to determine whether a training program can be considered good. In this study, ratings of 3.0 and above were considered satisfactory, corresponding to an above-average rating on a five-point Likert scale.

Study 1: The Humanity of Technologies

Study 1 analyses Training Module 1, “The Humanity of Technologies”, which addresses the first key point for comprehensive persuasion literacy: understanding the persuasive human-like characteristics of technologies such as IVAs and their influence on user perceptions and behaviors.

Materials

Training Module 1, “The Humanity of Technologies”¹, explains why people perceive IVAs and other technologies as human or social and shows the potential influences on perception and use. The module begins with a video of Heider and Simmel (1944), which shows that people anthropomorphize even simple, moving geometric figures. It then introduces the CASA concept (Computers are Social Actors), which explains unconscious social perceptions and reactions to technology (Nass & Moon, 2000). To illustrate the social nature of human-computer interaction, classic social-psychological phenomena that occur in human-human interactions are described, which also manifest in the context of technology (e.g., gender stereotyping, reciprocity). The acquired knowledge is then applied to IVAs, describing their persuasive and anthropomorphic characteristics and highlighting the risks associated with anthropomorphization (e.g., increased trust or self-disclosure). The training includes multiple-choice, gap-filling, and free-text exercises (examples in Table A5).

Participants

A total of 36 participants ($n_{\text{female}} = 33$, $n_{\text{male}} = 3$) with a mean age of 20.78 years ($SD = 1.40$) were recruited from the university's participant pool and compensated with course credits. Participants who answered at least two control questions incorrectly (e.g., “Mark the middle of the scale”) and incomplete data sets were excluded from the data analysis ($n = 2$). Most participants stated that they used IVA less than once a month (never = 13.9%, less than monthly = 33.3%, monthly = 13.9%, 2–3 times a month = 16.7%, weekly = 11.1%, 2–3 times a week = 8.3%, daily = 2.8%).

Results

Table 1 provides an overview of the results. Consistent with the hypothesis, training significantly enhances **persuasion understanding**, both in terms of understanding the human likeness of IVAs and their potential to influence users. Consistent with the hypothesis, **persuasion knowledge** increases, as evidenced by higher awareness of IVAs' persuasion and selling intentions. Contrary to the hypothesis, a significant decrease in **persuasion literacy** is observed following the training. At the item level, participants are less convinced after the training that they can prevent persuasive influences through IVAs. In line with the hypothesis, a significant reduction in perceived agency is found for **anthropomorphism**. There is no training effect on the Sociability, Animacy, and Disturbance subscales. Results regarding the **privacy** hypotheses confirmed that training reduces trust significantly and willingness to disclose personal information marginally significantly. Contrary to the hypothesis, training does not significantly impact privacy control; however, it does increase privacy awareness marginally, as hypothesized. In line with the hypotheses for **self-determined interaction**, the training has a marginally significant and negative effect on Indulgence. However, the training does not affect reflection. Evaluation of the **training quality** shows that the criteria for satisfactory training quality are met: Subjective Fun ($M = 4.08$, $SD = 0.72$), Perceived Usefulness ($M = 3.93$, $SD = 0.73$), Perceived Difficulty ($M = 4.31$, $SD = 0.74$), Subjective Knowledge Growth ($M = 3.67$, $SD = 0.89$) and Attitude towards Training ($M = 3.85$, $SD = 0.94$).

Table 1. *Descriptive and Inferential Statistical Evaluation Results for Training Module 1.*

	Pre		Post		Statistic			Effect Size
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t-value</i>	<i>df</i>	<i>p^a</i>	<i>d</i>
Persuasion Understanding								
Overall Score	3.47	0.64	3.93	0.59	-3.66**	35	.007	-0.61
IVA Human Likeness	3.53	0.70	4.13	0.61	-4.76***	35	.007	-0.79
IVA Persuasion Type	3.40	0.72	3.74	0.70	-2.25*	35	.033	-0.38
Persuasion Knowledge								
Overall Score	2.75	0.84	3.21	1.14	-3.24**	35	.007	-0.54
Persuasion Intention	2.82	0.87	3.25	1.14	-2.97**	35	.009	-0.49
Sales Intention	2.69	0.98	3.17	1.21	-2.90**	35	.009	-0.48
Persuasion Literacy								
Overall Score	3.34	0.52	3.09	0.57	2.65*	35	.015	0.44
Resistance (Item 1)	3.31	0.92	3.19	0.75	0.68	35	.313	0.11
Avoid (Item 2)	3.42	0.91	3.00	0.96	3.10**	35	.009	0.52
Recognize (Item 3)	3.31	1.04	3.08	1.03	1.07	35	.195	0.18
Anthropomorphization								
Sociability	2.68	0.76	2.58	0.82	1.35	35	.132	0.23
Animacy	2.06	0.67	2.16	0.78	-1.16	35	.874	-0.19
Agency	3.64	0.78	3.47	0.80	2.14*	35	.040	0.36
Disturbance	2.75	1.01	2.81	1.18	-0.55	35	.744	-0.09
Privacy								
Trust	2.29	0.67	1.96	0.84	2.95**	35	.009	0.49
Self-Disclosure	2.38	0.80	2.19	0.84	1.91 [†]	35	.058	0.32
Privacy Control	2.23	0.87	2.23	0.90	0.00	35	.588	0.00
Privacy Awareness	2.70	0.68	2.85	0.83	-1.73 [†]	35	.071	-0.29
Self-Determined Interaction								
Reflection	3.61	0.68	3.58	0.73	0.36	35	.711	0.06
Indulgence	3.09	0.94	2.93	0.94	1.77 [†]	35	.071	0.29

Note. ^a corrected *p*-value. [†]*p* < .10. **p* < .05. ***p* < .01. ****p* < .001.

Discussion

Training Module 1, “The Humanity of Technologies”, explains why technologies like IVAs are anthropomorphized and how this affects user perception and behavior. The training enhances understanding and knowledge of persuasion related to IVAs (H1a and H1b confirmed). Surprisingly, training decreases persuasion literacy (H1c not confirmed) and reduces the belief that one can prevent IVA persuasion attempts. The training could highlight the challenge of preventing persuasion, as IVAs can persuade users without their awareness (Nass & Moon, 2000). The decrease in persuasion literacy is not necessarily negative but can serve as an ‘aha’ moment, helping users realize the true impact IVAs can have. However, this recognition does not simultaneously lead to the belief of control, that they can free themselves from it or retain it. This is also evident in the fact that the training did not affect privacy control, although the perception of privacy awareness increased afterward. It should be further investigated here how control beliefs can be strengthened after the “aha” moment. Further, the training reduces anthropomorphic perceptions of IVAs regarding agency (H2 partly confirmed). Expanding knowledge of IVA and highlighting the importance of the human-machine distinction may trigger disruptive processes that reduce anthropomorphic perception, aligning with other research (Epley et al., 2007; Morewedge et al., 2007). Knowledge of the social effects of IVAs does not affect the perceived sociability, animacy, or disturbance of IVAs. This underlines, on the one hand, the assumption of media education theory that it is difficult to overcome the anthropomorphic perception of technologies regardless of contextual factors (e.g., technological competence, level of education, age) and, on the other hand, that there are variables such as individual preferences that are more important for the perceived sociability, animacy and disturbance of IVAs than objective information

(Reeves & Nass, 1996). The training reduces trust and willingness to disclose personal data to IVAs (H3a & H3b confirmed). These results align with previous findings indicating that knowledge of the social effects of IVAs has privacy-friendly effects that can improve coping with persuasion attempts (Hallam & Zanella, 2017; Wright et al., 2005). Training does not have a positive impact on perceived privacy control (H4a not confirmed). A possible reason is that the training focuses on unconscious influences and risks for users but does not provide coping strategies for protecting privacy and managing the persuasive potential of IVAs (topics covered in Training Module 2). Training has a positive impact on privacy awareness (H4b confirmed), motivating individuals to engage with privacy issues and confirming previous research findings (Lim et al., 2023; Schaub et al., 2016). Training does not affect reflection (H5a not confirmed) but reduces indulgence towards IVAs (H5b confirmed). Discovering persuasive potentials may lead to negative attitudes towards IVAs, which could explain the lower indulgence (Fransen, Verlegh, et al., 2015; Friestad & Wright, 1994). Overall, the evaluation of the training quality indicates satisfactory results. In particular, the subjective fun during training and the appropriateness of the learning complexity were rated positively.

