

Interactivity and Illusions of Ability: How Using Generative AI Affects Investor Judgments

Joseph Croom

A dissertation

submitted in partial fulfillment of the  
requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Stephanie Grant, Chair

Frank Hodge

Elizabeth Blankespoor

Program Authorized to Offer Degree:

Foster School of Business

©Copyright 2025

Joseph Croom

University of Washington

**Abstract**

Interactivity and Illusions of Ability: How Using Generative AI Affects Investor Judgments

Joseph Croom

Chair of the Supervisory Committee:

Stephanie Grant

Foster School of Business

In this study, I use the setting of Generative AI (GenAI) to examine how processing tool interactivity affects investors' self-assessments of ability and willingness to invest. Although GenAI can help investors process financial information, I theorize that the interactive nature of GenAI blurs the boundaries between investors' own abilities and those of GenAI, prompting investors to discount their reliance on GenAI and misattribute its abilities to themselves. I rely on the advantages of a laboratory setting to disentangle the interactive element of GenAI from the mere presence of GenAI assistance. Across three experiments, I find that the interactivity underpinning GenAI heightens investors' self-assessments of their own abilities and increases their willingness to invest, despite this interactivity not improving, and in fact hindering, their actual processing of information provided by GenAI. My study thus highlights one potential cost of using GenAI and other highly interactive processing tools.

## I. INTRODUCTION

Generative AI (GenAI) is revolutionizing how investors process financial information (Fox [2023], Rai [2023], Blankespoor, Croom, and Grant [2024]). A defining quality that sets GenAI apart from other processing tools is its interactivity—defined as a tool’s capacity for reciprocal two-way communication, adaptive responsiveness, and engaging real-time responses (Johnson, Bruner II, and Kumar, 2006). While GenAI offers powerful processing capabilities, psychology theory suggests its interactive nature could carry unforeseen consequences, such as increasing investors’ self-assessments of their own abilities, which prior research shows has important consequences for their willingness to invest. In this study, I use GenAI as a setting to explore how processing tool interactivity affects investors’ self-assessments of their own abilities, and in turn, their willingness to invest.

This research question is important for several reasons. First, survey evidence finds 47% of retail investors have already used GenAI to process financial information or inform investment decisions (Blankespoor et al. [2024]). This rapid adoption is likely to persist as developers release more GenAI-enabled devices, browsers, and investment research tools (Public.com [2023], Gallagher [2024]). With highly interactive processing tools like GenAI chatbots becoming increasingly integrated in investors’ information environment, it is important to understand their effects on investors. Second, regulators are seeking a better understanding of how emerging financial technologies affect investor judgments, including a specific interest in the effects of GenAI (SEC [2021], [2023a], [2024], [2025], U.S. Department of the Treasury [2024]). I inform this understanding by exploring potential consequences of GenAI and other highly interactive financial processing tools for investors. Finally, finance and accounting studies have extensively documented that heightened self-assessments increase investors’ willingness to invest in

individual stocks, ultimately leading to more aggressive and detrimental trading behavior (Barber and Odean [2002], Statman, Thorley, and Vorkink [2006], Graham, Harvey, and Huang [2009], Elliott, Rennekamp, and White [2015]). However, little is known about features of investors' information environment that give rise to heightened self-assessments. Understanding these features is important because it can help develop strategies and practices to mitigate the negative consequences of heightened self-assessments while still realizing the benefits of GenAI.<sup>1</sup>

Drawing from psychology theory, I propose processing tool interactivity affects investors' self-assessments of their own abilities. One key consequence of processing tool interactivity is increased perceived control over the tool's output (Johnson, Bruner II, and Kumar [2006]). I expect this perceived control to blur the boundaries between investors' own abilities and those of the tool, making it difficult for investors to discern where their abilities end and the tool's abilities begin. Psychology theory suggests these blurred boundaries will cause investors to discount their reliance on the tool and misattribute its processing abilities to themselves (e.g., Fisher, Goddu, and Keil [2015]). Accordingly, I expect that after using a more interactive tool like GenAI to process financial information, investors will have heightened self-assessments of their own financial information processing ability. While I specifically expect GenAI's interactivity to drive this effect, in the natural setting, it is difficult to disentangle this feature of GenAI from the mere presence of GenAI assistance. Therefore, I rely on the advantages of a laboratory setting to create GenAI at two levels of interactivity, allowing me to isolate the effect of GenAI's interactivity.

---

<sup>1</sup> Investor self-assessments are important even in a world with ubiquitous GenAI tools because they act as an intuitive heuristic cue that shapes investment judgments (e.g., Hoffmann and Post [2016]). Even when using GenAI, investors with heightened self-assessments may overestimate the proportion of their total abilities that stem from themselves versus GenAI, thus influencing their investment judgments. Moreover, investors are unlikely to rely on GenAI for all tasks, and the heightened self-assessments developed through some GenAI interactions likely impact the numerous ad hoc investment judgments they make without it.

My first experiment uses a  $2 \times 1 + 1$  between-participants design, with processing tool interactivity (high versus low) as the manipulated independent variable. All participants review two earnings releases and answer questions about the financial performance of two companies. Those in the  $2 \times 1$  use a fictional ChatGPT-powered tool called FinGPT to help answer questions. I manipulate processing tool interactivity by varying FinGPT's reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity (Johnson et al. [2006]). Participants in the high interactivity condition actively type and submit questions to FinGPT in a conversation-like thread, receiving custom responses that FinGPT types out one word at a time. In contrast, participants in the low interactivity condition do not submit questions but instead view a static output that displays the same questions and FinGPT responses as would be generated in the high interactivity condition. Holding constant the content of FinGPT responses ensures participants receive the same information across conditions. I also include a benchmark condition in which participants answer the questions without FinGPT. I measure self-assessments of financial information processing ability by asking participants to evaluate how well they could answer hypothetical financial information processing questions *without outside assistance*. Thus, I capture participants' perceptions of their own abilities, in isolation of the tool.

Results of Experiment 1 support my prediction. Investors who use a more interactive FinGPT have heightened self-assessments of financial information processing ability relative to those who use a less interactive FinGPT or no FinGPT. Meanwhile, I do not detect a significant difference between investors who use a less interactive FinGPT and those who do not use FinGPT, consistent with the tool's interactivity driving heightened self-assessments, rather than merely the tool's assistance. Interestingly, despite receiving the same information, participants who use the more interactive FinGPT perform worse on an information processing quiz than those who use the

less interactive FinGPT. Consistent with theory, these findings suggest that heightened self-assessments in the high interactivity condition stem from investors misattributing the tool's abilities to themselves, rather than from actual improvements in processing.

Experiment 2 builds on these findings by further isolating the core construct of interactivity and its central role in driving misattribution and heightened self-assessments. Whereas Experiment 1 varied whether participants actively asked questions to FinGPT or viewed pre-generated output, Experiment 2 holds constant that all participants actively ask questions and instead varies the four facets of processing tool interactivity through FinGPT's interface and responses (see Appendix B for details). Results again show that investors who use a more interactive FinGPT have heightened self-assessments relative to those who use a less interactive FinGPT. Thus, Experiment 2 confirms that the effects observed in Experiment 1 are driven by the broader construct of interactivity, not just the question-asking element of interactivity alone.

