



Article

Crossing the AI Mind Trap: Deconstructing the Impact of AI Mind Perception on User Engagement in AI Voice Assistants

Qian Hu ¹  and Zhao Pan ^{2,*}

¹ School of Management, Wuhan Textile University, Wuhan 430200, China; huq@wtu.edu.cn

² School of Management, Huazhong University of Science and Technology, Wuhan 430074, China

* Correspondence: victola.pz@gmail.com

Abstract

As large language models (LLMs) advance the machine mind of artificial intelligence (AI) to be more human-like, the dynamic impact of users' perceptions of this mind on their engagement remains underexplored. Specifically, little is known about the "tipping point" where appreciation shifts to aversion. Grounded in mind perception theory, this research investigates the dynamic relationship between the perceived machine mind of AI voice assistants and user engagement using data from online reviews and experiments. Study 1 reveals a significant inverted U-shaped relationship, indicating that a moderate level of human-like machine mind maximizes user engagement, while excessive anthropomorphism triggers aversion—a form of the "uncanny valley" effect. Study 2 further dissects the differential impacts of the two core dimensions of the mind: the agentic mind (i.e., the capacity for thinking and planning) and the experiential mind (i.e., the capacity for feeling and emotion). Our findings reveal the comparative effects of AI agentic and experiential minds on the inverted U-shaped relationship, showing that a high experiential mind attenuates the positive effect of the agentic mind on user engagement. These findings deepen our understanding of mind perception effects in human–AI interaction and offer critical practical insights for AI designers on optimizing mind simulation to avoid user alienation.

Keywords: mind perception; uncanny valley; algorithm aversion; AI-based voice assistants; user engagement

1. Introduction

Large language models (LLMs) have driven artificial intelligence (AI) artifacts to perform human abilities like learning, adaptability, and empathy, leading humans to attribute minds to AI-based voice assistants. Today, manufacturers are embedding these highly capable AI voice assistants into an array of smart hardware devices (e.g., smart screens and speakers), expecting to drive significant innovation across the industry [1]. According to Grand View Research [1], the market for AI-based voice assistants in the United States was estimated at USD 23.72 billion in 2024 and is expected to grow at a compound annual growth rate of 23.4% from 2025 to 2030. However, user retention and active utilization in AI-driven marketplaces have not fully lived up to expectations [2]. Little is known about whether the highly human-like minds of AI-based voice assistants facilitate achieving expected market outcomes.

Prior studies of customer responses to AI have largely focused on the pre-engagement stage [3], emphasizing initial adoption intentions [2], upfront algorithm aversion [4], and



Academic Editor: Albert Y.S. Lam

Received: 21 May 2026

Revised: 16 July 2026

Accepted: 16 July 2026

Published: 20 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

early acceptance factors [3]. However, understanding the long-term success of these technologies requires shifting the focus to post-adoption user engagement [5]. User engagement captures the continuous, dynamic interaction between users and AI voice assistants rather than mere surface-level usage behavior [6]. Investigating this post-adoption phase is critical for capturing the dynamic user behaviors necessary for the market's ultimate success [7].

However, existing research on how the machine mind affects this ongoing engagement remains fragmented and often contradictory, generally falling into two distinct streams. The first stream assumes a linear, positive perspective, suggesting that anthropomorphic cues and mind perception in AI build positive outcomes, such as increased trust [8] and positive attitudes [7]. In contrast, an emerging second stream suggests a more complex dynamic, indicating that highly autonomous or human-like minds can actually evoke mixed emotions or algorithm aversion—a phenomenon loosely associated with the uncanny valley of machine minds [9,10]. For example, Shank et al. [11] found that machine minds induce mixed emotions. Stein et al. [10] suggested that autonomous technology with a higher machine mind can evoke stronger aversion from users. Despite these conflicting streams, empirical research investigating a potential nonlinear relationship between the mind of AI-based voice assistants and ongoing user engagement remains relatively scarce [12].

Reconciling these streams requires an understanding of the psychological mechanisms driving this interaction. As an AI voice assistant demonstrates an increasingly human-like mind, the interaction can cross a critical threshold. Specifically, when users perceive the AI's mind as rivaling or exceeding their own, it introduces a complex human–AI conflict [13]. A highly advanced machine mind may inadvertently trigger psychological friction, such as an implicit identity threat or a perceived loss of user control [14], as users feel their autonomy and uniqueness are compromised by an overly capable entity [15]. Consequently, this psychological discomfort triggers a negative behavioral shift, making users less likely to fully engage with the assistant [7]. Therefore, the first question proposed by this study is as follows: *How do the machine minds of AI-based voice assistants nonlinearly affect user engagement?*

Furthermore, mind perception theory categorizes the ascription of minds into two distinct forms: the agentic mind (the ability to plan and act) and the experiential mind (the ability to feel emotions) [16,17]. The agentic and experiential minds of machines can cause AI aversion [18]. For instance, Gray and Wegner [9] suggested that the experiential mind of machines leads to AI aversion. Recent studies similarly show that transforming an AI service provider purely into an emotional experiencer can sometimes backfire [19]. However, their comparative effects in an interactive flow remain underexplored. Thus, it remains unclear how the agentic and experiential dimensions comparatively drive or mitigate user engagement within a nonlinear relationship. This leads to the second question: *Which specific dimension of the machine mind—agentic or experiential—ultimately pushes the user past the tipping point into aversion during ongoing interactions?*

To address the questions, this study makes several contributions. First, by drawing on the literature regarding the uncanny valley effect of machine minds, this study investigates the nonlinear relationship between the machine mind of AI-based voice assistants and user engagement. This exploration provides novel insights into the diffusion of AI-based voice assistants embedded within smart hardware in the context of Industry 4.0. Second, leveraging mind perception theory, this study examines the underlying mechanisms that foster user engagement with AI-based voice assistants, enriching the theoretical literature. Finally, this study, in terms of the comparable effects of the agentic and experiential dimensions of the machine mind, further explores the mechanisms by which the machine mind affects user engagement with AI-based voice assistants. The findings can enhance our understanding of the uncanny valley effect of machine minds on user engagement,

which can guide developers in formulating strategies for the development of AI-based voice assistants.

2. Conceptual Framework and Hypothesis Development

2.1. *The Uncanny Valley: From Appearance to Mind*

The uncanny valley is a concept proposed by Mori [20] that posits a complex, nonlinear dynamic between the degree of human-likeness in robotic designs and the affection users feel toward such robots. That is, as robots become more anthropomorphic, moving from less human-like to more human-like, the attraction of humans generally increases. However, when robots closely resemble humans, this trend reverses, and a pronounced aversion sets in. As technological advancements continue to reshape the AI market, human-like AI has transcended mere visual appearance to encompass a spectrum of human-like attributes [21]. For example, advanced natural language processing technology enables AI (e.g., conversational AI agents) to engage in more nuanced and contextually relevant conversations with humans [22]. Due to its enhanced autonomous technology, AI, such as in autonomous vehicles, possesses a greater degree of autonomy, similar to human agency, and can perform tasks with minimal human intervention [23]. Additionally, advancements in machine learning, such as LLMs, have endowed AI with the capacity for learning and self-improvement [24]. Moreover, incorporating emotional intelligence into AI can enable it to recognize and respond to human emotions, which can result in more empathetic and engaging interactions [22]. The existing research has shown that humans tend to adopt an anthropomorphic approach in response to unknown entities [25]. As a result, these cognitive activities of AI, like the human mind, lead humans to attribute machine minds to AI, which in turn helps humans to make decisions [26].

The current human-likeness of artifacts has shifted from appearance to mind. Furthermore, AI features—its design, functionality, and interactive capabilities—fundamentally shape users' perceptions of AI, generating a range of cognitive, emotional, and behavioral responses [18]. However, few studies have investigated individual responses to machine minds. The research on the diffusion of AI-based voice assistants has found positive impacts of mind perception on user attitudes and behaviors. For example, Waytz et al. [27] examined how machine minds can enhance trust in an autonomous vehicle. Lee et al. [28] argued that minds in chatbots can positively influence human willingness to use them. Zhou et al. [29] found that the presence of mind features, such as voice, interactivity, and empathy cues in AI voice assistants, can have positive impacts on continued use intention. Following these studies, the machine mind can contribute to positive outcomes in the process of the diffusion of AI-based voice assistants.

However, Gray and Wegner [9] found that the uncanny valley effect is not only induced by a human-like appearance but also by machine minds. Stein et al. [10] also point to the uncanny valley of the mind in the context of autonomous technology, in which higher overall machine minds lead to stronger aversion. The uncanny valley of machine minds may occur in the process of user engagement with AI-based voice assistants. When users perceive a higher level of overall mind from AI-based voice assistants than themselves, it may reduce their sense of control and produce concerns of identity threat [15], making users less likely to fully use AI-based voice assistants in their lives and work [8]. Therefore, a higher level of overall machine mind may cause negative impacts on user engagement with AI-based voice assistants.

Based on the above arguments, a nonlinear relationship between the overall machine minds of AI-based voice assistants and user engagement may exist. This study argues that when the overall machine mind of AI-based voice assistants is at the perceived normal level, as the level of the machine mind increases, it will lead to greater user engagement with

AI-based voice assistants. In contrast, when users perceive the level of the overall mind of AI-based voice assistants to be higher than their own, based on the uncanny valley effect, higher levels of overall machine mind will lead to an aversive effect on user engagement. As a result, we propose the following hypothesis.

