

Article

Not peer-reviewed version

Financial Digital Twins and Conversational AI in Robo- Advisory: Evidence from a Scenario-Based Randomized Experiment

[Marco I. Bonelli](#) *

Posted Date: 15 May 2026

doi: 10.20944/preprints202605.0983.v1

Keywords: robo-advisors; financial digital twins; conversational AI; personalization; privacy con-cern; FinTech



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

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.

Article

Financial Digital Twins and Conversational AI in Robo- Advisory: Evidence from a Scenario-Based Randomized Experiment

Marco I. Bonelli

Ca' Foscari University of Venice, Italy; mibonelli6@gmail.com

Abstract

Robo-advisors have expanded access to automated investment services, but many platforms continue to rely on relatively static onboarding procedures and limited forms of user interaction. This study examines how participants with investment experience respond to two next-generation robo-advisory design features: financial digital twins, understood as dynamic investor profiles that integrate goals, risk tolerance, cash-flow patterns, and anticipated life events, and conversational artificial intelligence (AI), understood as an interactive interface for explaining recommendations. Using a scenario-based randomized 2×2 online experiment, 336 adult respondents with self-reported investment experience, recruited through professional and academic networks, were assigned to one of four robo-advisor scenarios that varied the personalization architecture, standard versus digital twin, and the interface style, plain dashboard versus conversational AI, while holding the portfolio recommendation constant. The results show that digital-twin personalization increases perceived personalization and privacy concern, indicating that more adaptive advisory architectures may be viewed as both more relevant and more data-intensive. Conversational AI increases the perceived interactive quality of the advisory experience, while the clearest adoption-related patterns emerge when it is combined with digital-twin personalization, particularly for selected indicators of stated behavioral willingness. Given the limited internal consistency of several secondary composite measures, the findings are best interpreted as evidence of scenario-based investor responses rather than as validated evidence of actual adoption behavior or confirmed psychological mechanisms. The study contributes to behavioral FinTech research by clarifying the personalization-privacy tension in AI-enabled robo-advisory services and by offering design implications for more transparent, interactive, and responsibly personalized digital wealth management systems.

Keywords: robo-advisors; financial digital twins; conversational AI; personalization; privacy concern; FinTech

JEL Classification: G11; G24; G41; O33; C91

1. Introduction

The financial advisory industry is being reshaped by the transition from human-centered advisory relationships to algorithmic and data-driven portfolio management. Robo-advisors represent one of the clearest expressions of this transition because they automate investor profiling, portfolio recommendation, rebalancing, and portfolio monitoring at relatively low cost and with broad scalability [1–4]. Their growth matters not only because they widen access to investment services, but also because they relocate parts of financial judgment into software systems that investors must evaluate for competence, credibility, transparency, and alignment with their interests.

Despite this progress, many first-generation robo-advisors still rely on relatively static onboarding questionnaires built around a limited set of variables, such as age, income, investment horizon, financial objectives, and risk tolerance. This structure supports efficiency and standardization, but it may also produce a partial representation of the investor and limit the perceived relevance of the recommendation. In a setting where investment advice depends not only on portfolio optimization but also on perceived suitability, explainability, and confidence in the advisory process, the design of the robo-advisory architecture becomes central to how users evaluate automated financial advice [1–4].

Recent advances in artificial intelligence have opened a plausible next stage in robo-advisory design. One development is the financial digital twin, understood here as a dynamic, permission-based digital representation of the investor's financial life that can incorporate goals, income patterns, savings behavior, liabilities, account structure, household circumstances, and anticipated life events [5–7]. A second development is conversational artificial intelligence (AI), which can move the advisory interface from a static dashboard toward a more interactive explanation of the recommendation [7,8]. These features may allow robo-advisors to become more adaptive, more explanatory, and more responsive to individual financial circumstances.

These innovations are attractive because they address two persistent limitations of conventional robo-advisory services: shallow personalization and limited interaction. A richer investor model may make recommendations appear more relevant and better aligned with personal circumstances, while a conversational interface may make the advisory experience more accessible and easier to interpret. At the same time, these same features raise important concerns. A financial digital twin requires broader personal and financial data, which may increase privacy concern even when the recommendation appears more personalized. Similarly, a more interactive AI interface may improve the perceived quality of the advisory experience, but it may also encourage users to rely on automated recommendations without fully evaluating their assumptions, limitations, or suitability [8–10].

The current literature explains parts of this problem, but not the problem as a whole. Prior studies show that robo-advisor acceptance is associated with perceived usefulness, ease of use, trust, financial literacy, familiarity with digital finance, and preferences for human or hybrid advice [2–4,11,12]. Other research has emphasized personalization, recommendation-agent familiarity, privacy calculus, perceived transparency, and the role of conversational interfaces in shaping user experience [8–10,13]. More recent conceptual work has identified digital twins and natural-language technologies as likely directions for the future development of robo-advisory services [5–7]. However, empirical evidence remains limited on how users respond when these design features are experimentally varied within the same robo-advisory setting.

This study addresses that gap through a scenario-based randomized 2×2 online experiment in which 336 adult respondents with self-reported investment experience were exposed to one of four robo-advisor scenarios. The scenarios varied the personalization architecture, standard profile versus financial digital twin, and the interface style, plain dashboard versus conversational AI, while holding the portfolio recommendation constant. This design allows the study to isolate how specific advisory features shape perceived personalization, privacy concern, and selected indicators of stated behavioral willingness.

The contribution of the study should be understood within the boundaries of its design. The purpose is not to measure actual adoption behavior in a live robo-advisory environment, but to examine how participants respond to described advisory scenarios under controlled experimental conditions. The study therefore provides early experimental evidence on scenario-based investor responses rather than field evidence of real platform use, real investment behavior, or confirmed psychological mechanisms. This distinction is important because the outcomes capture perceptions, intentions, and hypothetical willingness rather than observed financial decisions.

The results indicate that digital-twin personalization increases perceived personalization and privacy concern, suggesting that more adaptive advisory architectures may be viewed as both more relevant and more data-intensive. Conversational AI appears to shape participants' experience of the

advisory interface, with the clearest adoption-related patterns emerging in selected behavioral-willingness outcomes when conversational AI is paired with digital-twin personalization. However, these adoption-related findings are more limited and outcome-specific than the effects observed for perceived personalization and privacy concern. Accordingly, the study interprets the evidence cautiously and avoids treating the results as proof of actual robo-advisor adoption or as confirmation of a general trust-based mechanism.

The significance of the study lies in showing that the next stage of robo-advisory competition may depend not only on portfolio automation, but also on how advisory systems represent the investor, explain recommendations, and manage the tension between personalization and privacy.

Financial digital twins may increase the perceived relevance of automated advice, but their value depends on whether users regard the additional data requirements as acceptable and transparent. Conversational AI may make robo-advisory interaction feel more accessible, but its role appears strongest when it complements a richer personalization architecture.

These findings contribute to behavioral FinTech research by clarifying how emerging AI-enabled advisory designs are perceived in a controlled scenario-based setting and by identifying design implications for more transparent, interactive, and responsibly personalized digital wealth-management systems.

2. Literature Review and Hypotheses Development

2.1. Robo-Advisors and AI-Driven Financial Advice

Robo-advisors are digital financial-advice platforms that collect investor information, assess risk tolerance, recommend portfolios, and automate investment management. Their appeal lies in scalability, relatively low fees, automated rebalancing, broader access to financial advice, and a more disciplined investment process for retail users who may not otherwise rely on traditional advisors [1–4].

At the same time, robo-advisory raises a persistent challenge because investors are asked to rely on algorithmic recommendations in decisions involving personal wealth, risk exposure, and long-term financial goals, often without the relational reassurance or contextual interpretation provided by human advisors. For this reason, adoption depends not only on functionality, but also on whether the system is perceived as useful, understandable, credible, and aligned with investor interests [2–4]. Technology-adoption research provides the basic framework for explaining these responses.

The Technology Acceptance Model emphasizes perceived usefulness and ease of use [11], while UTAUT broadens the explanation by incorporating social influence and facilitating conditions [14]. In financial services, however, these models are informative but incomplete, because they do not fully capture the trust, expertise, and risk judgments that accompany delegated investment decisions. A meta-analytical study confirms this broader view, showing that trust, attitude, perceived usefulness, and ease of use remain central drivers of robo-advisor adoption, while also indicating that institutional and cultural context can condition these effects [15]. The implication is that the next stage of research should move beyond general acceptance and examine how particular design architectures shape investor responses.

2.2. Trust, Expertise, and Adoption in Automated Financial Advice

Trust occupies a central place in robo-advisory because users delegate part of their financial judgment to a non-human system. General trust theory defines trustworthiness through ability, benevolence, and integrity [16], and AI-trust research extends this logic by emphasizing competence, reliability, transparency, and the cues embedded in human-machine interaction [17]. In the robo-advisory context, these issues become even more important because users are not only evaluating an algorithm, but also evaluating the firm behind it, the broader advisory environment, and the expected consequences of relying on automated advice.