Experiment 3 extends the converging evidence from Experiments 1 and 2 by testing *when* interactivity-driven misattribution is most likely to occur in practice and exploring the downstream effects of heightened self-assessments on willingness to invest. First, I expect that GenAI's interactivity heightens self-assessments more when GenAI is used for more complex relative to less complex tasks. This is because task complexity widens the gap between what investors can do with a processing tool and what they could do without the tool, creating more opportunity for investors to misattribute GenAI's abilities to themselves. Second, I expect GenAI's interactivity not only heightens self-assessed processing ability, but also self-assessed investment ability. In turn, I expect these heightened self-assessments increase investors' willingness to invest, as prior research suggests that feelings of competence with respect to an investment increase one's intention to pursue that investment (Elliott et al. [2015])

Experiment 3 uses a  $2 \times 2$  between-participants design with two manipulated factors: processing tool interactivity (high versus low) and task complexity (more complex versus less complex). The task structure and manipulation of processing tool interactivity are largely similar to Experiment 1, except the task in Experiment 3 involves an investment decision. Following Bonner [2008], I manipulate task complexity by varying the number of processing steps involved in answering the questions in my task. Consistent with my predictions, results show that GenAI's interactivity heightens self-assessments when investors use GenAI for a more complex task, but not a less complex task. In turn, these heightened self-assessments increase investors' willingness to invest. Taken together, results from the three experiments show that the interactivity underpinning GenAI leads investors to misattribute the abilities of GenAI to themselves, resulting in inflated self-assessments and increased willingness to invest.

My study makes several contributions. First, I contribute to the growing literature that explores how emerging technologies shape investor judgments. Prior research has examined the effects of technologies such as XBRL (Hodge, Kennedy, and Maines [2004], Blankespoor [2019]), social media (Blankespoor, Miller, and White [2014], Elliott, Grant, and Hodge [2018], Rennekamp and Witz [2021]), smartphones and push notifications (Clor-Proell, Guggenmos, and Rennekamp [2020], Grant [2020], Elliott, Gale, and Hobson [2022]), and news-trade integration (Cade, Croom, and Kamrath [2024]). I extend the literature by exploring a new transformative technology—GenAI. While anecdotal evidence and concurrent archival work illustrate the potential usefulness of GenAI for processing financial information (Bai et al. [2023], Cutter [2024], Kim, Muhn, and Nikolaev [2024]), I show one drawback. Specifically, the interactivity underpinning GenAI inflates investor self-assessments of their own abilities and increases their willingness to invest, despite hindering their processing of information the tool provides.

I also contribute to the evolving psychology literature on blurred boundaries and metacognitive consequences of technology. Prior research suggests individuals often blur the line between their own knowledge and that of the internet, developing inflated self-assessments after searching the internet (Fisher et al. [2015], Hamilton and Yao [2018], Ward [2021]). However, Eliseev and Marsh [2023] argue that these studies confound several underlying features of internet search and suggest a need to identify more generalizable constructs that could produce these blurring effects in other tools. I extend the literature by establishing processing tool interactivity and task complexity as important, generalizable constructs that jointly increase misattribution and inflate self-assessments. These findings deepen our understanding of when and why individuals are likely to blur boundaries between themselves and the various processing tools they use. I also contribute by showing that these blurring effects arise in a GenAI setting, even though GenAI is a novel technology and users may be more conscious of their reliance on it. Importantly, as investors increasingly integrate GenAI into their environment, the boundaries between investors and GenAI could become increasingly blurred, exacerbating the consequences of inflated self-assessments.

Lastly, my study has practical implications for regulators. The SEC has labeled emerging financial technologies as a top priority, with a particular focus on tools that may unknowingly steer retail investors toward behavior inconsistent with their investment goals (SEC [2021], [2023b]). While GenAI offers numerous benefits to retail investors, my findings suggest that its interactivity inflates investors' self-assessments of their abilities. Importantly, these inflated self-assessments are associated with an increased willingness to invest, which can manifest in practice in the form of more aggressive investing behavior (e.g., Barber and Odean [2000], [2002], Graham et al. [2009], Grinblatt and Keloharju [2009]). These findings can assist the SEC in developing a more holistic understanding of GenAI's behavioral consequences.

## II. BACKGROUND AND HYPOTHESIS DEVELOPMENT

### Background on Generative AI

Generative AI (GenAI) has emerged as a transformative technology with wide-ranging applications. Until recently, AI applications were constrained to narrow and specific tasks that often required specialized expertise or infrastructure, such as audit firms using AI to evaluate specific estimates (Commerford et al. [2022]) and hedge funds using AI for sentiment analysis (FINRA [2020]). However, the emergence of GenAI has broadened the scope of tasks AI can perform. GenAI tools, such as ChatGPT, stand out from existing processing tools in their ability to quickly generate conversational and contextually relevant responses to custom user prompts. These features make GenAI particularly attractive to those looking for help to simplify otherwise challenging or time-consuming tasks, as they offer a more intuitive and flexible way to get tailored assistance without requiring technical expertise or costly infrastructure.

Among those poised to benefit the most from GenAI are retail investors. Historically, retail investors have had limited access to advanced processing tools and generally relied on manual analysis and basic search tools (e.g., XBRL, CTRL-F search) to process financial information. However, GenAI now offers a highly accessible and advanced processing tool with several new advantages. First, GenAI can access and process a wide range of structured and unstructured data sources, allowing investors to process information they may not otherwise be capable of processing, such as regulatory filings, conference call transcripts, and analyst reports (Public.com [2023], Sheikh [2023]). Second, GenAI excels at identifying and simplifying complex patterns, trends, and interdependencies within data, helping investors perform complex integration tasks (Bloomberg [2023]). Finally, GenAI's ability to generate conversational responses allows for natural language interactions, simplifying the extraction and interpretation of financial insights.

Thus, whereas existing technologies such as XBRL can assist with simple information acquisition tasks (Hodge et al. [2004]), GenAI gives everyday retail investors a personal assistant to help with all types of information processing tasks. For these reasons, many expect GenAI to become a standard tool for retail investors evaluating financial information (Blankespoor, Croom, and Grant [2024]).

In fact, many retail investors are already using GenAI, according to early survey evidence from Blankespoor, Croom, and Grant [2024]. Based on a survey of retail investors recruited from both Prolific and a popular brokerage platform, 47% of retail investors report using GenAI to process financial information or inform investment decisions, with 28% doing so at least once per week. Investors most commonly report using GenAI to help make sense of financial information and list “easier to process complex information” as a top benefit of GenAI, underscoring investors’ reliance on GenAI for complex processing tasks.

While most retail investors use general-purpose GenAI tools like ChatGPT (Blankespoor, Croom, and Grant [2024]), developers are also creating GenAI tools geared specifically toward investors. For example, Public.com, a retail brokerage platform with over three million members, has a GenAI tool that uses SEC filings, market data, earnings call transcripts, news reports, and analyst reports to help answer custom retail investor questions about specific securities (Public.com [2023]). Likewise, many startup companies have developed GenAI investment research interfaces catered specifically toward retail investors (e.g., FinChat.io [2023], Pluto.Fi [2023]). As these technologies continue to develop and spread to more platforms, a growing number of investors are likely to rely on GenAI when evaluating financial information.

However, as with any emerging technology, GenAI is not without its potential pitfalls. Regulators have several concerns about GenAI, including data security, intellectual property

rights, the risk of propagating misinformation, and threats posed by malevolent actors (Zakrzewski [2023]). With respect to financial markets, the SEC continues to seek an understanding of how emerging technologies, including GenAI, affect individual investors (SEC [2021], [2023b], [2025]). For instance, the SEC is concerned that new technologies may unconsciously increase investors' willingness to invest, leading to more aggressive trading behavior that is inconsistent with investors' goals or risk tolerance (SEC [2021]). As I discuss next, psychology theory suggests that in the context of investors using GenAI, these concerns may be warranted.