In summary, Training Module 1, "The Humanity of Technologies", increases persuasion understanding and knowledge of IVAs and sensitizes people to the difficulty of avoiding persuasive influence. The training reduces anthropomorphic perceptions and indulgence towards IVAs. Additionally, it decreases privacy-related trust in IVAs and the willingness to disclose personal information to them, which can improve privacy protection (Hallam & Zanella, 2017; Tam, 2015; Wright et al., 2005). Ultimately, the training offers a compelling incentive to pay closer attention to privacy issues. The training module can be considered a considerable success, as it marks the first successful training to signal significant progress in raising users' awareness of the persuasive tactics used by IVAs.

Study 2: Understand and Avoid Persuasion

Study 2 analyses Training Module 2, "Understand and Avoid Persuasion", which addresses the second key point of comprehensive persuasion literacy: Promoting the ability to recognize and avoid persuasive attempts by IVAs.

Materials

Training Module 2, "Understand and Avoid Persuasion"², explains 1) which design features of technologies promote anthropomorphization (e.g., similarity, emotion, personalization), 2) how user personality traits influence susceptibility to persuasion through anthropomorphization (e.g., loneliness, attachment style) and 3) which strategies can be used to control persuasion, based on Fransen, Smit, et al. (2015) and Zuwerink Jacks and Cameron (2003). The pedagogical basis for recognizing persuasion attempts is elucidating technical design features and personality traits that promote anthropomorphization. Then, effective persuasion coping strategies were presented. In an interactive exercise, participants analyzed anthropomorphic design features in IVAs via case studies. They then identified potential persuasion tactics and intentions, devising coping strategies based on previously presented defenses. For instance, it has been demonstrated that IVAs can prolong conversations with users by praising them and asking questions (a persuasion technique), which makes them appear more human-like and increases the risk of unconsciously disclosing personal data (persuasion intention; Ischen et al., 2020; Munnukka et al., 2022). Awareness of this anthropomorphic, persuasive power of IVA's dialogue capability could help to counteract such risks. The training then provides exercises to consolidate the acquired knowledge through multiple-choice and drag-and-drop tasks (examples in Table A5).

Participants

The study included 34 participants ($n_{\text{female}} = 27$, $n_{\text{male}} = 7$) with a mean age of 20.91 years ($SD = 2.85$). Participants were recruited from the university's pool of test subjects and compensated with course credits. Participants who answered at least two control questions incorrectly (e.g., 'Mark the middle of the scale') and those with incomplete data sets were excluded from data analysis ($n = 3$). The majority of participants reported using IVAs less frequently (never = 23.5%, less than monthly = 23.5%, monthly = 5.9%, 2–3 times a month = 14.7%, weekly = 14.7%, 2–3 times a week = 11.8%, daily = 5.9%).

Results

Table 2 presents an overview of the results. According to the hypothesis, training results in a significant increase in **understanding of persuasion** related to IVA human likeness and IVA persuasion type. As suggested, training has a positive effect on **persuasion knowledge**, increasing awareness of persuasion and sales intentions arising from IVAs. Contrary to the hypothesis, training has no significant effect on **persuasion literacy**. However, an analysis of the individual items reveals that the training reinforces the ability to recognize persuasion. Regarding the **anthropomorphism** hypothesis, the training has a marginally significant negative effect on sociability and animacy. There are no effects for the subscales Agency and Disturbance. Consistent with the hypotheses on **privacy**, the training shows a significant negative effect on trust in IVAs and positive effects on privacy control and privacy awareness (marginal effect). Contrary to the hypothesis, the training does not affect self-disclosure. Regarding self-determined interaction, the training has a positive effect on reflection, as predicted; however, it does not affect indulgence, contrary to expectations. All quality criteria relating to **training quality** fulfill the defined satisfaction threshold: Fun ($M = 3.99$, $SD = 0.77$), Perceived Usefulness ($M = 3.84$, $SD = 0.98$), Perceived Difficulty ($M = 4.44$, $SD = 0.54$), Subjective Knowledge Growth ($M = 3.83$, $SD = 1.00$) and Attitude towards Training ($M = 3.91$, $SD = 0.73$).

Table 2. Descriptive and Inferential Statistical Evaluation Results for Training Module 2.

	Pre		Post		Statistic			Effect Size
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t-value</i>	<i>df</i>	<i>p</i> ^a	<i>d</i>
Persuasion Understanding								
Overall Score	3.37	0.70	4.08	0.79	-5.31**	33	.003	-0.91
IVA Human Likeness	3.43	0.82	4.22	0.80	-5.49**	33	.003	-0.94
IVA Persuasion Type	3.31	0.78	3.94	0.84	-4.02	33	.003	-0.69
Persuasion Knowledge								
Overall Score	2.98	1.05	3.73	1.13	-4.79**	33	.003	-0.82
Persuasion Intention	3.06	1.09	3.78	1.09	-4.02**	33	.003	-0.69
Sales Intention	2.90	1.19	3.68	1.22	-4.84**	33	.003	-0.83
Persuasion Literacy								
Overall Score	3.30	0.63	3.36	0.65	-0.44	33	.471	-0.08
Resistance (Item 1)	3.29	0.94	3.29	0.87	0.00	33	.588	0.00
Avoid (Item 2)	3.62	0.92	3.35	0.98	1.20	33	.881	0.21
Recognize (Item 3)	3.00	1.10	3.44	0.79	-2.68*	33	.013	-0.46
Anthropomorphization								
Sociability	2.77	0.73	2.60	0.72	1.72 [†]	33	.078	0.30
Animacy	2.07	0.70	1.94	0.73	1.64 [†]	33	.085	0.28
Agency	3.74	0.78	3.71	0.73	0.36	33	.480	0.06
Disturbance	2.65	0.78	2.74	0.99	-0.76	33	.814	-0.13
Privacy								
Trust	2.35	0.70	1.92	0.69	3.42**	33	.003	0.59
Self-Disclosure	2.43	1.03	2.40	0.94	0.26	33	.498	0.05
Privacy Control	2.04	0.66	2.42	0.83	-2.86*	33	.010	-0.49
Privacy Awareness	2.79	0.91	2.96	0.84	-1.84 [†]	33	.067	-0.32
Self-Determined Interaction								
Reflection	3.45	0.82	3.70	0.71	-2.29*	33	.028	-0.39
Indulgence	2.93	0.97	2.85	1.04	0.68	33	.814	0.12

Note. ^a corrected *p*-values. [†] $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Discussion

Training Module 2, “Understand and Avoid Persuasion”, explains the anthropomorphic design features of IVAs, introduces personality traits that promote susceptibility to anthropomorphism, and offers persuasion protection strategies. The training enhances understanding and knowledge of persuasion and IVAs (H1a & H1b confirmed). In addition, the training promotes the belief that one can recognize persuasive attempts by IVAs (H1c partially confirmed). As other studies have demonstrated, the training may have enhanced the ability to identify ulterior motives in persuasion attempts, thereby making it easier to recognize persuasion attempts by IVAs (Campbell & Kirmani, 2000). Training reduces anthropomorphic perceptions of IVAs in terms of sociality and animacy (marginally; H2 confirmed) but not in terms of agency and disturbance. Explicitly explaining why IVAs appear human-like when interacting with users could contribute to weakening the concept of “human”, reducing anthropomorphic tendencies in line with other research (Epley et al., 2007; Morewedge et al., 2007). The results show that training significantly reduces trust in IVAs (H3a confirmed), which may help regulate usage risks, such as over-trust (Hallam & Zanella, 2017; Wienrich et al., 2021). There is no effect on the intention to disclose personal information (H3b is not confirmed). Training increases data protection awareness and perceived privacy control (H4a & b confirmed). In line with previous research, knowing how IVAs can persuade and what counter-strategies individuals can use increases awareness and interest in privacy and its protection (Ham et al., 2015; Lim et al., 2023; Schaub et al., 2016). The training has no effect on indulgence (H5b not confirmed) but has a positive effect on users' intention to reflect on their interaction with IVA, e.g., regarding disadvantages and risks (H5a confirmed). This finding aligns with other studies, which show that the greater the understanding of persuasion, the greater the willingness to question the potential for persuasion (Boerman et al., 2017; Obermiller et al., 2005; van Reijmersdal et al., 2016). The training quality is rated as satisfactory by the participants. In particular, the fun during the training and the comprehensibility of the learning content are positively emphasized.

In summary, Training Module 2, “Understand and Avoid Persuasion”, improves understanding and knowledge of persuasion in IVAs and clarifies their ulterior motives and effects. The training enhances the ability to recognize persuasive attempts by IVAs while reducing anthropomorphic perceptions and trust in them—a crucial contribution to privacy protection (Wienrich et al., 2021). Training enhances privacy awareness and perceived control over one's privacy, underscoring the importance of training in a digitized world.