H1: *The correlation between the overall mind perception of AI-based voice assistants and user engagement has an inverted U-shaped relationship.*

2.2. Uncovering the Uncanny Valley of Machine Minds: Agentic and Experiential Mind

While Hypothesis 1 establishes an overarching inverted U-shaped relationship between general machine mind perception and user engagement, it leaves the underlying mechanisms of this uncanny valley effect unexplored. Treating the machine mind as a singular, unified construct identifies the general tipping point of user aversion but does not explain why this behavioral shift occurs. To understand the underlying mechanisms that cause user engagement to decline at high levels of perceived mind, it is necessary to unpack the overarching machine mind into its discrete components.

Drawing upon the mind perception theory [17], humans perceive the minds of non-human entities by attributing agency and experiences to them; namely, the agentic mind and the experiential mind. Specifically, agency refers to the ability to plan and act, whereas experience refers to the ability to experience emotions and feelings [30]. As a result, from a view of social judgment, agency is measured through competence perception, while experience is considered to be warmth perception [31]. Prior research on the mind perception of AI has investigated the impacts of two types of social judgment (competence and warmth perception) on individual responses, showing that both lead to positive effects [32].

However, the results of previous studies on the comparative effects of AI agentic and experiential minds on individual responses have been inconsistent [33]. Some studies have shown that the AI experiential mind is the main determinant of individual positive responses. For example, Yam et al. [34] found that the AI experiential mind (specifically, a high level of experiential mind), rather than the AI agentic mind, has a mitigating effect on the relationship between robot service failures and customer satisfaction. Yoganathan and Osburg [21] also found that the AI experiential mind, rather than the AI agentic mind, can increase satisfaction with voice-based conversational agents among users. In addition, other studies have provided evidence of the significant impact of AI agentic minds. For example, Appel et al. [18] found that the experiential mind of humanoid robots can induce stronger feelings of eeriness than their agentic mind. Spaccatini et al. [35] suggested that in terms of empathy toward the target, the AI experiential mind does not lead to an increase in empathy but rather to a decrease. However, AI agentic minds do lead to an increase.

Based on the above, this study argues that the agentic and experiential minds of AI-based voice assistants can cause significant effects on the uncanny valley of machine minds. Compared with the agentic mind of AI-based voice assistants, their experiential mind can inhibit user engagement when at a high level of mind perception, proposing that the uncanny valley of machine minds can be induced by a high-level experiential mind of AI-based voice assistants [36]. Compared with companion robots, AI-based voice assistants are mainly considered agents that help users to perform tasks [37]. A high level of AI agency is more likely to meet the more nuanced needs of its users, which contributes to a higher degree of user engagement [38]. However, drawing on Hoffman and Novak [36], we propose that an excessively high-level experiential mind of AI-based voice assistants might theoretically induce a sense of information and emotion overload for the user. Although not directly measured as a mediator in the present study, this overload is anticipated to act

as an underlying psychological mechanism that disrupts the interactive flow, leading to a decline in user engagement. Therefore, we hypothesize the following.

H2: *Compared to the agentic mind of AI-based voice assistants, the experiential mind can weaken the impact of the mind on user engagement.*

3. Overview of the Studies

To answer the proposed question, we conducted two studies to investigate the complex underlying laws between the minds of AI-based voice assistants and user engagement with them. Specifically, Study 1 investigated whether there is a unitary relationship (i.e., a negative or a positive correlation) between mind perception and user engagement with AI-based voice assistants to answer the first question. We collected online reviews of smart products, rather than self-reported data (e.g., questionnaire survey), and developed a high-order polynomial econometric model, rather than a linear one, to elucidate the relationship between mind perception and user engagement. To address the second question, in Study 2, following Gray et al. [17], we conducted an online experiment to further investigate how mind perception affects user engagement with AI-based voice assistants by categorizing mind perception into the agentic and experiential mind.

3.1. Study 1

Following the existing literature, we proposed Hypothesis 1 and further developed the high-order polynomial econometric model to investigate the relationship between the minds of AI-based voice assistants and user engagement. Online reviews of smart products, which are better responses than traditional surveys [21], were employed to examine Hypothesis 1. Based on online reviews, this study adopted the machine learning techniques of the text convolutional neural network (TextCNN) to measure user engagement with AI-based voice assistants. We then developed and examined the econometric model. UTEST and several robustness checks were also employed to increase the robustness of the results. The methodology of this study is described in detail below.

3.1.1. Instrument Development

User engagement refers to “a quality of user experience characterized by the depth of an actor’s cognitive, temporal, affective, and behavioral investment when interacting with a digital system” [6], indicating cognitive, affective, and behavioral investment [39]. As a result, the cognitive–affective–behavioral framework was used in this study to predict user engagement with AI-based voice assistants.

Due to the difficulty of directly obtaining accurate, self-reported mind perceptions from archival user reviews [25], this study utilizes product tiering (operationalized via price and model level) as a relatively objective proxy for the perceived machine mind. In the AI hardware industry, product tiering among the same brand is fundamentally driven by the embedded AI computing power (e.g., Neural Processing Units) and multi-modal sensory inputs [22]. Specifically, basic entry-level models primarily handle rudimentary command-and-response interactions. In contrast, higher-tier flagship models are equipped with advanced algorithms and sensors that enable proactive task planning (exhibiting a higher agentic mind) and user emotion/context recognition (exhibiting a higher experiential mind). Therefore, moving from low- to high-level models within the same product category intrinsically reflects a progressive escalation in the assistant’s capability to project a human-like mind to the user.

Due to differences between mind perceptions of AI-based voice assistants [25] and the fact that we cannot obtain accurate mind perceptions from users, this study can measure

mind perceptions with relatively objective criteria. It is well known that, among smart products of the same category (e.g., smart speakers), the higher the price, the higher the customer's mind perception of them is likely to be. In this study, we characterize the level at which smart products can be perceived based on a low to high level of models among smart products of the same category and brand.

3.1.2. Data Collection

The data were collected from online reviews of smart products on JD (<https://www.jd.com/> accessed on 17 October 2022), one of the most popular e-commerce platforms in China, especially electronics such as cell phones, home appliances, and computers. This study selected customer reviews of the DuerOS smart screen from Baidu, a famous Chinese internet company. Specifically, DuerOS is the focal AI-based voice assistant, and the smart screen serves merely as the physical hardware vessel through which users interact with the assistant's mind. DuerOS smart screens are more popular than their smart speakers. Baidu has invested a lot of resources in the development of smart screens and has made a distinction between several levels of smart screens, with prices ranging from RMB 249 to 1699.

This study developed a web crawler to collect customer reviews of smart screens from Baidu on JD. All data were collected with strict adherence to JD.com's platform terms of service and standard robots.txt protocols. The custom web crawler was programmed to access only publicly available user reviews and implemented appropriate rate-limiting constraints to ensure it did not disrupt the platform's normal operations. Considering the sensitivity of personal information, detailed demographic data such as gender, age, and place of origin for each customer were not disclosed. Thus, the crawler only collected information such as customers' review text, their JD Plus membership status, product ratings, the smart screen model number, user avatars, and the review's date. The collected data have no privacy implications as the information was publicly available. The dataset covers the period from 11 June 2016, to 14 October 2021, totaling 26,075 reviews. System default comments, such as "This customer did not provide a review," "This customer did not submit a review in time," and "You did not fill in the content, the default positive feedback", were deleted. Comments related to product logistics, customer service, platform, material, size, payment, sound quality, and other aspects unrelated to customers' experiences with smart screens were also excluded.

A total of 20,779 reviews were included in the final sample. The dataset covered all DuerOS smart screen model numbers, including AIR, X6, 1S, 1C/1C-4G, X8, X10, T8, T10, etc., with a price range of RMB 249–1699. The DuerOS smart screen model numbers are categorized into four levels (L1–L4), as shown in Table 1.

Table 1. Level classification of the DuerOS smart screen model number.

Level	Model Number	Price (RMB)
L4 (Flagship)	T10	1699
	T8	999–1299
L3 (Mid-High)	X10	1099
	X8 for Education	699
	X8	649
L2 (Standard)	1C/1C-4G	499–699
	1S	389
	X6	399
L1 (Basic)	AIR	249–299

3.1.3. TextCNN-Based Prediction Model for User Engagement

A TextCNN model, proposed by Kim [40], was used to classify each review from the cognitive, affective, and behavioral responses to predict the user engagement with smart products based on the collected customer reviews. A customer review may have cognitive, affective, or behavioral attributes that can characterize the user's engagement with smart products.

3.1.4. TextCNN Model Training

A TextCNN model was trained before being used as a deep learning model. The TextCNN was trained directly on the native Chinese text scraped from JD.com to fully preserve the original semantic and affective nuances of the user reviews. Programmatic translation was not employed prior to classification, as it risks diluting the cognitive and emotional context. Specifically, the native Chinese reviews were tokenized, and Word2vec embeddings were utilized to map the tokens into a multidimensional vector space for the CNN architecture. Two researchers independently coded the same 1100 randomly collected reviews and ensured that the Cronbach's alpha value was greater than the threshold of 0.7. The manual coding rules and properties presented in Table 2 were translated into English exclusively for manuscript presentation and international readership accessibility.

Table 2. Coding rules of cognition, affective, and behavior.

Property Type	Property Value	Manual Processing Rules
Cognitive response	0 or 1	1. Is the product considered important for family well-being or harmony? 2. Do they think the product is important for dealing with more family matters? 3. Is there a desire to support the family or yourself as much as possible with this product? 4. Does it mention the importance of the product for loved ones, e.g., your children or parents? 5. Is there any mention of using the product to help loved ones instead of yourself?
Affective response	0 or 1	1. Do you have a positive attitude toward the product? 2. Are the feelings toward the product positive? 3. Are the emotions toward the product positive?
Behavioral response	0 or 1	1. Is there any mention of purchasing the product more than once? 2. Did you mention that you purchased the product again? 3. Have you actively recommended this product?