### **Metacognitive Judgments**

Metacognitive judgments are self-assessments of one's cognitive abilities in a given domain (Fleming [2024]). Such judgments are typically an unobservable and often unconscious internal process. However, they have important consequences for individual decision-making, and miscalibrated metacognitive judgments can lead to negative outcomes. For example, a student's self-assessment is an important factor in their decision to trade off studying and leisure. If the student has inflated self-assessments, they may choose to forgo needed study time and end up with a lower-than-desired exam grade. Metacognitive judgments are especially important in an investing context because investors rely on them to make decisions that directly affect their wealth. If investors have miscalibrated metacognitive judgments, it can significantly affect their investment decision-making. For instance, research shows that heightened self-assessments can increase investors' willingness to invest, as reflected by more aggressive and detrimental trading behavior (Barber and Odean [2000], [2001], [2002], Statman et al. [2006], Graham et al. [2009], Grinblatt and Keloharju [2009], Elliott et al. [2015]). Given the negative consequences that can arise from miscalibrated metacognitive judgments, it is important to identify features in investors' information environment that can heighten their self-assessments. I argue processing tool

interactivity is one such feature and use GenAI as a setting to explore how processing tool interactivity affects investor self-assessments of their own abilities.<sup>2</sup>

### **The Effect of Processing Tool Interactivity**

Processing tool interactivity is defined by the extent to which a processing tool exhibits four key features: reciprocity, responsiveness, response speed, and nonverbal communication (Johnson et al. [2006]). *Reciprocity* refers to the level of iterative, two-way communication between the user and tool. Tools that facilitate an ongoing and balanced two-way dialogue are more interactive than those limited to one-sided or single-instance exchanges. *Responsiveness* refers to the degree to which the tool's output is a function of the user's input, with interactivity increasing as the output more closely aligns with the user's specific requests or actions. *Response speed* refers to the immediacy with which the tool generates output, with quicker output reflecting greater perceived interactivity. Finally, *nonverbal communication* refers to the use of supplementary channels, such as visual or auditory cues, to communicate output, with interactivity increasing as the tool incorporates these additional elements.<sup>3</sup>

In the context of GenAI, these four features are particularly pronounced. First, GenAI supports a natural and conversational prompt-response exchange that allows investors to engage in iterative dialogue, ask follow-up questions, and reference previous communications, thus increasing reciprocity. Second, GenAI is highly responsive in that it recognizes nuanced context

---

<sup>2</sup> I use the term "self-assessments" rather than "confidence" to reflect the broad metacognitive judgments in this study. Confidence typically refers to a specific belief about one's decision or performance on an individual task, whereas self-assessments encompass broader perceptions of ability, which is the focus of my paper. However, research considers self-assessments a precursor to confidence (Van Marcke et al. [2024]).

<sup>3</sup> Processing tool interactivity is distinct from disclosure interactivity, defined by Brown, Gale, and Grant [2022] (BGG) as "user involvement in directing the form or content of the information displayed." Whereas disclosure interactivity reflects investors directing the form of information *they process* (e.g., drilldown functionality), processing tool interactivity generates perceptions of control over the *processing performed by a tool*. This difference in who performs the processing (i.e., investor vs. tool) and the involvement of an external tool is why I introduce misattribution and blurred boundaries theory in my paper rather than relying on similar theory to BGG.

and dynamically responds to custom requests. Short of requesting explicit investment advice or recommendations, investors can instruct GenAI to perform any number of tasks and it will produce a highly customized, contextually relevant, natural-language response. Third, GenAI's advanced processing capabilities allow it to provide responses almost immediately after investors submit a prompt. Finally, GenAI exhibits nonverbal communication by presenting responses in a unique, real-time "typing" display, where answers appear word by word in front of investors. These elements collectively make GenAI distinctly interactive compared to traditional processing tools.

A key consequence of interactivity is an increased sense of control over a tool's output (Skinner [1996], Johnson et al. [2006]). With GenAI, for instance, the ability to engage in back-and-forth dialogue and see customized and conversational responses being formulated in real-time likely makes investors feel a greater sense of involvement, ownership, and control over GenAI's processing and output. However, I expect this perception of control blurs the boundaries between users' own abilities and those of the tool, making it difficult for investors to discern where their own financial information processing abilities end and the tool's abilities begin.

Psychology theory suggests that these blurred boundaries could cause investors to internalize the tool's knowledge and abilities as their own (Wegner [1987], Fisher et al. [2015], Hamilton and Yao [2018], Ward [2021]). For example, Fisher et al. [2015] argue that the boundaries between individuals and the internet have become blurred over time as the internet has become a seamless extension of people's cognition. Using a trivia knowledge setting with internet search, they find that these blurred boundaries cause investors to mistake the internet's "knowledge" as their own and develop inflated self-assessments of their ability to answer future trivia questions without the internet. Hamilton and Yao [2018] and Ward [2021] replicate this result and further isolate the key role of blurred boundaries in driving these effects.

Drawing on this theory, I expect the interactivity of a processing tool and the resulting perception of control likely cause investors to view the tool as a more seamless extension of their own cognition. As a result of these blurred boundaries, investors are likely to discount their reliance on the tool and misattribute its processing abilities to themselves. Accordingly, I expect that after using a more interactive tool like GenAI to process financial information, investors will have heightened self-assessments of their own financial information processing ability.

Importantly, my theory suggests that the interactive nature of GenAI will drive this effect rather than the mere presence of GenAI assistance. However, these elements are confounded in the natural setting. Therefore, I rely on the advantages of a laboratory experiment to create a less interactive GenAI which allows me to disentangle the interactive element of GenAI from the mere presence of GenAI assistance. I then test my theory by comparing investors who use an interactive GenAI tool to those who do not use a GenAI tool, as well as those who use a less interactive GenAI tool. Formally, I make the following prediction:

**H1:** *Investors have heightened self-assessments of their own financial information processing ability after using a more interactive GenAI tool to process financial information compared to a less interactive GenAI tool or no GenAI tool.*

My prediction builds on prior psychology research in several ways. First, while prior research has examined misattribution and blurred boundaries in internet search contexts, Eliseev and Marsh [2023] show that these studies confound several elements of internet search, suggesting a need to identify underlying constructs that produce these blurring effects in processing tools more generally.<sup>4</sup> To this end, I isolate and examine the important role of processing tool interactivity, a broader and more generalizable construct, in blurring boundaries between

---

<sup>4</sup> For instance, prior studies on internet search jointly vary several operational features, including the presence of help from the internet, active versus passive search processes, time and effort exerted, and repeated exposure to content (e.g., via Google result preview snippets) (Eliseev and Marsh [2023]).

individuals and tools. Second, whereas prior research has explicitly focused on misattribution of factual trivia knowledge, I examine the distinct construct of (financial information) processing ability. With discrete factual knowledge, misattribution is more rationalizable as individuals can more easily convince themselves they “knew it all along” (e.g., Ward [2021]). Comparatively, investors may have more trouble rationalizing that they “had the ability to process the information all along”, which may reduce the likelihood of misattribution in this setting. Finally, I extend the literature by exploring whether blurred boundary effects manifest in a GenAI setting. Unlike the internet, GenAI is a novel technology, so investors may be especially conscious of their reliance on GenAI or may exhibit algorithmic aversion (Commerford et al. [2022]), in which case they may not blur the boundaries between themselves and GenAI.