General Discussion

In this work, two training modules were developed (Training Module 1: “The Humanity of Technologies”; Training Module 2: “Understand and Avoid Persuasion”), which were evaluated for their effectiveness in increasing persuasion literacy and their effects on anthropomorphization, privacy, and self-determined interaction. Most results are consistent with the hypotheses (see Table 3).

Table 3. Overview of the Confirmed (✓) and Unconfirmed (X) Hypotheses of the Study.

Hypotheses			Training Module 1: The Humanity of Technologies	Training Module 2: Understand and Avoid Persuasion
1	Increasing	a) Persuasion Understanding	✓	✓
		b) Persuasion Knowledge	✓	✓
		c) Persuasion Literacy	X	✓ (partly)
2	Reducing	Anthropomorphization	✓ (partly)	✓ (partly)
3	Reducing	a) Trust	✓	✓
		b) Self-Disclosure	✓	X
4	Increasing	a) Privacy Control	X	✓
		b) Privacy Awareness	✓	✓
5	Increasing	a) Reflection	X	✓
	Reducing	b) Indulgence	✓	X

Both persuasion literacy training modules promote understanding and knowledge of IVA persuasion. Compared to Training Module 2, Training Module 1 reduces anthropomorphic agency perceptions of IVAs. High agency associations occur when technologies are associated with unique human-like characteristics and abilities, giving

the impression that they are real entities or persons (Haslam, 2006; Heil, 2019). Explaining the CASA concept in Training Module 1 could help users better understand the emergence of such associations. Training Module 2, on the other hand, addresses the intention-driven and intelligent persuasion tactics of IVAs, thus contributing to maintaining their perceived human-like competence level. Training Module 2 also tended to decrease the social perception of IVAs, suggesting, and consistent with other research, that the realization that IVAs use persuasive tactics could induce negative emotions or attitudes that could decrease their social perception (Darke & Ritchie, 2007; Fransen, Verlegh, et al., 2015; Spatola & Wykowska, 2021). There are further differences in the training effects regarding reducing the intention to disclose personal information (a marginal effect only in Training Module 1) and increasing privacy control (affecting only Training Module 2). The specific effects indicate that the training courses do not only have a general activation effect. Training Module 1 focuses on the risks of use through persuasion, whereas Training Module 2 focuses on defending against persuasion while neglecting awareness of threats. However, emphasizing the risks of usage could be crucial for reducing the disclosure of personal information (Robinson, 2018). In line with previous research, the positive effect of Training Module 2 on privacy control could be explained by acquiring concrete coping strategies that make persuasion controllable, a topic neglected in Training Module 1 (Ham et al., 2015). Other differences in training effects can be seen in increased reflection (Training Module 2 only) and decreased indulgence (Training Module 1 only) when using IVAs. Training Module 1 highlights the persuasive risks of IVAs, which other studies have shown can lead to negative emotional responses and reduced indulgence, decreasing willingness to reflect (Friestad & Wright, 1994; Schmeichel & Tang, 2015). Training Module 2 presents coping strategies that protect against IVA persuasion. Although indulgence scores also decrease, the effect is not significant. The decrease in indulgence may suggest a more neutral emotional response, as participants potentially feel better protected against persuasive influences. Compensation of negative emotions, combined with acquired knowledge about persuasion protection, could encourage greater reflection on IVA use (Desimpelaere et al., 2020; Igartua, 2010).

To summarize, both training modules promote understanding and knowledge of persuasion of IVAs and reduce anthropomorphic perceptions, which is important in the context of self-determined interactions (Carolus, Augustin, et al., 2023). From a data protection perspective, both trainings are essential to promote the questioning of data-related trust in IVAs and tend to stimulate greater privacy awareness. Interestingly, neither of the training modules increases the perceived resistance or avoidance of IVA persuasion. The ability to resist persuasion depends on experience and practice (Friestad & Wright, 1994). Practical exercises with IVA-specific application scenarios may be necessary to improve persuasion skills in these areas. In addition, the training modules have individual strengths. For example, Training Module 1 reduces the intention to disclose and raises awareness of the dangers of persuasion, while Training Module 2 strengthens privacy control and reflection and promotes the perceived ability to recognize persuasive attempts. The synthesis of both training modules forms a holistic training program that develops AI-related skills for self-determined interaction with IVAs. Combining these complementary strengths by integrating both modules into a single training approach is a promising best practice for empowering users. Future research should empirically evaluate the effectiveness of such combined training to inform evidence-based recommendations for user empowerment in AI interactions.

Implications

The training modules have varying effects on levels of persuasion literacy, underscoring the multidimensionality of this construct and highlighting the need for further theoretical clarification regarding its precise structure. On the other hand, the findings indicate that training modules have a multifaceted impact and must be carefully designed to address various dimensions, thereby comprehensively strengthening users' persuasion literacy. These results offer valuable insights for both practice and research aimed at enhancing various facets of persuasion literacy. The current work demonstrates that the developed training modules and applied approach effectively improve users' understanding of persuasion and their skills in interacting with IVAs. For instance, from a data protection perspective, both training modules play a crucial role in fostering critical trust toward IVAs regarding data handling and increasing awareness of privacy issues. Thus, training modules are essential for enabling self-determined interaction. Learners rate the training approach positively, for example, concerning perceived comprehensibility and appropriate complexity, which indicates a successful instructional design—e.g., avoiding the seductive detail effect (Harp & Mayer, 1997)—and that the learning content did not lead to cognitive overload, which promotes positive learning processing processes (Sweller, 2005). As AI-related skills become increasingly important, it is desirable to apply the training approach to other key competencies, such as privacy literacy, algorithm awareness, and handling (Carolus, Augustin, et al., 2023).

The online format of the training modules ensures accessibility to different target groups, promoting inclusion and reducing the digital divide. The tested design and effects of training are based on young, highly educated adults, but the training could also benefit a broader range of groups, including children, professionals, and older users of IVAs, each with varying needs. Young people familiar with technology may require support in critically evaluating the persuasive strategies of IVAs to maintain their autonomy in decision-making. Professionals using IVAs in the workplace can learn to recognize and question persuasive content, while older adults, who may not be tech-savvy, can learn how to protect themselves from manipulative influences. To address the diverse needs of these groups, the training's accessibility should be adapted, such as using simple language for children, ensuring visual accessibility for seniors, or incorporating job-relevant practical examples for professionals. To integrate digital literacy into educational institutions, the developed training modules could be incorporated into curricula and evaluated for effectiveness, suitability, and possible adaptations. The training modules could also be beneficial for older people with physical limitations or visual impairments, as IVAs are particularly important for them in everyday life, for example, to increase their independence (Vieira et al., 2022). The training modules developed could improve the quality of life for these individuals, for example, by enabling them to assess the persuasion risks associated with IVAs more effectively and thus manage them more self-determiningly. The digital transformation presents challenges and opportunities. Therefore, it is crucial to systematically improve digital interaction literacy, including persuasion literacy (Carolus, Augustin, et al., 2023). Training in this area can enhance people's ability to interact with voice-based AI systems in a self-determined manner, reducing misunderstandings, improving their ability to resist persuasion, and protecting their privacy competently. Given the rapid advancement of large-scale language models such as ChatGPT that exhibit increasingly sophisticated and human-like persuasive capabilities, these training programs are particularly timely and valuable. Teachers and trainers can leverage the findings of this study to design effective modules aimed at improving persuasion literacy and fostering autonomous use of intelligent voice assistants. Thus, this work lays the foundation for further development of educational strategies that empower diverse user groups to navigate the evolving digital technology landscape with confidence and critical awareness.

