Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The stock market hallucination paradox: explanation depth, conversational naturalness and epistemic trust in GenAI assistants

Upayana Wiguna Eka Saputra
upayanawiguna@pnb.ac.id
Politeknik Negeri Bali
I Putu Okta Priyana
Politeknik Negeri Bali
Ida Ayu Putri Widiasuari Riyasa
Politeknik Negeri Bali
Gede Sedana Wibawa Yasa
Politeknik Negeri Bali
Ida Bagus Gede Dananjaya
Politeknik Negeri Bali
Putu Sandra Putri Astariani
Politeknik Negeri Bali
Anak Agung Ayu Mirah Kencanawati
Politeknik Negeri Bali
Ketut Vini Elfarosa
Politeknik Negeri Bali

Research Article

Keywords: Generative AI, Epistemic trust, Hallucination risk, Value co-creation, Retail investors

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11748493/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

The stock market hallucination paradox: explanation depth, conversational naturalness and epistemic trust in GenAI assistants

ABSTRACT

Purpose – This study examines how the explanation depth and conversational naturalness of generative AI (GenAI) stock assistants relate to retail investors' perceived epistemic trust, perceived control, continuance intention and self-reported value co-creation, and how hallucination risk, overload and literacy condition these associations.

Design/methodology/approach – Drawing on cognitive appraisal and epistemic trust theories, a cross-sectional survey of 386 Indonesian retail investors was analyzed with PLS-SEM, Gaussian copulas, PLSpredict, CVPAT and importance–performance map analysis (IPMA).

Findings – Both design features were positively associated with trust and control, which were associated with continuance and co-creation. Perceived hallucination risk strengthened the explanation depth–trust association but weakened the naturalness–trust association. Overload was negatively related to control without significantly moderating the design features. Literacy moderated the trust–outcome associations in opposite directions. The explanation depth–control path was not robust to endogeneity correction.

Research limitations/implications – The cross-sectional, self-reported, single-country design precludes causal inference; trust calibration against actual AI accuracy was not assessed.

Practical implications – Providers could prioritize verifiable, layered explanations over conversational polish; regulators could consider disclosure and verification-literacy guidance.

Originality/value – The study introduces perceived stock data hallucination risk as a context-specific construct and links perceived trust and self-reported engagement to managerial priorities.

Keywords – Generative AI, Epistemic trust, Hallucination risk, Value co-creation, Retail investors

INTRODUCTION

Mobile brokerage applications, near-zero commissions and fractional trading have drawn millions of first-time retail investors into stock markets, whose participation affected liquidity during COVID-19 (Ozik *et al.*, 2021) and whose attention-driven trading is related to prices and returns (Barber *et al.*, 2022; Welch, 2022). Indonesia exemplifies this shift: capital market

investors grew from 12.17 million at the end of 2023 to about 31.1 million by August 2026 (KSEI, 2026), and individual investors are predominantly young (55.25% aged 30 or younger) (KSEI, 2026).

Retail investors increasingly ask generative artificial intelligence (GenAI) tools built on large language models (LLMs), to summarize financial statements, interpret ratios and recommend buy, sell or hold decisions. Their analytical capability is real: LLM-derived news sentiment predicts stock returns (Lopez-Lira and Tang, 2023), ChatGPT-4 stock ratings correlate with later returns (Pelster and Val, 2024), LLMs support asset selection (Ko and Lee, 2024) and give advice comparable to robo-advisors (Oehler and Horn, 2024; see also Zaremba and Demir, 2023). Yet LLMs remain prone to "hallucination", generating fluent but non-factual content, including fabricated references and figures (Alkaissi and McFarlane, 2023; Dowling and Lucey, 2023; Huang *et al.*, 2025; Ji *et al.*, 2023), which can translate directly into financial loss.

Robo-advisory research examines how investors perceive robo-advisors (Zhang *et al.*, 2021) and how design shapes reliance (Back *et al.*, 2023), but robo-advisors are rule-based and do not hallucinate. Research on algorithmic decision making finds trust central to AI acceptance (Choung *et al.*, 2023; Vereschak *et al.*, 2021), including ChatGPT adoption (Choudhury and Shamszare, 2023; Menon and Shilpa, 2023). Studies of explainable AI (XAI) and conversational agents report that explainability enhances trust (Haque *et al.*, 2023; Shin, 2021), especially in finance (Černevičienė and Kabašinskas, 2024; Weber *et al.*, 2024), and that interaction quality and human-likeness increase acceptance (Blut *et al.*, 2021; Pelau *et al.*, 2021; Rapp *et al.*, 2021). These streams share an implicit assumption: deeper explanations and smoother interactions uniformly increase trust and usage.

This assumption is problematic for GenAI stock assistants, for three reasons. First, fluency is decoupled from accuracy: conversational presentation raises perceived credibility (Anderl *et al.*, 2024) and users overestimate the accuracy of longer LLM answers (Steyvers *et al.*, 2025), so whether naturalness still relates to trust under high perceived hallucination risk is unknown. Second, explanations can increase acceptance of AI advice regardless of correctness (Bansal *et al.*, 2021), and lengthy explanations may cause overload that impairs decision making (Arnold *et al.*, 2023; Graf and Antoni, 2021). Third, prior studies treat trust as a general belief, whereas a fallible knowledge source calls for epistemic trust calibrated to actual trustworthiness (Ferrario, 2024; Heersmink *et al.*, 2024; Magnus, 2025) and for value co-creation rather than blind compliance (Bartelheimer *et al.*, 2025; Zhu *et al.*, 2022); importance–

performance map analysis (IPMA) can show where providers should invest (Hair *et al.*, 2024; Hauff *et al.*, 2024).

To address these gaps, this study integrates cognitive appraisal theory with epistemic trust theory. Explanation depth and conversational naturalness are opportunity cues, perceived hallucination risk and information overload are threat conditions, and epistemic trust and perceived control are appraisals associated with continuance and co-creation, with literacy conditioning these associations. All constructs are self-reported perceptions, so the study addresses perceived trust and self-reported engagement, not whether trust is warranted or investors verify outputs.

Perceived stock data hallucination risk differs from generic perceived risk, which captures expected loss from technology use, and from general AI trust, which captures willingness to rely: it is a threat appraisal about one failure mode, fabricated figures, signals or sources checkable against official records, that can coexist with trust. The construct is provisional and context-specific.

LITERATURE REVIEW AND HYPOTHESES DEVELOPMENT

Cognitive Appraisal Theory

Cognitive appraisal theory posits that responses to a stimulus depend on a primary appraisal of whether it is an opportunity or a threat and a secondary appraisal of whether one can cope. Challenge and threat appraisals are associated with GenAI use intention (Kinskofer and Tulis, 2025). In GenAI-assisted investing, analytical richness and conversational ease are opportunity cues, whereas possible fabrication and excessive information are threat cues.

Epistemic Trust in Large Language Models

Epistemic trust is the willingness to accept information from a source as a basis for one's own beliefs. Users should match their trust in LLMs to actual trustworthiness, but opacity and fluent, human-like interaction invite unwarranted trust (Heersmink *et al.*, 2024), and justified trust must be anchored in evidence of reliability (Ferrario, 2024). Epistemic trust should therefore rise with verifiable cues and fall with cues of unreliability; this study measures perceived epistemic trust only.

Value Co-creation in Human–AI Investing

Service-dominant logic holds that value is co-created through actors' resource integration, and hybrid intelligent service ecosystems configure human and artificial agencies to do so

(Bartelheimer *et al.*, 2025). In GenAI-assisted investing, co-creation occurs when investors contribute context, verify outputs against official sources and combine AI analysis with their own judgment; here it is self-reported.

Explanation Depth, Conversational Naturalness, and Epistemic Trust

Explanation depth is the extent to which a GenAI assistant discloses the data, calculations, reasoning and risks underlying its stock analysis. Users can calibrate trust in an opaque LLM only with cues about the source's competence (Heersmink *et al.*, 2024), and explanations supply such cues by making an output inspectable. Explainability and causability are associated with higher trust (Haque *et al.*, 2023; Shin, 2021), and in finance with accountability (Černevičienė and Kabašinskas, 2024; Weber *et al.*, 2024).

H1. Explanation depth is positively associated with epistemic trust in stock AI.

Conversational naturalness is the perceived human-likeness, coherence and contextual responsiveness of the interaction. Fluent LLMs are experienced as a "quasi-other" (Heersmink *et al.*, 2024), so communicative competence is read as epistemic competence. Anthropomorphism is associated with greater intention to use AI service agents (Blut *et al.*, 2021), and conversational presentation with higher credibility judgments (Anderl *et al.*, 2024).

H2. Conversational naturalness is positively associated with epistemic trust in stock AI.

Explanation Depth, Conversational Naturalness, and Perceived Control

Perceived control is the investor's belief that they retain the authority and capacity to decide while using a GenAI assistant, the secondary appraisal of whether they can cope with the technology rather than be displaced by it (Kinskofer and Tulis, 2025). Explanation depth should strengthen it by making AI output decomposable: investors who see which assumptions drive a valuation can accept some elements and reject others.

H3. Explanation depth is positively associated with perceived control.

Conversational naturalness may contribute to perceived control through steerability: a system that understands follow-up questions lets investors direct the analysis, for example by rerunning a valuation under a lower growth assumption (Rapp *et al.*, 2021).

H4. Conversational naturalness is positively associated with perceived control.

Epistemic Trust, Perceived Control, Continuance, and Value Co-creation

Trust is among the most consistent correlates of AI use: it is associated with usage intention (Choung *et al.*, 2023), ChatGPT use (Choudhury and Shamszare, 2023) and continuance in digital financial services (Nguyen and Dao, 2024). Because stock analysis is recurrent, epistemic trust lowers the cost of drawing on the assistant in each episode.

> **H5.** Epistemic trust is positively associated with continuance intention to use stock AI.

Because epistemic trust should be calibrated rather than blind, its consequence may extend beyond continued use. An investor who regards the assistant as credible but fallible has reason to treat its output as an input to be checked (Ferrario, 2024), which is rational given that ChatGPT's stock ratings are informative (Pelster and Val, 2024) yet LLM output can be fabricated (Huang *et al.*, 2025). Such investors are expected to supply context, verify figures and integrate the AI's analysis with their own, that is, to co-create value (Bartelheimer *et al.*, 2025; Zhu *et al.*, 2022).

> **H6.** Epistemic trust is positively associated with self-reported value co-creation behavior.

Perceived control should likewise sustain use: when investors appraise the assistant as augmenting rather than usurping their judgment, challenge appraisals are positively associated with GenAI use intention (Kinskofer and Tulis, 2025).