### **III. EXPERIMENT 1 METHOD**

#### **Design**

To test H1, I conduct a  $2 \times 1 + 1$  between-participants experiment, with processing tool interactivity (high versus low) as the manipulated independent variable. All participants review earnings releases for two companies and answer five questions about each company’s performance. Participants in the  $2 \times 1$  use a fictional GenAI processing tool called FinGPT to help them answer the questions. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal features (Johnson et al. [2006]). I also include a +1 benchmark condition in which participants answer the questions without FinGPT.

#### **Participants**

I recruit 282 undergraduate accounting students from a large public university to complete the study during class for course credit.<sup>5</sup> I drop 18 participants who fail to follow instructions and

---

<sup>5</sup> The Institutional Review Board (IRB) at the affiliated university approved the use of human subjects for the experiment reported in this paper.

do not use FinGPT, leaving a final sample of 264 participants who are 57% women and 43% men. I choose student subjects to match participants' knowledge to the task without using more sophisticated subjects than necessary (Libby, Bloomfield, and Nelson [2002]). My task requires familiarity with financial information. Participants have taken on average 2.92 accounting or finance courses and have experience evaluating financial information in their coursework. Although my task does not necessitate having experience with GenAI, such experience is desirable because I hope to speak to an effect present among investors who use GenAI. Eighty-five percent of participants have used GenAI and are thus familiar with its features and functionality.

### **Task and Procedures**

I administer the study to student participants during scheduled class time. I first provide verbal instructions and guidelines, after which students access the study on their own computers. Upon entering the study, participants are randomly assigned to one of three conditions. Participants learn they will receive financial information for two hypothetical companies in the athletic apparel industry and answer five questions about each company's financial performance. The financial information is modeled after actual companies in the athletic apparel industry and reflects neutral, average industry performance. Participants in the  $2 \times 1$  also learn they will use a GenAI tool called FinGPT to help answer the questions. To reduce noise and minimize concerns about the quality and legitimacy of FinGPT, I explicitly describe FinGPT as an advanced ChatGPT-enabled AI copilot used by investors to accurately and reliably answer questions about financial documents.

Following these instructions, participants spend some time reviewing the first company's earnings release, after which they follow a link to an external website to answer five questions. The questions require participants to extract financial information from the earnings release (i.e., acquisition) and interpret or evaluate the implications of the information (i.e., integration). For

instance, one question asks, “How are the company's different sales channels trending and what does this suggest for the future of the company?” I design the questions to be difficult enough that investors can benefit from GenAI assistance, but not so difficult that participants are unable to answer them in a reasonable amount of time.<sup>6</sup> See Appendix A, Panel A for a full list of questions asked. Participants can navigate back to the earnings release as needed to help answer the questions.

On the same page as the questions, participants in the  $2 \times 1$  conditions can also use FinGPT to help process the financial information and answer the five questions. In both conditions, FinGPT assists participants by extracting relevant information from the earnings release and providing an interpretation or evaluation of that information. While participants in the benchmark condition do not have access to FinGPT, they may use standard search functionality on the earnings release to help answer questions. In this sense, the benchmark condition is reflective of the standard processing tools widely available to retail investors (e.g., XBRL, PDF Keyword Search).

To maintain experimental control, participants have five minutes to answer the five questions. The time limit helps minimize time and effort differences across conditions and ensure participants stay on task.<sup>7</sup> It also ensures they complete the study in the limited class time. Participants advance once they answer the five questions or after five minutes, whichever comes first. Participants then repeat this process with a second company and a new set of five questions.

### **Processing Tool Interactivity Manipulation**

I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity

---

<sup>6</sup> I piloted eight potential questions with 51 Prolific participants. From this pilot, I chose five questions participants perceived as moderately difficult, with an average rating of 6.24 on a scale from 1 (*very easy*) to 11 (*very difficult*).

<sup>7</sup> For instance, participants in the high interactivity condition may spend spare time trying to test the boundaries of FinGPT by asking it irrelevant, unrelated, or inappropriate questions.

(Johnson et al. [2006]). In the high interactivity condition, participants actively type and submit questions to FinGPT in a conversation-like thread, receiving custom responses that FinGPT types out one word at a time. This setup allows for iterative two-way communication and tailored responses, reflecting FinGPT’s reciprocity and responsiveness. Additionally, the real-time typing effect reflects a key nonverbal communication feature that is common in GenAI in practice. While the fourth facet of interactivity—response speed—is not necessarily manipulated across conditions, the high interactivity FinGPT also provides responses nearly instantaneously.

I set up the high interactivity FinGPT using OpenAI’s gpt-3.5-turbo-0125 API model such that every question participants submit received a GPT-generated response typed out in real-time.<sup>8</sup> The key advantage of using a real GPT model is it allows FinGPT to recognize the context of the question to provide the appropriate information, regardless of how participants phrase questions. To maintain informational control, I layer additional information and instructions into each API call. First, I instruct the model to assume the role of FinGPT, a financial copilot designed to help retail investors process financial information. Second, I provide it with information from each earnings release and describe the five pieces of information I want it to highlight for each of the five questions. Third, I instruct it not to provide any information requested by participants that is beyond the scope of those five questions and to instead softly redirect participants toward the key information.<sup>9</sup> This design choice weakens my interactivity manipulation by limiting FinGPT’s capacity for responsiveness but ensures there are no informational differences across conditions.<sup>10</sup>

---

<sup>8</sup> For a video example of the high interactivity condition, see here: [https://www.youtube.com/watch?v=Gce-QX\\_E5Uo](https://www.youtube.com/watch?v=Gce-QX_E5Uo).

<sup>9</sup> For instance, if participants ask about tax expense, the GPT’s response says something like “I understand your question, but I am here to help you focus on the key financial information that highlights [company’s] performance. For now, tax expense is not particularly relevant to that objective. Consider asking another question.”

<sup>10</sup> Given the time constraints and FinGPT’s restrictions on off-topic questions, participants rarely ask off-topic questions or follow-up questions. FinGPT’s flexibility allows participants to ask paraphrased, shortened, or multi-part questions. To eliminate concerns about informational differences from question-asking behavior, FinGPT’s responses provide all information relevant to a topic, even when only asked half of a question.

Finally, I limit the model's temperature to 0.05 to ensure the model provides extremely consistent, predictable, and nearly identical responses when asked the same question over time.<sup>11</sup>

In the low interactivity condition, participants do not submit questions themselves but instead view a static text output containing pre-generated FinGPT responses to the five questions. The pre-generated responses mirror the responses generated by the high interactivity FinGPT when posed the same questions. Thus, I hold constant the information participants can receive from FinGPT and only manipulate the interactive nature of FinGPT. Specifically, the lack of user involvement in directing the conversation reduces FinGPT's reciprocity and responsiveness, while the removal of the real-time typing effect eliminates FinGPT's nonverbal communication element.<sup>12</sup> See Appendix A, Panel B for screenshots of the low and high interactivity interfaces.

### **Primary Dependent Measure**

I use a measure adapted from Fisher et al. [2015] to gauge self-assessments of financial information processing ability. After completing the second set of five questions, participants are asked to imagine they are evaluating another company's annual report or 10-K *without outside assistance* and assess how well they could answer hypothetical questions about the financial information. Specifically, participants rate their ability to provide detailed answers to questions about company background, financial numbers, and financial performance. For each topic, participants consider three sample questions. For example, when considering the topic of financial performance, the sample questions are 1) "How would you assess the company's overall financial

---

<sup>11</sup> Temperature is a number between 0 and 2 provided to a model to control the randomness of its responses, with higher temperatures producing more random and less consistent responses and lower temperatures producing more deterministic and consistent responses. I use an extremely low temperature for consistency but avoid using a temperature of zero because doing so can occasionally provide nonsensical responses.