Limitations and Future Directions

The study sample comprises young adults aged 18 to 31, representing the largest segment of active IVA users (Wardini, 2024). However, due to the sample's specific characteristics, including its narrow age range and predominantly female composition, generalizing the findings to broader populations is limited. Future research should include more diverse samples in terms of age, gender, education, technical background, and IVA usage patterns. Considering that gender affects the effectiveness of persuasion, investigating the effects of gender-specific training becomes especially important (Jacks & Lancaster, 2015). Greater heterogeneity in samples would enable a more comprehensive analysis of potential differences in training outcomes across various user groups. A longitudinal study could make an essential contribution to the present work by evaluating the sustainability of training effects. In this context, a follow-up measurement would be particularly valuable for assessing whether observed effects persist over time and for examining changes in interaction behavior (e.g., usage frequency, employed features, privacy settings), which can serve as a basis for evidence-based recommendations for future training design. In addition, a detailed analysis of the individual components (e.g., individual text passages or exercises) within the training modules could be conducted in the future to optimize the training further and adapt it to the individual needs of the participants. The effectiveness of the training in promoting persuasion literacy was measured using subjective measures such as persuasion understanding. Future studies should validate training effects at the behavioral level. Lee et al. (2020) integrated the persuasion principle of reciprocal self-disclosure into a chatbot that interacted with users. The results showed that participants disclosed more personal information if the chatbot had previously disclosed it. Such persuasive principles could be applied to practical interaction scenarios with IVAs to investigate the behavioral intervention effects of the training presented, such as resistance, recognition, and avoidance. Intentions can predict behavior changes and have been used to measure self-determined interaction with subjective measures (Chiaburu et al., 2011; McEachan et al., 2011). However, a discrepancy often exists between intentions and actual behavior, known as the "Intention-Behaviour Gap" (Sheeran, 2002). To analyze the extent to which the training contributes to self-determined IVA interactions, suitable behavioral indicators should be used (e.g., privacy settings, type and amount of personal information disclosed). The Digital Interaction Literacy Model, which encompasses the necessary skills for competent IVA use (Carolus, Augustin, et al., 2023), could serve as a basis for determining these indicators. The results indicate a negative effect of the presented training on indulgence towards IVAs. However, indulgence is desirable for a better

user experience, so specific training should be developed to promote it (Carolus, Augustin, et al., 2023; Schweitzer et al., 2019). For example, learning about how IVAs function helps users grasp the processes involved in handling voice commands (Markus et al., 2024). Such understanding can raise awareness of potential challenges and sources of error in voice processing, which may positively influence users' willingness to indulge (Okada, 2005; Xu & Schwarz, 2009).

Conclusion

The results demonstrate that the developed training modules improve the understanding of persuasion by IVAs and promote the recognition of their persuasive attempts. The training reduces the tendency to anthropomorphize IVAs and reinforces essential aspects of data protection. The present study highlights the importance of persuasion literacy for a self-determined interaction with IVAs, aiming to minimize potential risks and misconceptions. The results of this study represent a significant milestone in promoting and researching persuasion literacy in the use of IVAs. The developed training modules offer new perspectives for competent and safe interaction with IVAs while providing an impetus for promoting responsible use of AI-based technologies in digital education.

Footnotes

¹ Training Module 1 (Study Version): https://motiv.professor-x.de/train/PL1_v1.5JS.zip/index.html

² Training Module 2 (Study Version): https://motiv.professor-x.de/train/PL3_v2.0JS.zip/index.html

Conflict of Interest

The authors have no conflicts of interest to declare.

Use of AI Services

During the preparation of this work, the authors utilized Grammarly and DeepL to enhance the readability and language of individual sentences. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

Authors' Contribution

André Markus: conceptualization, methodology, validation, formal analysis, investigation, writing—original draft, visualization, writing—review & editing. **Maximilian Baumann:** conceptualization, methodology, validation, formal analysis, investigation, visualization, software, writing—review & editing. **Jan Pfister:** conceptualization, software, data curation. **Astrid Carolus:** writing—review & editing, supervision, project administration, funding acquisition. **Andreas Hotho:** resources, project administration, funding acquisition. **Carolin Wienrich:** writing—review & editing, supervision, project administration, funding acquisition.

Acknowledgement

This research project was supported by the Bavarian Research Institute for Digital Transformation (bidt), an institute of the Bavarian Academy of Sciences and Humanities. The authors are responsible for the content of this publication. We thank our research assistants, Jule Puchner, Pascal Jeck, and Johannes Aumiller, for their support in the technical and content-related creation of the training modules.

References

- Adobe Systems. (2019). *Adobe captivate classics* (Version 11) [Computer software]. Adobe Systems.
- Aïmeur, E., Díaz Ferreyra, N., & Hage, H. (2019). Manipulation and malicious personalization: Exploring the self-disclosure biases exploited by deceptive attackers on social media. *Frontiers in Artificial Intelligence*, 2, Article 26. <https://doi.org/10.3389/frai.2019.00026>
- Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, and crowding*. Brooks/Cole Publishing Company.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Boerman, S., Willemsen, L., & Van Der Aa, E. (2017). “This post is sponsored” effects of sponsorship disclosure on persuasion knowledge and electronic word of mouth in the context of Facebook. *Journal of Interactive Marketing*, 38(1), 82–92. <https://doi.org/10.1016/j.intmar.2016.12.002>
- boyd, d. (2010). Social network sites as networked publics: Affordances, dynamics, and implications. In Z. Papacharissi (Ed.), *A networked self* (pp. 47–66). Routledge. <https://doi.org/10.4324/9780203876527-8>
- boyd, d., & Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday*, 15(8). <https://doi.org/10.5210/fm.v15i8.3086>
- boyd, d., & Marwick, A. E. (2011). Social privacy in networked publics: Teens’ attitudes, practices, and strategies. In *A decade in Internet time: Symposium on the dynamics of the Internet and society*. <https://ssrn.com/abstract=1925128>
- Buijzen, M., Van Reijmersdal, E. A., & Owen, L. H. (2010). Introducing the PCMC model: An investigative framework for young people's processing of commercialized media content. *Communication Theory*, 20(4), 427–450. <https://doi.org/10.1111/j.1468-2885.2010.01370.x>
- Campbell, M. C., & Kirmani, A. (2000). Consumers' use of persuasion knowledge: The effects of accessibility and cognitive capacity on perceptions of an influence agent. *Journal of Consumer Research*, 27(1), 69–83. <https://doi.org/10.1086/314309>
- Carolus, A., Augustin, Y., Markus, A., & Wienrich, C. (2023). Digital interaction literacy model–conceptualizing competencies for literate interactions with voice-based AI systems. *Computers and Education: Artificial Intelligence*, 4, Article 100114. <https://doi.org/10.1016/j.caeai.2022.100114>
- Carolus, A., Koch, M., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS - Meta AI Literacy Scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies. *Computers in Human Behavior: Artificial Humans*, 1(2), Article 100014. <https://doi.org/10.1016/j.chbah.2023.100014>
- Chiaburu, D., Oh, I.-S., Berry, C., Li, N., & Gardner, R. (2011). The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, 96(6), 1140–1166. <https://doi.org/10.1037/a0024004>
- Cho, E., & Sundar, S. S. (2022). Should Siri be a source or medium for ads? The role of source orientation and user motivations in user responses to persuasive content from voice assistants. In *CHI conference on human factors in computing systems extended abstracts* (pp. 1–7). Association for Computing Machinery. <https://doi.org/10.1145/3491101.3519667>
- Couper, M., Tourangeau, R., & Steiger, D. (2001). Social presence in web surveys. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 412–417). Association for Computing Machinery. <https://doi.org/10.1145/365024.36530>
- Dai, Y., Lee, J., & Kim, J. W. (2023). Ai vs. human voices: How delivery source and narrative format influence the effectiveness of persuasion messages. *International Journal of Human–Computer Interaction*, 40(24), 8735–8749. <https://doi.org/10.1080/10447318.2023.2288734>