> **H7.** Perceived control is positively associated with continuance intention to use stock AI.

Perceived control may also be a precondition for co-creation: contesting an AI's conclusion or requesting an alternative scenario requires confidence that one is entitled and able to do so, and analytical engagement is associated with less overreliance (Buçinca *et al.*, 2021).

> **H8.** Perceived control is positively associated with self-reported value co-creation behavior.

Mediating Roles of Epistemic Trust and Perceived Control

Cognitive appraisal theory implies a stimulus–appraisal–response sequence (Kinskofer and Tulis, 2025), consistent with evidence that anthropomorphic design is related to usage through users' perceptions of the agent (Blut *et al.*, 2021) and that trust links AI characteristics to

acceptance (Choung *et al.*, 2023). The design features are therefore expected to relate to continuance and co-creation via the two appraisals.

> **H9a–d.** Epistemic trust statistically mediates the associations of (a) explanation depth and (b) conversational naturalness with continuance intention, and of (c) explanation depth and (d) conversational naturalness with value co-creation behavior.

> **H10a–b.** Perceived control statistically mediates the associations of (a) explanation depth and (b) conversational naturalness with value co-creation behavior.

The Hallucination Paradox

Perceived stock data hallucination risk is an investor's belief that a GenAI assistant may generate fabricated or inaccurate financial figures, technical signals, sources or corporate events, reflecting factuality hallucinations documented in LLM research (Alkaissi and McFarlane, 2023; Huang *et al.*, 2025; Ji *et al.*, 2023). As a threat cue, it should alter how the two opportunity cues are appraised, in opposite directions.

When hallucination risk is salient, coping requires verification, and explanations that disclose sources and calculations make claims checkable: they reduce overreliance when they lower verification costs (Vasconcelos *et al.*, 2023). Under low perceived risk, explanations tend to act as persuasive rather than diagnostic cues, increasing acceptance regardless of correctness (Bansal *et al.*, 2021).

> **H11.** Perceived stock data hallucination risk strengthens the positive association between explanation depth and epistemic trust.

Conversational naturalness should follow the opposite pattern. Conversational presentation raises credibility (Anderl *et al.*, 2024), and users overestimate the accuracy of elaborate LLM answers (Steyvers *et al.*, 2025); investors aware that LLMs can be confidently wrong should recognize that fluency says nothing about accuracy, and human-like cues can backfire (Crolic *et al.*, 2022).

> **H12.** Perceived stock data hallucination risk weakens the positive association between conversational naturalness and epistemic trust.

The Overload Paradox

Information overload occurs when information exceeds processing capacity, impairing decision quality (Arnold *et al.*, 2023); quantity, complexity and poor structure are among its strongest antecedents (Graf and Antoni, 2021). Longer LLM explanations raise confidence without improving discrimination between correct and incorrect answers (Steyvers *et al.*, 2025).

> **H13.** Information overload weakens the positive association between explanation depth and epistemic trust.

Overload also depletes the coping resources on which perceived control rests: an investor facing more ratios and scenarios than they can integrate is less able to decide what to accept, and an unstructured stream of dialogue can undermine the sense of directing the analysis.

> **H14.** Information overload weakens the positive association between explanation depth and perceived control.

> **H15.** Information overload weakens the positive association between conversational naturalness and perceived control.

The Role of Financial and Stock Literacy

Financial literacy, the knowledge and skills needed to process financial information, is associated with better financial outcomes (Lusardi and Mitchell, 2023) and with reliance on algorithms (Mahmud *et al.*, 2022), and is especially relevant in Indonesia's young and educationally diverse investor base (KSEI, 2026).

Literate investors possess alternative analytical resources, so their continued use should depend less on trust, whereas less literate investors, who cannot readily assess output themselves, should rely more on trust.

> **H16.** Financial and stock literacy weakens the positive association between epistemic trust and continuance intention.

Conversely, literacy supplies the skills needed to turn trust into co-creation: verifying an earnings figure or challenging a valuation assumption requires financial knowledge.

> **H17.** Financial and stock literacy strengthens the positive association between epistemic trust and value co-creation behavior.

Figure 1 presents the conceptual model.

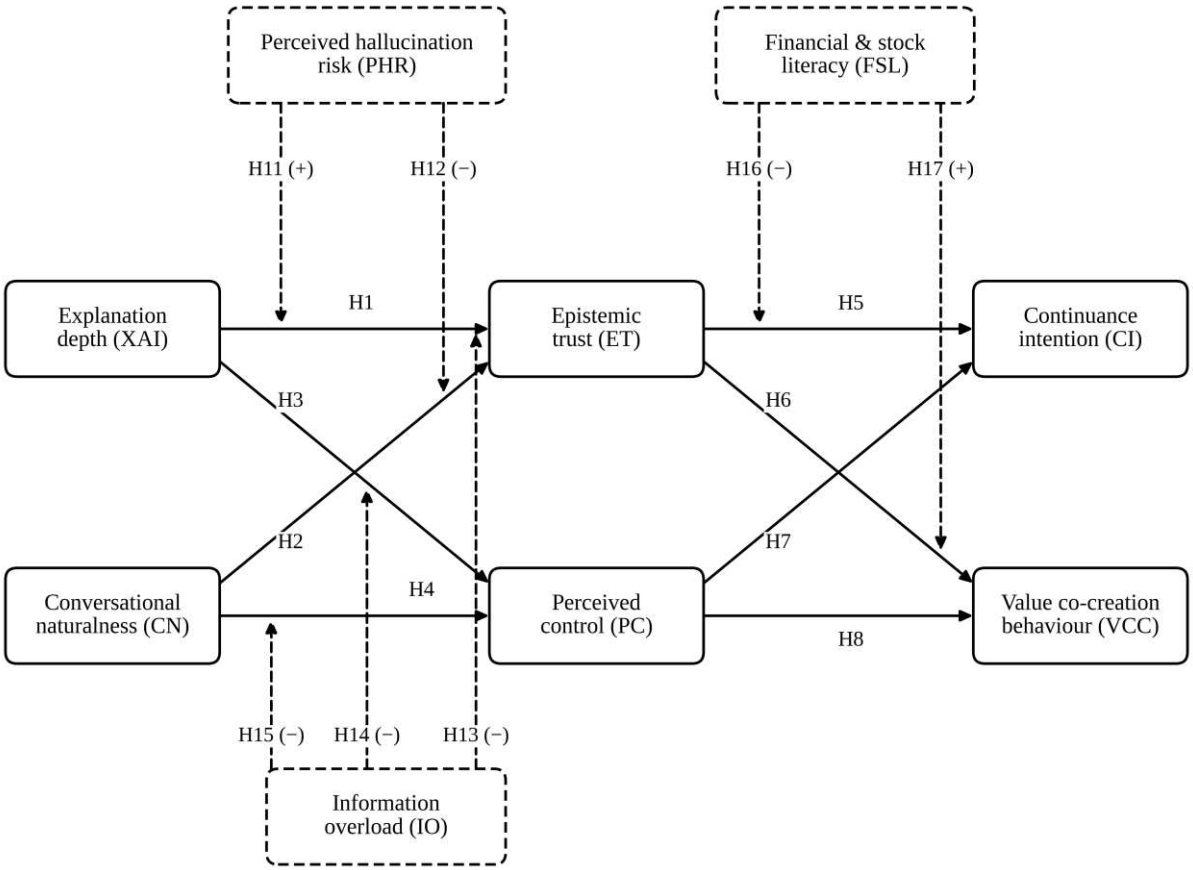

Figure 1. Conceptual model

Note: Solid arrows denote direct effects; dashed arrows denote moderating effects. Mediation hypotheses (H9a–d and H10a–b) operate through epistemic trust and perceived control

METHODS

Research Design

This study adopts a quantitative, explanatory, cross-sectional survey design in which all variables are self-reported at one time point. PLS-SEM suits the explanatory and predictive objective, the multiple interaction terms and the new construct, and IPMA builds on its latent variable scores (Becker *et al.*, 2023; Hair *et al.*, 2022, 2024; Sarstedt *et al.*, 2022). "Effect" denotes a standardized path or conditional coefficient, that is, a conditional association, not a demonstrated causal effect.

Population, Sampling, and Data Collection

The target population comprises Indonesian retail stock investors who use GenAI tools for stock analysis. Without a sampling frame, purposive sampling recruited respondents through investor communities, capital market school programs, university investment galleries and social media. Eligible respondents were aged 18 or above, held an active securities account,

had traded stocks in the past three months and had used GenAI for stock analysis at least twice a month for three months; licensed professionals were excluded (Table A1). Data were collected between December 2025 and August 2026.

Sample Size

Sample size was determined with two methods. A power analysis in G*Power (Kang, 2021) for the endogenous construct with the most predictors (epistemic trust: seven, including interaction terms), with $f^2 = 0.15$, $\alpha = 0.05$ and power = 0.95, yielded a minimum of 153 respondents; the inverse square root method (Hair *et al.*, 2022), assuming a minimum path coefficient of 0.15, required $n > 275$. The final sample of 386 exceeds *both. A sensitivity computation (F test of one coefficient, seven predictors,* $\alpha = 0.05$*, power = 0.80) indicates that the sample could detect* $f^2 \approx 0.020$ *or larger.*

Measures

All constructs were measured reflectively on five-point Likert scales (1 = strongly disagree; 5 = strongly agree). Items were developed for this study from construct definitions in recent literature (Table A2). Perceived stock data hallucination risk is a new construct grounded in the hallucination taxonomies of Huang et al. (2025) and Ji et al. (2023); it was examined only within this sample and is treated as provisional. Items were translated between English and Indonesian following back-translation recommendations (Klotz *et al.*, 2023), and a pilot test with 30 respondents preceded the main survey.

Common Method Bias

Because all constructs were self-reported by the same respondents, common method bias was addressed procedurally (anonymity, separate sections for predictors and criteria, randomized item order, pilot-refined wording and attention checks) and statistically (Podsakoff *et al.*, 2024): full collinearity VIFs (1.099–1.701) were below 3.3, although residual bias in same-source self-reports cannot be excluded.

Data Analysis

All analyses were performed in SmartPLS 4.3 (Ringle *et al.*, 2024) with the path weighting scheme. The measurement model was evaluated through loadings, reliability, convergent validity and discriminant validity (HTMT and HTMT2 (Hair *et al.*, 2022; Roemer *et al.*, 2021)). The structural model was assessed through collinearity, percentile bootstrapping with 5,000 subsamples, R^2 and f^2; effects were deemed significant when the 95% interval excluded zero,

and nonsignificant results were interpreted through confidence intervals, not observed power. Mediation was tested with bootstrapped specific indirect effects, moderation with the two-stage approach and simple slopes at ±1 SD (Becker *et al.*, 2023; Hair *et al.*, 2022), endogeneity with Gaussian copulas (Becker *et al.*, 2022; Hair *et al.*, 2024) and predictive power with PLSpredict and CVPAT (Liengaard *et al.*, 2021; Sarstedt *et al.*, 2022) (Table 1).