Recent studies reinforce this multi-layered understanding of trust, and Nourallah et al. show that young retail investors build initial trust in financial robo-advisors through a combination of perceived benefits, risk evaluation, and confidence in the system [18]. Cao et al. distinguish between technology trust, firm-specific trust, and system trust, showing that investors evaluate not only the tool itself but also the provider and the wider financial setting in which the advice is embedded [19]. Lourenço et al. similarly show that the type of firm offering robo-advice shapes perceptions of trust and expertise, which in turn influence acceptance [13]. Luo et al. add that provider reputation, fiduciary signals, customer education, privacy protection, and trust all contribute to willingness to use robo-advisors and to allocate assets through them [20], while evidence from India confirms that trust, performance expectancy, and preferences for human advice remain important in explaining adoption [21].

What emerges from this literature is a consistent point: robo-advisor adoption is not explained by usability alone, but by a broader judgment about competence, credibility, and the acceptability of delegating financial decisions to an algorithmic system.

2.3. Financial Digital Twins in Robo-Advisory Services

Financial digital twins extend this debate because they promise a richer and more adaptive form of personalization than conventional robo-advisory profiles. In this context, a financial digital twin can be understood as a dynamic, permission-based representation of an investor's financial life, incorporating risk tolerance, financial goals, income patterns, spending behavior, liabilities, account structure, household needs, tax context, and expected life events. Rather than relying on a one-time onboarding questionnaire, a digital-twin system can update the investor profile as circumstances change and, in principle, generate more context-sensitive recommendations.

Anshari et al. identify digital twins as a potential next frontier for robo-advisory because they may enable more customized and adaptive financial management [5]. Bhat and Krishnan place digital twins within a wider transformation of financial systems, emphasizing their role in simulation, personalization, and decision support [22], while earlier work on robo-advisory innovation also highlights digital-twin integration as a plausible strategic path for improving personalization and service quality [6,7]. Yet this line of work remains largely conceptual, the causal evidence is still limited, and it is not yet clear whether investors respond more favorably to a robo-advisor that presents itself as a dynamic financial twin than to one based on a standard risk-profile questionnaire.

The likely attraction of digital twins lies in their promise of deeper personalization, because recommendation-agent research has long shown that personalization and familiarity can strengthen trust and adoption when users perceive the advice as more relevant and user-centered [12]. Recent evidence from Germany also shows that investor profiling and portfolio recommendations remain heterogeneous across robo-advisory platforms, which keeps questions of suitability and transparency open [23].

In the investment setting, that mechanism should be especially important, since suitability depends on understanding the investor's goals, constraints, and life circumstances rather than a narrow set of profile variables. At the same time, a more intensive personalization architecture introduces a competing mechanism, because broader personalization requires broader data access.

Privacy-calculus research suggests that users weigh expected benefits against privacy costs [9], and in the robo-advisory setting those costs may be particularly salient when data extend to income, spending, tax, household, and account information. Ethical research on AI-based financial advice further reinforces this tension by highlighting concerns around explainability, vulnerability, responsibility, and the possibility that users may overtrust systems whose operations they do not fully understand [26]. Digital-twin personalization may therefore increase perceived usefulness and trust, while also increasing privacy concern.

2.4. Conversational AI and Social Presence in Financial Advice

Conversational AI addresses a different limitation of conventional robo-advisors, namely the weak interaction that characterizes many dashboard-based platforms. A static interface may present a portfolio allocation and a brief explanation, but it does not necessarily help the investor understand why the recommendation is suitable, what assumptions underpin it, or how it relates to personal goals. Conversational systems can make the advisory experience feel more responsive by explaining the rationale of the recommendation, answering likely follow-up questions, and presenting the advice in a more interactive way.

The literature increasingly suggests that this relational layer matters, a point illustrated by Hildebrand and Bergner, who show that conversational robo-advisors can act as surrogates of trust by improving onboarding experience, firm perception, and consumer financial decision-making [8]. Tang and Son further show that the sociality of AI digital human advisors can increase user experience value through social presence in digital finance [24], while Wang and Aw show that recommendation adoption and stickiness to the robo-advisor are shaped by trust, perceptions of the bot, and the broader interaction with the advisory system [25].

These studies indicate that conversational features are not merely cosmetic but can influence how investors interpret the advice, how much confidence they place in the system, and how willing they are to continue relying on it over time. At the same time, the literature does not present conversational design as an unqualified benefit. Back et al. show that social design elements in robo-advisors can alter investor behavior, suggesting that a more human-like presentation may shape reliance rather than simply improve understanding [10]. This is an important tension for financial advice, because a system that feels more social and responsive may encourage stronger compliance even when the investor has not fully evaluated the recommendation. Comparable evidence from Italy further suggests that financial literacy and the continuing demand for human financial advice shape the use of robo-advisory services, reinforcing the importance of context in adoption research [27].

The value of conversational AI therefore lies not only in higher engagement, but in whether the added social presence improves informed understanding rather than passive acceptance.

2.5. Conceptual Model and Hypotheses

The literature reviewed above points to a clear design-based gap, one already identified in the introduction but not yet resolved empirically. Research has shown that trust, usefulness, personalization, expertise, and privacy matter for robo-advisor adoption, but it has said much less about how investors respond when two next-generation design features, deeper personalization through financial digital twins and more interactive explanation through conversational AI, are varied directly while the underlying portfolio recommendation is held constant, that is the gap addressed in this study.

To organize these expected relationships, **Figure 1** presents the conceptual model guiding the analysis.

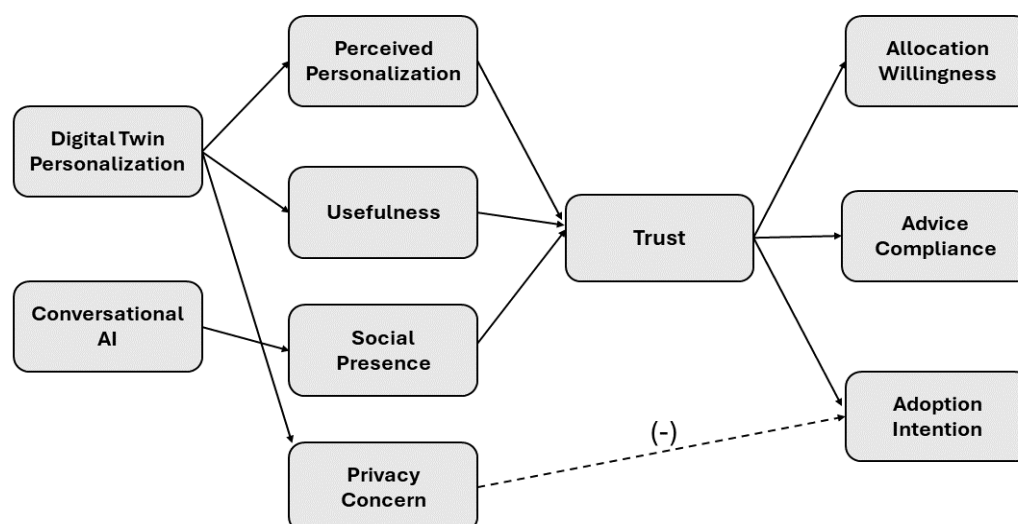


Figure 1. Conceptual Model: Responses to Financial Digital-Twin and Conversational-AI Robo-Advisory Designs.

The model links digital-twin personalization to perceived personalization, usefulness, and privacy concern; conversational AI to social presence; perceived personalization, usefulness, and social presence to trust; and trust to adoption intention, advice compliance, and allocation willingness. Privacy concern is modeled as a competing mechanism that may reduce adoption intention, while the digital-twin and conversational-AI combination is expected to produce the strongest behavioral-reliance outcomes.

Against this background, the study is guided by the following **research questions**:

RQ1: Does financial digital-twin personalization increase perceived personalization, usefulness, trust, and adoption intention?

RQ2: Does conversational AI increase social presence, trust, and willingness to follow robo-advisor recommendations?

RQ3: Does the richer data requirement of financial digital twins increase privacy concern?

RQ4: Does the combination of digital twins and conversational AI produce the strongest adoption and behavioral-reliance outcomes?

From this logic, the following testable **hypotheses** are derived:

H1. Financial digital-twin robo-advisors will produce higher perceived personalization than standard robo-advisors.

H2. Financial digital-twin robo-advisors will produce higher perceived usefulness than standard robo-advisors.

H3. Financial digital-twin robo-advisors will produce higher privacy concern than standard robo-advisors.

H4. Conversational AI interfaces will produce higher social presence than plain dashboard interfaces.