- Darke, P., & Ritchie, R. (2007). The defensive consumer: Advertising deception, defensive processing, and distrust. *Journal of Marketing research*, 44(1), 114–127. <https://doi.org/10.1509/jmkr.44.1.114>
- De Jans, S., Hudders, L., & Cauberghe, V. (2017). Advertising literacy training: The immediate versus delayed effects on children's responses to product placement. *European Journal of Marketing*, 51(11/12), 2156–2174. <https://doi.org/10.1108/EJM-08-2016-0472>
- Dehnert, M., & Mongeau, P. A. (2022). Persuasion in the age of artificial intelligence (AI): Theories and complications of AI-based persuasion. *Human Communication Research*, 48(3), 386–403. <https://doi.org/10.1093/hcr/hqac006>
- Desimpelaere, L., Hudders, L., & Van de Sompel, D. (2020). Knowledge as a strategy for privacy protection: How a privacy literacy training affects children's online disclosure behavior. *Computers in Human Behavior*, 110, Article 106382. <https://doi.org/10.1016/j.chb.2020.106382>
- Desimpelaere, L., Hudders, L., & Van de Sompel, D. (2024). Children's hobbies as persuasive strategies: The role of literacy training in children's responses to personalized ads. *Journal of Advertising*, 53(1), 70–85. <https://doi.org/10.1080/00913367.2022.2102554>
- Diederich, S. (2020). *Designing anthropomorphic conversational agents in enterprises: A nascent theory and conceptual framework for fostering a human-like interaction* (Vol. 102). Cuvillier Verlag.
- Duggal, R. (2020, March 23). *Is voice the next big thing to transform consumer behavior?* Forbes. <https://www.forbes.com/sites/forbescommunicationscouncil/2020/03/23/is-voice-the-next-big-thing-to-transform-consumer-behavior/>
- Eagle, L. (2007). Commercial media literacy: What does it do, to whom—and does it matter? *Journal of Advertising*, 36(2), 101–110. <https://doi.org/10.2753/JOA0091-3367360207>
- Epley, N., Waytz, A., & Cacioppo, J. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. <https://doi.org/10.1037/0033-295X.114.4.864>
- Ermakova, T., Fabian, B., & Zarnekow, R. (2014). Acceptance of health clouds—a privacy calculus perspective. In *Proceedings of the 22nd European Conference on Information Systems (ECIS)*, Article 11. <https://aisel.aisnet.org/ecis2014/proceedings/track09/11/>
- Fogg, B. (1997). *Charismatic computers: Creating more likable and persuasive interactive technologies by leveraging principles from social psychology*. Stanford University.
- Fransen, M. L., Smit, E. G., & Verlegh, P. W. J. (2015). Strategies and motives for resistance to persuasion: An integrative framework. *Frontiers in Psychology*, 6, Article 1201. <https://doi.org/10.3389/fpsyg.2015.01201>
- Fransen, M. L., Verlegh, P. W. J., Kirmani, A., & Smit, E. G. (2015). A typology of consumer strategies for resisting advertising, and a review of mechanisms for countering them. *International Journal of Advertising*, 34(1), 6–16. <https://doi.org/10.1080/02650487.2014.995284>
- Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of Consumer Research*, 21(1), 1–31. <https://doi.org/10.1086/209380>
- Gaiser, F., & Utz, S. (2023). Is hearing really believing? The importance of modality for perceived message credibility during information search with smart speakers. *Journal of Media Psychology: Theories, Methods, and Applications*, 36(2), 93–106. <https://doi.org/10.1027/1864-1105/a000384>
- Google. (2024). *Learn what your Google assistant is capable of*. <https://assistant.google.com/>
- Guthrie, S. E. (1993). *Faces in the clouds: A new theory of religion*. Oxford University Press. <https://doi.org/10.1093/oso/9780195069013.001.0001>
- Haas, M., & Keller, A. (2021). 'Alexa, adv (ert) ise us!': How smart speakers and digital assistants challenge advertising literacy amongst young people. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung*, 43, 19–40. <https://doi.org/10.21240/mpaed/43/2021.07.23.X>
- Hage, H., Aïmeur, E., & Guedidi, A. (2020). Understanding the landscape of online deception. In *Navigating Fake News, Alternative Facts, and Misinformation in a Post-Truth World* (pp. 290–317). IGI Global. <https://doi.org/10.4018/978-1-7998-2543-2.ch014>

- Hallam, C., & Zanella, G. (2017). Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards. *Computers in Human Behavior*, 68, 217–227.
<https://doi.org/10.1016/j.chb.2016.11.033>
- Ham, C.-D. (2017). Exploring how consumers cope with online behavioral advertising. *International Journal of Advertising*, 36(4), 632–658. <https://doi.org/10.1080/02650487.2016.1239878>
- Ham, C.-D., Nelson, M., & Das, S. (2015). How to measure persuasion knowledge. *International Journal of Advertising*, 34(1), 17–53. <https://doi.org/10.1080/02650487.2014.994730>
- Harmsen, W. N., Van Waterschoot, J., Hendrickx, I., & Theune, M. (2023). Eliciting user self-disclosure using reciprocity in human-voicebot conversations. In *Proceedings of the 5th International Conference on Conversational User Interfaces*, Article 50. Association for Computing Machinery. <https://doi.org/10.1145/3571884.3604301>
- Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89(1), 92–102.
<https://doi.org/10.1037/0022-0663.89.1.92>
- Hasan, M., Prajapati, N., & Vohara, S. (2010). Case study on social engineering techniques for persuasion. *International Journal on Applications of Graph Theory in Wireless Ad Hoc Networks and Sensor Networks*, 2(2), 17–23.
<https://doi.org/10.5121/jgraphoc.2010.2202>
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review*, 10(3), 252–264. https://doi.org/10.1207/s15327957pspr1003_4
- Hatlevik, I. K. R. (2012). The theory-practice relationship: Reflective skills and theoretical knowledge as key factors in bridging the gap between theory and practice in initial nursing education. *Journal of Advanced Nursing*, 68(4), 868–877. <https://doi.org/10.1111/j.1365-2648.2011.05789.x>
- Heil, J. (2019). The intentional stance. In J. Heil (Ed.), *Philosophy of mind* (4th ed., pp. 129–144). Routledge.
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2), 243–259. <https://doi.org/10.2307/1416950>
- Hudders, L., Cauberghe, V., & Panic, K. (2016). How advertising literacy training affect children's responses to television commercials versus advergames. *International Journal of Advertising*, 35(6), 909–931.
<https://doi.org/10.1080/02650487.2015.1090045>
- Igartua, J.-J. (2010). Identification with characters and narrative persuasion through fictional feature films. *Communications*, 35(4), 347–373. <https://doi.org/10.1515/comm.2010.019>
- Im, H., Sung, B., Lee, G., & Kok, K. Q. X. (2023). Let voice assistants sound like a machine: Voice and task type effects on perceived fluency, competence, and consumer attitude. *Computers in Human Behavior*, 145, Article 107791. <https://doi.org/10.1016/j.chb.2023.107791>
- Ischen, C., Araujo, T., Voorveld, H., van Noort, G., & Smit, E. (2020). Privacy concerns in chatbot interactions. In A. Følstad, T. Araujo, S. Papadopoulos, E. L.-C. Law, O.-C. Granmo, E. Luger, & P. B. Brandtzaeg (Eds.), *Chatbot research and design: CONVERSATIONS 2019, lecture notes in computer science* (Vol. 11970, pp. 34–48). Springer.
https://doi.org/10.1007/978-3-030-39540-7_3
- Jacks, J. Z., & Lancaster, L. C. (2015). Fit for persuasion: The effects of nonverbal delivery style, message framing, and gender on message effectiveness. *Journal of Applied Social Psychology*, 45(4), 203–213.
<https://doi.org/10.1111/jasp.12288>
- Jin, H., & Wang, S. (2018). *Voice-based determination of physical and emotional characteristics of users*. (U.S. Patent No. 10,096,319). U.S. Patent and Trademark Office. <https://patents.google.com/patent/US10096319B1/en>
- Kember, D., McKay, J., Sinclair, K., & Wong, F. K. Y. (2008). A four-category scheme for coding and assessing the level of reflection in written work. *Assessment & Evaluation in Higher Education*, 33(4), 369–379.
<https://doi.org/10.1080/02602930701293355>
- Ki, C.-W. C., Cho, E., & Lee, J.-E. (2020). Can an intelligent personal assistant (IPA) be your friend? Para-friendship development mechanism between IPAs and their users. *Computers in Human Behavior*, 111, Article 106412.
<https://doi.org/10.1016/j.chb.2020.106412>