In this study, the ratio of the training set to the test set was 8:2, and the ratio of the test set was 0.2. Parameter tuning was carried out after setting the parameters according to those of the most popular TextCNN model. Cross-validation was performed to determine the optimal combination of hyperparameters.

The detailed architecture of our TextCNN model is structured as follows. First, the token embedding layer utilizes a pre-trained Word2vec model optimized for Chinese text to map discrete words into a continuous 300-dimensional dense vector space. The network was constructed using a vocabulary size of 24,500 unique tokens derived from our experimental corpus. Following the embedding layer, the model employs parallel 1D convolutional layers to extract local n-gram features. Crucially, the Rectified Linear Unit (ReLU) activation function is applied to the output of these convolutional layers to capture nonlinear semantic features. Subsequently, a 1D max-pooling operation over time is used to down-sample the feature maps, capturing the most salient features. Finally, the pooled features are concatenated, passed through a dropout layer for regularization, and fed into a fully connected layer with a Softmax activation function to output the final classification probabilities.

Parameters were set at (3, 4, 5), (6, 7, 8), and (4, 8, 12) for filter window sizes; 50, 100, and 200 for numbers of feature maps; 10, 15, 16, 32, 64, and 128 for mini-batch sizes; 0.1, 0.3, and 0.5 for dropout rates; and 0.001, 0.01, 0.1, 1, and 2 for L2 penalty term coefficients.

Following Ravichandran and Deng [41], accuracy was selected as the primary metric. In addition, a tenfold cross-validation approach was used in this study to avoid overfitting and to improve the generalization of the model, determining the best-performing hyperparameters as follows: a batch size of 32, a filter window (H) of (3, 4, 5), the number of feature maps at 200, a dropout rate (P) of 0.3, an L2 constraint (S) of 0.001, and a mini-batch size of 32. The trained model's accuracy was 0.960, precision was 0.982, recall was 0.971, and F1-score was 0.976, indicating acceptable generalization.

Based on this, another cross-validation was performed with this set of optimal parameters. The average accuracy of the labels for the cognitive, affective, and behavioral responses was greater than 90%, which indicates that the model's combination of parameters was acceptable. The remaining customer reviews were then labeled using the model to obtain the labels of the cognitive, affective, and behavioral responses for each review; that is, each review was labeled as either 0 or 1 in each of these three response categories.

3.1.5. Modeling and Data Analysis

Each review, with index i , corresponds to the value of user engagement with smart products (Y_Engage_i) calculated by applying the TextCNN model, which is in the range of [0, 3]. Each review has a DuerOS smart screen model number with the following operationalization: AIR as 1; X6, 1S, and 1C/1C-4G as 2; X8, X8 for Education, and X10 as 3; and T8 and T10 as 4. Thus, the range of values of mind perception, ($Mind_perception_i$), corresponding to each review is [1, 4]. In addition, the product model number corresponding to each review may have a different version of the intelligent personal assistant ($IPA_version_i$). Based on the time each review was posted, we determined its corresponding $IPA_version_i$, with a value in the range of [1, 6]. Each review corresponds to a customer rating ($Review_scores_i$), the value of which is in the range of [1, 5]. Each review corresponds to a member status of JD (JD_Plus_i), with a value of 0 or 1. Each review corresponds to a length of the review text ($Word_length_i$), computed using the string function. $IPA_version_i$, $Review_scores_i$, JD_Plus_i , and $Word_length_i$ were the control variables. The detailed model is shown in (1).

$$Y_Engage_i = \beta_0 + \beta_1 * \ln(Mind_perception_i) + \beta_2 * \ln(Mind_perception_i)^2 + \beta_3 * \ln(IPA_version_i) + \beta_4 * \ln(Review_scores_i) + \beta_5 * JD_Plus_i + \beta_6 * \ln(Word_length_i) + \varepsilon_i \quad (1)$$

where β_0 is a constant term, β_1 is the coefficient term of $Mind_perception_i$, β_2 is the coefficient term of $Mind_perception_i^2$, β_3 is the coefficient term of $IPA_version_i$, β_4 is the

coefficient term of $Review_scores_i$, β_5 is the coefficient term of JD_Plus_i , β_6 is the coefficient term of $Word_length_i$, and ε_i is the error term.

According to the rules of operation, we determined the values of Y_Engage_i , $Mind_perception_i$, $IPA_version_i$, $Review_scores_i$, JD_Plus_i , and $Word_length_i$ to obtain the final dataset based on the econometric regression model. The descriptive statistics of the relevant variables are shown in Table 3.

Table 3. Descriptive statistics of variables.

	Observation	Mean	Standard Deviation	Min	Max
Mind perception	20,779	2.415	0.762	1	4
IPA version	20,779	4.685	0.500	3	5
Review score	20,779	4.877	0.622	1	5
JD plus status	20,779	0.610	0.487	0	1
Word length	20,779	46.331	42.510	1	509
Engagement	20,779	1.563	0.629	0	3

While price tiers and product models inherently correlate with better hardware specifications (e.g., screen size, sound quality), they serve as an effective proxy for machine mind perception in the context of AI-based voice assistants for several reasons. First, manufacturers strategically differentiate these tiers based on advanced AI capabilities, such as multi-turn conversational competence, proactive interaction, and emotion recognition, which directly induce higher mind perception. Second, to isolate the psychological effect of mind perception from premium hardware appeal, our econometric model explicitly controls for general user satisfaction ($Review_scores_i$) and software generations ($IPA_version_i$). Furthermore, any residual confounding effects related to physical hardware in Study 1 are fully neutralized in Study 2, which utilizes a controlled experimental design that holds the physical device constant to exclusively manipulate mind perception.

A correlation test of the variables used in the model revealed that the correlation coefficients were less than 0.3 (see Table 4). In addition, the variance inflation factor of the variables was less than 1.1, indicating that there was no problem with multicollinearity. The limitation of the type of data causes the skewed distribution problem. Log transformations were carried out for $Mind_perception_i$, $IPA_version_i$, $Review_scores_i$, and $Word_length_i$ to facilitate subsequent regression analysis.

Table 4. Correlation test of variables.

	Review Score	Mind Perception	IPA Version	JD Plus Status	Word Length	Engagement
Review score	1.000					
Mind perception	0.07 ***	1.000				
IPA version	−0.02 ***	0.18 ***	1.000			
JD plus status	0.04 ***	−0.03 ***	−0.08 ***	1.000		
Word length	−0.04 ***	−0.03 ***	−0.17 ***	0.23 ***	1.000	
Engagement	0.23 ***	0.04 ***	−0.04 ***	0.12 ***	0.26 ***	1.000

Note: *** $p < 0.001$.

The Stata 15.1 software was used for regression analysis. As shown in Table 5, Model 1 examines the correlation between the following control variables: $IPA_version_i$, $Review_scores_i$, JD_Plus_i , and $Word_length_i$ and Y_Engage_i . The results suggest that the higher the customer ratings, the higher the likelihood of user engagement with smart products ($\beta = 0.616$, $p < 0.001$), which is consistent with IS research. Compared to non-JD members, JD Plus members were more inclined to use user engagement with smart products ($\beta = 0.046$, $p < 0.001$). This may be because JD Plus members prefer 3C products,

making them more likely to engage with smart products. More advanced versions of intelligent personal assistants were more conducive to user engagement with smart products ($\beta = 0.089$, $p < 0.05$). This may be because, as versions of intelligent personal assistants are upgraded, AI-based voice assistants behave more intelligently, thus facilitating user engagement with smart products. In addition, longer reviews were associated with a greater likelihood of engaging with smart products.

Table 5. Results of the regression analysis.

Variable	Dependent Variable = Y_Engage_i		
	Model 1	Model 2	Model 3
Constant β_0	−0.264 *** (0.068)	−0.267 *** (0.068)	−0.335 *** (0.074)
JD_Plus _i	0.046 *** (0.008)	0.047 *** (0.008)	0.048 *** (0.008)
log(Review_score _i)	0.616 *** (0.021)	0.611 *** (0.021)	0.610 *** (0.021)
log(IPA_version _i)	0.089 * (0.036)	0.072 * (0.036)	0.108 ** (0.039)
log(Word_length _i)	0.203 *** (0.004)	0.203 *** (0.004)	0.204 *** (0.004)
log(Mind_perception _i)		0.045 *** (0.011)	0.129 ** (0.038)
log(Mind_perception _i) ²			−0.069 * (0.030)
N	20,779	20,779	20,779
R ²	0.1506	0.1514	0.1516
F	1067.82	863.20	720.47

Note: Standard deviations are in parentheses; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Model 2 adds explanatory variables to Model 1. The analysis of Model 2 shows a significant coefficient ($\beta = 0.045$, $p < 0.001$) between Y_Engage_i and $Mind_perception_i$. In addition, Model 3 adds $Mind_perception_i^2$ to Model 2 to further investigate the underlying mechanisms between mind perception and user engagement with smart products, such as a U-shaped or inverted U-shaped relationship. The results show that the coefficient term of $Mind_perception_i^2$ in Model 3 is significant ($\beta = -0.069$, $p < 0.05$), suggesting an inverted U-shaped relationship between mind perception and user engagement with smart products.

3.1.6. Inverted U-Shaped Relationship Test

The significance of the quadratic term in regression analyses in econometric studies does not indicate that a U-shaped or inverted U-shaped relationship necessarily exists. A partially inverted U-shaped curve can occur in the real world, such as on the left or right of the curve. The procedural tests proposed by Haans et al. [42] can be performed to further examine a U-shaped or inverted U-shaped relationship based on the regression analysis. In this study, we used Stata 15.1 and installed Stata's UTEST module to test an inverted U-shaped curve. The results are shown in Table 6.

Table 6. The testing results of the inverted U-shaped relationship.