H5. Higher perceived personalization, usefulness, and social presence will be positively associated with trust, and trust will positively predict adoption intention and advice compliance.

H6. The digital-twin and conversational-AI condition will produce the highest adoption intention, advice compliance, and hypothetical allocation willingness.

As discussed in the measurement diagnostics below, hypotheses involving usefulness, trust, adoption intention, and advice compliance are interpreted as exploratory because the corresponding composite indices display limited internal consistency in this sample. The strongest inferential emphasis is therefore placed on perceived personalization and privacy concern, while results involving weaker composite indicators are treated as suggestive rather than confirmatory.

These hypotheses convert a largely conceptual debate into a set of observable effects on personalization, usefulness, social presence, privacy concern, trust, and behavioral reliance, while

directly addressing the unresolved question of how investors respond to richer personalization and more interactive explanation in automated financial advice. The next section describes the experimental design, survey instrument, sample, and analytical strategy used to test these propositions.

3. Materials and Methods

3.1. Research Design and Sample

This study used a randomized 2×2 between-subjects online experiment administered through the Qualtrics survey platform. Each participant evaluated one fictional robo-advisor scenario only, while the experiment manipulated two design features, namely personalization architecture, standard robo-advisor versus financial digital twin, and interface style, plain dashboard versus conversational AI chatbot.

The recommended portfolio was held constant across all four conditions at 60% global equity ETFs, 35% bond ETFs, and 5% cash equivalents, so that any observed differences across conditions could be attributed to the personalization architecture and interface style rather than to differences in the financial recommendation.

Participants were recruited online through the author's professional and academic network, yielding a non-probability convenience sample of adults with varying levels of investment experience and familiarity with digital financial services. Eligible respondents were required to be at least 18 years old and to report that they had personally invested or made investment decisions during the previous 24 months.

A total of 411 responses were obtained, and after pre-randomization screen-outs together with post-randomization exclusions for failed recall checks, incomplete responses, and excessive missingness, the final analytical sample consisted of 336 valid respondents. Because recruitment relied on professional and academic networks, the sample should not be interpreted as fully representative of the broader retail-investor population. In addition, investment experience was self-reported rather than independently verified, and the sample was concentrated in Italy, which further limits the generalizability of the findings.

The experimental conditions and valid cell sizes are reported in **Table 1**.

Table 1. Experimental conditions and valid sample sizes.

Condition	Digital twin	Conversational AI	Description	Valid N
C1	No	No	Standard robo-advisor + plain dashboard	82
C2	No	Yes	Standard robo-advisor + conversational AI chatbot	85
C3	Yes	No	Digital-twin robo-advisor + plain dashboard	83
C4	Yes	Yes	Digital-twin robo-advisor + conversational AI chatbot	86
Total				336

3.2. Procedure and Measures

After confirming eligibility, participants completed baseline questions covering investment experience, robo-advisor familiarity, prior robo-advisor use, chatbot familiarity, financial knowledge, and baseline privacy concern. They were then randomly assigned to one of the four robo-advisor

scenarios, completed manipulation-check items, answered the post-exposure measures, made a hypothetical allocation decision out of USD 10,000, and finally provided demographic information.

All perceptual items were measured on seven-point Likert scales, where higher values indicated stronger agreement, while willingness-to-share-data items used the same seven-point format to capture greater or lower willingness to allow the specified data source. Multi-item construct scores were calculated as the arithmetic mean of their respective items, allocation willingness was measured as the dollar amount from USD 0 to USD 10,000 that the respondent would allow the robo-advisor to manage, and portfolio acceptance was derived from the portfolio-action item by coding “accept as shown” and “accept with small changes” as acceptance, while all other responses were coded as non-acceptance. The full survey instrument is reported in Appendix A, Tables A1–A3, the experimental stimuli are summarized in Appendix B, Table A4, and the analytical variables are listed in Appendix C, Table A5. The main constructs and their analytical role are summarized in **Table 2**.

Table 2. Measurement overview.

Construct	Items	Main role in the model
Digital-twin manipulation check	2	Validates digital-twin treatment
Conversational AI manipulation check	2	Validates chatbot treatment
Perceived personalization	4	Mechanism
Social presence	3	Mechanism
Trust	4	Exploratory mediator / outcome
Privacy concern	4	Competing mechanism
Data control / transparency	3	Exploratory safeguard perception
Usefulness / benefits	4	Exploratory mechanism
Fiduciary quality	6	Exploratory trust-related design perception
Adoption intention	3	Exploratory stated-willingness outcome
Advice compliance	3	Exploratory behavioral-reliance outcome
Data-sharing willingness	7	Exploratory privacy-related outcome
Allocation willingness	1	Hypothetical behavioral outcome
Portfolio action	1	Acceptance / rejection behavior

3.3. Analytical Strategy

After screening and cleaning, the analytical file retained only valid cases that reached randomization, completed the post-exposure measures, passed the portfolio-recall check, and did not exhibit excessive missingness, resulting in a final sample of 336 observations. Data were analyzed using IBM SPSS Statistics for Windows, Version 29.0 (IBM Corp., Armonk, NY, USA); reliability was assessed using Cronbach’s alpha, and the reporting of reliability and model diagnostics followed standard recommendations for quantitative survey-based analysis where applicable [29].

The empirical analysis proceeded through descriptive statistics, randomization balance checks, reliability analysis, manipulation checks, treatment-effect regressions, mechanism tests, and robustness checks.

The primary treatment-effect model was specified as follows:

$$Y_i = b_0 + b_1 DT_i + b_2 CAI_i + b_3 (DT_i \times CAI_i) + g X_i + e_i \quad (1)$$

where Y_i represents the outcome of interest, DT_i indicates assignment to a digital-twin condition, CAI_i indicates assignment to a conversational-AI condition, $DT_i \times CAI_i$ captures the interaction between the two experimental factors, X_i represents optional covariates, and e_i is the error term.

The main outcomes were perceived personalization, social presence, trust, privacy concern, usefulness, adoption intention, advice compliance, allocation willingness, portfolio acceptance, and

data-sharing willingness, while the covariate-adjusted models included age, gender, education, assets bracket, investment experience, financial knowledge, baseline privacy concern, and robo-advisor familiarity.

Treatment effects on continuous outcomes were estimated using two-way ANOVA and OLS regression with robust standard errors, portfolio acceptance was analyzed using logistic regression, allocation willingness was also examined using a nonparametric test because the variable was bounded between USD 0 and USD 10,000, and mediation effects were examined using bootstrapped indirect effects.

Participation was voluntary and anonymous, no personally identifying data were collected, and respondents were informed that their responses would be used solely for academic research and that they could withdraw at any time before submission.

The study followed institutional ethical guidelines and the principles of responsible research conduct; because it relied on anonymous adult responses to hypothetical scenarios and did not collect personally identifying financial information, formal ethical approval was not required under the applicable institutional guidelines.

The anonymized dataset supporting this study has been deposited in Mendeley Data under the title “Digital Twin Robo-Advisor Adoption Dataset V_2” and is publicly available at <https://doi.org/10.17632/39rw5ywj8r.1> [28]; the survey instrument, experimental vignettes, and codebook are provided in the appendices.

Generative AI tools were not employed in the generation, collection, analysis, or interpretation of the data, and their use was limited to language polishing, formatting, and structural editing without influencing the substance, empirical results, or conclusions of the research.

3.4. Measurement Reliability and Interpretation of Composite Indicators

Before interpreting the treatment effects, internal consistency diagnostics were used to assess whether the multi-item indices could be treated as reliable composite measures. Perceived personalization ($\alpha = .77$) and privacy concern ($\alpha = .70$) showed acceptable internal consistency, while social presence ($\alpha = .69$) was close to the conventional threshold and is interpreted cautiously. By contrast, trust ($\alpha = .43$), data control/transparency ($\alpha = .13$), usefulness ($\alpha = .50$), fiduciary quality ($\alpha = .34$), adoption intention ($\alpha = .26$), advice compliance ($\alpha = .31$), and data-sharing willingness ($\alpha = .40$) were below conventional thresholds for internally consistent scales.

These indices are therefore not interpreted as validated latent constructs. They are retained as exploratory descriptive indicators of participants’ responses to the experimental scenarios, mainly to preserve transparency and comparability with the original research design. Accordingly, the main inferential emphasis is placed on perceived personalization and privacy concern, while findings based on weaker composite indicators are interpreted cautiously and, where relevant, treated as secondary or exploratory evidence.

4. Results

4.1. Sample, Balance Checks, and Measurement Reliability

The final analyzed sample consisted of 336 respondents distributed relatively evenly across the four experimental conditions. There were 82 respondents in the standard robo-advisor and plain-dashboard condition, 85 in the standard robo-advisor and chatbot condition, 83 in the digital-twin and plain-dashboard condition, and 86 in the digital-twin and chatbot condition. Participants were adults with self-reported investment experience, recruited through the author’s professional and academic networks. The sample was concentrated in Italy, which accounted for 202 respondents, or 60.1% of the sample, while the remaining respondents came from the United States, India, other European countries, and selected Asian countries. Most participants were between 25 and 44 years old, and most reported at least some investment experience.