- Kim, D., Park, K., Park, Y., & Ahn, J.-H. (2019). Willingness to provide personal information: Perspective of privacy calculus in IoT services. *Computers in Human Behavior*, 92, 273–281. <https://doi.org/10.1016/j.chb.2018.11.022>
- Koo, M., Oh, H., & Patrick, V. M. (2019). From oldie to goldie: Humanizing old produce enhances its appeal. *Journal of the Association for Consumer Research*, 4(4), 337–351. <https://doi.org/10.1086/705032>
- Kunkel, D. (2010). Commentary mismeasurement of children's understanding of the persuasive intent of advertising. *Journal of Children and Media*, 4(1), 109–117. <https://doi.org/10.1080/17482790903407358>
- Langer, I., von Thun, F. S., & Tausch, R. (2019). *Sich verständlich ausdrücken* [Expressing oneself clearly]. Ernst Reinhardt Verlag.
- Lapierre, M. A. (2015). Development and persuasion understanding: Predicting knowledge of persuasion/selling intent from children's theory of mind. *Journal of Communication*, 65(3), 423–442. <https://doi.org/10.1111/jcom.12155>
- Lee, Y.-C., Yamashita, N., Huang, Y., & Fu, W. (2020). "I hear you, I feel you": Encouraging deep self-disclosure through a chatbot. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1–12). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376175>
- Lim, R. E., Sung, Y., & Hong, J. (2023). Online targeted ads: Effects of persuasion knowledge, coping self-efficacy, and product involvement on privacy concerns and ad intrusiveness. *Telematics and Informatics*, 76, Article 101920. <https://doi.org/10.1016/j.tele.2022.101920>
- Litt, E. (2013). Understanding social network site users' privacy tool use. *Computers in Human Behavior*, 29(4), 1649–1656. <https://doi.org/10.1016/j.chb.2013.01.049>
- Lou, C., Ma, W., & Feng, Y. (2020). A sponsorship disclosure is not enough? How advertising literacy intervention affects consumer reactions to sponsored influencer posts. *Journal of Promotion Management*, 27(2), 278–305. <https://doi.org/10.1080/10496491.2020.1829771>
- Mangleburg, T. F., & Bristol, T. (2013). Socialization and adolescents' skepticism toward advertising. *Journal of Advertising*, 27(3), 11–21. <https://doi.org/10.1080/00913367.1998.10673559>
- Markus, A., Pfister, J., Carolus, A., Hotho, A., & Wienrich, C. (2024). Empower the user-the impact of functional understanding training on usage, social perception, and self-determined interactions with intelligent voice assistants. *Computers and Education: Artificial Intelligence*, 6, Article 100229. <https://doi.org/10.1016/j.caeai.2024.100229>
- Marwick, A. E., & boyd, d. (2014). Networked privacy: How teenagers negotiate context in social media. *New media & society*, 16(7), 1051–1067. <https://doi.org/10.1177/1461444814543995>
- Masur, P. K. (2019). Privacy and self-disclosure in the age of information. In P. K. Masur (Ed.), *Situational privacy and self-disclosure: Communication processes in online environments* (pp. 105–129). Springer. https://doi.org/10.1007/978-3-319-78884-5_6
- Mayer, R. (2014). Multimedia instruction. In M. Spector, D. Merrill, J. Elen, & M. Bishop (Eds.), *Handbook of research on educational communications and technology* (4th ed., pp. 385–399). Springer. <https://doi.org/10.1007/978-1-4614-3185-5>
- Mayer, R. E., & Fiorella, L. (2014). 12 principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (Vol. 279–315). Cambridge University Press.
- McEachan, R. R. C., Conner, M., Taylor, N. J., & Lawton, R. J. (2011). Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. *Health Psychology Review*, 5(2), 97–144. <https://doi.org/10.1080/17437199.2010.521684>
- McLean, G., & Osei-Frimpong, K. (2019). Hey Alexa... examine the variables influencing the use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, 99, 28–37. <https://doi.org/10.1016/j.chb.2019.05.009>
- Mesch, G. S. (2012). Is online trust and trust in social institutions associated with online disclosure of identifiable information online? *Computers in Human Behavior*, 28(4), 1471–1477. <https://doi.org/10.1016/j.chb.2012.03.010>

- Mikołajczak-Degrauwe, K., & Brengman, M. (2014). The influence of advertising on compulsive buying—the role of persuasion knowledge. *Journal of Behavioral Addictions*, 3(1), 65–73. <https://doi.org/10.1556/jba.2.2013.018>
- Morewedge, C., Preston, J., & Wegner, D. (2007). Timescale bias in the attribution of mind. *Journal of Personality and Social Psychology*, 93(1), 1–11. <https://doi.org/10.1037/0022-3514.93.1.1>
- Morimoto, M. (2021). Privacy concerns about personalized advertising across multiple social media platforms in Japan: The relationship with information control and persuasion knowledge. *International Journal of Advertising*, 40(3), 431–451. <https://doi.org/10.1080/02650487.2020.1796322>
- Munnukka, J., Talvitie-Lamberg, K., & Maity, D. (2022). Anthropomorphism and social presence in human–virtual service assistant interactions: The role of dialog length and attitudes. *Computers in Human Behavior*, 135, Article 107343. <https://doi.org/10.1016/j.chb.2022.107343>
- Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. MIT press Cambridge.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- Obermiller, C., Spangenberg, E., & MacLachlan, D. L. (2005). Ad skepticism: The consequences of disbelief. *Journal of Advertising*, 34(3), 7–17. <https://doi.org/10.1080/00913367.2005.10639199>
- Okada, E. M. (2005). Justification effects on consumer choice of hedonic and utilitarian goods. *Journal of Marketing research*, 42(1), 43–53. <https://doi.org/10.1509/jmkr.42.1.43.56889>
- Pal, D., Arpniakanondt, C., & Razzaque, M. A. (2020). Personal information disclosure via voice assistants: The personalization–privacy paradox. *SN Computer Science*, 1, Article 280. <https://doi.org/10.1007/s42979-020-00287-9>
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.
- Rhee, C. E., & Choi, J. (2020). Effects of personalization and social role in voice shopping: An experimental study on product recommendation by a conversational voice agent. *Computers in Human Behavior*, 109, Article 106359. <https://doi.org/10.1016/j.chb.2020.106359>
- Ritzmann, S., Hagemann, V., & Kluge, A. (2020). The Training Evaluation Inventory (TEI) – evaluation of training design and measurement of training outcomes for predicting training success. *Vocations and Learning*, 7(1), 41–73. <https://doi.org/10.1007/s12186-013-9106-4>
- Robinson, S. C. (2018). Factors predicting attitude toward disclosing personal data online. *Journal of Organizational Computing and Electronic Commerce*, 28(3), 214–233. <https://doi.org/10.1080/10919392.2018.1482601>
- Rzepka, C., Berger, B., & Hess, T. (2020). Why another customer channel? Consumers' perceived benefits and costs of voice commerce. In *Proceedings of the 53rd Hawaii International Conference on System Sciences*, Article 4. https://aisel.aisnet.org/hicss-53/in/electronic_marketing/4/
- Saffarizadeh, K., Keil, M., Boodraj, M., & Alashoor, T. (2023). “My Name is Alexa. What’s your name?” The impact of reciprocal self-disclosure on post-interaction trust in conversational agents. *Journal of the Association for Information Systems*, 25(3), 528–568. <https://doi.org/10.17705/1jais.00839>
- Sah, Y. J., & Peng, W. (2015). Effects of visual and linguistic anthropomorphic cues on social perception, self-awareness, and information disclosure in a health website. *Computers in Human Behavior*, 45, 392–401. <https://doi.org/10.1016/j.chb.2014.12.055>
- Schaub, F., Marella, A., Kalvani, P., Ur, B., Pan, C., Forney, E., & Cranor, L. F. (2016). Watching them watching me: Browser extensions' impact on user privacy awareness and concern. In *Proceedings of the NDSS Workshop on Usable Security (USEC 2016)*. <https://doi.org/10.14722/usec.2016.23017>
- Schmeichel, B. J., & Tang, D. (2015). Individual differences in executive functioning and their relationship to emotional processes and responses. *Current Directions in Psychological Science*, 24(2), 93–98. <https://doi.org/10.1177/0963721414555178>