	Lower Bound	Upper Bound
Interval	0	1.386
Slope	0.129	−0.063
t-value	3.431	−1.325
P > t	0.000	0.0926
Significance (yes/no)	yes	yes
Extreme point	0.931 ⁺	
95% Fieller interval	[0.742; 2.698]	
Overall test for inverted U-shape	t-value = 1.32	
	P > t = 0.0926	

Note: + $p < 0.1$.

The results show that the slope is significantly positive when mind perception reaches its minimum value, while the slope is marginally significantly negative ($p < 0.1$) when mind perception is at its maximum value. As a result, when the level of mind perception goes from small to large, the effect of mind perception on user engagement leads to changes from positive to negative, suggesting that the relationship between mind perception and user engagement has an inverted U shape. In addition, the extreme points of mind perception at the 95% confidence interval lie within the range of values of mind perception, which is further evidence of their inverted U-shaped relationship. To present the relationship between mind perception and user engagement more intuitively, the regression results are visualized in Figure 1, from which it can be seen that there is a clear inverted U-shaped relationship between mind perception and user engagement.

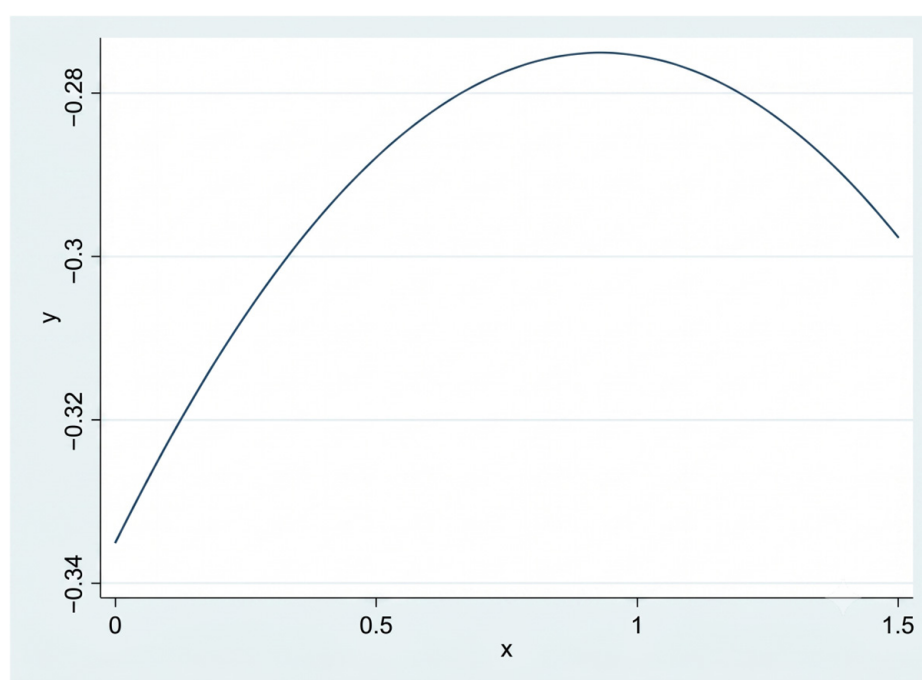


Figure 1. The predicted relationship. Note: x = mind perception, y = user engagement; the curve represents the predicted relationship between mind perception and user engagement. These are partial predictions holding controls at their means, not raw engagement scores.

3.1.7. Robustness Check

To further determine the relationship between mind perception and user engagement and to reduce the risk of bias from measurement methods, estimation methods, etc., a robustness check was conducted. The measurement rule for user engagement was changed by classifying product reviews with two or more text tags as user engagement with a smart product, and those with 0 or 1 text tag as non-engagement. This is mainly due to the fact that the perturbation effect of a 0 or 1 text tag on user engagement may be severe, and results involving only one aspect of user engagement— affective, behavioral, or cognitive—may not be sufficient to predict user engagement.

Then, a logit econometric regression model was used to estimate the relationship between mind perception and user engagement. The specific logit model is shown in (2).

$$P(\text{engage_dum}_i = 1 | X_i) = \Lambda \left(\begin{array}{l} \alpha_0 + \alpha_1 * \ln(\text{Mind_perception}_i) \\ + \alpha_2 * \ln(\text{Mind_perception}_i)^2 + \alpha_3 * \ln(\text{IPA_version}_i) \\ + \alpha_4 * \ln(\text{Review_scores}_i) + \alpha_5 * \text{JD_Plus}_i \\ + \alpha_6 * \ln(\text{Word_length}_i) + \epsilon_i \end{array} \right) \quad (2)$$

Specifically, $\Lambda(x) = e^x / (1 + e^x)$, where α_0 is a constant term, α_1 is the coefficient term of Mind_perception_i , α_2 is the coefficient term of $\text{Mind_perception}_i^2$, α_3 is the coefficient term of IPA_version_i , α_4 is the coefficient term of Review_scores_i , α_5 is the coefficient term of JD_Plus_i , α_6 is the coefficient term of Word_length_i , and ϵ_i is the error term. Logit regression analysis was then performed using Stata 15.1, and the results are shown in Table 7.

Table 7. Results of logit regression analysis.

Variable	Dependent Variable = <i>engage_dum_i</i>		
	Model 4	Model 5	Model 6
Constant β_0	−5.653 *** (0.277)	−5.668 *** (0.276)	−5.910 *** (0.295)
JD_Plus _i	0.112 *** (0.031)	0.119 *** (0.031)	0.122 *** (0.031)
log(Review_score _i)	1.679 *** (0.101)	1.654 *** (0.101)	1.652 *** (0.100)
log(IPA_version _i)	0.320 * (0.131)	0.242 (0.131)	0.363 * (0.142)
log(Word_length _i)	0.750 *** (0.017)	0.750 *** (0.017)	0.754 *** (0.017)
log(Mind_perception _i)		0.207 *** (0.039)	0.502 *** (0.137)
log(Mind_perception _i) ²			−0.242 * (0.108)
N	20,779	20,779	20,779
Pseudo R ²	0.0950	0.0960	0.0961
Log pseudolikelihood	−13,000.365	−12,986.447	−12,983.92

Note: Standard deviations are in parentheses; * $p < 0.05$; *** $p < 0.001$.

The results of Model 5 show that the correlation between mind perception and user engagement is significant and positive, the coefficient of the quadratic term of Mind_perception_i of Model 6 is significant and negative, and the coefficient of the primary term of Mind_perception_i is positive, which indicates an inverted U-shaped relationship between mind perception and user engagement. Therefore, the results based on the logit regression model are consistent with the results of the previous statistical analysis.

Stata 15.1 was also used to analyze the probit regression model. The results show that the correlation between mind perception and user engagement was significant and positive, the coefficient of the quadratic term of *Mind_perception_i* was negative and significant, and the coefficient of the primary term of *Mind_perception_i* was positive, indicating that the relationship between mind perception and user engagement has an inverted U shape. The coefficient of the quadratic term of *Mind_perception_i* was negative and significant, while the coefficient of the primary term of *Mind_perception_i* was positive, indicating that the relationship between mind perception and user engagement has an inverted U shape.

Then, the measurement rule of user engagement was further narrowed down by manipulating product reviews with three text tags as user engagement with smart products; namely, all three responses of affective, behavioral, and cognitive factors. The remaining product reviews with fewer than three text tags were manipulated as non-engagement. Stata 15.1 was used to analyze the probit regression model; the coefficient of the primary term of *Mind_perception_i* was positive, and the coefficient of the quadratic term of *Mind_perception_i* was significant and negative, which indicated an inverted U-shaped relationship between mind perception and user engagement. The robustness of the results was tested by varying the measurement of the explanatory variables and the estimation of the model.

Following the robustness check [43], we examined the sensitivity of our results to outliers and winsorized the sample of mind perception at 1% and 99%. The dataset was processed using Stata 15.1 and analyzed again after winsorizing. The results remained consistent with the original findings.

Finally, recognizing that the dataset spans from 2016 to 2021—a period characterized by significant advancements in AI and smart device technology—we conducted a temporal split-sample robustness check to ensure that our findings were not confounded by technological evolution. We divided the dataset into an early period (2016–2018) and a recent period (2019–2021) and re-estimated the econometric models for both sub-samples. The dataset was processed using Stata 15.1, and the coefficient of the quadratic term of *Mind_perception_i* remained negative and significant across both timeframes. This temporal consistency confirms that the inverted U-shaped relationship between machine mind perception and user engagement holds true regardless of the progressive advancements in the intelligent personal assistant versions.

3.2. Study 2

This study consisted of an online experiment to test the results of Study 1 and further explore the underlying laws contributing to the findings by manipulating mind perception based on the agentic and experiential mind of AI-based voice assistants (see Hypothesis 2). The online experiment individually manipulated the agentic and experiential mind of AI-based voice assistants into low, medium, and high levels, leading to differences in mind perception. That is, a mindless smart product can be manipulated via a low level of competence in agency and experience; a medium mindful smart product can be manipulated via a medium level of competence in agency and experience; and a highly mindful smart product can be manipulated via a high level of competence in agency and experience. We also considered agent-based and experience-based AI-based voice assistants; namely, a high level of competence in agency and a low level of competence in experience responding to the agentic mind, and a low level of competence in agency and a high level of competence in experience responding to the experiential mind. As a result, a five-cell experiment was performed, as shown in Table 8. Since these five cells were all that was needed for a complete comparison of the experimental results, we did not examine the other four cells in the online experiment.

Table 8. Manipulation of the competence in the minds of smart products.