Table 3 reports the sample profile and baseline balance diagnostics.

Table 3. Sample profile and baseline balance.

Variable	Full sample statistic	P-value
Valid analyzed sample	N = 336	n.a.
Condition sizes	C1 = 82; C2 = 85; C3 = 83; C4 = 86	n.a.
Age group	25-34: 100 (29.8%); 35-44: 98 (29.2%)	0.271
Gender	Female: 166 (49.4%); Male: 155 (46.1%)	0.011
Education	Bachelor degree or above: 237 (70.5%)	0.651
Employment	Full-time: 173 (51.5%)	0.437
Country / region	Italy: 202 (60.1%); other countries: 134 (39.9%)	0.849
Assets bracket	Mean bracket: 2.82 (1.30)	0.914
Investment experience	2.73 (1.09)	0.002
Robo-advisor familiarity	3.90 (1.30)	0.221
Chatbot familiarity	4.99 (1.15)	0.193
Financial knowledge	3.99 (1.06)	0.694
Baseline privacy concern	4.27 (1.23)	0.031
Prior robo-advisor use	No prior use: 189 (56.2%)	0.406

Note: C1 = standard robo-advisor with plain dashboard; C2 = standard robo-advisor with chatbot interface; C3 = digital-twin robo-advisor with plain dashboard; C4 = digital-twin robo-advisor with chatbot interface.

Balance p-values are based on ANOVA for ordinal or scale variables and chi-square tests for nominal variables. For nominal variables, the descriptive entry reports a compact summary, while the p-value is calculated using the full category distribution. Accordingly, the gender balance test includes all reported gender categories.

The balance checks indicate that the treatment groups were broadly similar on age, education, employment status, country or region, assets bracket, robo-advisor familiarity, chatbot familiarity, financial knowledge, and prior robo-advisor use. However, statistically significant differences were observed for gender, investment experience, and baseline privacy concern. These differences are relevant because they concern variables that may plausibly influence responses to automated financial advice, privacy-sensitive personalization, and AI-mediated interaction. For this reason, the main treatment-effect estimates are supplemented by covariate-adjusted models that include gender, investment experience, and baseline privacy concern among the baseline controls. These additional

models are used to assess whether the main findings remain stable after accounting for observed baseline differences across conditions.

Before interpreting the treatment effects, internal-consistency diagnostics were also examined to assess whether the multi-item indices could be treated as reliable composite measures. Perceived personalization ($\alpha = .77$) and privacy concern ($\alpha = .70$) showed acceptable internal consistency. Social presence ($\alpha = .69$) was close to the conventional threshold and is therefore interpreted cautiously. By contrast, several other indices displayed low internal consistency, including trust ($\alpha = .43$), data control/transparency ($\alpha = .13$), usefulness ($\alpha = .50$), fiduciary quality ($\alpha = .34$), adoption intention ($\alpha = .26$), advice compliance ($\alpha = .31$), and data-sharing willingness ($\alpha = .40$). These measures are therefore not interpreted as validated latent constructs. They are retained as exploratory descriptive indicators of participants' responses to the experimental scenarios, mainly to preserve transparency and comparability with the original research design.

Accordingly, the results are interpreted with different levels of inferential weight, and findings for perceived personalization and privacy concern are treated as the most reliable evidence because these constructs display acceptable internal consistency. Findings involving social presence are interpreted cautiously, while findings involving trust, data control/transparency, usefulness, fiduciary quality, adoption intention, advice compliance, and data-sharing willingness are treated as exploratory descriptive patterns rather than as evidence from validated latent constructs. This distinction is important because the study aims to provide evidence on scenario-based investor responses, not to validate a full psychometric model of robo-advisor adoption.

4.2. Manipulation Checks and Descriptive Outcomes

The manipulation checks indicate that respondents recognized the intended experimental features. Digital-twin recognition was higher in the two digital-twin conditions than in the two standard robo-advisor conditions, with an average digital-twin effect of 2.30 points, $p < 0.001$. Chatbot recognition was higher in the two conversational-AI conditions than in the two plain-dashboard conditions, with an average conversational-AI effect of 2.08 points, $p < 0.001$.

The descriptive outcome pattern is reported in **Table 4**.

Table 4. Manipulation checks and outcome means by condition.

Outcome	C1 Standard + Plain	C2 Standard + Chatbot	C3 Digital Twin + Plain	C4 Digital Twin + Chatbot
Digital-twin recognition	2.82	3.03	5.09	5.35
Chatbot recognition	3.23	5.36	3.34	5.37
Perceived personalization	3.72	3.96	5.15	5.40
Social presence	3.38	4.84	3.65	5.01
Trust	4.12	4.41	4.40	4.78
Privacy concern	3.91	4.03	4.83	4.74
Usefulness	4.13	4.35	4.80	4.99
Adoption intention	4.08	4.32	4.31	4.66
Advice compliance	4.00	4.16	4.11	4.61
Allocation willingness, USD	3,612	4,304	4,735	5,933
Portfolio acceptance rate	43.9%	52.9%	60.2%	72.1%

Note: Scale outcomes are condition means on 1–7 scales. Portfolio acceptance includes respondents who accepted the recommendation as shown or with small changes. Because some composite outcomes displayed limited internal consistency, descriptive means for trust, usefulness, adoption intention, advice compliance, and related constructs should be interpreted as exploratory.

Consistent with the experimental manipulations, the digital-twin conditions showed higher mean levels of perceived personalization and privacy concern, the two composite outcomes with acceptable internal consistency. Usefulness also followed a positive descriptive pattern in the digital-twin conditions, although this result is interpreted cautiously because the usefulness index displayed limited internal consistency. The conversational-AI conditions showed higher social presence, which is also interpreted cautiously because its reliability estimate was close to, but slightly below, the conventional threshold. Secondary design-perception and privacy-related outcomes, including fiduciary quality, data control/transparency, and data-sharing willingness, are reported in Appendix C, Table A8.

4.3. Treatment Effects

Table 5 reports the two-way treatment-effect regressions.

Table 5. Two-way treatment-effect regressions.

Predictor	Perceived personalization	Social presence	Trust	Privacy concern	Usefulness	Adoption intention	Advice compliance	Allocation willingness
DT effect	1.43*** (0.05)	0.22*** (0.05)	0.33*** (0.05)	0.81*** (0.07)	0.66*** (0.05)	0.28*** (0.06)	0.28*** (0.06)	1,376*** (114)
CAI effect	0.24*** (0.05)	1.41*** (0.05)	0.33*** (0.05)	0.01 (0.07)	0.20*** (0.05)	0.29*** (0.06)	0.33*** (0.06)	945*** (114)
DT x CAI	0.01 (0.10)	-0.09 (0.11)	0.09 (0.10)	-0.22 (0.14)	-0.03 (0.10)	0.11 (0.12)	0.34** (0.12)	506* (228)
R ²	0.725	0.683	0.206	0.284	0.375	0.123	0.155	0.400
N	336	336	336	336	336	336	336	336

Note: Coefficients are reported with robust standard errors in parentheses. DT = digital twin; CAI = conversational AI. Treatment indicators were mean-centered, so the DT and CAI coefficients represent average factorial effects across the other treatment factor. All outcomes except allocation willingness are measured on 1–7 scales. Because several composite outcomes displayed low internal consistency, coefficients for trust, usefulness, adoption intention, advice compliance, and related composite measures should be interpreted as exploratory treatment patterns rather than as effects on validated latent constructs. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Digital-twin personalization had a strong positive effect on perceived personalization, $b = 1.43$, and also increased privacy concern. These two findings carry the strongest inferential weight because both constructs displayed acceptable internal consistency. Usefulness, trust, adoption intention, and advice compliance also showed positive treatment patterns, but these composite-based results should be interpreted as exploratory where internal consistency was weak. Hypothetical allocation willingness also increased under the digital-twin condition. This pattern supports H1 and H3, while providing only exploratory support for H2. More broadly, the results suggest that the perceived benefits of richer personalization appear alongside higher privacy concern.

Conversational AI had its largest effect on social presence, $b = 1.41$; this provides cautious support for H4, given that the social-presence index was close to, but slightly below, the conventional reliability threshold. Additional positive patterns were observed for trust, adoption intention, advice compliance, and allocation willingness, but these findings are interpreted cautiously because several of the relevant composite measures displayed weak internal consistency.

The interaction term was positive and significant for advice compliance and allocation willingness, providing partial support for H6. This pattern suggests that the combined condition was

associated with stronger selected behavioral-willingness responses, especially allocation willingness and the exploratory advice-compliance index.