Finally, an IPMA contrasted each predecessor's importance (total effect) with its performance (mean latent variable score rescaled to 0–100) for continuance intention and value co-creation behavior (Hair *et al.*, 2024; Hauff *et al.*, 2024). Quadrant thresholds are sample means, so quadrants indicate relative priorities only; the negatively valenced hallucination risk and overload were excluded from the quadrants.

Table 1. Summary of assessment criteria

Stage	Criterion	Threshold
Measurement model	Indicator loading	≥ 0.708
	Internal consistency (α, ρA, ρC)	0.70–0.95
	Convergent validity (AVE)	≥ 0.50
	Discriminant validity (HTMT/HTMT2)	< 0.85 (conservative) or < 0.90
Structural model	Inner VIF	< 3
	Path, indirect and conditional effects	95% percentile bootstrap CI excludes zero (5,000 subsamples)
	f^2 (main effects)	0.02 / 0.15 / 0.35
	f^2 (interaction effects)	0.005 / 0.01 / 0.025
	Minimum detectable effect (sensitivity)	$f^2 \approx 0.020$ at power = 0.80
Robustness	Endogeneity (Gaussian copula)	Copula term not significant
Prediction	Predictive power	PLSpredict (Q^2predict > 0; RMSE vs LM); CVPAT (ALD < 0)
Prioritization	IPMA	Importance = total effect; performance = rescaled latent variable score (0–100)

Source: Adapted from Hair *et al.* (2022, 2024), Becker *et al.* (2022, 2023), Roemer *et al.* (2021), Liengaard *et al.* (2021) and Hauff *et al.* (2024)

RESULTS

Sample Profile and Data Screening

Of the 498 questionnaires returned, 112 (22.49%) were excluded, leaving 386 valid responses: n = 62 (55.36% of exclusions) failed the screening questions (Table A1), 33 (29.46%) failed the attention check and 17 (15.18%) were straight-liners; each excluded case is counted once, and none was excluded as a multivariate outlier. Straight-lining suggests inattentive responding (Hair *et al.*, 2022) and can distort correlations and add shared method variance (Podsakoff *et al.*, 2024). The retained sample exceeds both minimum sample sizes.

Table 2 presents the respondent profile. Most respondents were men (64.51%), aged 30 or younger (55.44%; mean age = 32.8, SD = 9.4) and held a bachelor's degree (44.30%); 53.37% had over three years of investment experience. Claude (33.68%) and ChatGPT (31.09%) were the most used GenAI tools.

The data contained no missing values; skewness and excess kurtosis lay within ±1, but Cramér–von Mises tests rejected normality for all latent variable scores ($p \leq 0.024$), supporting bootstrap-based inference (Hair *et al.*, 2022; Sarstedt *et al.*, 2022) and the Gaussian copula requirement (Becker *et al.*, 2022).

Table 2. Respondent profile (n = 386)

Category	*n*	%
Gender		
Male	249	64.51
Female	137	35.49
Age (years); mean = 32.8, SD = 9.4, range 18–67		
18–30	214	55.44
31–40	98	25.39
41–50	50	12.95
> 50	24	6.22
Education		
Junior high school or below	12	3.11
Senior high school	119	30.83
Diploma (D1–D3)	52	13.47
Bachelor's degree (S1/D4)	171	44.30
Master's degree	28	7.25
Doctoral degree	4	1.04
Stock investment experience		
< 1 year	70	18.13
1–3 years	110	28.50
3–5 years	100	25.91
> 5 years	106	27.46
Portfolio value		
< IDR 10 million	100	25.91
IDR 10–50 million	125	32.38
IDR 50–100 million	95	24.61
> IDR 100 million	66	17.10
Investment style		
Fundamental	145	37.56
Technical	110	28.50
Mixed	131	33.94

Category	*n*	%
GenAI tool used most often for stock analysis		
ChatGPT	120	31.09
Gemini	50	12.95
Claude	130	33.68
AI feature embedded in a brokerage app	86	22.28
Total	**386**	**100.00**

Note: SD = standard deviation.

Measurement Model

All nine constructs were reflective. Loadings exceeded 0.708 (0.738–0.877; Table 3), Cronbach's α (0.792–0.910), ρA and ρC (0.802–0.933) lay within 0.70–0.95 (Hair *et al.*, 2022), AVE values were 0.615–0.735 and outer VIFs 1.469–2.592.

Discriminant validity was assessed with the HTMT ratio (Hair *et al.*, 2022): all values were below 0.85 (maximum 0.500, ET–VCC; Table 4), HTMT2 (Roemer *et al.*, 2021) agreed (maximum 0.493) and the Fornell–Larcker criterion was met. Perceived hallucination risk correlated only weakly with epistemic trust ($r = -0.159$) and was virtually uncorrelated with conversational naturalness and perceived control (both $r = 0.015$), suggesting a threat appraisal distinct from low trust (within-sample evidence only).

Table 3. Measurement model results

Construct	Item	Loading	VIF	α	ρA	ρC	AVE
Explanation depth (XAI)	XAI1	0.851	2.476	0.910	0.912	0.933	0.735
	XAI2	0.843	2.358				
	XAI3	0.865	2.521				
	XAI4	0.862	2.524				
	XAI5	0.867	2.592				
Conversational naturalness (CN)	CN1	0.738	1.559	0.854	0.859	0.896	0.632
	CN2	0.818	1.890				
	CN3	0.769	1.725				
	CN4	0.832	2.051				
	CN5	0.815	2.006				
Perceived stock data hallucination risk (PHR)	PHR1	0.788	1.769	0.885	0.925	0.912	0.676
	PHR2	0.856	2.415				
	PHR3	0.829	2.225				
	PHR4	0.756	2.092				
	PHR5	0.876	2.198				
Information overload (IO)	IO1	0.829	1.600	0.808	0.840	0.871	0.629
	IO2	0.750	1.652				
	IO3	0.824	1.969				

Construct	Item	Loading	VIF	α	ρA	ρC	AVE
	IO4	0.766	1.558				
Epistemic trust in stock AI (ET)	ET1	0.833	2.089	0.876	0.881	0.909	0.668
	ET2	0.817	2.009				
	ET3	0.795	1.859				
	ET4	0.798	1.942				
	ET5	0.842	2.111				
Perceived control (PC)	PC1	0.794	1.615	0.792	0.802	0.864	0.615
	PC2	0.829	1.692				
	PC3	0.742	1.543				
	PC4	0.767	1.469				
Continuance intention (CI)	CI1	0.852	2.138	0.859	0.865	0.905	0.704
	CI2	0.853	2.168				
	CI3	0.877	2.331				
	CI4	0.770	1.638				
Value co-creation behavior (VCC)	VCC1	0.821	1.983	0.858	0.865	0.898	0.637
	VCC2	0.841	2.136				
	VCC3	0.792	1.870				
	VCC4	0.791	1.731				
	VCC5	0.743	1.696				
Financial and stock literacy (FSL)	FSL1	0.813	2.007	0.877	0.888	0.910	0.670
	FSL2	0.855	2.226				
	FSL3	0.803	2.000				
	FSL4	0.849	2.211				
	FSL5	0.770	1.770				

Note: All loadings are significant at $p < 0.001$ (bootstrapping, 5,000 subsamples). α = Cronbach's alpha; ρA and ρC = composite reliability; AVE = average variance extracted

Table 4. Descriptive statistics, correlations and discriminant validity

Construct	Mean	SD	XAI	CN	PHR	IO	ET	PC	CI	VCC	FSL
XAI	3.751	0.836	**.858**	*.343*	*.153*	*.170*	*.416*	*.385*	*.273*	*.307*	*.153*
CN	3.974	0.708	.304	**.795**	*.061*	*.050*	*.370*	*.342*	*.160*	*.222*	*.079*
PHR	3.979	0.759	.135	.015	**.822**	*.198*	*.161*	*.056*	*.041*	*.036*	*.133*
IO	3.416	0.793	.148	.011	.159	**.793**	*.061*	*.119*	*.067*	*.098*	*.162*
ET	3.670	0.816	.375	.327	−.159	.011	**.817**	*.290*	*.487*	*.500*	*.159*
PC	4.016	0.672	.336	.286	.015	−.098	.244	**.784**	*.328*	*.444*	*.200*
CI	3.855	0.809	.241	.141	−.011	−.039	.424	.277	**.839**	*.271*	*.140*
VCC	3.736	0.769	.273	.192	−.022	−.068	.441	.371	.234	**.798**	*.235*
FSL	3.664	0.808	.140	.063	.125	−.127	.142	.168	.120	.209	**.819**

Note: Bold diagonal values are the square roots of the AVE; values below the diagonal are latent variable correlations; italic values above the diagonal are HTMT ratios. Means and standard deviations are based on unweighted item averages (five-point scale). Maximum HTMT2 = 0.493 (ET–VCC)

Structural Model

Inner VIFs (1.010–1.159) indicated no collinearity (Becker *et al.*, 2023), and SRMR was 0.045 (saturated) and 0.049 (estimated), acceptable but interpreted cautiously (Hair *et al.*, 2022). The

model explained 26.6% of the variance in epistemic trust, 17.8% in perceived control, 23.3% in continuance intention and 32.9% in value co-creation (Figure 2), modest values from a parsimonious model.

All eight direct-effect hypotheses were supported in the baseline model (Table 5), but H3 was not robust to the endogeneity correction. Explanation depth was positively associated with trust ($\beta = 0.356$; $f^2 = 0.149$) and control ($\beta = 0.299$; $f^2 = 0.096$), and naturalness with trust ($\beta = 0.210$; $f^2 = 0.054$) and control ($\beta = 0.192$; $f^2 = 0.041$; all $p < 0.001$).

Epistemic trust showed the strongest associations with continuance ($\beta = 0.380$) and co-creation ($\beta = 0.353$), and perceived control with continuance ($\beta = 0.184$) and co-creation ($\beta = 0.257$; all $p < 0.001$; $f^2 = 0.041$–0.175), consistent with H5–H8.

Among the moderators' direct associations, hallucination risk was negatively related to trust ($\beta = -0.191$, $p < 0.001$), overload to control ($\beta = -0.142$, $p = 0.005$) but not trust, and literacy positively to co-creation ($\beta = 0.135$, $p = 0.004$) but not continuance.