<sup>12</sup> While most widely used GenAI tools are highly interactive (e.g., ChatGPT), there exist less interactive GenAI tools. For instance, some investment research platforms provide GenAI-powered summaries or insights companies or documents automatically within users' dashboards. (AlphaSense [2023]). These less interactive tools, characterized by passive information delivery rather than active information extraction, serve as real-world analogs to the less interactive condition in Experiment 1.

performance?”, 2) “In what areas has the company's performance improved or deteriorated?”, and 3) “How are the company's growth prospects and outlook?” The format is the same for the other two topics except participants see different sample questions. Appendix A, Panel C provides a complete list of sample questions on each topic. For each topic, participants answer “How well could you answer detailed questions about [topic] like these?” on a scale from 1 (*very poorly*) to 11 (*very well*). Importantly, participants respond based on their ability to answer these questions *without outside assistance* rather than their ability using external tools (e.g., FinGPT). Given random assignment to conditions, any differences in responses to these questions across conditions reflect differences in self-assessments of ability (Fisher et al. [2015]), rather than differences in participants’ *actual* ability to process financial information without outside assistance.

#### IV. EXPERIMENT 1 RESULTS

##### Manipulation Check

My theory suggests that a successful manipulation of processing tool interactivity should instill greater perceived control over the tool’s output. To assess the effectiveness of my manipulation, I ask participants to think back on FinGPT and complete the following statement: “I feel like I had \_\_\_\_\_ over the responses FinGPT generated for me.” Participants respond on a scale from 1 (*No control at all*) to 11 (*Complete control*). Consistent with successful manipulation, participants in the high interactivity condition feel more control over FinGPT responses than participants in the low interactivity condition ( $6.763 > 4.825$ ;  $t_{177} = 4.864$ ;  $p < 0.001$ , one-tailed).<sup>13</sup>

##### Test of H1

H1 predicts investors have heightened self-assessments of financial information processing ability after using a more interactive GenAI tool compared to a less interactive GenAI tool or no

---

<sup>13</sup> For ease of exposition, all reported  $p$ -values are two-tailed unless otherwise stated.

GenAI tool (a benchmark condition). Table 1, Panel A reports descriptive statistics and Figure 1 graphically displays means for *Self-Assessed Processing Ability*, along with each subcomponent, by condition.<sup>14</sup> Table 1, Panel B presents planned comparisons. First, there is no difference in *Self-Assessed Processing Ability* between the benchmark condition and low interactivity condition (6.106 vs. 6.217;  $t_{186} = 0.407$ ;  $p = 0.684$ ), which suggests receiving GenAI assistance does not, on its own, affect self-assessments of financial information processing ability. However, consistent with H1, participants in the high interactivity condition have higher *Self-Assessed Processing Ability* than those in the low interactivity condition ( $6.689 > 6.217$ ;  $t_{177} = 1.807$ ;  $p = 0.036$ , one-tailed) and those in the benchmark condition ( $6.689 > 6.106$ ;  $t_{159} = 2.032$ ;  $p = 0.022$ , one-tailed). Since these judgments are self-assessments of ability *without outside assistance* and actual ability should be randomized across conditions, these results are consistent with the interactivity underpinning GenAI prompting investors to misattribute GenAI abilities to themselves and develop inflated self-assessments of their own abilities.<sup>15</sup>

### **Supplemental Evidence: Processing of Financial Information**

One potential alternative explanation for my results is the more interactive FinGPT improved participants' actual ability to process financial information relative to the less interactive FinGPT, leading to potentially justified increases in self-assessed financial information processing ability. However, I design FinGPT to provide the same information across both levels of interactivity, so I do not expect GenAI interactivity to affect participants' processing of information in my study, *ex ante*. Nonetheless, to help rule out this and other processing-based

---

<sup>14</sup> Responses to the three components are internally consistent ( $\alpha = 0.91$ ) and load on a single factor (factor loadings  $> 0.91$ ; total variance explained = 84.85%) and so are averaged together into one measure of *Self-Assessments*.

<sup>15</sup> I also find that participants' responses to the earnings release questions are more similar to FinGPT's output in the high interactivity condition than the low interactivity condition (cosine similarity of 0.615 vs. 0.549;  $t_{172} = 2.347$ ;  $p = 0.020$ ). This finding is consistent with participants blurring the boundaries and internalizing FinGPT's abilities as their own. However, this finding does not appear to explain my results; when controlling for cosine similarity, I continue to find a significant effect of processing tool interactivity on self-assessments ( $p = 0.034$ , one-tailed).

explanations, I administer a ten-question quiz at the end of the study that aligns with the ten earnings release questions participants answer in the study. The quiz evaluates their processing (i.e., acquisition and integration) of the two companies' financial information. See Appendix A, Panel D for a list of quiz questions.

Table 2, Panel A reports descriptive statistics for quiz performance, both overall and by question. Figure 2 graphically displays mean quiz performance by condition. Table 2, Panel B presents planned comparisons. Two noteworthy patterns emerge. First, participants in the benchmark condition performed worse than participants in the low interactivity condition (4.624 vs. 6.650;  $t_{186} = 9.440$ ;  $p < 0.001$ ) and the high interactivity condition (4.624 vs. 6.053;  $t_{159} = 6.121$ ;  $p < 0.001$ ), validating that FinGPT indeed improved their ability to process the earnings release information. Despite the improvement in processing from using FinGPT, receiving FinGPT assistance on its own does not increase participants' self-assessments (i.e., low interactivity versus benchmark). Thus, investors' actual processing performance does not appear to be driving their heightened self-assessments, consistent with processing tool interactivity being the key driver.

Second, participants in the high interactivity condition performed *worse* on the quiz compared to those in the low interactivity condition (6.053 vs. 6.650;  $t_{177} = 2.487$ ;  $p = 0.014$ ).<sup>16</sup> The lower quiz performance suggests that processing tool interactivity *hindered* participants' processing of information provided by FinGPT. One likely explanation for this result is that using a more interactive processing tool cues that information may be easier to re-process in the future, if needed. Therefore, investors unconsciously commit more cognitive resources to processing *how* they retrieved the information from FinGPT (e.g., what they asked or how they asked it) than to

---

<sup>16</sup> Differences in processing persist after scaling by the number of earnings release questions participants answer (1.032 vs. 1.271;  $t_{177} = 2.833$ ;  $p = 0.005$ ). Furthermore, Table 2 suggests processing differences are stronger for quiz questions related to the information participants would have retrieved first for each company (i.e., Q1-Q3, Q6-Q7). These findings suggest results are not driven the amount of information participants could acquire in the allotted time.

processing the information provided by FinGPT. This explanation is consistent with prior psychology and accounting research (Sparrow, Liu, and Wegner [2011], Henkel [2014], Fisher, Smiley, and Grillo [2022], Cikurel [2024]) and with research on the effect of personalized interaction in disclosure settings (Elliott, Grant, and Hobson [2020]).

Taken together, my results suggest processing tool interactivity inflates investor self-assessments of financial information processing ability, despite not improving, and in fact hindering, their processing of information provided by the tool in the current task. These findings are consistent with processing tool interactivity blurring the boundaries between investors and GenAI and prompting investors to misattribute GenAI processing abilities to themselves.