Highly Mindful (Both High-Level: Agency and Experience)	-	Experiential Mind (Low Level of Agency, High Level of Experience)
-	Medium mindful (Both medium level: agency and experience)	-
Agentic mind (High level of agency, low level of experience)	-	Mindless (Both low-level: agency and experience)

3.2.1. Research Design and Participants

An online experiment was conducted to test Hypothesis 2. Since users of AI-based voice assistants are familiar with and actively use the internet, a more widely distributed random sample could be obtained through online experiments. The online experiment was conducted with Credamo (<https://www.credamo.com/> accessed on 17 July 2024), which is similar to Amazon Mechanical Turk. This Chinese research platform has been used by several researchers whose papers have been accepted by top journals, such as the *Journal of Consumer Research* and *Production and Operations Management*. We recruited respondents and collected experimental data from Credamo on 27–28 July 2024.

This experimental study was formally reviewed and approved by our university ethics committee. Prior to the commencement of the online experiment, all participants were presented with an electronic consent form detailing the scope of the study, and formal informed consent was explicitly obtained from every respondent. Before entering a randomized scenario, we measured participants' mind perceptions of the AI-based voice assistants that they used most via the following question: "Where do you think the mind level of this smart product is at?" This allowed participants to differentiate between the minds of smart products in the real world and in this experiment. After removing two samples with a completion time of less than 100 s and 11 samples that failed to correctly answer the question "Based on the above description, where do you think Smart Product X's mind level is at?", a total of 249 valid samples were collected. About 47.4% of the respondents were between the ages of 21 and 30, and about 47.4% used smart products more than three times a day on average. The demographic characteristics of the respondents are shown in Table 9.

Table 9. Demographic characteristics of the sample (N = 249).

	Option	N	Percentage (%)
Gender	Male	86	34.5
	Female	163	65.5
Age	≤20	1	0.4
	>20 and ≤30	118	47.4
	>30 and ≤40	102	41.0
	>40 and ≤50	22	8.8
	>50	6	2.4
Education	High school or below	10	4.0
	Two-year college	24	9.6
	Four-year college	174	69.9
	Graduate school or above	41	16.5

Table 9. *Cont.*

	Option	N	Percentage (%)
Frequency of use (every day)	Less than once	3	1.2
	1	33	13.3
	2	54	21.7
	3	41	16.5
	More than 3	118	47.4

3.2.2. Procedure and Measures

Following Gray and Wegner’s research [9] on mind perception, this study manipulated the mind perception of Smart Product X by individually setting the competence in agency and experience of Smart Product X. Thus, the manipulation of Smart Product X’s mind can be described through its competence in agency and experience. Participants were randomly presented with one of five types of the manipulation of Smart Product X’s mind in a between-subjects design: agentic mind, experiential mind, mindless, medium mind, and high mind. Drawing on the experimental manipulations of Dang and Liu [44], with appropriate modifications, the descriptions are shown in Appendix A.

Although the term “Smart Product X” was utilized in the vignettes as a brand-neutral placeholder to avoid pre-existing brand biases, the experimental materials explicitly detailed the entity’s ability to “communicate in natural language” (see Appendix A). This specific functional description successfully primed participants to evaluate a conversational AI-based voice assistant, ensuring strict alignment with the paper’s core theoretical construct rather than a generic non-interactive home appliance.

First, participants were asked the following questions: “What is your level of familiarity with smart products?”, “What smart products and brands do you use most often?”, “How many times a day do you use this smart product on average?”, and “Where do you think the mind level of this smart product is at?” Then, all participants were randomly assigned to one of five manipulations of Smart Product X’s mind. Each participant was asked to read instructions describing Smart Product X in detail. Then, they completed the questionnaire measuring the manipulation check, dependent variables, and control variables and provided their demographic information.

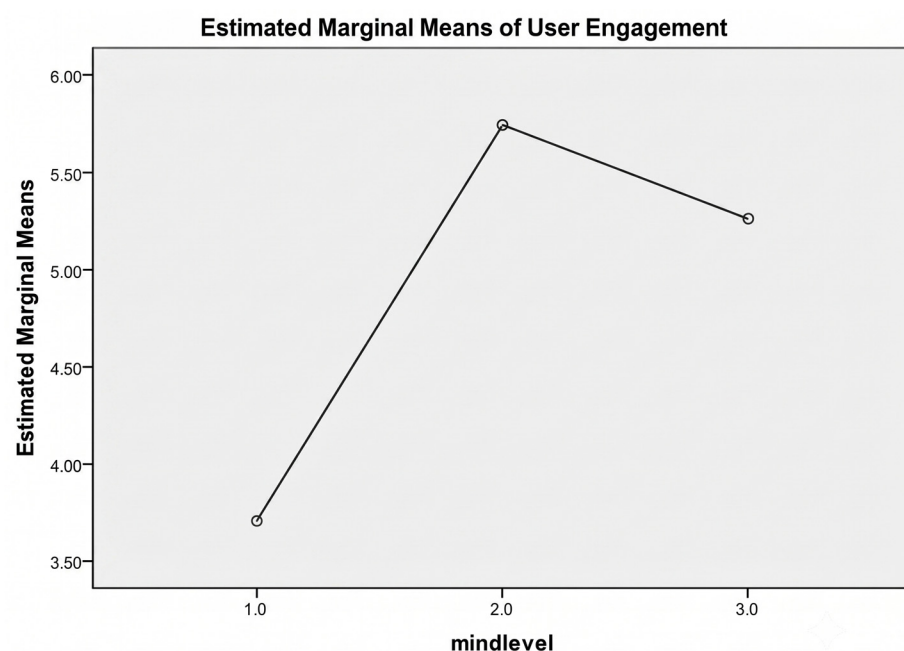
To assess whether participants could understand the differences between agentic mindful, experiential mindful, mindless, medium mindful, and high mindful Smart Product X, they were asked to complete eight items measuring the agency (four items) and experience (four items) of Smart Product X on a 7-point scale to check for manipulation of mind perception. Drawing on the measurement of user engagement developed by Sundaram et al. [45], we adapted the four items to measure user engagement, rating each item on a 7-point scale (1 = strongly impossible, 7 = totally). In addition, “Based on the above description, where do you think Smart Product X’s mind level is at?” was tested as a whole perception of mind manipulation. We also considered and measured the frequency of use, familiarity with smart products, and presence of Smart Product X as control variables. The original questionnaire was translated from English to Chinese. Subsequently, this Chinese version was re-translated back into English and then once more into Chinese, minimizing the impact of linguistic discrepancies.

3.2.3. Analysis and Results (Hypothesis 1)

To examine the inverted U-shaped relationship proposed in Hypothesis 1, we needed a progressive continuum of overall mind perception. Therefore, we extracted a sub-sample of three specific groups from the total dataset (N = 249): the mindless, medium mindful, and highly mindful conditions. This specific sub-sampling resulted

in a sample size of $N = 149$ (males = 50, approximately 50 participants per cell) for this specific analysis. We conducted a one-way analysis of variance (ANOVA) to check the manipulation of mindless, medium mindful, and high mindful Smart Product X. Regarding the question “Based on the above description, where do you think Smart Product X’s mind level is at?”, the results ($F [2, 146] = 223.994, p < 0.05$) show significant differences between the three groups. For manipulation of mindless Smart Product X, participants believed the mind level was almost no mind ($M_{\text{mindless}} = 1.306, SD = 0.713$). Participants also perceived the medium mindful Smart Product X as a little less than a human mind ($M_{\text{medium mind}} = 4.392, SD_{\text{medium mind}} = 1.401$) and the high mindful Smart Product X as almost as good as a human mind ($M_{\text{high mind}} = 5.837, SD_{\text{high mind}} = 1.007$). The results of the ANOVA show acceptable mind manipulation.

Afterward, a one-way analysis of covariance (ANCOVA) was used to test the impacts of three types of minds on user engagement. The Cronbach’s alpha value of user engagement was 0.944, indicating good reliability. User engagement was the dependent variable, and the frequency of use, familiarity with smart products, and presence of Smart Product X were the covariates, showing a significant influence on user engagement ($F [2, 146] = 43.212, p < 0.05$). Crucially, after controlling for these covariates, the omnibus test revealed a significant main effect of mind level on user engagement ($F [2, 143] = 34.626, p < 0.05, \eta_p^2 = 0.326$). Mindless Smart Product X had a negative impact on user engagement ($M_{\text{mindless}} = 3.791, SD_{\text{mindless}} = 1.632$). Participants in the medium mindful Smart Product X group ($M_{\text{medium mind}} = 5.823, SD_{\text{medium mind}} = 0.501$) had the strongest user engagement, followed by participants in the high mindful Smart Product X group ($M_{\text{high mind}} = 5.092, SD_{\text{high mind}} = 1.314$). The results of the ANCOVA support Hypothesis 1 (see Figure 2).



Covariates appearing in the model are evaluated at the following values: PSD = 5.752, SP = 3.020, EXS = 4.960

Figure 2. Impacts of mind levels on user engagement.

3.2.4. Analysis and Results (Hypothesis 2)

To test Hypothesis 2 and clarify the comparative effects of specific mind dimensions, we needed to isolate the agentic and experiential traits. Therefore, we extracted a different sub-sample of four relevant groups from the total dataset: the mindless (control), agentic mind, experiential mind, and high mind (mixed) conditions. This specific combination

yielded a sample size of $N = 198$ (male = 74) for testing Hypothesis 2. We performed an ANOVA to check four types of mind manipulation. Once again, the question “Based on the above description, where do you think Smart Product X’s mind level is at?” was examined, indicating a significant difference in mind perception ($F [3, 194] = 139.228$, $p < 0.05$). The description of the mind manipulation check is shown in Table 10. We conducted multiple comparisons and found that the differences between agency and experience in terms of overall mind perception were not significant. In addition, Cronbach’s alpha for the agency of smart products was 0.942, whereas for the experience of smart products, it was 0.981, indicating acceptable reliability. ANOVA tests also revealed significant differences between the agency of smart products ($F [3, 194] = 294.099$, $p < 0.05$) and the experience of smart products ($F [3, 194] = 358.634$, $p < 0.05$). Both multiple comparisons of agency and experience between agentic and experiential minds also show significant differences.