Table 5. Structural model results: direct and interaction effects

Hyp.	Path	*β*	SD	*t*	*p*	95% CI	f^2	Decision
Direct effects								
H1	XAI → ET	0.356	0.044	8.121	< 0.001	[0.266, 0.443]	0.149	Supported
H2	CN → ET	0.210	0.047	4.445	< 0.001	[0.117, 0.301]	0.054	Supported
H3	XAI → PC	0.299	0.048	6.181	< 0.001	[0.203, 0.393]	0.096	Supported[a]
H4	CN → PC	0.192	0.052	3.713	< 0.001	[0.092, 0.298]	0.041	Supported
H5	ET → CI	0.380	0.044	8.541	< 0.001	[0.292, 0.467]	0.175	Supported
H6	ET → VCC	0.353	0.045	7.859	< 0.001	[0.263, 0.440]	0.173	Supported
H7	PC → CI	0.184	0.046	4.000	< 0.001	[0.092, 0.274]	0.041	Supported
H8	PC → VCC	0.257	0.047	5.496	< 0.001	[0.165, 0.350]	0.091	Supported
Interaction effects								
H11	PHR × XAI → ET	0.119	0.037	3.213	0.001	[0.047, 0.193]	0.021	Supported
H12	PHR × CN → ET	−0.137	0.050	2.754	0.006	[−0.234, −0.038]	0.020	Supported
H13	IO × XAI → ET	−0.099	0.054	1.848	0.065	[−0.192, 0.020]	0.011	Not supported
H14	IO × XAI → PC	−0.078	0.059	1.323	0.186	[−0.187, 0.044]	0.006	Not supported
H15	IO × CN → PC	−0.036	0.051	0.711	0.477	[−0.129, 0.073]	0.001	Not supported
H16	FSL × ET → CI	−0.149	0.049	3.043	0.002	[−0.244, −0.052]	0.026	Supported
H17	FSL × ET → VCC	0.232	0.044	5.319	< 0.001	[0.141, 0.309]	0.071	Supported
Direct effects of moderators								
–	PHR → ET	−0.191	0.039	4.908	< 0.001	[−0.272, −0.126]	0.047	–
–	IO → ET	−0.005	0.065	0.084	0.933	[−0.132, 0.109]	0.000	–
–	IO → PC	−0.142	0.051	2.779	0.005	[−0.232, −0.026]	0.024	–
–	FSL → CI	0.024	0.050	0.473	0.636	[−0.076, 0.121]	0.001	–
–	FSL → VCC	0.135	0.047	2.898	0.004	[0.045, 0.229]	0.026	–

Note: β = standardized path coefficient; SD = bootstrap standard deviation; CI = percentile bootstrap confidence interval (5,000 subsamples, two-tailed). R^2: ET = 0.266, PC = 0.178, CI = 0.233, VCC = 0.329; adjusted R^2: 0.252, 0.167, 0.225 and 0.322, respectively. SRMR = 0.045 (saturated model). [a]Supported in the baseline model but not robust to the Gaussian-copula correction (see Endogeneity Assessment)

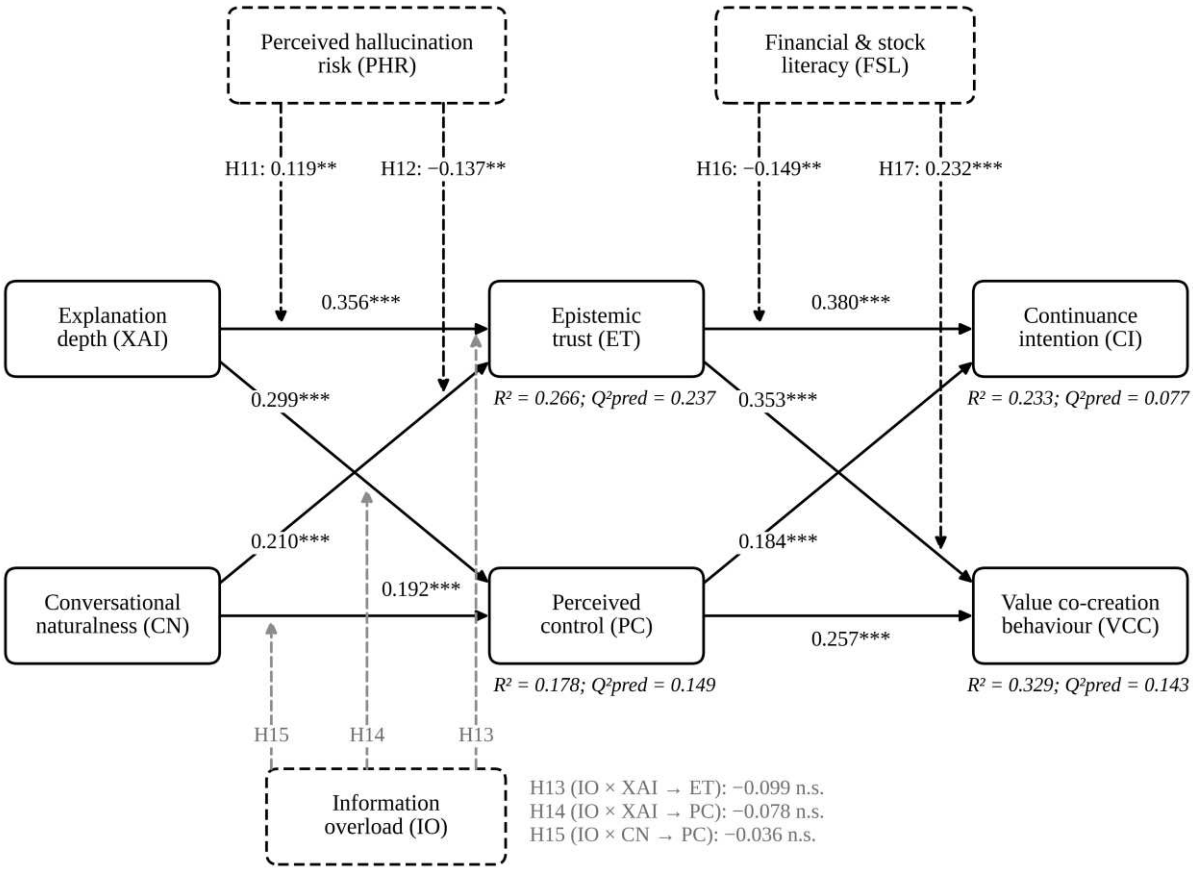

Figure 2. Results of the structural model

Note: Standardized path coefficients; ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$; n.s. = not significant (gray). Solid arrows denote direct effects; dashed arrows denote interaction effects. R^2 = coefficient of determination; Q^2pred = construct-level Q^2predict. Direct effects of the moderators are reported in Table 5

Mediation Analysis

All six hypothesized indirect effects were significant (Table 6), consistent with H9a–d and H10a–b, although H10a depends on the H3 path. Trust carried the associations of explanation depth (continuance: β = 0.135; co-creation: 0.126) and naturalness (0.080; 0.074) with the outcomes, and control those with co-creation (0.077; 0.050). Because the model has no direct paths from the design features to the outcomes, indirect effects equal total effects.

Table 6. Mediation analysis: specific indirect effects

Hyp.	Indirect path	*β*	SD	*t*	*p*	95% CI	Decision
Via epistemic trust							
H9a	XAI → ET → CI	0.135	0.023	5.849	< 0.001	[0.093, 0.182]	Supported
H9b	CN → ET → CI	0.080	0.020	3.987	< 0.001	[0.042, 0.121]	Supported
H9c	XAI → ET → VCC	0.126	0.023	5.369	< 0.001	[0.084, 0.174]	Supported
H9d	CN → ET → VCC	0.074	0.019	3.999	< 0.001	[0.040, 0.112]	Supported
Via perceived control							
H10a	XAI → PC → VCC	0.077	0.019	4.012	< 0.001	[0.043, 0.117]	Supported[a]
H10b	CN → PC → VCC	0.050	0.015	3.209	0.001	[0.023, 0.084]	Supported
–	XAI → PC → CI	0.055	0.017	3.273	0.001	[0.025, 0.090]	Not hypothesized
–	CN → PC → CI	0.035	0.013	2.783	0.005	[0.014, 0.064]	Not hypothesized

Note: Percentile bootstrap confidence intervals (5,000 subsamples); indirect effects via epistemic trust are evaluated at the mean of financial and stock literacy. [a]Depends on the H3 path, which was not robust to the Gaussian-copula correction

Moderation and Moderated Mediation

Interactions were probed with simple slopes at ±1 SD of each moderator (Becker *et al.*, 2023) (Table 7; Figure 3).

The hallucination paradox. Both hallucination-risk hypotheses were supported. Risk strengthened the explanation depth–trust association ($\beta = 0.119$, $p = 0.001$; $f^2 = 0.021$), from 0.237 at low risk to 0.475 at high risk, and weakened the naturalness–trust association ($\beta = -0.137$, $p = 0.006$; $f^2 = 0.020$), from 0.347 at low risk to a nonsignificant 0.073 at high risk (Figure 3a–b). Trust was thus more strongly associated with how the assistant talks than with what it shows at low perceived risk, and this ordering inverted at high perceived risk.

The overload paradox. None of the three overload hypotheses was supported (H13: $\beta = -0.099$, $p = 0.065$; H14: $\beta = -0.078$, $p = 0.186$; H15: $\beta = -0.036$, $p = 0.477$). These results are inconclusive rather than evidence of no moderation: the sample could reliably detect only $f^2 \geq 0.020$, whereas the observed interaction effects were smaller ($f^2 = 0.001$–0.011).

Literacy as a moderator. Both literacy hypotheses were supported. Literacy weakened the trust–continuance association ($\beta = -0.149$, $p = 0.002$; $f^2 = 0.026$) and strengthened the trust–co-creation association ($\beta = 0.232$, $p < 0.001$; $f^2 = 0.071$). At low literacy, trust was strongly associated with continuance ($\beta = 0.528$) but not distinguishably with co-creation ($\beta = 0.121$); at high literacy the pattern reversed (0.231 and 0.585; Figure 3f–g).

Moderated mediation. Four of ten indices were significant (Table 7, Panel B), all involving hallucination risk, which strengthened the indirect associations of explanation depth via trust (continuance: 0.045; co-creation: 0.042) and weakened those of naturalness (−0.052; −0.048). Literacy changed the conditional indirect associations (Panel C): explanation depth with co-

creation rose from a nonsignificant 0.043 at low literacy to 0.209 at high literacy, whereas with continuance it fell from 0.188 to 0.082.