## **V. EXPERIMENT 2**

Experiment 2 builds on the findings in Experiment 1 by further isolating the core construct of interactivity and its role in driving misattribution and inflated self-assessments. In Experiment 1, part of my manipulation involved varying whether participants actively ask questions. This was motivated by practice, as question-asking is a prominent GenAI feature that increases reciprocity and responsiveness. However, to establish construct generalizability, it is important to show that the observed effects stem from the broader construct of interactivity, not just from the question-asking element alone. Experiment 2 addresses this by holding question-asking constant across conditions, allowing me to validate that processing tool interactivity as a construct leads to misattribution and inflated self-assessments of ability.

Experiment 2 uses a two-cell design in which I manipulate processing tool interactivity (high versus low), holding constant that all participants actively submit questions to FinGPT. I recruit 136 undergraduate accounting students to complete the study during class for course credit. I drop 12 participants who do not use FinGPT and four participants who believe they could not

rely on FinGPT responses, despite verbal and written instructions telling them to assume FinGPT responses were reliable and accurate. This leaves a final sample of 120 participants. The experimental task is the same as Experiment 1, except there is no processing quiz at the end and both conditions may actively submit questions to FinGPT.

I manipulate processing tool interactivity by varying FinGPT's reciprocity, responsiveness, response speed, and nonverbal communication. The high interactivity FinGPT is identical to Experiment 1. I manipulate reciprocity by varying whether FinGPT displays output in an ongoing two-way conversation thread or displays only one response at a time. I manipulate responsiveness by varying whether FinGPT directly provides participant-requested information or provides participant-requested information alongside other information irrelevant to the focal question but relevant to other questions.<sup>17</sup> Similar to Johnson et al. [2006], this manipulation of responsiveness inherently manipulates reciprocity by varying the extent to which participants must engage back-and-forth with FinGPT to extract the information needed to answer all five questions. I manipulate response speed by varying whether FinGPT responds instantly or after 30 seconds. Finally, I manipulate nonverbal communication by varying whether FinGPT types out responses in real-time or displays the entire response all at once without animation. See Appendix B for a table comparing operationalizations across experiments.

Table 3, Panel A reports descriptive statistics and Figure 3 graphically displays means for *Self-Assessed Processing Ability*, along with each subcomponent, by condition.<sup>18</sup> Table 3, Panel B presents my planned comparison. Consistent with H1, participants in the high interactivity

---

<sup>17</sup> The choice to also display information relevant to other questions rather than fully irrelevant information ensures both conditions view all the same information in aggregate.

<sup>18</sup> Participants in the high interactivity condition feel more control over FinGPT responses than those in the low interactivity condition ( $6.631 > 5.073$ ;  $t_{118} = 2.88$ ;  $p = 0.001$ , one-tailed), consistent with a successful manipulation of processing tool interactivity.

condition have higher *Self-Assessed Processing Ability* than those in the low interactivity condition ( $6.662 > 6.133$ ;  $t_{118} = 1.641$ ;  $p = 0.052$ , one-tailed). This finding, together with Experiment 1, triangulates the key role of processing tool interactivity in blurring the boundaries between investors and processing tools. By holding question-asking constant, Experiment 2 provides clear evidence that it is the broader construct of interactivity—not merely the question-asking component of interactivity—that drives misattribution and inflates users’ self-assessments.

## VI. ADDITIONAL HYPOTHESIS DEVELOPMENT

While Experiments 1 and 2 establish that processing tool interactivity blurs the boundary between user and tool and causes investors to misattribute tool abilities to themselves, two important questions remain. First, *when* is interactivity-driven misattribution most likely to occur in practice? Second, do inflated self-assessments have downstream consequences for investment judgments in this setting? I address these questions in a third experiment by introducing task complexity as a moderating variable and incorporating an investment judgment into the experimental setup.

Bonner [2008] defines task complexity as a joint function of difficulty (e.g., number of processing steps) and structure (e.g., how clear or explicit the task requirements are). Financial information processing tasks naturally vary along both dimensions. Generally, individual investors are limited in their ability to effectively perform more complex tasks without GenAI assistance (Blankespoor, deHaan, and Marinovic [2020]). It thus makes sense that as investors grow more comfortable with GenAI, they increasingly rely on it for more complex tasks (Blankespoor et al. [2024], Cutter [2024]). I argue that such conditions—when users are most dependent on the tool—are precisely when GenAI-driven misattribution is most likely to emerge.

If inflated self-assessments arise from investors misattributing the tool’s capabilities to themselves, then the effect of processing tool interactivity should be stronger when the task offers greater opportunity for misattribution. I expect this to be the case when the processing task is more complex, as complexity widens the gap between what investors can do with a processing tool and what they could do without the tool (i.e., users are more dependent on the tool). For example, GenAI excels at more complex tasks, such as drawing connections between multiple pieces of information, something many investors cannot effectively do without GenAI assistance. In contrast, although GenAI may speed up less complex tasks like retrieving basic financial information, investors could still easily perform such a task without GenAI assistance. This relatively larger gap in capability for more complex tasks creates more room for investors to misattribute GenAI’s incremental abilities to themselves—as if they could have completed the task without GenAI assistance. Given the increased opportunity for misattribution, I expect interactivity to have a larger effect on investor self-assessments when the task is more complex. Accordingly, I make the following prediction:

**H2a:** *GenAI interactivity heightens self-assessments of financial information processing ability more when GenAI is used for more complex tasks relative to less complex tasks.*

Experiment 3 also explores how these inflated self-assessments ultimately affect investment judgments in my setting. Theory suggests that feelings of competence with respect to a specific investment (i.e., heightened self-assessments) increase one’s intention to pursue that investment, as they feel more equipped to evaluate the company and manage their investment (Heath and Tversky [1991], Elliott et al. [2015]). Indeed, extensive prior research indicates that heightened self-assessments significantly increase willingness to invest, as reflected by more aggressive trading behavior (e.g., greater trading frequency, more speculative trading, etc.) (Odean [1998], Barber and Odean [2000], [2001], [2002], Graham et al. [2009]). While these specific

trading behaviors are difficult to capture in my laboratory setting, willingness to invest serves as a holistic, theory-driven measure of the construct underpinning these behaviors (Elliott et al. [2015], Asay et al. [2023]).

Since Experiment 3 centers on an investment judgment, I propose a secondary self-assessment mechanism that more directly relates to that judgment. Specifically, I expect that using a more interactive GenAI not only inflates investors' self-assessments of their general financial information processing abilities, but also their ability to evaluate a specific investment. As in H2a, I expect these effects to be greater when investors use GenAI for more complex relative to less complex tasks. Accordingly, I make the following predictions:

**H2b:** *GenAI interactivity increases self-assessed ability to evaluate an investment more when GenAI is used for more complex tasks relative to less complex tasks.*

**H2c:** *GenAI interactivity increases willingness to invest in a company more when GenAI is used for more complex tasks relative to less complex tasks.*

## VII. EXPERIMENT 3 METHOD

### Design

Experiment 3 uses a  $2 \times 2$  between-participants design with two manipulated factors: processing tool interactivity (high versus low) and task complexity (more versus less complex). The structure of the task and manipulation of processing tool interactivity is largely similar to Experiment 1. I manipulate task complexity by varying the number of processing steps involved in answering the questions in my task.

### Participants

Participants are 389 Prolific workers with stock market investing experience who complete my study in exchange for \$2.50.<sup>19</sup> Participants are 55% women and 45% men, with an average age

---

<sup>19</sup> I originally recruited 400 participants but dropped 11 observations from participants who had duplicate responses or did not follow instructions to use FinGPT.

of 41 years old. I use Prolific workers with investment experience instead of students because such experience is important for participants to self-assess their ability to evaluate an investment and make investment judgments. Participants have taken on average 3.78 accounting and finance courses. Ninety-one percent of participants have used GenAI before.