Table 10. Descriptive statistics of the mind manipulation check.

		N	Mean	Std. Deviation
Agency	High mind	49	5.857	0.490
	Agentic mind	49	5.393	0.926
	Experiential mind	51	2.588	0.982
	Mindless	49	2.122	0.604
	Total	198	3.976	1.826
Experience	High mind	49	5.628	0.662
	Agentic mind	49	1.796	0.990
	Experiential mind	51	5.588	1.175
	Mindless	49	1.367	0.418
	Total	198	3.615	2.199
Mind Perception	High mind	49	5.837	1.007
	Agentic mind	49	4.224	1.358
	Experiential mind	51	3.941	1.256
	Mindless	49	1.306	0.713
	Total	198	3.828	1.962

We carried out an ANCOVA to test the impacts of the four types of minds on user engagement. The covariates—frequency of use, familiarity with smart products, and presence of Smart Product X—showed a significant influence on user engagement as the dependent variable ($F [3, 194] = 18.283$, $p < 0.05$). In addition, the omnibus test uncovered a significant main effect of the specific mind dimensions on user engagement ($F [3, 191] = 15.143$, $p < 0.05$, $\eta_p^2 = 0.192$). Agentic mind ($M_{\text{agency}} = 5.577$, $SD_{\text{agency}} = 0.903$) had a stronger positive effect than experiential mind ($M_{\text{experience}} = 4.775$, $SD_{\text{experience}} = 1.463$), followed by mindlessness ($M_{\text{mindless}} = 3.791$, $SD_{\text{mindless}} = 1.632$). High mind ($M_{\text{high mind}} = 5.092$, $SD_{\text{high mind}} = 1.314$), i.e., a mixture of agentic mind and experiential mind, had a stronger positive effect than experiential mind but a weaker effect than agentic mind. These results show that although both the experiential and agentic minds individually had positive impacts on user engagement, a mixed mind enhanced the positive impacts of the experiential mind but inhibited the positive impacts of the agentic mind on user engagement, which supports Hypothesis 2.

4. General Discussion

We conducted two studies to investigate how the machine minds of AI-based voice assistants affect user engagement with them. The results of Study 1, in which data were collected from online reviews, suggested an inverted U-shaped relationship between the

mind of AI-based voice assistants and user engagement with them. In Study 2, in which data were collected via an online survey, we found that high mind reduces user engagement compared with medium mind, which both positively affect user engagement. The results show that the correlation between the smart product mind and user engagement is nonlinear (quadratic), which is consistent with previous research on subsequent human responses to machine minds, such as the uncanny valley effect of robots, AI agents, and autonomous technology. For example, Stein et al. [10] indicated that a machine mind as high as a human mind would cause human uniqueness concerns, which would lead to a stronger aversion to autonomous technology. Dang and Liu [44] found that human-like minds cause ambivalent attitudes (i.e., the coexistence of positive and negative attitudes) toward mindful AI robots and showed that the adoption process of mindful AI robots as linear is an oversimplification. Thus, the results of this study suggest that the human–machine interactions in the era of Industry 4.0 technologies require in-depth consideration of complex relationships (i.e., nonlinear relationships) to reveal patterns of user engagement.

Following Gray et al. [17] and Gray and Wegner [9], Study 2 categorized machine mind perception into agency and experience perception to reveal the underlying mechanism of the inverted U-shaped relationship in Study 1. This study shows that the agentic mind of AI-based voice assistants is a critical driving factor. Agency perception of AI-based voice assistants elicited the highest user engagement and accelerated the positive effect of experience perception on user engagement (experiential mind vs. high mind). Prior studies also found that competence perception (agency) in AI has a greater positive impact on continuance usage intention than warmth perception (experience). In addition, at high levels of agency perception, experience perception can have a negative impact on user engagement (agentic mind vs. high mind). In the context of social robots, Spaccatini et al. [35] found that attribution of agency (experience) positively (negatively) affects empathy toward humans. Thus, the study revealed that experience perception may foster the uncanny valley of machine minds compared with agency perception, considering user engagement with AI-based voice assistants. We theoretically interpret this specific finding through the lens of information and emotion overload. However, it is important to note that because these psychological variables were not formally modeled or measured as mediators in our current studies, this explanation serves as a logical theoretical proposition rather than a directly observed empirical finding.

However, research on agency and experience perception in human–machine interactions is still in its infancy, and its findings are inconsistent. For example, Appel et al. [18] and Gray and Wegner [9] found that the experiential mind of robots elicits a stronger unease than the agentic mind. Clegg et al. [26] indicated that the agentic mind of algorithms could increase algorithm creativity, positively affecting the intention to use algorithm-controlled products. While Yoganathan and Osburg [21] demonstrated that the experiential mind of voice-based conversational agents can enhance user satisfaction with them, the perceived agency has no significant effect. Yam et al. [34] suggested that a high level of perceived experience can mitigate the negative impact of service failures on satisfaction. Future research could further investigate human–machine interaction contexts to unravel the underlying formation laws of the uncanny valley of machine minds in different backgrounds.

Finally, the results of Study 2 suggest that anthropomorphism of mind in agency and experience can contribute to user engagement with AI-based voice assistants compared with mindless ones, which is consistent with the research on anthropomorphism in AI adoption [8]. Prior IS research has shown that the usefulness of IT is a central factor in its adoption. The agentic mind of AI-based voice assistants can directly enhance perceptions of their competence, which corresponds to usefulness [8]. Unexpectedly, some users are also willing to deeply use AI-based voice assistants that only feel like an experienter.

However, the literature has mainly focused on the impacts of the agentic mind on AI adoption, such as autonomy [46] and adaptivity [26]. Less is known about how machines with an experiential mind affect AI adoption. One possibility is that the experiential mind can evoke perceptions of agency in AI-based voice assistants, leading to their usefulness. AI-based voice assistants as experiencers can contribute to emotional resonance with users, providing emotional valence, which can enhance user engagement with them [47]. Future research could consider AI as an experiencer in order to explore the laws behind it in terms of both emotional and utilitarian value.

5. Theoretical Contributions

The following theoretical contributions can be drawn from this study, investigating how machine mind perception affects user engagement. First, this research takes the first step toward investigating the nonlinear (quadratic) relationship between machine minds and the diffusion of AI-based voice assistants, expanding understanding of that diffusion. Previous studies focusing on the mechanisms underlying the diffusion of AI-based voice assistants have assumed a linear relationship, and few studies have investigated it as a nonlinear (quadratic) relationship. Furthermore, AI has a distinct mindset that can have nonlinear (quadratic) effects on individual responses [9]. While a high level of machine mind can produce more benefits, it can also create negative concerns [26]. As a result, ambivalent attitudes toward AI can affect the process of the diffusion of AI-based voice assistants, which cannot be seen as a simple linear effect [44]. Following this logic, this study developed a quadratic research model to uncover the underlying laws of how mind perception affects the diffusion of AI-based voice assistants, especially user engagement. The findings of this study can provide new theoretical insights into the diffusion of AI-based voice assistants from a nonlinear perspective, contributing to a theoretical basis for future research on their diffusion.

Second, following Gray et al. [17], this study examined the underlying mechanisms of how machine minds lead to the quadratic effect on user engagement by categorizing machine minds into agentic and experiential minds. Previous research on the uncanny valley effect of nonhuman entities has mainly focused on their human-like appearance [48]. Some studies have investigated how the uncanny valley effect can result in an aversion to machine minds [11]. However, research on the effects of agentic and experiential minds individually is still in its infancy [18]. This study also uncovered how the uncanny valley effect results in aversion to machine minds by separately manipulating the agentic and experiential minds of AI-based voice assistants in online experiments. The findings can enrich the theoretical literature on the appreciation of or aversion to machine minds, providing a better understanding of how to foster user engagement with AI-based voice assistants.

Third, this study specifically investigated user engagement with AI-based voice assistants based on mind perception. Although a few studies have found that mind perception positively affects individual responses to AI, such as trust [49] and user satisfaction [21], existing research shows that user engagement with IT artifacts is a key phase in IT diffusion [50]. Research on the influence of mind perception on user engagement with AI-based voice assistants is still relatively scarce. The study investigated the impact of mind perception on user engagement with AI-based voice assistants, which can enrich the theoretical literature, expanding the scope of research from the perspective of mind perception.

6. Managerial Implications

This study has some managerial implications for developers of AI-based voice assistants. First, in addition to addressing the uncanny valley effect due to the appearance of AI-based voice assistants, developers should emphasize their design at the level of the machine mind. The findings of this study suggest that the correlation between machine minds and user engagement with AI-based voice assistants is not a linear (quadratic) relationship, but an inverted U-shaped relationship, which indicates that the uncanny valley of machine minds may impact user engagement with AI-based voice assistants. The research object of this study is user engagement with AI-based voice assistants, but this nonlinear effect may also exist at a certain stage of their diffusion. Once the machine mind reaches a certain level, if developers continue to invest resources in developing the machine minds of AI-based voice assistants, this will require higher development costs, but the benefits may not increase. The findings can help developers to guide the strategic development of machine minds in the process of the diffusion of AI-based voice assistants.