- Schneider, S., Dyrna, J., Meier, L., Beege, M., & Rey, G. D. (2018). How affective charge and text–picture connectedness moderate the impact of decorative pictures on multimedia learning. *Journal of Educational Psychology*, 110(2), 233–249. <https://doi.org/10.1037/edu0000209>
- Schneider, S., Nebel, S., & Rey, G. D. (2016). Decorative pictures and emotional design in multimedia learning. *Learning and Instruction*, 44, 65–73. <https://doi.org/10.1016/j.learninstruc.2016.03.002>
- Schweitzer, F., Belk, R., Jordan, W., & Ortner, M. (2019). Servant, friend or master? The relationships users build with voice-controlled smart devices. *Journal of Marketing Management*, 35(7–8), 693–715. <https://doi.org/10.1080/0267257X.2019.1596970>
- Sheeran, P. (2002). Intention—behavior relations: A conceptual and empirical review. *European Review of Social Psychology*, 12(1), 1–36. <https://doi.org/10.1080/14792772143000003>
- Singh, R. (2022). Hey Alexa—order groceries for me –the effect of consumer–VAI emotional attachment on satisfaction and repurchase intention. *European Journal of Marketing*, 56(6), 1684–1720. <https://doi.org/10.1108/EJM-12-2019-0942>
- Spatola, N., Kühnlenz, B., & Cheng, G. (2021). Perception and evaluation in human–robot interaction: The Human–Robot Interaction Evaluation Scale (HRIES)—A multicomponent approach of anthropomorphism. *International Journal of Social Robotics*, 13(7), 1517–1539. <https://doi.org/10.1007/s12369-020-00667-4>
- Spatola, N., & Wykowska, A. (2021). The personality of anthropomorphism: How the need for cognition and the need for closure define attitudes and anthropomorphic attributions toward robots. *Computers in Human Behavior*, 122, Article 106841. <https://doi.org/10.1016/j.chb.2021.106841>
- Storey, J. D. (2002). A direct approach to false discovery rates. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 64(3), 479–498. <https://doi.org/10.1111/1467-9868.00346>
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge University Press. <https://doi.org/10.1017/CBO9780511816819.003>
- Tam, K.-P. (2015). Are anthropomorphic persuasive appeals effective? The role of the recipient's motivations. *British Journal of Social Psychology*, 54(1), 187–200. <https://doi.org/10.1111/bjso.12076>
- Tassiello, V., Tillotson, J. S., & Rome, A. S. (2021). “Alexa, order me a pizza!”: The mediating role of psychological power in the consumer–voice assistant interaction. *Psychology & Marketing*, 38(7), 1069–1080. <https://doi.org/10.1002/mar.21488>
- Touré-Tillery, M., & McGill, A. L. (2015). Who or what to believe: Trust and the differential persuasiveness of human and anthropomorphized messengers. *Journal of Marketing*, 79(4), 94–110. <https://doi.org/10.1509/jm.12.0166>
- Tutaj, K., & van Reijmersdal, E. A. (2012). Effects of online advertising format and persuasion knowledge on audience reactions. *Journal of Marketing Communications*, 18(1), 5–18. <https://doi.org/10.1080/13527266.2011.620765>
- van Reijmersdal, E. A., Fransen, M. L., Van Noort, G., Oprea, S. J., Vandeberg, L., Reusch, S., Van Lieshout, F., & Boerman, S. C. (2016). Effects of disclosing sponsored content in blogs: How the use of resistance strategies mediates effects on persuasion. *American Behavioral Scientist*, 60(12), 1458–1474. <https://doi.org/10.1177/0002764216660141>
- Vieira, A. D., Leite, H., & Lachowski Volochchuk, A. V. (2022). The impact of voice assistant home devices on people with disabilities: A longitudinal study. *Technological Forecasting and Social Change*, 184, Article 121961. <https://doi.org/10.1016/j.techfore.2022.121961>
- Voorveld, H. A. M., & Araujo, T. (2020). How social cues in virtual assistants influence concerns and persuasion: The role of voice and a human name. *Cyberpsychology, Behavior, and Social Networking*, 23(10), 689–696. <https://doi.org/10.1089/cyber.2019.0205>
- Wardini, J. (2024). *Voice search statistics: Smart speakers, voice assistants, and users in 2024*. Serpwatch. <https://serpwatch.io/blog/voice-search-statistics/>

- Wienrich, C., Carolus, A., Markus, A., Augustin, Y., Pfister, J., & Hotho, A. (2023). Long-term effects of perceived friendship with intelligent voice assistants on usage behavior, user experience, and social perceptions. *Computers*, 12(4), Article 77. <https://doi.org/10.3390/computers12040077>
- Wienrich, C., Reitelbach, C., & Carolus, A. (2021). The trustworthiness of voice assistants in the context of healthcare investigating the effect of perceived expertise on the trustworthiness of voice assistants, providers, data receivers, and automatic speech recognition. *Frontiers in Computer Science*, 3, Article 685250. <https://doi.org/10.3389/fcomp.2021.685250>
- Wisniewski, P., Lipford, H. R., & Wilson, D. (2012). Fighting for my space: Coping mechanisms for SNS boundary regulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 609–618). <https://doi.org/10.1145/2207676.2207761>
- Wright, P., Friestad, M., & Boush, D. M. (2005). The development of marketplace persuasion knowledge in children, adolescents, and young adults. *Journal of Public Policy & Marketing*, 24(2), 222–233. <https://doi.org/10.1509/jppm.2005.24.2.222>
- Xu, H., Dinev, T., Smith, H. J., & Hart, P. (2008). Examining the formation of individual's privacy concerns: Toward an integrative view. In *Proceedings of the International Conference on Information Systems (ICIS 2008)*, Article 6. <https://aisel.aisnet.org/icis2008/6>
- Xu, H., Dinev, T., Smith, J., & Hart, P. (2011). Information privacy concerns: Linking individual perceptions with institutional privacy assurances. *Journal of the Association for Information Systems*, 12(12), 1. <https://doi.org/10.17705/1jais.00281>
- Xu, J., & Schwarz, N. (2009). Do we really need a reason to indulge? *Journal of Marketing Research*, 46(1), 25–36. <https://doi.org/10.1509/jmkr.46.1.25>
- Youn, S. (2009). Determinants of online privacy concern and its influence on privacy protection behaviors among young adolescents. *Journal of Consumer Affairs*, 43(3), 389–418. <https://doi.org/10.1111/j.1745-6606.2009.01146.x>
- Zhang, A., & Rau, P.-L. P. (2023). Tools or peers? Impacts of anthropomorphism level and social role on emotional attachment and disclosure tendency towards intelligent agents. *Computers in Human Behavior*, 138, Article 107415. <https://doi.org/10.1016/j.chb.2022.107415>
- Zhao, L., Lu, Y., & Gupta, S. (2012). Disclosure intention of location-related information in location-based social network services. *International Journal of Electronic Commerce*, 16(4), 53–90. <https://doi.org/10.2753/JEC1086-4415160403>
- Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media & Society*, 8(3), 401–419. <https://doi.org/10.1177/1461444806061951>
- Zhou, X., Kim, S., & Wang, L. (2019). Money helps when money feels: Money anthropomorphism increases charitable giving. *Journal of Consumer Research*, 45(5), 953–972. <https://doi.org/10.1093/jcr/ucy012>
- Zuwerink Jacks, J., & Cameron, K. A. (2003). Strategies for resisting persuasion. *Basic and Applied Social Psychology*, 25(2), 145–161. https://doi.org/10.1207/S15324834BASP2502_5

Appendix

Table A1. Extracts of Exercises Sorted by Training Modules. Answers Marked With a Star (*) are Correct.

Module	Exercise Type	Question	Answer Options
1	MC	What risks can arise for users if voice assistants appear human to them in the interaction?	a) Answers from voice assistants are unconsciously perceived as more credible.* b) The promotion of impulsive online shopping behavior.* c) Unconscious disclosure of personal information and data.* d) Encouraging the user to become dependent on the voice assistant.
1	MC	What effect could a male and female voice of a navigation device have on the evaluation of navigation competence?	a) The female navigation system is less trusted when determining the route.* b) The female navigation system is considered to have greater navigational competence. c) The male navigation system's navigational competence is questioned more. d) The male navigation system is seen as having less doubt when determining the route.*
2	MT	Match the following anthropomorphic design features to their description. - Similarity - Emotion - Personalization - Language	a) A robot uses human gestures to signalize production workers "good" or "bad". / <i>Similarity*</i> b) A voice assistant has 23 different speech styles and greets the user when they come home. / <i>Language*</i> c) A chatbot communicates that it is annoyed by the user's rude spelling. / <i>Emotion*</i> d) A smartphone learns the user's sleeping habits and always sets the alarm clock at the right time / <i>Personalization*</i>
2	MC	When asked for a recommendation for a cough medicine, the voice assistant recommends a specific preparation. This is followed by further questions about health status. What persuasive intention could the voice assistant be pursuing with this?	a) The more precisely the voice assistant knows the health situation, the more targeted the advertising can be personalized*. b) The voice assistant passes the information on to the manufacturer of the medication in order to check the effectiveness of the product. c) With more knowledge about the person's state of health, the voice assistant can recommend other products in the future*. d) The voice assistant wants to know more about the person's state of health in order to make holistic recommendations for a healthy lifestyle.

Note. MC = Multiple Choice, MT = Matching Task.