4.4. Mechanism Tests and Sensitivity Checks

The mechanism evidence is interpreted as exploratory because the trust, adoption-intention, and advice-compliance indices displayed weak internal consistency. In the supplemental model using the exploratory trust index, perceived personalization, usefulness, and social presence were positively associated with trust-related responses, with coefficients of $b = 0.14$, $p < 0.001$; $b = 0.12$, $p = 0.019$; and $b = 0.20$, $p < 0.001$, respectively. The exploratory trust index was also positively associated with adoption intention, $b = 0.24$, $p < 0.001$, and advice compliance, $b = 0.15$, $p = 0.019$, when the main mechanisms were estimated jointly. These associations are therefore reported as suggestive pathway evidence rather than as confirmation of a validated psychological mechanism. The bootstrapped indirect effects in **Table 6** provide exploratory evidence consistent with a positive digital-twin pathway.

Table 6. Mediation and mechanism results.

Model	Pathway	Indirect effect	Bootstrap SE	95% CI	Interpretation
A	DT -> personalization -> trust -> adoption intention	0.060	0.025	[0.017, 0.111]	Exploratory support
B	DT -> privacy concern -> adoption intention	-0.068	0.039	[-0.146, 0.006]	Directional, not confirmatory
C	CAI -> social presence -> trust -> advice compliance	0.050	0.024	[0.008, 0.103]	Exploratory support

Note: Indirect effects were estimated using 5,000 bootstrap resamples. Because the trust, adoption-intention, and advice-compliance indices displayed weak internal consistency, mediation results are interpreted as exploratory pathway evidence rather than as confirmation of validated psychological mechanisms.

The privacy pathway was negative, as expected, but its confidence interval crossed zero, so it is treated as directional evidence rather than a confirmatory mediation result. The robustness checks are reported in Appendix D, Table A9. These checks are especially important because the randomization diagnostics identified baseline imbalances in gender, investment experience, and baseline privacy concern. The covariate-adjusted specification therefore estimates each outcome as a function of the digital-twin condition, the conversational-AI condition, their interaction, and baseline controls including gender, investment experience, baseline privacy concern, age, education, assets bracket, financial knowledge, and robo-advisor familiarity.

The main effects for perceived personalization and privacy concern remain substantively similar after adjustment, suggesting that the central findings are not driven solely by the observed baseline imbalances. The broader pattern of findings also remains broadly stable for social presence and allocation willingness, while adoption-related, trust-related, and mechanism results continue to be interpreted as exploratory because of the measurement limitations discussed above.

5. Discussion

The results clarify how two AI-enabled robo-advisory design features shape participants' responses when the underlying portfolio recommendation is held constant. The clearest evidence concerns digital-twin personalization, which increased perceived personalization and privacy concern, the two composite outcomes with acceptable internal consistency. This pattern suggests that participants recognized the added relevance of a richer investor profile, but also perceived the broader data requirements of the digital-twin architecture as more privacy-sensitive. Conversational

AI mainly affected the perceived interactional quality of the advisory experience, with its strongest and most direct effect observed for social presence. Because social presence was close to, but slightly below, the conventional reliability threshold, this result is interpreted cautiously, although it remains consistent with the intended manipulation.

The findings for trust, usefulness, adoption intention, advice compliance, and related composite indicators require a more restrained interpretation. These outcomes showed positive descriptive or regression-based patterns in several models, but their internal consistency was weak. They are therefore best interpreted as exploratory indicators of scenario-based responses rather than as validated psychological constructs. This distinction is important because the study does not provide evidence of actual robo-advisor adoption in a live advisory environment. It provides experimental evidence on how participants with self-reported investment experience evaluated described robo-advisory scenarios that varied personalization architecture and interface style.

The combined digital-twin and conversational-AI condition produced the highest descriptive levels of several adoption-related and reliance-related indicators, including hypothetical allocation willingness and portfolio acceptance. However, the regression evidence for the interaction was clearest for allocation willingness and the exploratory advice-compliance index, rather than across all adoption-related outcomes. This suggests that the combination of richer personalization and conversational explanation may be especially relevant when participants evaluate whether to follow or allocate funds to a specific recommendation, but the evidence does not support a broad claim that the combined design uniformly increases adoption across all outcomes.

These results extend research on robo-advisor acceptance by showing that responses to automated financial advice are shaped not only by general attitudes toward digital finance, but also by the architecture through which advice is generated and explained. Prior studies have emphasized usefulness, trust, familiarity, provider credibility, and perceived expertise as important drivers of robo-advisor acceptance [1–4,13,18–21].

The present study adds that the source of perceived suitability also matters, because a digital-twin profile may be evaluated more favorably than a standard profile when it appears to capture a broader and more adaptive representation of the investor's financial life. However, this favorable response is accompanied by higher privacy concern, indicating that deeper personalization is not perceived as an unqualified benefit.

The study also contributes to the emerging literature on financial digital twins. Existing work has discussed digital twins as a promising direction for personalization, simulation, and decision support in financial services [5–7,22], but empirical evidence on user responses to this architecture remains limited. The present experiment suggests that participants can distinguish between a conventional onboarding profile and a more data-rich financial twin, and that this distinction affects perceived personalization and privacy concern.

The evidence for usefulness, trust-related responses, and adoption-related willingness is more exploratory, but it points to a plausible interpretation: financial digital twins may improve the perceived relevance of automated advice, while simultaneously increasing concern about the amount and sensitivity of personal data required to support that relevance.

This personalization-privacy tension is consistent with privacy-calculus logic [9]. In ordinary digital services, users may tolerate data collection when the service is convenient, entertaining, or low risk. In robo-advisory services, the threshold for acceptability is likely higher because the platform may request information about income, assets, liabilities, spending patterns, goals, household circumstances, tax position, and future life events. The results therefore suggest that financial digital twins are likely to be more credible when users can understand what data are collected, why those data are needed, how they influence the recommendation, and how data access can be reviewed, corrected, limited, or withdrawn. Without these safeguards, deeper personalization may increase perceived relevance while also making users feel overexposed.

The conversational-AI findings also contribute to research on social presence and automated financial advice. The chatbot condition increased perceived social presence and was associated with

more favorable patterns in several stated willingness outcomes. This is consistent with evidence that conversational robo-advisors can make onboarding and recommendation explanation feel more interactive [8,24,25]. At the same time, the results should not be interpreted as evidence that conversational AI reliably increases trust or adoption in a general sense, because several downstream composite measures displayed weak internal consistency. A more defensible interpretation is that conversational AI appears to strengthen the interactional experience and may support selected willingness responses, especially when paired with a richer personalization architecture.

This distinction matters because prior research has warned that social design elements in robo-advisory can shape investor behavior and may increase reliance on automated recommendations [10]. In this study, the combined digital-twin and conversational-AI condition was associated with stronger allocation willingness and a positive pattern in advice compliance. This may indicate that conversational explanation can support confidence in the recommendation, but it also raises a design responsibility. In financial advice, the objective should not be to maximize comfort, fluency, or compliance for its own sake. The objective should be conversational explainability, where the interface helps users understand why a recommendation is suitable, what assumptions it depends on, what risks it involves, and what alternatives remain available.

The exploratory mechanism tests provide suggestive, but not confirmatory, evidence regarding the role of trust-related responses. Perceived personalization, usefulness, and social presence were positively associated with the exploratory trust index, and the trust index was positively associated with adoption intention and advice compliance. However, because the trust, adoption-intention, and advice-compliance indices displayed weak internal consistency, these associations should not be interpreted as evidence of a validated psychological mechanism. They are better viewed as preliminary pathway evidence that can inform future research using stronger measurement instruments. Similarly, the privacy pathway was negative but not statistically confirmatory, which means that the study supports the coexistence of personalization benefits and privacy concern, but not a confirmed privacy-mediated adoption mechanism.

The practical implication is that robo-advisory providers should not introduce financial digital twins as a simple expansion of data collection. Digital-twin features should be presented as controlled personalization tools that improve suitability, transparency, and user understanding. Users should be able to inspect the financial profile used by the system, see which assumptions influence the recommendation, update or correct profile information, and manage data permissions in a visible and reversible way. The more comprehensive the digital twin becomes, the more important it is to make the profile auditable from the user's perspective.

Conversational AI should follow the same principle, since its role should be to clarify the recommendation, explain portfolio allocation, identify relevant risks, and support comparison with alternatives. A chatbot that only increases comfort or stated willingness would be problematic in a financial-advice setting if it does not also improve understanding. The better design objective is to make conversational AI an explanation layer rather than a persuasion layer. This is particularly important because the strongest willingness-related responses appeared when conversational AI was combined with digital-twin personalization.