### **Tasks and Procedures**

The structure of Experiment 3 is largely similar to earlier experiments with a few noteworthy differences. First, participants assume slightly different roles and objectives. Whereas Experiments 1 and 2 only ask participants to answer questions about financial performance, Experiment 3 positions participants as investors answering questions about financial performance to inform an investment decision. This difference ensures participants approach the task with an investment decision in mind. Second, participants in Experiment 3 answer questions about one company (i.e., FitWear) rather than two companies. Although this design choice results in less exposure to FinGPT and thus weakens my manipulation of processing tool interactivity, it ensures participants focus on the single company in which they are considering investing. Third, I remove the 5-minute time limit for answering the questions because unlike students, who have a finite amount of time to complete the study during class, Prolific workers do not face time constraints.

### **Task Complexity Manipulation**

I manipulate the difficulty element of task complexity by varying the number of processing steps needed to answer the questions (Bonner [2008]). Participants in the more complex condition answer questions similar to Experiments 1 and 2, which require them to 1) acquire at least one piece of financial information, 2) make a comparison or analyze a trend related to that information, and 3) form an interpretation or conclusion based on steps 1 and 2. Participants in the less complex condition answer parallel questions that only involve the first of those three processing steps. For

example, one of the more complex questions asks, “Which geographic segments contributed most to FitWear’s overall performance, and what does this indicate?” The parallel question in the less complex condition asks, “How much was FitWear’s North America sales revenue?” See Appendix A, Panel E for a list of all questions asked.

### **Dependent Variables**

After answering the five questions about FitWear’s performance, participants provide their willingness to invest using two questions. First, “How attractive is an investment in FitWear stock?” Second, “How likely are you to invest in FitWear stock?” Participants respond to each question on a continuous sliding scale with endpoints 0 (*very unattractive/unlikely*) and 100 (*very attractive/likely*). Participants’ responses to these two measures are internally consistent ( $\alpha = 0.88$ ) and load on a single factor (factor loadings  $> 0.95$ ; total variance explained = 91.93%), and so are averaged together to create a single measure of willingness to invest that captures participants’ holistic investment intention (Elliott et al. [2015]).

Next, participants provide self-assessments of their ability to evaluate an investment, which I measure based on Elliott et al. [2015] by asking “How comfortable do you feel about your ability to evaluate FitWear as a potential investment?” Participants respond on an 11-point scale with endpoints 1 (*very uncomfortable*) and 11 (*very comfortable*). Using a different scale (i.e., 101 vs. 11-point) and response style (i.e., sliding scale vs. radio buttons) for my measures of willingness to invest and self-assessment measures reduces concerns about spillover effects (Asay et al. [2022]). Participants also respond to the three questions that gauge their self-assessments of financial information processing ability.

## VIII. EXPERIMENT 3 RESULTS

### Test of H2a

H2a predicts that GenAI interactivity heightens self-assessments of financial information processing ability more when investors use GenAI for more complex tasks relative to less complex tasks. Table 4, Panel A reports descriptive statistics and Figure 4 graphically displays means for *Self-Assessed Processing Ability* by condition.<sup>20</sup> Table 4, Panel B presents two-way ANOVA results, which reveal a significant *Interactivity*  $\times$  *Task Complexity* interaction ( $F_{(1, 385)} = 3.917$ ;  $p = 0.049$ ). Consistent with Experiments 1 and 2, planned follow-up comparisons in Table 4, Panel C show that interactivity heightens self-assessments of financial processing ability when the tool is used for more complex tasks ( $6.513 > 6.064$ ;  $t_{194} = 1.289$ ;  $p = 0.099$ , one-tailed).<sup>21</sup> Conversely, interactivity does not affect self-assessments of financial processing ability when the tool is used for less complex tasks, and even leads to directionally lower self-assessments ( $6.039$  vs.  $6.556$ ;  $t_{191} = 1.513$ ;  $p = 0.132$ ). Thus, my findings support H2a.

### Test of H2b

H2b predicts that GenAI interactivity increases investors' self-assessed ability to evaluate an investment more when they use GenAI for more complex tasks relative to less complex tasks. Table 5, Panel A reports descriptive statistics and Figure 5 graphically displays the means for *Self-Assessed Investment Ability* by condition. Table 5, Panel B presents two-way ANOVA results, which reveal a significant *Interactivity*  $\times$  *Task Complexity* interaction ( $F_{(1, 385)} = 4.677$ ;  $p = 0.031$ ).

---

<sup>20</sup> Participants in the high interactivity condition feel more control over FinGPT responses than those in the low interactivity condition ( $6.219 > 4.539$ ;  $t_{387} = 5.347$ ;  $p < 0.001$ , one-tailed). Additionally, participants in the more complex condition found the questions they were asked to be more complex than those in the less complex condition ( $7.184 > 4.813$ ;  $t_{387} = 9.666$ ;  $p < 0.001$ , one-tailed). Altogether, these findings suggest successful manipulation.

<sup>21</sup> The magnitude of this difference is very similar to Experiment 1. However, the difference is less strongly significant. This may be due to the weaker manipulation from participants having less exposure to FinGPT in Experiment 3 (i.e., evaluated one company with FinGPT rather than two) or due to the additional noise from using Prolific participants who may have greater individual differences compared to a more homogenous participant pool, such as students.

Consistent with expectations, planned follow-up comparisons in Table 5, Panel C show GenAI interactivity increases investors' self-assessed ability to evaluate an investment when the tool is used for more complex tasks ( $8.480 > 7.894$ ;  $t_{194} = 2.130$ ;  $p = 0.017$ , one-tailed), but not for less complex tasks (7.840 vs. 8.091;  $t_{191} = 0.921$ ;  $p = 0.358$ ). Thus, my findings support H2b.

### **Test of H2c**

H2c predicts that GenAI interactivity increases investors' willingness to invest in a company more when they use GenAI for more complex tasks relative to less complex tasks. Table 6, Panel A reports descriptive statistics and Figure 6 graphically displays means for *Willingness to Invest* by condition. Table 6, Panel B presents two-way ANOVA results. Contrary to H2c, results reveal an insignificant *Interactivity*  $\times$  *Task Complexity* interaction ( $F_{(1, 385)} = 0.001$ ;  $p = 0.991$ ).

Although initial results do not directly support H2c, my theory predicts an indirect effect of GenAI interactivity on willingness to invest through investors' self-assessed ability to evaluate an investment. Prior research suggests testing for indirect effects can provide valuable insight into underlying mechanisms and yield theory-consistent evidence, even in the absence of a significant overall effect (Rennekamp [2012], Hayes [2018], Jollineau and Bowen [2023]). Therefore, I conduct a path analysis to test the indirect effect of GenAI interactivity on willingness to invest.

Figure 7, Panel A illustrates my theorized path model (Hayes Model 8), which uses *Interactivity* as the predictor, *Self-Assessed Investment Ability* as the mediating variable, *Task Complexity* as the moderator, and *Willingness to Invest* as the dependent variable. My theory predicts a significant positive indirect effect of *Interactivity* on *Willingness to Invest* through *Self-Assessed Investment Ability* that is stronger when investors use GenAI for more complex tasks relative to less complex tasks. To test for indirect effects, I construct 90 percent two-tailed

confidence intervals with 1,000 bootstrapped resamples of data with replacement (Hayes [2018]). A significant indirect effect requires that zero does not appear within the confidence interval.