Table 7. Conditional effects and indices of moderated mediation

Effect	**Moderator at −1 SD**	**Moderator at mean**	**Moderator at +1 SD**
Panel A: conditional direct effects (simple slopes)			
XAI → ET by PHR (H11)	0.237 [0.140, 0.329]	0.356 [0.266, 0.443]	0.475 [0.344, 0.606]
CN → ET by PHR (H12)	0.347 [0.224, 0.460]	0.210 [0.117, 0.301]	0.073 [−0.076, 0.219][a]
XAI → ET by IO (H13)	0.456 [0.292, 0.587]	0.356 [0.266, 0.443]	0.257 [0.145, 0.392]
XAI → PC by IO (H14)	0.377 [0.220, 0.528]	0.299 [0.203, 0.393]	0.221 [0.083, 0.373]
CN → PC by IO (H15)	0.229 [0.091, 0.359]	0.192 [0.092, 0.298]	0.156 [0.016, 0.315]
ET → CI by FSL (H16)	0.528 [0.412, 0.646]	0.380 [0.292, 0.467]	0.231 [0.093, 0.376]
ET → VCC by FSL (H17)	0.121 [−0.006, 0.261][a]	0.353 [0.263, 0.440]	0.585 [0.467, 0.690]
Panel B: indices of moderated mediation			
Moderated indirect path	*Index*	*95% CI*	*Conclusion*
PHR × XAI → ET → CI	0.045	[0.018, 0.077]	Significant
PHR × XAI → ET → VCC	0.042	[0.016, 0.073]	Significant
PHR × CN → ET → CI	−0.052	[−0.094, −0.014]	Significant
PHR × CN → ET → VCC	−0.048	[−0.085, −0.013]	Significant
IO × XAI → ET → CI	−0.038	[−0.076, 0.008]	Not significant
IO × XAI → ET → VCC	−0.035	[−0.070, 0.007]	Not significant
IO × XAI → PC → CI	−0.014	[−0.037, 0.008]	Not significant
IO × XAI → PC → VCC	−0.020	[−0.050, 0.012]	Not significant
IO × CN → PC → CI	−0.007	[−0.026, 0.014]	Not significant
IO × CN → PC → VCC	−0.009	[−0.037, 0.018]	Not significant
Panel C: conditional indirect effects via ET at levels of FSL (PHR and IO at mean)			
XAI → ET → CI	0.188 [0.128, 0.254]	0.135 [0.093, 0.182]	0.082 [0.032, 0.139]
XAI → ET → VCC	0.043 [−0.002, 0.097][a]	0.126 [0.084, 0.174]	0.209 [0.146, 0.275]
CN → ET → CI	0.111 [0.059, 0.167]	0.080 [0.042, 0.121]	0.048 [0.018, 0.090]
CN → ET → VCC	0.025 [−0.001, 0.059][a]	0.074 [0.040, 0.112]	0.123 [0.067, 0.183]

Note: Conditional effects with 95% percentile bootstrap CIs in brackets. In Panel A, for XAI → ET, the second moderator is held at its mean. [a]Not significant (CI includes zero). Indices in Panel B are evaluated at the mean of financial and stock literacy

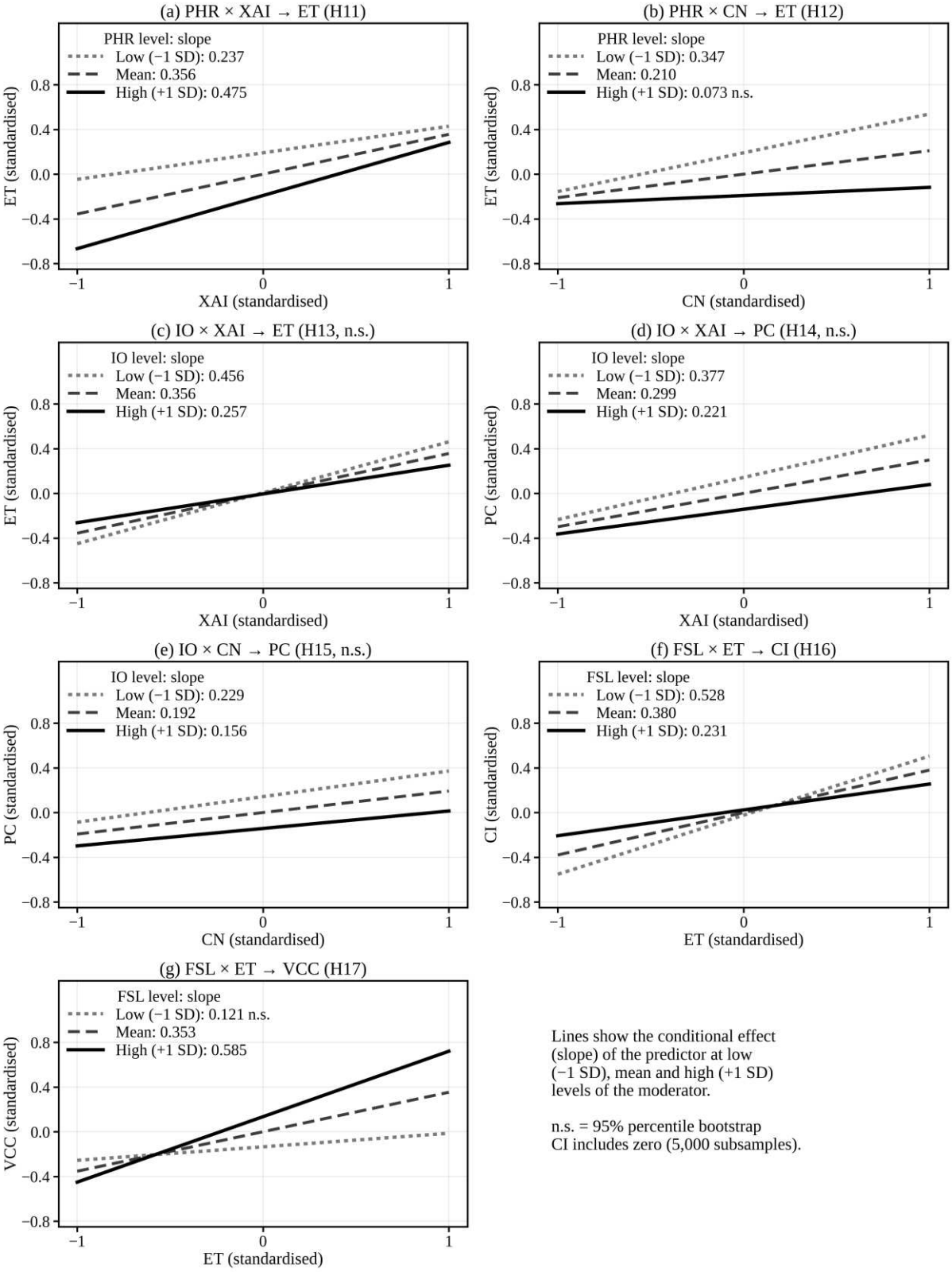

Figure 3. Simple slope analysis

Note: Lines show the effect of the predictor at low (−1 SD), mean and high (+1 SD) levels of the moderator; slopes are the conditional effects in Table 7. Panels (c)–(e) show non-significant interactions, and differences between their lines must not be interpreted as moderation; the fitted lines include the direct effect of the moderator

Endogeneity Assessment

Because cross-sectional paths could be biased by omitted variables such as self-efficacy, Gaussian copula terms were added for the eight paths linking design features, appraisals and outcomes (Becker *et al.*, 2022; Hair *et al.*, 2024); non-normality of all regressors satisfied the identification requirement. Seven copula terms were not significant, but the term for explanation depth → perceived control was (0.380, [0.089, 0.602]), and the corrected path was no longer significant ($\beta = -0.033$, [−0.254, 0.239]). This estimate is imprecise; support for H3, and for H10a, which depends on that path, is weakened, and both are treated as conditionally supported only.

Predictive Power

PLSpredict (10 folds, 10 repetitions) gave positive Q^2predict values for all 18 indicators (0.041–0.201), and PLS-SEM beat the linear benchmark on RMSE and MAE for all of them (Hair et al., 2022; Table 8); construct-level Q^2predict ranged from 0.077 (CI) to 0.237 (ET). CVPAT showed significantly lower average loss than the indicator average (−0.093) and the linear benchmark (−0.062; both $p < 0.001$) (Liengaard et al., 2021; Table 9).

Table 8. PLSpredict results

Indicator	Q^2predict	RMSE			MAE		
		PLS-SEM	LM	Δ	PLS-SEM	LM	Δ
ET1	0.178	0.932	0.965	−0.033	0.731	0.774	−0.042
ET2	0.164	0.884	0.914	−0.030	0.709	0.721	−0.012
ET3	0.147	0.917	0.950	−0.034	0.735	0.762	−0.027
ET4	0.091	0.942	0.983	−0.041	0.747	0.784	−0.037
ET5	0.201	0.916	0.946	−0.029	0.732	0.763	−0.032
PC1	0.071	0.807	0.842	−0.035	0.642	0.674	−0.032
PC2	0.121	0.848	0.874	−0.026	0.653	0.674	−0.021
PC3	0.043	0.819	0.853	−0.034	0.641	0.667	−0.027
PC4	0.122	0.794	0.812	−0.017	0.619	0.647	−0.029
CI1	0.065	0.951	0.963	−0.012	0.756	0.774	−0.018
CI2	0.046	0.904	0.934	−0.030	0.694	0.735	−0.041
CI3	0.066	0.939	0.978	−0.039	0.736	0.790	−0.054
CI4	0.041	0.964	0.994	−0.029	0.735	0.778	−0.044
VCC1	0.092	0.906	0.948	−0.042	0.715	0.748	−0.033
VCC2	0.106	0.897	0.935	−0.039	0.736	0.772	−0.037
VCC3	0.063	0.951	0.992	−0.040	0.765	0.794	−0.029
VCC4	0.095	0.954	1.002	−0.048	0.769	0.803	−0.034
VCC5	0.097	0.893	0.943	−0.050	0.730	0.751	−0.021

Note: LM = linear model benchmark; Δ = PLS-SEM minus LM (negative values favor PLS-SEM). Construct-level Q^2predict: ET = 0.237, PC = 0.149, CI = 0.077, VCC = 0.143

Table 9. Cross-validated predictive ability test (CVPAT)

Construct	PLS-SEM vs indicator average			PLS-SEM vs linear model		
	ALD	*t*	*p*	ALD	*t*	*p*
ET	−0.157	4.949	< 0.001	−0.063	3.341	< 0.001
PC	−0.066	3.674	< 0.001	−0.047	3.869	< 0.001
CI	−0.051	3.125	0.002	−0.052	2.487	0.013
VCC	−0.084	4.928	< 0.001	−0.083	5.065	< 0.001
Overall	−0.093	6.380	< 0.001	−0.062	6.700	< 0.001

Note: ALD = average loss difference; negative values indicate that PLS-SEM has a lower prediction loss than the benchmark

Importance–Performance Map Analysis

The IPMA translates the structural results into priorities (Table 10; Figure 4). Epistemic trust had the largest total effect on both outcomes (continuance = 0.380; co-creation = 0.353) yet the lowest performance of all positively valenced predictors (66.64), placing it in the "concentrate here" quadrant. Explanation depth combined above-average importance for continuance (0.190 vs a mean of 0.179) with below-average performance (68.76). Conversational naturalness performed well (74.28) but had the lowest importance among the design features. Quadrants are relative to sample means and should not be over-read near the boundaries.