The findings also have regulatory and consumer-protection implications, because financial digital twins depend on broader and potentially more continuous data flows than standard robo-advisory questionnaires. This makes consent, data minimization, explainability, and user control central to responsible deployment. Regulators and professional bodies should therefore pay attention not only to the algorithm that generates the recommendation, but also to the data architecture that feeds it and the interface that frames it. Conversational AI raises a related issue because a more socially present interface may increase stated willingness to follow automated advice. Disclosure rules and interface standards may need to distinguish between explanation that improves understanding and interaction design that encourages passive acceptance.

Several limitations should guide the interpretation of the findings, most importantly that the experiment used scenario-based descriptions rather than a live robo-advisory platform, which

allowed clean manipulation of digital-twin personalization and conversational AI but could not reproduce all features of real platform use. Allocation willingness was hypothetical, and portfolio acceptance was based on stated response rather than actual investment behavior. The sample was recruited through professional and academic networks, investment experience was self-reported rather than independently verified, and the sample was concentrated in Italy, although it included respondents from several countries. The findings should therefore be interpreted as evidence from a digitally exposed convenience sample rather than as representative evidence for the broader retail-investor population.

A further limitation concerns measurement reliability, since several multi-item indices displayed low internal consistency, which limits their interpretation as validated constructs. This is especially relevant for trust, usefulness, fiduciary quality, adoption intention, advice compliance, data control/transparency, and data-sharing willingness. The revised interpretation therefore places the greatest weight on perceived personalization and privacy concern, while treating the remaining composite indicators as exploratory. Future studies should use more extensively validated scales, conduct confirmatory factor analysis where sample size permits, and supplement stated willingness measures with behavioral or incentivized outcomes.

Future research should test these design features in functioning robo-advisory environments where users can interact with dashboards, chatbots, account-linking tools, consent interfaces, and explanation modules. Field experiments with FinTech providers would be especially useful for observing real data-sharing choices, portfolio acceptance, allocation decisions, and continued use. Further studies could also compare different consent dashboards, privacy labels, explanation formats, and warnings against overreliance. Cross-country research would help determine whether responses to financial digital twins and conversational AI vary across regulatory regimes, financial-literacy levels, and cultural attitudes toward automated advice.

6. Conclusions

This study shows that robo-advisory design architecture matters for perceived personalization, privacy concern, and selected indicators of stated behavioral willingness in a scenario-based experimental setting. When the same portfolio recommendation was presented through different advisory designs, participants responded more favorably to richer personalization and more interactive explanation, although the strength of the evidence varied across outcomes. Digital-twin personalization produced the clearest effects on perceived personalization and privacy concern, indicating that a more adaptive investor profile can make automated advice appear more relevant while also making the underlying data requirements more salient. Conversational AI strengthened the perceived interactional quality of the advisory experience, with its clearest role emerging through social presence and selected willingness-related responses.

The central conclusion is that financial digital twins and conversational AI may improve how participants evaluate robo-advisory services, but their value depends on transparency, user control, and responsible explanation. A richer investor profile can increase perceived relevance, yet it also raises privacy concern. A conversational interface can make advice easier to engage with, yet it may also encourage stronger reliance on automated recommendations. The strongest willingness-related patterns appeared when digital-twin personalization and conversational AI were combined, especially for allocation willingness and the exploratory advice-compliance index. However, these findings should be interpreted as scenario-based responses rather than as evidence of actual adoption behavior in a live robo-advisory environment.

The study contributes to behavioral FinTech research by providing experimental evidence on two next-generation robo-advisory features that have often been discussed conceptually but tested less directly. Its broader implication is that the future of automated financial advice will depend not only on better algorithms, but also on whether AI-enabled platforms can make personalization useful, explanation credible, and data use acceptable to investors. Future robo-advisors should therefore be

designed not only for optimization and convenience, but also for explainability, consent, data control, and active investor understanding.

Supplementary Materials: The following supporting materials are available with the article or in the associated data repository: the survey instrument, experimental vignettes, codebook, anonymized dataset, and analysis files.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing, original draft preparation, writing, review and editing, visualization, supervision, and project administration, Marco I. Bonelli. The author has read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study under applicable institutional requirements for anonymous survey research involving adult participants and no collection of personally identifying information.

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

Data Availability Statement: The anonymized dataset supporting the findings of this study is publicly available in Mendeley Data under the title "Digital Twin Robo-Advisor Adoption Dataset V_2" at <https://doi.org/10.17632/39rw5ywj8r.1>.

Acknowledgments: The author thanks the respondents who participated in the survey and the professional contacts who supported its dissemination.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AI Artificial intelligence

ANOVA Analysis of variance

CAI Conversational artificial intelligence

CI Confidence interval

DT Digital twin

ETF Exchange-traded fund

OLS Ordinary least squares

SE Standard error

SPSS Statistical Package for the Social Sciences

USD United States dollar

UTAUT Unified Theory of Acceptance and Use of Technology

Appendix A. Survey Questionnaire

Unless otherwise stated, all perceptual items used seven-point Likert scales, where 1 = strongly disagree and 7 = strongly agree.

Table A1. Eligibility and baseline items.

Block	Item	Response format
Consent	Do you agree to participate in this study?	1 = Yes; 0 = No
Eligibility	Are you at least 18 years old?	1 = Yes; 0 = No

Eligibility	Have you personally invested or made investment decisions during the last 24 months?	1 = Yes; 0 = No
Baseline	How familiar are you with robo-advisors or automated investment platforms?	1 = Not at all familiar; 7 = Very familiar
Baseline	Have you ever used a robo-advisor or automated investment platform?	1 = No; 2 = Tried once; 3 = Previously used; 4 = Currently use
Baseline	How would you rate your investment experience?	1 = Novice; 2 = Beginner; 3 = Intermediate; 4 = Advanced; 5 = Professional
Baseline	How familiar are you with AI chatbots or conversational AI tools?	1 = Not at all familiar; 7 = Very familiar
Baseline	How would you rate your overall financial knowledge?	1 = Very low; 7 = Very high
Baseline	In general, I am concerned about how financial apps use my personal data.	1 = Strongly disagree; 7 = Strongly agree

Table A2. Manipulation-check and post-exposure perceptual items.

Construct	Item
Digital-twin recognition	The advisor appeared to create a detailed digital profile of the investor.
Digital-twin recognition	The advisor seemed to use ongoing personal financial data to personalize the recommendation.
Conversational-AI recognition	The advisor interacted in a conversational way.
Conversational-AI recognition	The advisor felt more like a chat-based assistant than a static dashboard.
Perceived personalization	The recommendation felt tailored to the investor's specific financial situation.
Perceived personalization	The advisor appeared to consider future life changes, not only current risk tolerance.
Perceived personalization	The advisor seemed capable of updating advice as the investor's circumstances change.
Perceived personalization	The advisor seemed to understand the investor as a whole financial person rather than as a simple questionnaire profile.
Social presence	The advisor's communication felt interactive.
Social presence	The advisor made the recommendation easier to understand.
Social presence	The advisor felt responsive to possible investor questions.
Trust	I would trust this robo-advisor to provide reliable investment guidance.

Trust	This robo-advisor seems competent in making portfolio recommendations.
Trust	This robo-advisor seems honest about how it generates advice.
Trust	I would feel confident relying on this robo-advisor for basic investment guidance.
Privacy concern	I would be concerned about how much personal data this service uses.
Privacy concern	The personalization seems to require sensitive financial information.
Privacy concern	I would worry that my data could be misused or shared in ways I do not expect.
Privacy concern	The benefits of personalization may not justify the privacy risks.
Data control / transparency	The service seems transparent about what data it uses.
Data control / transparency	I would feel in control of the data used to generate recommendations.
Data control / transparency	The service seems to allow users to review or change the data used for personalization.
Usefulness	This robo-advisor would make investment planning more useful.
Usefulness	This robo-advisor would help investors make better long-term financial decisions.
Usefulness	This robo-advisor would save time compared with managing investments alone.
Usefulness	This robo-advisor offers clear value compared with a basic investment app.
Fiduciary quality	The advisor appears capable of managing portfolio risk responsibly.
Fiduciary quality	The advisor seems to consider the investor's financial well-being.
Fiduciary quality	The advisor appears to use enough information to make a suitable recommendation.
Fiduciary quality	The advisor seems aligned with the investor's interests.
Fiduciary quality	The advisor appears transparent about why the recommendation was made.
Fiduciary quality	The advisor does not seem designed mainly to benefit the platform at the investor's expense.

Table A3. Adoption, behavioral, data-sharing, and demographic items.