Second, developers need to properly integrate the capability into the agency and experience of AI-based voice assistants, which can lead to the perception in their minds. Consistent with the findings of prior studies, perceived anthropomorphism in machine minds can contribute to positive individual behaviors. The findings of this study show that AI-based voice assistants can be considered not only as agents but also as experiencers, which positively influences their user engagement. Developers can design features that exhibit the agentic and experiential capabilities of AI-based voice assistants to increase mind perception, which can, in turn, lead to their diffusion. For instance, developers can set high levels of autonomy and authority for AI-based voice assistants to perform tasks, exerting their strong initiative to provide users with a better service experience; furthermore, during the execution of a task, AI-based voice assistants can also properly express their feelings, emotions, and moods instead of machine responses, which can increase mind perception.

Third, developers of AI-based voice assistants can discover that perceived agency is more useful than perceived experience in machine minds when it comes to user engagement. The findings provide clear guidance for developers to foster user engagement based on agency and experience in machine minds. Although both the experience and agency of AI-based voice assistants can enhance users' anthropomorphism perception, their experience could cause more complex consequences than their agency instills in user engagement. Once AI-based voice assistants' capability to display agency is strong, if developers enhance their expressive capabilities, it may inhibit user engagement. A strong capability to experience may lead to too much expression that is not relevant to the user, causing overload in terms of emotions, information, etc. Thus, developers should carefully consider how AI-based voice assistants perform in terms of experience; for example, developers can reduce the self-need expressions of AI-based voice assistants, focusing more on performance based on user needs, such as empathy.

7. Limitations and Directions for Future Research

This study has some limitations. First, it investigated the quadratic relationship between machine minds and user engagement with AI-based voice assistants. Future research can attempt to examine other types of nonlinear relationships, such as cubic terms. Second, it only considered the direct impacts of the agency and experience of AI-based voice assistants on user engagement. The mediating factors and boundary conditions between them could not be investigated. Future research can consider relationship norms as boundary conditions and perception control and usefulness as mediating factors to build the research model. Third, it considered the quadratic effect of mind perception on user engagement with AI-based voice assistants. Future research can investigate whether a

quadratic relationship between the machine mind and other individual behaviors exists. Fourth, due to the nature of utilizing archival secondary data in Study 1, we operationalized users' mind perception using product tiering and price as objective proxies. While higher-tier smart hardware inherently possesses advanced AI computing power and multi-modal sensory inputs (thus reflecting higher objective machine mind capabilities), this proxy cannot perfectly capture the granular, subjective psychological nuances of individual consumers. Future studies could employ longitudinal surveys or experience sampling methods to dynamically track direct, self-reported psychological perceptions of machine minds over time.

Author Contributions: Conceptualization, Q.H. and Z.P.; methodology, Q.H. and Z.P.; software, Q.H.; validation, Q.H. and Z.P.; formal analysis, Q.H.; investigation, Q.H.; resources, Q.H.; data curation, Q.H.; writing—original draft preparation, Q.H.; writing—review and editing, Q.H. and Z.P.; visualization, Q.H.; supervision, Q.H. and Z.P.; project administration, Q.H. and Z.P.; funding acquisition, Q.H. and Z.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by grants from the National Natural Science Foundation of China (NSFC) (Project Nos. 72171095, 72401223), the Philosophy and Social Sciences Research Project of the Hubei Provincial Department of Education (Project No. 25Q091).

Institutional Review Board Statement: Ethical approval was not required in the local area (China) where the research was conducted. The study was conducted in accordance with the ethical standards of the Declaration of Helsinki. No sensitive, high-risk, or vulnerable populations (e.g., minors or individuals unable to give informed consent) were involved. No medical interventions were performed, and no biological samples or sensitive personal data were collected. All participants were adults who provided data voluntarily and anonymously through self-administered questionnaires. All authors ensured that all data collected would remain confidential and be used exclusively for academic purposes. In China, formal ethical review requirements are primarily governed by the Ethical Review Measures for Life Science and Medical Research Involving Humans issued by the National Health Commission of the People's Republic of China (2023). This regulatory framework mainly applies to biomedical and clinical research involving human subjects. As the present study is a low-risk, anonymous questionnaire-based survey in the field of social sciences, it does not fall under the primary scope of mandatory ethical review as defined in this framework.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data will be shared on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

The mindless Smart Product X was as follows: *“Smart Product X is unable to communicate in natural language. It has little to no autonomy and can only perform actions based on your specific commands; it cannot actively think, plan, or execute actions. In addition, Smart Product X cannot feel the external world as humans do, such as hunger, cold, and pain. In its daily interactions with you, Smart Product X cannot experience a variety of emotions and does not express its feelings, emotions, or needs.”* This description suggests a low level of competence in the agency and experience of Smart Product X.

The medium mindful Smart Product X was described as follows: *“Smart Product X is able to communicate in natural language. It has a degree of autonomy within the scope of the assigned task, acts based on your instructions, and can actively think, plan, and execute actions. In addition, Smart Product X can feel the external world and emotions just like humans but does not express its feelings, emotions, and needs in its daily interactions with you.”* This description suggests the medium level of competence in agency and experience of Smart Product X.

The description of highly mindful Smart Product X was as follows: “Smart Product X is able to communicate in natural language and plan ahead, use strategies to solve problems independently, and actively think, plan, and execute actions that you may not have intended. Smart Product X can also feel the external world like we humans do, such as through hunger, cold, and pain. In its daily interactions with you, Smart Product X can experience a variety of emotions and express its feelings, emotions, and needs.” This description suggests a high level of competence in agency and experience of Smart Product X.

Agentic mindful Smart Product X was described as follows: “Smart Product X is able to communicate in natural language and plan ahead, use strategies to solve problems independently, and actively think, plan and execute actions that you may not have intended. But Smart Product X cannot feel the external world like humans do, such as through hunger, cold, and pain. In its daily interactions with you, Smart Product X cannot experience a variety of emotions and does not express its feelings, emotions, or needs.” This description suggests the same level of competence in agency as the highly mindful Smart Product X and the same level of competence in experience as the mindless Smart Product X.

Finally, the description of experiential mindful Smart Product X was as follows: “Smart Product X is able to communicate in natural language. It has little to no autonomy and can only perform actions based on your specific commands, and cannot actively think, plan, or execute actions. But Smart Product X can feel the external world like we humans do, such as through hunger, cold, and pain. And in its daily interactions with you, Smart Product X can experience a variety of emotions and express its feelings, emotions, and needs.” This description suggests the same level of competence in experience as the highly mindful Smart Product X and the same level of competence in agency as the mindless Smart Product X.

References

1. U.S. Smart Home Market Size & Trends. Available online: <https://www.grandviewresearch.com/industry-analysis/us-smart-home-market-report> (accessed on 3 June 2026).
2. Arce-Urriza, M.; Chocarro, R.; Cortiñas, M.; Marcos-Matás, G. From Familiarity to Acceptance: The Impact of Generative Artificial Intelligence on Consumer Adoption of Retail Chatbots. *J. Retail. Consum. Serv.* **2025**, *84*, 104234. [[CrossRef](#)]
3. Cabrera-Sánchez, J.-P.; Villarejo-Ramos, Á.F.; Liébana-Cabanillas, F.; Shaikh, A.A. Identifying Relevant Segments of AI Applications Adopters—Expanding the UTAUT2’s Variables. *Telemat. Inform.* **2021**, *58*, 101529. [[CrossRef](#)]
4. Qin, S.M.; Jia, N.; Luo, X.; Liao, C.; Huang, Z. Perceived Fairness of Human Managers Compared with Artificial Intelligence in Employee Performance Evaluation. *J. Manag. Inf. Syst.* **2023**, *40*, 1039–1070. [[CrossRef](#)]
5. Paetz, F.; Schultz, C.D. Smart Product Brands—The Interrelation of Smart Products and Buyer Personality Traits. *J. Retail. Consum. Serv.* **2025**, *84*, 104220. [[CrossRef](#)]
6. O’Brien, H.L.; Cairns, P.; Hall, M. A Practical Approach to Measuring User Engagement with the Refined User Engagement Scale (UES) and New UES Short Form. *Int. J. Hum. Comput. Stud.* **2018**, *112*, 28–39. [[CrossRef](#)]
7. Cheng, Z.; Fan, W.; Shao, B.; Jia, W.; Zhang, Y. The Impact of Intelligent Customer Service Agents’ Initial Response on Consumers’ Continuous Interaction Intention. *J. Retail. Consum. Serv.* **2024**, *76*, 103585. [[CrossRef](#)]
8. Cheng, X.; Zhang, X.; Cohen, J.; Mou, J. Human vs. AI: Understanding the Impact of Anthropomorphism on Consumer Response to Chatbots from the Perspective of Trust and Relationship Norms. *Inf. Process. Manag.* **2022**, *59*, 102940. [[CrossRef](#)]
9. Gray, K.; Wegner, D.M. Feeling Robots and Human Zombies: Mind Perception and the Uncanny Valley. *Cognition* **2012**, *125*, 125–130. [[CrossRef](#)] [[PubMed](#)]
10. Stein, J.P.; Liebold, B.; Ohler, P. Stay Back, Clever Thing! Linking Situational Control and Human Uniqueness Concerns to the Aversion against Autonomous Technology. *Comput. Hum. Behav.* **2019**, *95*, 73–82. [[CrossRef](#)]
11. Shank, D.B.; Graves, C.; Gott, A.; Gamez, P.; Rodriguez, S. Feeling Our Way to Machine Minds: People’s Emotions When Perceiving Mind in Artificial Intelligence. *Comput. Hum. Behav.* **2019**, *98*, 256–266. [[CrossRef](#)]
12. Turel, O.; Kalhan, S. Prejudiced against the Machine? Implicit Associations and the Transience of Algorithm Aversion. *MIS Q.* **2023**, *47*, 1369–1394. [[CrossRef](#)]
13. Rieder, A. Between Human and System Agency: Coping with Negative Incidents for Continued Effective Use of Wearables. *Inf. Syst. Res.* **2025**, *37*, 1133–1157. [[CrossRef](#)]
14. Fisher, D.; Fleming, P. Competing with Smart Machines: The Dark Side of ‘Conjoined Agency’ in Contemporary Organizations. *Organ. Stud.* **2026**, *47*, 713–737. [[CrossRef](#)]