Perceived hallucination risk, a negatively valenced threat appraisal, had a high score (74.37) and was negatively related to both outcomes through trust (−0.073; −0.067). Information overload had the lowest score (60.19) and small negative total effects (−0.028; −0.039).

At the indicator level (Table 11), ET5 (overall trustworthiness) combined low performance (65.41) with the highest importance, and XAI3 (explaining how ratios and indicators are calculated) had the lowest performance (66.77) and highest importance within explanation depth.

Table 10. Importance–performance map analysis (constructs)

Construct	Performance	Target: CI		Target: VCC	
		Importance	Quadrant	Importance	Quadrant
Epistemic trust (ET)	66.64	0.380	Concentrate here	0.353	Concentrate here
Explanation depth (XAI)	68.76	0.190	Concentrate here	0.203	Low priority
Perceived control (PC)	75.45	0.184	Keep up	0.257	Keep up
Conversational naturalness (CN)	74.28	0.115	Possible overkill	0.124	Possible overkill
Financial and stock literacy (FSL)	66.66	0.024	Low priority	0.135	Low priority
Perceived hallucination risk (PHR)[b]	74.37	−0.073	–	−0.067	–
Information overload (IO)[b]	60.19	−0.028	–	−0.039	–
Mean of positively valenced predictors	*70.36*	*0.179*		*0.214*	

Note: Importance = total effect on the target construct; performance = mean latent variable score rescaled to 0–100. Quadrants are defined relative to the mean importance and mean performance of the five positively

valenced predictors. Quadrants are sample-relative and indicative only. [b]Negatively valenced threat appraisal; higher performance indicates a stronger perceived threat, so no quadrant is assigned

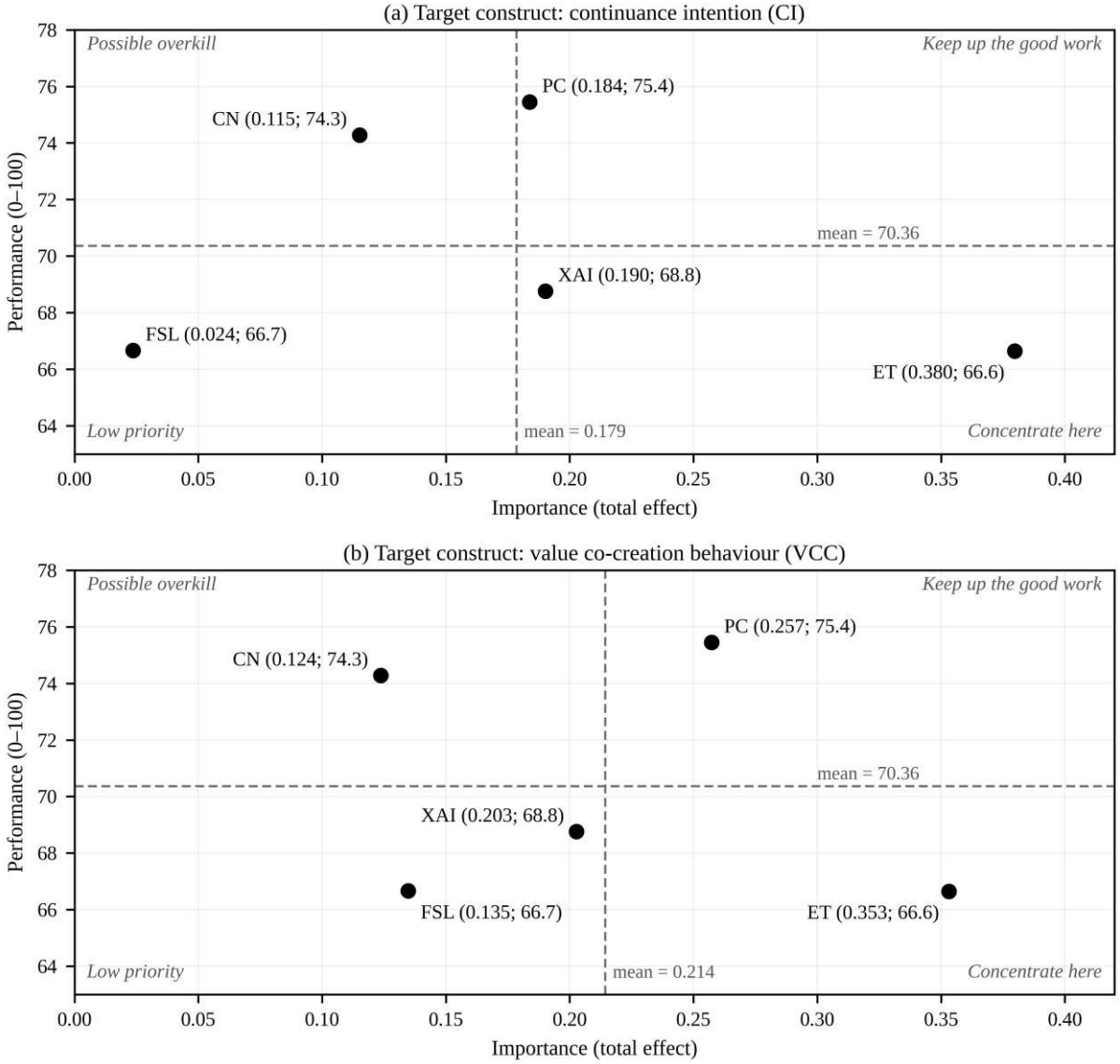

Figure 4. Importance–performance maps for continuance intention and value co-creation behavior

Note: Dashed lines mark the mean importance and mean performance of the five positively valenced predictors; quadrants are sample-relative and indicative only

Table 11. Importance–performance map analysis (indicators of epistemic trust and explanation depth)

Item	Content (abridged)	Performance	Importance (CI)	Importance (VCC)
ET1	Competent source of stock knowledge	65.93	0.098	0.091
ET2	Willing to base understanding on its information	64.18	0.091	0.084
ET3	Presents analysis honestly, incl. limitations	68.33	0.089	0.083
ET4	Can rely on it to learn about the market	69.95	0.081	0.076
ET5	Stock knowledge obtained is trustworthy overall	65.41	0.105	0.098
XAI1	Explains reasoning behind each analysis	70.98	0.041	0.043
XAI2	Shows the data behind its analysis	68.65	0.043	0.045
XAI3	Explains how ratios/indicators are calculated	66.77	0.047	0.050
XAI4	Explains risk factors that could make it wrong	67.68	0.046	0.049
XAI5	Explains why a stock is buy/sell/hold	69.75	0.045	0.048

Note: Full item wording is given in Table A2

DISCUSSION

Of the 21 hypotheses, 18 were supported in the baseline model (H3 and H10a only conditionally). Four findings stand out: perceived hallucination risk was associated with opposite weighting of the two design features; overload was negatively associated with perceived control but did not significantly moderate the design features; literacy moderated the trust–outcome associations in opposite directions; and epistemic trust has the most leverage in the IPMA but performs worst, whereas naturalness performs best but matters least.

Design Features, Appraisals, and Outcomes

Explanation depth had the stronger association with both appraisals, consistent with evidence that explainability is related to trust (Haque *et al.*, 2023; Shin, 2021), particularly in finance (Černevičienė and Kabašinskas, 2024; Weber *et al.*, 2024). Conversational naturalness was also positively associated with them (Anderl *et al.*, 2024; Blut *et al.*, 2021; Pelau *et al.*, 2021), but more weakly: what the assistant discloses appears to matter more than how fluently it speaks.

Epistemic trust had the strongest associations with both outcomes, in line with the central role of trust in AI acceptance (Choudhury and Shamszare, 2023; Choung *et al.*, 2023) and in continued use of digital financial services (Nguyen and Dao, 2024). It was also associated with self-reported verification and integration of AI output, as a calibrated conception of trust would imply (Ferrario, 2024), although calibration was not measured. Control was relatively more strongly associated with co-creation than with continuance, consistent with the argument that steering an AI's analysis requires agency (Buçinca *et al.*, 2021; Kinskofer and Tulis, 2025).

Explaining the Hallucination Paradox

The central finding is that perceived hallucination risk was associated with opposite weighting of the two design features: under high perceived risk, the association of explanation depth with trust roughly doubled, whereas naturalness was no longer positively associated with trust. This is consistent with vigilant investors using explanations diagnostically, since explanations reduce overreliance when they lower verification costs (Vasconcelos *et al.*, 2023), and discounting fluency as uninformative about accuracy (Crolic *et al.*, 2022; Steyvers *et al.*, 2025).

Investors who perceived little hallucination risk showed trust more strongly associated with conversational qualities than with explanatory substance, the configuration in which surface cues increase acceptance regardless of correctness (Bansal *et al.*, 2021). The modest negative association of hallucination risk with trust suggests that awareness of hallucination was not accompanied by algorithm aversion (Mahmud *et al.*, 2022) but by a different basis of trust, a pattern we label risk-contingent trust.

Information Overload: Related to Control Rather Than a Boundary Condition

Contrary to H13–H15, overload did not significantly moderate the design features, although the simple slopes moved in the hypothesized direction. The test is inconclusive: limited sensitivity may explain it, but so may the self-paced nature of dialogue, in which investors can request simpler answers, preserving steerability (Rapp *et al.*, 2021); if so, explanations depend more on the motivation to verify than on the capacity to process, qualifying expectations that longer explanations lose value under cognitive load (Steyvers *et al.*, 2025; Vasconcelos *et al.*, 2023).

Overload was nevertheless negatively associated with perceived control, consistent with evidence that overload impairs decision making (Arnold *et al.*, 2023; Graf and Antoni, 2021); trust was unrelated to overload, and intact trust with lower control may signal a pathway to dependence.