Block	Item	Response format
Adoption intention	I would consider using this robo-advisor.	1 = Strongly disagree; 7 = Strongly agree
Adoption intention	I would be willing to open an account with a service like this.	1 = Strongly disagree; 7 = Strongly agree
Adoption intention	I would recommend a service like this to someone with investment needs.	1 = Strongly disagree; 7 = Strongly agree
Advice compliance	I would be comfortable following this portfolio recommendation.	1 = Strongly disagree; 7 = Strongly agree
Advice compliance	I would be likely to accept the recommendation without major changes.	1 = Strongly disagree; 7 = Strongly agree
Advice compliance	I would rely on this advisor's guidance during normal market conditions.	1 = Strongly disagree; 7 = Strongly agree
Recall check	According to the scenario, what portfolio was recommended?	1 = 60% global equity ETFs, 35% bond ETFs, 5% cash equivalents; 2 = 35% global equity ETFs, 60% bond ETFs, 5% cash equivalents; 3 = 100% cash equivalents; 4 = Do not remember
Allocation willingness	Imagine you had USD 10,000 available to invest. How many dollars would you allow this robo-advisor to manage?	0 to 10,000
Portfolio action	What would you do with the recommended 60/35/5 portfolio?	1 = Accept as shown; 2 = Accept with small changes; 3 = Make major changes; 4 = Reject; 5 = Need human advisor first
Data-sharing willingness	Bank transaction data	1 = Definitely not allow; 7 = Definitely allow
Data-sharing willingness	Brokerage or retirement account data	1 = Definitely not allow; 7 = Definitely allow
Data-sharing willingness	Income or payroll pattern data	1 = Definitely not allow; 7 = Definitely allow
Data-sharing willingness	Tax profile information	1 = Definitely not allow; 7 = Definitely allow
Data-sharing willingness	Spending category data	1 = Definitely not allow; 7 = Definitely allow
Data-sharing willingness	Household or family-planning information	1 = Definitely not allow; 7 = Definitely allow

Data-sharing willingness	Social-media signals, such as major life events	1 = Definitely not allow; 7 = Definitely allow
Demographics	What is your age group?	1 = 18–24; 2 = 25–34; 3 = 35–44; 4 = 45–54; 5 = 55–64; 6 = 65+
Demographics	What is your gender?	1 = Female; 2 = Male; 3 = Non-binary; 4 = Prefer not to say; 5 = Self-describe
Demographics	What is your highest completed education?	1 = High school or less; 2 = Some college; 3 = Bachelor degree; 4 = Master degree; 5 = Doctoral/professional; 6 = Prefer not to say
Demographics	What is your current employment status?	1 = Full-time; 2 = Part-time; 3 = Self-employed; 4 = Student; 5 = Retired; 6 = Unemployed; 7 = Prefer not to say
Demographics	Approximately how much do you have in investable assets, excluding your primary residence?	1 = Less than USD 5,000; 2 = USD 5,000–24,999; 3 = USD 25,000–99,999; 4 = USD 100,000–249,999; 5 = USD 250,000 or more; 6 = Prefer not to say
Demographics	What is your country or region of residence?	Open text, cleaned into country-region groups

Appendix B. Experimental Stimuli

All four conditions presented the same recommended portfolio: 60% global equity ETFs, 35% bond ETFs, and 5% cash equivalents. Only the personalization architecture and interface style varied across conditions.

Table A4. Experimental conditions and scenario descriptions.

Condition	Profile description	Interface description
C1 Standard + Plain	The robo-advisor used a standard questionnaire based on age, investment horizon, risk tolerance, income range, and investment goals.	The recommendation was shown through a plain dashboard with the portfolio allocation, a short risk description, and expected diversification benefits.
C2 Standard + Chatbot	The robo-advisor used a standard questionnaire based on age, investment horizon, risk tolerance, income range, and investment goals.	The recommendation was explained through a conversational AI chatbot that described the role of equities, bonds, and cash equivalents in plain language.
C3 Digital Twin + Plain	The robo-advisor created a financial digital twin using a dynamic, permission-based profile that could include goals, income pattern,	The recommendation was shown through a plain dashboard with the portfolio allocation, a short risk description,

	spending behavior, liabilities, household needs, account structure, tax context, and expected life events.	and expected diversification benefits.
C4 Digital Twin + Chatbot	The robo-advisor created a financial digital twin using a dynamic, permission-based profile that could include goals, income pattern, spending behavior, liabilities, household needs, account structure, tax context, and expected life events.	The recommendation was explained through a conversational AI chatbot that described how the digital-twin profile informed the portfolio recommendation.

Appendix C. Codebook, Construct Diagnostics, and Secondary Outcomes

The main analytical variables, construct diagnostics, and selected secondary outcomes are summarized below. Table A5 reports the key analytical variables used in the experiment.

Table A5. Key analytical variables.

Variable	Description	Coding / scale
Condition	Experimental condition	1 = C1; 2 = C2; 3 = C3; 4 = C4
DT	Digital-twin treatment	1 = Digital twin; 0 = Standard profile
CAI	Conversational-AI treatment	1 = Chatbot; 0 = Plain dashboard
Investment experience	Self-rated investment experience	1 = Novice; 5 = Professional
Robo-advisor familiarity	Familiarity with robo-advisors	1 = Not at all familiar; 7 = Very familiar
Chatbot familiarity	Familiarity with AI chatbots	1 = Not at all familiar; 7 = Very familiar
Financial knowledge	Self-rated financial knowledge	1 = Very low; 7 = Very high
Baseline privacy concern	General concern financial-app data use	1 = Strongly disagree; 7 = Strongly agree
Allocation willingness	Amount assigned to the robo-advisor	USD 0–10,000
Portfolio acceptance	Accepted as shown or with small changes	1 = Accepted; 0 = Otherwise
Digital-twin recognition	Mean of two manipulation-check items	1–7
Chatbot recognition	Mean of two manipulation-check items	1–7
Perceived personalization	Mean of four personalization items	1–7
Social presence	Mean of three social-presence items	1–7
Trust	Mean of four trust items	1–7
Privacy concern	Mean of four privacy-concern items	1–7
Data control / transparency	Mean of three data-control items	1–7
Usefulness	Mean of four usefulness items	1–7
Fiduciary quality	Mean of six fiduciary-quality items	1–7
Adoption intention	Mean of three adoption-intention items	1–7
Advice compliance	Mean of three advice-compliance items	1–7
Data-sharing willingness	Mean of seven data-sharing items	1–7

The reliability diagnostics are reported to clarify how each composite indicator is interpreted in the analysis. Perceived personalization and privacy concern showed acceptable internal consistency and are therefore treated as the main interpretable composite outcomes. Social presence was close to the conventional reliability threshold and is interpreted cautiously. Several other indices, including

trust, data control/transparency, usefulness, fiduciary quality, adoption intention, advice compliance, and data-sharing willingness, displayed low alpha values.

These low values suggest either that the items did not capture a single underlying construct in this sample or that the relevant judgments were multidimensional in the scenario-based experimental context. For this reason, these indices are retained only as exploratory descriptive indicators of participants' responses to the experimental scenarios. The revised analysis therefore avoids relying on these weaker indices as primary evidence for mechanism claims and treats findings involving them as suggestive rather than confirmatory.

Table A6. Interpretation of composite indicators based on reliability diagnostics.

Category	Construct	Interpretation
Main interpretable composite outcomes	Perceived personalization; privacy concern	These constructs showed acceptable internal consistency and receive the strongest inferential weight in the analysis.
Secondary or borderline composite outcome	Social presence	This construct was close to the conventional reliability threshold and is interpreted cautiously.
Exploratory descriptive indicators	Trust; data control/transparency; usefulness; fiduciary quality; adoption intention; advice compliance; data-sharing willingness	These indices showed low internal consistency and are treated as exploratory descriptive indicators, not as validated latent constructs.

Table A7. Construct diagnostics.

Construct	Items	Mean	SD	Alpha
Perceived personalization	4	4.57	0.85	0.77
Social presence	3	4.23	0.87	0.69
Trust	4	4.43	0.52	0.43
Privacy concern	4	4.38	0.77	0.70
Data control/transparency	3	4.25	0.50	0.13
Usefulness	4	4.57	0.56	0.50
Fiduciary quality	6	4.37	0.40	0.34

Adoption intention	3	4.34	0.59	0.26
Advice compliance	3	4.22	0.59	0.31
Data-sharing willingness	7	3.52	0.45	0.40

Note: Alpha values are reported as diagnostics. Perceived personalization and privacy concern showed acceptable internal consistency, while social presence was close to the conventional threshold. Trust, data control/transparency, usefulness, fiduciary quality, adoption intention, advice compliance, and data-sharing willingness displayed low internal consistency and are therefore treated as exploratory descriptive indicators rather than validated latent constructs.

Table A8. Secondary outcomes.

Outcome	Full sample mean	SD	Interpretation
Data control/transparency	4.25	0.50	Moderate perceived control and transparency, interpreted as exploratory because of low internal consistency.
Fiduciary quality	4.37	0.40	Moderate perceived investor alignment, interpreted as exploratory because of low internal consistency.
Data-sharing willingness	3.52	0.45	Below-midpoint willingness to share additional data, interpreted as exploratory because of low internal consistency.