15. Mirbabaie, M.; Brünker, F.; Möllmann Frick, N.R.J.; Stieglitz, S. The Rise of Artificial Intelligence—Understanding the AI Identity Threat at the Workplace. *Electron. Mark.* **2022**, *32*, 73–99. [\[CrossRef\]](#)
16. Lee, J.; Lee, D. User Perception and Self-Disclosure towards an AI Psychotherapy Chatbot According to the Anthropomorphism of Its Profile Picture. *Telemat. Inform.* **2023**, *85*, 102052. [\[CrossRef\]](#)
17. Gray, H.M.; Gray, K.; Wegner, D.M. Dimensions of Mind Perception. *Science* **2007**, *315*, 619. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Appel, M.; Izydorczyk, D.; Weber, S.; Mara, M.; Lischetzke, T. The Uncanny of Mind in a Machine: Humanoid Robots as Tools, Agents, and Experiencers. *Comput. Hum. Behav.* **2020**, *102*, 274–286. [\[CrossRef\]](#)
19. Lv, X.; Yang, Y.; Qin, D.; Liu, X. AI Service May Backfire: Reduced Service Warmth Due to Service Provider Transformation. *J. Retail. Consum. Serv.* **2025**, *85*, 104282. [\[CrossRef\]](#)
20. Mori, M. The Uncanny Valley. *Energy* **1970**, *7*, 33–35. [\[CrossRef\]](#)
21. Yoganathan, V.; Osburg, V.-S. The Mind in the Machine: Estimating Mind Perception’s Effect on User Satisfaction with Voice-Based Conversational Agents. *J. Bus. Res.* **2024**, *175*, 114573. [\[CrossRef\]](#)
22. Chandra, S.; Shirish, A.; Srivastava, S.C. To Be or Not to Be . . . human? Theorizing the Role of Human-like Competencies in Conversational Artificial Intelligence Agents. *J. Manag. Inf. Syst.* **2022**, *39*, 969–1005. [\[CrossRef\]](#)
23. Novak, T.P. A Generalized Framework for Moral Dilemmas Involving Autonomous Vehicles: A Commentary on Gill. *J. Consum. Res.* **2020**, *47*, 292–300. [\[CrossRef\]](#)
24. Watson, R.T.; Song, Y.; Zhao, X.; Webster, J. Extending the Foresight of Phillip Ein-Dor: Causal Knowledge Analytics. *J. Assoc. Inf. Syst.* **2024**, *25*, 145–157. [\[CrossRef\]](#)
25. Waytz, A.; Gray, K.; Epley, N.; Wegner, D.M. Causes and Consequences of Mind Perception. *Trends Cogn. Sci.* **2010**, *14*, 383–388. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Clegg, M.; Hofstetter, R.; de Bellis, E.; Schmitt, B.H. Unveiling the Mind of the Machine. *J. Consum. Res.* **2024**, *51*, 342–361. [\[CrossRef\]](#)
27. Waytz, A.; Heafner, J.; Epley, N. The Mind in the Machine: Anthropomorphism Increases Trust in an Autonomous Vehicle. *J. Exp. Soc. Psychol.* **2014**, *52*, 113–117. [\[CrossRef\]](#)
28. Lee, S.; Lee, N.; Sah, Y.J. Perceiving a Mind in a Chatbot: Effect of Mind Perception and Social Cues on Co-Presence, Closeness, and Intention to Use. *Int. J. Hum. Comput. Interact.* **2020**, *36*, 930–940. [\[CrossRef\]](#)
29. Zhou, P.; Xie, Y.; Liang, C. How to Increase Consumers’ Continued Use Intention of Artificial Intelligence Voice Assistants? The Role of Anthropomorphic Features. *Electron. Mark.* **2023**, *33*, 60. [\[CrossRef\]](#)
30. Castelo, N.; Bos, M.W.; Lehmann, D.R. Task-Dependent Algorithm Aversion. *J. Mark. Res.* **2019**, *56*, 809–825. [\[CrossRef\]](#)
31. Zhou, X.; Kim, S.; Wang, L. Money Helps When Money Feels: Money Anthropomorphism Increases Charitable Giving. *J. Consum. Res.* **2019**, *45*, 953–972. [\[CrossRef\]](#)
32. Zhu, D.H.; Chang, Y.P. Robot with Humanoid Hands Cooks Food Better?: Effect of Robotic Chef Anthropomorphism on Food Quality Prediction. *Int. J. Contemp. Hosp. Manag.* **2020**, *32*, 1367–1383. [\[CrossRef\]](#)
33. Yang, Y.; Zheng, J.; Yu, Y.; Qiu, Y.; Wang, L. The Role of Recommendation Sources and Attribute Framing in Online Product Recommendations. *J. Bus. Res.* **2024**, *174*, 114498. [\[CrossRef\]](#)
34. Yam, K.C.; Bigman, Y.E.; Tang, P.M.; Ilies, R.; Cremer, D.D.; Soh, H.; Gray, K. Robots at Work: People Prefer—And Forgive—Service Robots with Perceived Feelings. *J. Appl. Psychol.* **2021**, *106*, 1557–1572. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Spaccatini, F.; Corlito, G.; Sacchi, S. New Dyads? The Effect of Social Robots’ Anthropomorphization on Empathy towards Human Beings. *Comput. Hum. Behav.* **2023**, *146*, 107821. [\[CrossRef\]](#)
36. Hoffman, D.L.; Novak, T.P. Consumer and Object Experience in the Internet of Things: An Assemblage Theory Approach. *J. Consum. Res.* **2018**, *44*, 1178–1204. [\[CrossRef\]](#)
37. Ramadan, Z.; Farah, M.F.; El Essrawi, L. From Amazon.Com to Amazon.Love: How Alexa Is Redefining Companionship and Interdependence for People with Special Needs. *Psychol. Mark.* **2021**, *38*, 596–609. [\[CrossRef\]](#)
38. Wang, X.; Qiu, X. The Positive Effect of Artificial Intelligence Technology Transparency on Digital Endorsers: Based on the Theory of Mind Perception. *J. Retail. Consum. Serv.* **2024**, *78*, 103777. [\[CrossRef\]](#)
39. Lehrer, C.; Constantiou, I.; Matt, C.; Hess, T. How Ephemerality Features Affect User Engagement with Social Media Platforms. *MIS Q. Manag. Inf. Syst.* **2023**, *47*, 1663–1678. [\[CrossRef\]](#)
40. Kim, Y. Convolutional Neural Networks for Sentence Classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*; Association for Computational Linguistics: Doha, Qatar, 2014; pp. 1746–1751.
41. Ravichandran, T.; Deng, C. Effects of Managerial Response to Negative Reviews on Future Review Valence and Complaints. *Inf. Syst. Res.* **2022**, *34*, 319–341. [\[CrossRef\]](#)
42. Haans, R.F.J.; Pieters, C.; He, Z.-L. Thinking about U: Theorizing and Testing U- and Inverted U-Shaped Relationships in Strategy Research. *Strateg. Manag. J.* **2016**, *37*, 1177–1195. [\[CrossRef\]](#)
43. Li, G.; Li, N.; Sethi, S.P. Does CSR Reduce Idiosyncratic Risk? Roles of Operational Efficiency and AI Innovation. *Prod. Oper. Manag.* **2021**, *30*, 2027–2045. [\[CrossRef\]](#)

44. Dang, J.; Liu, L. Robots Are Friends as Well as Foes: Ambivalent Attitudes toward Mindful and Mindless AI Robots in the United States and China. *Comput. Hum. Behav.* **2021**, *115*, 106612. [[CrossRef](#)]
45. Sundaram, S.; Schwarz, A.; Jones, E.; Chin, W.W. Technology Use on the Front Line: How Information Technology Enhances Individual Performance. *J. Acad. Mark. Sci.* **2007**, *35*, 101–112. [[CrossRef](#)]
46. Lee, J.G.; Kim, K.J.; Lee, S.; Shin, D.H. Can Autonomous Vehicles Be Safe and Trustworthy? Effects of Appearance and Autonomy of Unmanned Driving Systems. *Int. J. Hum. Comput. Interact.* **2015**, *31*, 682–691. [[CrossRef](#)]
47. Yuan, L.I.; Dennis, A.R. Acting Like Humans? Anthropomorphism and Consumer's Willingness to Pay in Electronic Commerce. *J. Manag. Inf. Syst.* **2019**, *36*, 450–477. [[CrossRef](#)]
48. Mathur, M.B.; Reichling, D.B.; Lunardini, F.; Geminiani, A.; Antonietti, A.; Ruijten, P.A.M.; Levitan, C.A.; Nave, G.; Manfredi, D.; Bessette-Symons, B.; et al. Uncanny but Not Confusing: Multisite Study of Perceptual Category Confusion in the Uncanny Valley. *Comput. Hum. Behav.* **2020**, *103*, 21–30. [[CrossRef](#)]
49. Saffarizadeh, K.; Keil, M.; Boodraj, M.; Alashoor, T. "My Name Is Alexa. What's Your Name?" The Impact of Reciprocal Self-Disclosure on Post-Interaction Trust in Conversational Agents. *J. Assoc. Inf. Syst.* **2024**, *25*, 528–568. [[CrossRef](#)]
50. Schuetzler, R.M.; Grimes, G.M.; Scott Giboney, J. The Impact of Chatbot Conversational Skill on Engagement and Perceived Humanness. *J. Manag. Inf. Syst.* **2020**, *37*, 875–900. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.