Literacy and the Trust–Outcome Associations

Literacy was associated with a crossover: among low-literacy investors, trust was associated with continued use but not co-creation, whereas among high-literacy investors it was associated mainly with co-creation. Consistent with research on literacy and algorithm reliance (Lusardi and Mitchell, 2023; Mahmud *et al.*, 2022), literacy may supply the operant resources that allow trust to translate into hybrid human–AI value creation (Bartelheimer *et al.*, 2025; Zhu *et al.*, 2022), a mechanism that was not observed. For less literate investors, the same trust may

resemble dependence, a risk worth noting given Indonesia's young and educationally diverse investor base (KSEI, 2026).

Priorities Suggested by the Importance–Performance Map

The IPMA suggests a mismatch between where providers appear to have invested and where the leverage lies: conversational naturalness performed well but mattered least, whereas epistemic trust was the most important yet worst-performing predictor. Because trust was more strongly associated with explanation depth than with fluency at high perceived risk, closing this gap may run through explanation depth, especially transparency about how ratios are calculated.

Theoretical Implications

This study makes four contributions, each bounded by its design. First, it extends cognitive appraisal theory to AI-assisted investing by showing that different threats show different patterns: hallucination risk moderated the design features, whereas overload was related directly to lower perceived control.

Second, it is consistent with conceptions of trust in LLMs as calibrated epistemic trust (Ferrario, 2024; Heersmink *et al.*, 2024; Magnus, 2025) and offers perceived stock data hallucination risk (Huang *et al.*, 2025; Ji *et al.*, 2023) as a provisional construct distinct from low trust; because only perceived trust was measured, it cannot show that trust is warranted.

Third, it qualifies the XAI assumption that deeper explanations uniformly increase trust (Haque *et al.*, 2023; Shin, 2021): the association of explanation depth with trust was conditional on perceived hallucination risk.

Fourth, it introduces self-reported value co-creation as an outcome of human–GenAI interaction in investing, extending robo-advisor research on adoption and reliance (Back *et al.*, 2023; Zhang *et al.*, 2021) and indicating that literacy moderates how trust relates to co-creation (Bartelheimer *et al.*, 2025); combining moderated mediation with IPMA translates a theory-testing model into a priority map (Hauff *et al.*, 2024).

Practical Implications

For securities firms and wealthtech providers, the results suggest four priorities for behavioral or experimental validation. First, earning epistemic trust may offer the greatest return, through verifiable explanation: the data source, reporting period and calculation behind each figure. Second, further investment in human-likeness may offer little return; assistants should instead

signal uncertainty, especially for numerical claims. Third, users should be able to rerun analyses under their own assumptions. Fourth, interfaces could prompt less experienced investors to check figures against issuer disclosures (Buçinca *et al.*, 2021).

For regulators and investor educators, the findings are consistent with requiring AI-assisted investment information to disclose data sources and warn about fabricated figures, and with integrating verification of AI-generated stock information into capital market education (KSEI, 2026).

CONCLUSION

GenAI stock assistants confront retail investors with a paradox: the fluency that makes them easy to use is unrelated to the accuracy of what they say. In this cross-sectional survey, investors who perceived higher hallucination risk reported trust more strongly associated with what the assistant shows than with how it talks, overload was associated with lower perceived control, and literacy was associated with whether trust related more to continued use or to self-reported co-creation. Whether such trust is warranted remains an open question.

LIMITATIONS AND FUTURE RESEARCH

This study has limitations. First, the cross-sectional, self-reported design precludes causal inference: common method variance and reverse relationships cannot be excluded, and the copula test indicated possible endogeneity in the explanation depth–control path; experiments seeding controlled hallucinations and longitudinal designs could test causality. Second, the study did not assess assistant accuracy, whether trust corresponded to reliability, or actual verification, so future research could add objective accuracy or trading data. Third, the hallucination-risk scale was validated only within this sample and needs independent replication. Fourth, excluding 112 of 498 questionnaires may have affected the estimates, and the nonsignificant overload interactions are inconclusive. Fifth, the purposive single-country sample limits generalizability, calling for multigroup analyses preceded by measurement invariance testing (Cheah *et al.*, 2023). Finally, IPMA quadrants are sample-relative; necessary condition analysis (Hauff *et al.*, 2024) could reveal minimum levels of trust or explanation depth required.

DECLARATIONS

Ethics approval: The study was a non-interventional, anonymous online survey of adult retail investors that collected no personally identifiable information. The research ethics review team

of the authors' institution issued an exemption from ethical review on 15 November 2026 (the issuing body is named on the title page to preserve anonymous review).

Informed consent: Before answering, respondents were informed of the purpose of the study, that participation was voluntary and that they could stop at any time; only respondents who agreed to participate proceeded to the questionnaire.

Conflict of interest: The authors declare no conflict of interest.

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability: The anonymized dataset of all 498 returned questionnaires, with an indicator of exclusion reason, is available from the corresponding author on reasonable request (contact details are on the title page).

Code and materials availability: The measurement items are provided in Table A2. The SmartPLS project file and output reports are available from the corresponding author on reasonable request.

Use of generative AI: Generative AI tools were used to assist with language editing, drafting and condensing parts of the text, formatting and reference management. The authors reviewed and edited all content and take full responsibility for the publication. No generative AI tool was used to collect or analyze the data; all statistical analyses were performed in SmartPLS.

REFERENCES

Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, *15*(2), Article e35179.

Anderl, C., Klein, S. H., Sarigül, B., Schneider, F. M., Han, J., Fiedler, P. L., & Utz, S. (2024). Conversational presentation mode increases credibility judgements during information search with ChatGPT. *Scientific Reports*, *14*, Article 17127.

Arnold, M., Goldschmitt, M., & Rigotti, T. (2023). Dealing with information overload: A comprehensive review. *Frontiers in Psychology*, *14*, Article 1122200.

Back, C., Morana, S., & Spann, M. (2023). When do robo-advisors make us better investors? The impact of social design elements on investor behavior. *Journal of Behavioral and Experimental Economics*, *103*, Article 101984.

Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M. T., & Weld, D. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary team performance. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–16.

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2022). Attention-induced trading and returns: Evidence from Robinhood users. *The Journal of Finance*, *77*(6), 3141–3190.

Bartelheimer, C., Heinz, D., Hönigsberg, S., Siemon, D., Li, M. M., Strohmann, T., Poeppelbuss, J., & Peters, C. (2025). Conceptualizing hybrid intelligent service ecosystems. *Electronic Markets*, *35*, Article 63.

Becker, J.-M., Cheah, J.-H., Gholamzade, R., Ringle, C. M., & Sarstedt, M. (2023). PLS-SEM's most wanted guidance. *International Journal of Contemporary Hospitality Management*, *35*(1), 321–346.

Becker, J.-M., Proksch, D., & Ringle, C. M. (2022). Revisiting Gaussian copulas to handle endogenous regressors. *Journal of the Academy of Marketing Science*, *50*(1), 46–66.

Bhatia, A., Chandani, A., Atiq, R., Mehta, M., & Divekar, R. (2021). Artificial intelligence in financial services: A qualitative research to discover robo-advisory services. *Qualitative Research in Financial Markets*, *13*(5), 632–654.

Blut, M., Wang, C., Wünderlich, N. V., & Brock, C. (2021). Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *Journal of the Academy of Marketing Science*, *49*(4), 632–658.

Buçinca, Z., Malaya, M. B., & Gajos, K. Z. (2021). To trust or to think: Cognitive forcing functions can reduce overreliance on AI in AI-assisted decision-making. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW1), Article 188.

Černevičienė, J., & Kabašinskas, A. (2024). Explainable artificial intelligence (XAI) in finance: A systematic literature review. *Artificial Intelligence Review*, *57*, Article 216.

Cheah, J.-H., Amaro, S., & Roldán, J. L. (2023). Multigroup analysis of more than two groups in PLS-SEM: A review, illustration, and recommendations. *Journal of Business Research*, *156*, Article 113539.

Choudhury, A., & Shamszare, H. (2023). Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. *Journal of Medical Internet Research*, *25*, Article e47184.

Choung, H., David, P., & Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human–Computer Interaction*, *39*(9), 1727–1739.

Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022). Blame the bot: Anthropomorphism and anger in customer–chatbot interactions. *Journal of Marketing*, *86*(1), 132–148.

Dowling, M., & Lucey, B. (2023). ChatGPT for (finance) research: The Bananarama Conjecture. *Finance Research Letters*, *53*, Article 103662.

Ferrario, A. (2024). Justifying our credences in the trustworthiness of AI systems: A reliabilistic approach. *Science and Engineering Ethics*, *30*(6), Article 55.

Graf, B., & Antoni, C. H. (2021). The relationship between information characteristics and information overload at the workplace – a meta-analysis. *European Journal of Work and Organizational Psychology*, *30*(1), 143–158.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.

Hair, J. F., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2024). *Advanced issues in partial least squares structural equation modeling* (2nd ed.). Sage.

Haque, A. K. M. B., Islam, A. K. M. N., & Mikalef, P. (2023). Explainable artificial intelligence (XAI) from a user perspective: A synthesis of prior literature and

problematizing avenues for future research. *Technological Forecasting and Social Change*, *186*, Article 122120.

Hauff, S., Richter, N. F., Sarstedt, M., & Ringle, C. M. (2024). Importance and performance in PLS-SEM and NCA: Introducing the combined importance-performance map analysis (cIPMA). *Journal of Retailing and Consumer Services*, *78*, Article 103723.

Heersmink, R., de Rooij, B., Clavel Vázquez, M. J., & Colombo, M. (2024). A phenomenology and epistemology of large language models: Transparency, trust, and trustworthiness. *Ethics and Information Technology*, *26*(3), Article 41.

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, *43*(2), 1–55.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, *55*(12), 1–38.

Kang, H. (2021). Sample size determination and power analysis using the G*Power software. *Journal of Educational Evaluation for Health Professions*, *18*, Article 17.

Kim, A., Muhn, M., & Nikolaev, V. V. (2024). *Financial statement analysis with large language models* [Chicago Booth Research Paper No. 24-11]. University of Chicago.

Kinskofer, F., & Tulis, M. (2025). Motivational and appraisal factors shaping generative AI use and intention in Austrian higher education students and teachers. *Frontiers in Education*, *10*, Article 1677827.

Klotz, A. C., Swider, B. W., & Kwon, S. H. (2023). Back-translation practices in organizational research: Avoiding loss in translation. *Journal of Applied Psychology*, *108*(5), 699–727.

Ko, H., & Lee, J. (2024). Can ChatGPT improve investment decisions? From a portfolio management perspective. *Finance Research Letters*, *64*, Article 105433.