Appendix D. Robustness Checks

This appendix summarizes additional analyses used to assess the stability of the main findings. The first robustness check directly addresses the observed baseline imbalances in gender, investment experience, and baseline privacy concern.

Table A9. Summary of robustness checks.

Check	Specification	Main result
Covariate-adjusted models addressing baseline imbalance	Controls for baseline imbalances and key respondent characteristics.	The adjusted results remained similar for perceived personalization and privacy concern, and broadly consistent for social presence and allocation willingness. Results for weaker composites remain exploratory.
No prior robo-advisor use	Subsample restricted to respondents with no prior use (N = 189)	DT and CAI effects remained positive for the main perceptual outcomes, adoption intention, and allocation willingness.

Italy-only subsample	Subsample restricted to Italy respondents (N = 202)	DT increased perceived personalization and privacy concern, CAI increased social presence, and both treatments remained positively associated with allocation willingness.
Allocation willingness	Kruskal–Wallis test across the four conditions	H = 139.26, p < 0.001; the highest allocation willingness remained in C4.
Portfolio acceptance	Logistic regression on binary acceptance of the recommended portfolio	OR_DT = 2.11, p = 0.001; OR_CAI = 1.57, p = 0.048; OR_interaction = 1.19, p = 0.706.
Mechanism tests	Bootstrapped indirect effects with 5,000 resamples	The pathway evidence is exploratory, and the privacy pathway remained negative but non-significant.

References

- Jung, D.; Dorner, V.; Glaser, F.; Morana, S. Robo-advisory: Digitalization and automation of financial advisory. *Bus. Inf. Syst. Eng.* **2018**, *60*, 81–86. doi:10.1007/s12599-018-0521-9.
- Belanche, D.; Casaló, L.V.; Flavián, C. Artificial intelligence in FinTech: Understanding robo-advisors adoption among customers. *Ind. Manag. Data Syst.* **2019**, *119*, 1411–1430. doi:10.1108/IMDS-08-2018-0368.
- D'Acunto, F.; Prabhala, N.; Rossi, A.G. The promises and pitfalls of robo-advising. *Rev. Financ. Stud.* **2019**, *32*, 1983–2020. doi:10.1093/rfs/hhz014.
- Brenner, L.; Meyll, T. Robo-advisors: A substitute for human financial advice? *J. Behav. Exp. Finance* **2020**, *25*, 100275. doi:10.1016/j.jbef.2020.100275.
- Anshari, M.; Almunawar, M.N.; Masri, M. Digital twin: Financial technology's next frontier of robo-advisor. *J. Risk Financ. Manag.* **2022**, *15*, 163. doi:10.3390/jrfm15040163.
- Bonelli, M.I.; Döngül, E.S. Robo-advisors in the financial services industry: Recommendations for full-scale optimization, digital twin integration, and leveraging natural language processing trends. In *Proceedings of the 2023 9th International Conference on Virtual Reality (ICVR)*, Xianyang, China, 14 May 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 268–275. doi:10.1109/ICVR57957.2023.10169615.
- Bonelli, M.I.; Liu, J. Revolutionizing robo-advisors: Unveiling global financial markets, AI-driven innovations, and technological landscapes for enhanced investment decisions. *Adv. Sci. Technol. Eng. Syst. J.* **2024**, *9*, 33–44. doi:10.25046/aj090205.
- Hildebrand, C.; Bergner, A. Conversational robo advisors as surrogates of trust: Onboarding experience, firm perception, and consumer financial decision making. *J. Acad. Mark. Sci.* **2021**, *49*, 659–676. doi:10.1007/s11747-020-00753-z.
- Bol, N.; Dienlin, T.; Kruikemeier, S.; Sax, M.; Boerman, S.C.; Strycharz, J.; Helberger, N.; de Vreese, C.H. Understanding the effects of personalization as a privacy calculus: Analyzing self-disclosure across health, news, and commerce contexts. *J. Comput.-Mediat. Commun.* **2018**, *23*, 370–388. doi:10.1093/jcmc/zmy020.
- Back, C.; Morana, S.; Spann, M. When do robo-advisors make us better investors? The impact of social design elements on investor behavior. *J. Behav. Exp. Econ.* **2023**, *103*, 101984. doi:10.1016/j.socec.2023.101984.
- Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q.* **1989**, *13*, 319–340. doi:10.2307/249008.
- Komiak, S.Y.X.; Benbasat, I. The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Q.* **2006**, *30*, 941–960. doi:10.2307/25148760.
- Lourenço, C.J.S.; Dellaert, B.G.C.; Donkers, B. Whose algorithm says so: The relationships between type of firm, perceptions of trust and expertise, and the acceptance of financial robo-advice. *J. Interact. Mark.* **2020**, *49*, 107–124. doi:10.1016/j.intmar.2019.10.003.

14. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User acceptance of information technology: Toward a unified view. *MIS Q.* **2003**, *27*, 425–478. doi:10.2307/30036540.
15. de Oliveira Santini, F.; Hasan Jafar, S.; Junior Ladeira, W.; Rasul, T. Drivers of robo-advisors adoption in financial services: A meta-analytical study. *Int. J. Bank Mark.* **2026**, *44*, 439–453. doi:10.1108/IJBM-05-2025-0350.
16. Mayer, R.C.; Davis, J.H.; Schoorman, F.D. An integrative model of organizational trust. *Acad. Manag. Rev.* **1995**, *20*, 709–734. doi:10.5465/amr.1995.9508080335.
17. Glikson, E.; Woolley, A.W. Human trust in artificial intelligence: Review of empirical research. *Acad. Manag. Ann.* **2020**, *14*, 627–660. doi:10.5465/annals.2018.0057.
18. Nourallah, M.; Öhman, P.; Amin, M. No trust, no use: How young retail investors build initial trust in financial robo-advisors. *J. Financ. Report. Account.* **2023**, *21*, 60–82. doi:10.1108/JFRA-12-2021-0451.
19. Cao, X.; de Zwaan, L.; Wong, V. Building trust in robo-advisory: Technology, firm-specific and system trust. *Qual. Res. Financ. Mark.* **2026**, *18*, 408–428. doi:10.1108/QRFM-02-2024-0033.
20. Luo, H.; Liu, X.; Lv, X.; Hu, Y.; Ahmad, A.J. Investors' willingness to use robo-advisors: Extrapolating influencing factors based on the fiduciary duty of investment advisors. *Int. Rev. Econ. Financ.* **2024**, *94*, 103411. doi:10.1016/j.iref.2024.103411.
21. Fatima, S.; Chakraborty, M. Adoption of artificial intelligence in financial services: The case of robo-advisors in India. *IIMB Manag. Rev.* **2024**, *36*, 113–125. doi:10.1016/j.iimb.2024.04.002.
22. Bhat, A.K.; Krishnan, G. A synergistic review: Revolutionizing financial systems through digital twins. In *Proceedings of the 2025 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, Coimbatore, India, 2025; pp. 939–944. doi:10.1109/ICoICI65217.2025.11253753.
23. Tahvildari, M. Analysis of investor profiling and portfolio recommendations by robo-advisors in Germany. *Glob. Bus. Finance Rev.* **2025**, *30*, 19–33. doi:10.17549/gbfr.2025.30.8.19.
24. Tang, Y.; Son, H. How the sociality of AI digital human advisors shapes user experience value in digital finance: The mediating role of social presence. *J. Theor. Appl. Electron. Commer. Res.* **2026**, *21*, 79. doi:10.3390/jtaer21030079.
25. Wang, C.; Aw, E.C. Trust (or not) the bot? Understanding recommendation adoption and stickiness to the robo-advisor. *Int. J. Bank Mark.* **2026**, in press. doi:10.1108/IJBM-06-2025-0420.
26. Brügggen, L.; Gianni, R.; de Haan, F.; Hogleve, J.; Meacham, D.; Post, T.; van der Werf, M. AI-based financial advice: An ethical discourse on AI-based financial advice and ethical reflection framework. *J. Public Policy Mark.* **2025**, *44*, 436–456. doi:10.1177/07439156241302279.
27. Aristei, D.; Gallo, M. Financial literacy, robo-advising, and the demand for human financial advice: Evidence from Italy. *J. Behav. Exp. Finance* **2026**, *49*, 101125. doi:10.1016/j.jbef.2025.101125.
28. Bonelli, M.I. Digital twin robo-advisor adoption dataset V_2. *Mendeley Data* **2026**, **V1**. doi:10.17632/39rw5ywj8r.1.
29. Hair, J.F.; Risher, J.J.; Sarstedt, M.; Ringle, C.M. When to use and how to report PLS-SEM. *Eur. Bus. Rev.* **2019**, *31*, 2–24. doi:10.1108/EBR-11-2018-0203.

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.