Table A2. *Developed Items, Scales, and Subscales (in Bold) for Persuasion in German and English.*

Scales / Subscales	Item German (English)	
Persuasion Understanding (α = .87)		
IVA Human Likeness	1	Ich verstehe, dass Sprachassistenten unbewusst menschlich wirken können. (I understand that voice assistants can unconsciously appear human.)
	2	Ich verstehe, aus welchen Gründen Sprachassistenten menschlich wirken können. (I understand the reasons why voice assistants can seem human.)
	3	Ich verstehe, wie Sprachassistenten mich durch ihre menschliche Wirkung beeinflussen können. (I understand how voice assistants can influence me through their human impact.)
	4	Ich verstehe, dass ich Sprachassistenten unbewusst menschlich behandeln könnte. (I understand that I might unconsciously treat voice assistants as human.)
	5	Ich verstehe, dass Sprachassistenten mich durch ihre sozialen Merkmale unbemerkt manipulieren können. (I understand that voice assistants can manipulate me unnoticed through their social characteristics.)
IVA Persuasion Type	1	Ich verstehe, dass Sprachassistenten mich unbewusst zu Dingen verleiten können, die ich eigentlich nicht möchte. (I understand that voice assistants can unconsciously tempt me to do things that I don't actually want to do.)
	2	Ich verstehe, dass Sprachassistenten Einfluss auf meine Entscheidungen, Einstellungen und Verhalten nehmen können. (I understand that voice assistants can influence my decisions, attitudes, and behavior.)
	3	Ich verstehe, dass Sprachassistenten beeinflussen können, was ich glaube. (I understand that voice assistants can influence what I believe.)
	4	Ich verstehe, dass Sprachassistenten beeinflussen können, was ich fühle. (I understand that voice assistants can influence how I feel.)
	5	Ich verstehe, dass Sprachassistenten beeinflussen können, was ich tue. (I understand that voice assistants can influence what I do.)
Persuasion Knowledge (α = .83)		
Persuasion Intention	1	Das Ziel von Sprachassistenten ist es, meine Meinung zu beeinflussen. (The aim of voice assistants is to influence my opinion.)
	2	Das Ziel von Sprachassistenten ist es, Menschen dazu zu bringen, bestimmte Produkte zu mögen. (The aim of voice assistants is to get people to like certain products.)
Sales Intention	3	Das Ziel von Sprachassistenten ist es, Produkte zu verkaufen. (The aim of voice assistants is to sell products.)
	4	Das Ziel von Sprachassistenten ist es, den Verkauf von Produkten zu fördern. (The aim of voice assistants is to promote the sale of products.)
AI Persuasion Literacy (α = .75)		
	1	Ich lasse mich in meinen Entscheidungen nicht durch Sprachassistenten beeinflussen. (I don't let voice assistants influence my decisions.)
	2	Ich kann verhindern, dass Sprachassistenten mich in meinen Entscheidungen beeinflussen. (I can prevent voice assistants from influencing my decisions.)
	3	Ich merke, wenn ich in meinen Entscheidungen von Sprachassistenten beeinflusst werde. (I notice when my decisions are influenced by voice assistants.)

Table A3. Sample and Adapted Privacy Scales Items in German (Used in This Study) and English (Original).

Privacy Scales	Sample Item German (English)	Author
Trust	Sprachassistenten würden kompetent mit persönlichen Informationen umgehen. (<i>Voice assistants would handle personal information in a competent fashion.</i>)	Kim et al. (2019)
Disclose Personal Information	Ich bin bereit, Sprachassistenten meine persönlichen Daten preiszugeben. (<i>I am willing to disclose voice assistants my personal information.</i>)	Pal et al. (2020)
Privacy Control	Ich glaube, dass ich die Kontrolle darüber habe, wie persönliche Informationen von Sprachassistenten verwendet werden. (<i>I believe I have control over how personal information is used by voice assistants.</i>)	Xu et al. (2011)
Privacy Awareness	Ich verfolge die Nachrichten und Entwicklungen zu Datenschutzfragen und Datenschutzverletzungen im Zusammenhang mit Sprachassistenten. (<i>I follow the news and developments about the privacy issues and privacy violations in the context of voice assistants.</i>)	Xu et al. (2008)

Table A4. Items from the Reflection and Indulgence Scale in German and English.

Scale	Item	Item German (English)
Instruction		In Zukunft... / In the future...
Reflection	1	...interagiere ich mit Sprachassistenten mit Bedacht. (<i>...I interact with voice assistants carefully.</i>)
	2	...reflektiere ich den Umgang mit Sprachassistenten. (<i>...I reflect on my interaction with voice assistants.</i>)
	3	...wäge ich die Stärken und Schwächen von Sprachassistenten ab. (<i>...I weigh the strengths and weaknesses of voice assistants.</i>)
	4	...denke ich darüber nach, was Sprachassistenten leisten können und was nicht. (<i>...I think about what voice assistants can and cannot do.</i>)
	5	...hinterfrage ich meine Erwartungen an Sprachassistenten. (<i>...I question my expectations of voice assistants.</i>)
	6	...überdenke ich meine Ansprüche an Sprachassistenten. (<i>...I rethink my expectations of voice assistants.</i>)
	7	...denke ich über die Gefahren und Risiken in Bezug auf Sprachassistenten nach. (<i>...I think about the dangers and risks in relation to voice assistants.</i>)
	8	...denke ich über Potenziale und Vorteile in Bezug auf Sprachassistenten nach. (<i>...I think about the potential and advantages of voice assistants.</i>)
Indulgence	1	...nehme ich Rücksicht auf Sprachassistenten, wenn sie nicht wie angenommen funktionieren. (<i>...I will be considerate of voice assistants when they do not work as expected.</i>)
	2	...interagiere ich geduldig, wenn Sprachassistenten anders reagieren als erwartet. (<i>...I interact patiently when voice assistants react differently than expected.</i>)
	3	...bin ich nachsichtig, wenn Sprachassistenten Fehler machen. (<i>...I am indulgent when voice assistants make mistakes.</i>)

Table A5. Sample and Adapted Items of the Training Evaluation Inventory by Ritzmann et al. (2020).

Training Quality	Sample Item German (English)
Subjective Fun	Das Training hat Spaß gemacht. (<i>The learning was enjoyable.</i>)
Perceived Usefulness	Das Training hat für mich einen persönlichen Wert. (<i>I find personal value in this training.</i>)
Perceived Difficulty	Der Inhalt war leicht zu verstehen. (<i>The content was easy to comprehend.</i>)
Subjective Knowledge Growth	Ich denke, dass ich mir die neuen Informationen gut merken werde. (<i>I believe I will retain the new information well.</i>)
Attitude towards Training	Ich würde dieses Training meinen Freunden und Kollegen empfehlen. (<i>I would recommend this training to my friends or colleagues.</i>)

About Authors

André Markus is a PhD student in the Psychology of Intelligent Interactive Systems group at the University of Würzburg. His research focuses on the development and psychometric assessment of AI literacy, user experience, and the psychology of video games.

<https://orcid.org/0000-0003-0632-6397>

Maximilian Baumann worked as a research assistant at the Chair of Media Psychology at the University of Würzburg. His research focuses on the training and assessment of AI-related skills, trust in AI, and interaction paradigms between humans and digital entities.

<https://orcid.org/0009-0004-8042-4555>

Jan Pfister is a PhD student at the Chair for Data Science at the University of Würzburg. He is researching new methods in natural language processing, particularly aspect-based sentiment analysis, using large language models and pointer networks.

<https://orcid.org/0009-0006-8071-635X>

Astrid Carolus is an academic advisor at the Chair of Media Psychology at the University of Würzburg. Her research adopts a psychological perspective on human interaction with media and digital technologies, particularly AI-based systems, focusing on the conditions for competent, autonomous, and reflective use, as well as the development of innovative teaching and learning settings that foster AI skills and an AI mindset.

<https://orcid.org/0000-0003-2206-2715>

Andreas Hotho is a professor at the University of Würzburg, leading the Chair for Data Science and the Center for Artificial Intelligence and Data Science. His research focuses on developing machine learning algorithms for environmental data and text mining, combining large language models with knowledge graphs for applications such as sentiment analysis and plot development in novels.

<https://orcid.org/0000-0002-0483-5772>

Carolyn Wienrich is a professor at the University of Würzburg and head of the Psychology of Intelligent Interactive Systems. Her research focuses on human cognition, emotion, and behavior in digital interactions, emphasizing multimodal interaction paradigms, virtual reality, and change experiences during and after digital interventions, as well as their psychological and societal implications.

<https://orcid.org/0000-0003-3052-7172>

✉ Correspondence to

André Markus, Professorship Psychology of Intelligent Interactive Systems, Institute Human-Computer-Media, Julius-Maximilians-University, Emil-Fischer-Straße 50, 97074 Würzburg, Germany, andre.markus@uni-wuerzburg.de

© Author(s). The articles in Cyberpsychology: Journal of Psychosocial Research on Cyberspace are open access articles licensed under the terms of the [Creative Commons BY-SA 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/) which permits unrestricted use, distribution and reproduction in any medium, provided the work is properly cited and that any derivatives are shared under the same license.