KSEI. (2026). *Statistik Pasar Modal Indonesia: Agustus 2026 [Indonesian Capital Market Statistics: August 2026]*. PT Kustodian Sentral Efek Indonesia. https://www.ksei.co.id

Liengaard, B. D., Sharma, P. N., Hult, G. T. M., Jensen, M. B., Sarstedt, M., Hair, J. F., & Ringle, C. M. (2021). Prediction: Coveted, yet forsaken? Introducing a cross-

validated predictive ability test in partial least squares path modeling. *Decision Sciences*, *52*(2), 362–392.

Lopez-Lira, A., & Tang, Y. (2023). *Can ChatGPT forecast stock price movements? Return predictability and large language models* [Working paper]. SSRN.

Lusardi, A., & Mitchell, O. S. (2023). The importance of financial literacy: Opening a new field. *Journal of Economic Perspectives*, *37*(4), 137–154.

Magnus, P. D. (2025). On trusting chatbots. *Episteme*, 1–11. https://doi.org/10.1017/epi.2024.29

Mahmud, H., Islam, A. K. M. N., Ahmed, S. I., & Smolander, K. (2022). What influences algorithmic decision-making? A systematic literature review on algorithm aversion. *Technological Forecasting and Social Change*, *175*, Article 121390.

Menon, D., & Shilpa, K. (2023). "Chatting with ChatGPT": Analyzing the factors influencing users' intention to use the Open AI's ChatGPT using the UTAUT model. *Heliyon*, *9*(11), Article e20962.

Nguyen, G.-D., & Dao, T.-H. T. (2024). Factors influencing continuance intention to use mobile banking: An extended expectation-confirmation model with moderating role of trust. *Humanities and Social Sciences Communications*, *11*, Article 276.

Niszczota, P., & Abbas, S. (2023). GPT has become financially literate: Insights from financial literacy tests of GPT and a preliminary test of how people use it as a source of advice. *Finance Research Letters*, *58*, Article 104333.

Oehler, A., & Horn, M. (2024). Does ChatGPT provide better advice than robo-advisors? *Finance Research Letters*, *60*, Article 104898.

Ozik, G., Sadka, R., & Shen, S. (2021). Flattening the illiquidity curve: Retail trading during the COVID-19 lockdown. *Journal of Financial and Quantitative Analysis*, *56*(7), 2356–2388.

Pelau, C., Dabija, D.-C., & Ene, I. (2021). What makes an AI device human-like? The role of interaction quality, empathy and perceived psychological anthropomorphic characteristics in the acceptance of artificial intelligence in the service industry. *Computers in Human Behavior*, *122*, Article 106855.

Pelster, M., & Val, J. (2024). Can ChatGPT assist in picking stocks? *Finance Research Letters*, *59*, Article 104786.

Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior*, *11*, 17–61.

Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies*, *151*, Article 102630.

Ringle, C. M., Wende, S., & Becker, J.-M. (2024). *SmartPLS 4* (Version 4.3) [Computer software]. SmartPLS. https://www.smartpls.com

Roemer, E., Schuberth, F., & Henseler, J. (2021). HTMT2–an improved criterion for assessing discriminant validity in structural equation modeling. *Industrial Management & Data Systems*, *121*(12), 2637–2650.

Sarstedt, M., Hair, J. F., Pick, M., Liengaard, B. D., Radomir, L., & Ringle, C. M. (2022). Progress in partial least squares structural equation modeling use in marketing research in the last decade. *Psychology & Marketing*, *39*(5), 1035–1064.

Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, *146*, Article 102551.

Steyvers, M., Tejeda, H., Kumar, A., Belem, C., Karny, S., Hu, X., Mayer, L. W., & Smyth, P. (2025). What large language models know and what people think they know. *Nature Machine Intelligence*, *7*(2), 221–231.

Vasconcelos, H., Jörke, M., Grunde-McLaughlin, M., Gerstenberg, T., Bernstein, M. S., & Krishna, R. (2023). Explanations can reduce overreliance on AI systems during decision-making. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW1), Article 129.

Vereschak, O., Bailly, G., & Caramiaux, B. (2021). How to evaluate trust in AI-assisted decision making? A survey of empirical methodologies. *Proceedings of the ACM on Human-Computer Interaction*, *5*(CSCW2), Article 327.

Weber, P., Carl, K. V., & Hinz, O. (2024). Applications of explainable artificial intelligence in finance – a systematic review of finance, information systems, and computer science literature. *Management Review Quarterly*, *74*(2), 867–907.

Welch, I. (2022). The wisdom of the Robinhood crowd. *The Journal of Finance*, *77*(3), 1489–1527.

Zaremba, A., & Demir, E. (2023). ChatGPT: Unlocking the future of NLP in finance. *Modern Finance*, *1*(1), 93–98.

Zhang, L., Pentina, I., & Fan, Y. (2021). Who do you choose? Comparing perceptions of human vs robo-advisor in the context of financial services. *Journal of Services Marketing*, *35*(5), 634–646.

Zhu, Y., Wang, P., & Duan, W. (2022). Exploration on the core elements of value co-creation driven by AI – measurement of consumer cognitive attitude based on Q-methodology. *Frontiers in Psychology*, *13*, Article 791167.

APPENDIX

Table A1. Screening questions

No.	Question	Qualifying response
S1	What is your age?	≥ 18 years
S2	Do you hold an active securities account in your own name?	Yes
S3	Have you bought or sold stocks in the past three months?	Yes
S4	Are you a licensed securities analyst, investment manager or broker-dealer representative?	No
S5	Do you use GenAI tools (e.g. ChatGPT, Gemini, Claude, Copilot or AI features in brokerage apps) to analyze stocks?	Yes
S6	How long have you used GenAI for stock analysis?	≥ 3 months
S7	How often do you use it?	≥ twice per month
S8	Have you used GenAI output when considering a buy, sell or hold decision?	Yes

Source: Authors' own work

Table A2. Measurement items

All items refer to "the GenAI assistant you use most often for stock analysis" and use a five-point scale (1 = strongly disagree; 5 = strongly agree).

Construct	Code	Item	Conceptual basis
Explanation depth (XAI)	XAI1	The GenAI assistant explains the reasoning behind each stock analysis or recommendation it provides.	Shin (2021); Haque et al. (2023)
	XAI2	The GenAI assistant shows the data (e.g. earnings, financial ratios, prices) on which its analysis is based.	
	XAI3	The GenAI assistant explains how the ratios or technical indicators it presents are calculated.	
	XAI4	The GenAI assistant explains the risk factors that could make its stock analysis wrong.	
	XAI5	The GenAI assistant's explanations help me understand why a stock is considered a buy, sell or hold.	
Conversational naturalness (CN)	CN1	Interacting with the GenAI assistant about stocks feels like discussing with a human.	Pelau et al. (2021); Anderl et al. (2024)
	CN2	The GenAI assistant understands my follow-up questions about stocks without my repeating the context.	
	CN3	The language used by the GenAI assistant feels natural and easy to follow.	
	CN4	The GenAI assistant responds to my questions relevantly and in line with the flow of the conversation.	
	CN5	Conversations with the GenAI assistant about stock analysis run smoothly.	
Perceived stock data hallucination risk (PHR)	PHR1	I worry that the GenAI assistant presents financial figures (e.g. EPS, PER, ROE) that do not match actual data.	New; grounded in Huang et al. (2025); Ji et al. (2023)
	PHR2	I worry that the GenAI assistant reports technical signals (e.g. golden cross, support/resistance levels) that did not actually occur.	
	PHR3	The GenAI assistant may deliver incorrect stock information in a very convincing tone.	
	PHR4	The GenAI assistant may fabricate sources, news or corporate actions that do not exist.	
	PHR5	Using stock data from the GenAI assistant without verification could cause me financial loss.	
Information overload (IO)	IO1	The GenAI assistant often provides more stock information than I can process.	Arnold et al. (2023); Graf and Antoni (2021)

Construct	Code	Item	Conceptual basis
	IO2	I find it difficult to identify which information in the GenAI assistant's answers is important.	
	IO3	Long answers from the GenAI assistant confuse me when making buy or sell decisions.	
	IO4	I feel overwhelmed by the number of ratios, indicators and scenarios the GenAI assistant presents.	
Epistemic trust in stock AI (ET)	ET1	I regard the GenAI assistant as a competent source of knowledge about stock analysis.	Heersmink et al. (2024); Choung et al. (2023)
	ET2	I am willing to use information from the GenAI assistant as a basis for my understanding of a stock.	
	ET3	I believe the GenAI assistant presents stock analysis honestly, including acknowledging its limitations.	
	ET4	I can rely on the GenAI assistant to learn new things about the stock market.	
	ET5	Overall, the stock knowledge I obtain from the GenAI assistant is trustworthy.	
Perceived control (PC)	PC1	With the help of the GenAI assistant, I feel in control of my stock investment decisions.	Kinskofer and Tulis (2025)
	PC2	I am able to judge for myself whether the GenAI assistant's analysis is worth following.	
	PC3	The final buy or sell decision remains entirely mine.	
	PC4	I can direct the GenAI assistant to produce analysis that suits my needs.	
Continuance intention (CI)	CI1	I intend to continue using the GenAI assistant for stock analysis in the future.	Nguyen and Dao (2024)
	CI2	I would rather continue using the GenAI assistant than return to stock research without AI.	
	CI3	I will use the GenAI assistant for stock analysis as often as possible.	
	CI4	I will recommend using the GenAI assistant for stock analysis to other investors.	
Value co-creation behavior (VCC)	VCC1	I provide specific context (e.g. goals, time horizon, risk profile) so that the GenAI assistant's analysis is more relevant.	Bartelheimer et al. (2025)
	VCC2	I verify data from the GenAI assistant against official sources (e.g. financial statements and disclosures on the stock exchange website).	
	VCC3	I combine the GenAI assistant's analysis with my own fundamental or technical analysis.	
	VCC4	I correct or challenge the GenAI assistant when its analysis conflicts with my findings.	
	VCC5	My stock decisions result from combining the GenAI assistant's input with my own judgment.	
Financial and stock literacy (FSL)	FSL1	I understand how to read an issuer's financial statements (income statement, balance sheet, cash flow).	Lusardi and Mitchell (2023)
	FSL2	I can interpret stock valuation ratios such as PER, PBV and ROE.	
	FSL3	I understand diversification and the relationship between risk and return in stocks.	
	FSL4	I can read basic technical indicators (e.g. moving averages, volume, RSI).	
	FSL5	I understand the impact of corporate actions (dividends, stock splits, rights issues) on stock prices.	

Source: Authors' own work; items developed for this study based on the sources listed