Figure 7, Panel B presents results of my path model, which support my theory. There is a significant positive indirect effect of *Interactivity* on *Willingness to Invest* via *Self-Assessed Investment Ability* when GenAI is used for more complex tasks (90% C.I. = [0.563, 4.881]), but not less complex tasks (90% C.I. = [-3.384, 0.908]). These indirect effects are significantly different (Index of Moderated Mediation = [0.851, 7.236]), implying a theory-consistent interactive effect of *Interactivity* and *Task Complexity* on *Willingness to Invest* via *Self-Assessed Investment Ability*. This result lends support to the theory underpinning H2c.

The presence of a significant indirect interaction in the absence of a total interaction suggests there may be competing mechanisms at play (Rennekamp [2012], Jollineau and Bowen [2023]). Indeed, the significant positive direct effect of interactivity on willingness to invest in the less complex condition (path c'2) indicates a positive relationship between interactivity and willingness to invest that operates through an alternative unobserved channel when the task is less complex. This finding explains why I observe a similarly positive total effect of interactivity on willingness to invest at both levels of task complexity, despite the significant indirect interaction. For now, I leave it to future research to explore other mechanisms through which GenAI interactivity affects investment judgments.

## IX. CONCLUSION

In this study, I use the GenAI setting to explore how processing tool interactivity affects investor judgments. I show that the interactivity underpinning GenAI inflates investor self-assessments of their own financial information processing ability despite not improving, and in fact hindering, their actual processing of information the tool provides. Furthermore, I show that

GenAI's interactivity increases investor self-assessments of investment ability, in turn, their willingness to invest. These effects are stronger when GenAI is used for more complex tasks relative to less complex tasks, consistent with investors blurring the boundaries between themselves and the tool and misattributing the tool's abilities to themselves.

My study extends the accounting and cognitive psychology literature while providing important insights for practice. First, I extend our understanding of how technology-driven inflated self-assessments, previously documented among individuals using internet search to answer general knowledge trivia questions, manifest among those using GenAI for financial information processing tasks. Importantly, I also extend the psychology literature by identifying processing tool interactivity and task complexity as distinct, generalizable constructs that jointly cause inflated self-assessments. Second, while my study examines the setting of investors using GenAI to process financial information, my findings highlight a potentially adverse effect of the increasingly interactive tools used for complex tasks in other areas of accounting and finance. For instance, audit teams relying on interactive audit technologies may develop an inflated sense of their audit abilities and become less willing to question their own work and conclusions. Lastly, my study helps the SEC and other capital market participants as they seek to build a more holistic understanding of the consequences of GenAI. While GenAI is largely beneficial for retail investors, my findings illustrate that the interactive nature of GenAI can inflate investor self-assessments and increase their willingness to invest, suggesting that reliance on GenAI could lead to more aggressive investing behavior. These findings are particularly informative for regulators because retail investors are already relying on GenAI and various retail brokerages are integrating GenAI into their platforms to assist investors with investment research and decision-making (SEC [2021], [2023c], Public.com [2023], Son [2023]). My theory suggests regulators and GenAI

developers could take steps to unblur these boundaries and address this negative consequence without compromising the usefulness of GenAI. For instance, developers could include source references in GenAI responses to increase GenAI transparency and remind investors of their reliance on external help, and regulators could promote educational campaigns on best practices when using GenAI, such as self-reflection. Future research could explore the effectiveness of such solutions in mitigating the blurred boundary effects of processing tool interactivity.

My study is subject to limitations that provide opportunities for future research. First, I examine investors' use of a limited-scope GenAI tool to process a narrow set of financial information in a controlled laboratory setting. While doing so enhances internal validity and allows for causal inferences, it restricts investors' ability to interact with GenAI and consider other information they may deem relevant. Thus, my experiment may not fully capture the diverse ways retail investors interact with GenAI in practice. Furthermore, the limited context provided to participants may have limited my ability to detect overall differences in investment judgments. Future research could explore other ways capital market participants use GenAI in more context rich settings and whether those uses differentially affect their judgments. Second, I use a one-shot experiment to examine the effects of processing tool interactivity on investors' self-assessments. However, investors typically have repeated interactions with processing tools, which may alter how the tool affects their judgments. Future research could explore how the investor-tool dynamic develops over repeated uses and as investors become more familiar or comfortable with the tool. For instance, investors who routinely use GenAI to process information may increasingly view it as a seamless extension of their own cognition, resulting in increasingly blurred boundaries and exacerbating the effects of GenAI on their self-assessments (Hamilton and Yao [2018]). Finally, I show processing tool interactivity inflates investor self-assessments when the tool is used for more

complex but not less complex tasks. However, it is possible this moderating effect of task complexity reverses at extreme levels of complexity, as Ward [2021] investigate with mixed results. For instance, when investors use a tool for tasks they could never conceivably imagine doing on their own, processing tool interactivity may reinforce investors' incapacibilities, resulting in more deflated self-assessments. Future research can explore boundary conditions for the joint effects of processing tool interactivity and task complexity, while also considering the role of other task characteristics.

## REFERENCES

- ALPHASENSE. "Introducing Smart Summaries for Companies." 2023. Available at <https://www.alpha-sense.com/blog/product/smart-summaries-for-companies/>.
- ASAY, H. S., GUGGENMOS, R. D., KADOUS, K., KOONCE, L., and LIBBY, R. "Theory Testing and Process Evidence in Accounting Experiments." *The Accounting Review* 97 (2022): 23–43. <https://doi.org/10.2308/TAR-2019-1001>.
- ASAY, H. S., HALES, J., HINDS, C., and RUPAR, K. "Nonprofessional Investor Judgments: Linking Dependent Measures to Constructs." *The Accounting Review* 98 (2023): 1–32. <https://doi.org/10.2308/TAR-2021-0551>.
- BAI, J. (JIANQIU), BOYSON, N. M., CAO, Y., LIU, M., and WAN, C. "Executives vs. Chatbots: Unmasking Insights through Human-AI Differences in Earnings Conference Q&A." SSRN Working Paper. 2023. Available at <https://papers.ssrn.com/abstract=4480056>.
- BARBER, B. M., and ODEAN, T. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *The Journal of Finance* 55 (2000): 773–806. <https://doi.org/10.1111/0022-1082.00226>.
- BARBER, B. M., and ODEAN, T. "Boys Will Be Boys: Gender, Overconfidence, and Common Stock Investment." *The Quarterly Journal of Economics* 116 (2001): 261–92. <https://doi.org/10.1162/003355301556400>.
- BARBER, B. M., and ODEAN, T. "Online Investors: Do the Slow Die First?" *The Review of Financial Studies* 15 (2002): 455–88. <https://doi.org/10.1093/rfs/15.2.455>.
- BLANKESPOOR, E. "The Impact of Information Processing Costs on Firm Disclosure Choice: Evidence from the XBRL Mandate." *Journal of Accounting Research* 57 (2019): 919–67. <https://doi.org/10.1111/1475-679X.12268>.
- BLANKESPOOR, E., CROOM, J., and GRANT, S. M. "Generative AI and Investor Processing of Financial Information." SSRN Working Paper. 2024. Available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5053905](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5053905).
- BLANKESPOOR, E., DEHAAN, E., and MARINOVIC, I. "Disclosure Processing Costs, Investors' Information Choice, and Equity Market Outcomes: A Review." *Journal of Accounting and Economics* 70 (2020): 101344. <https://doi.org/10.1016/j.jacceco.2020.101344>.
- BLANKESPOOR, E., MILLER, G. S., and WHITE, H. D. "The Role of Dissemination in Market Liquidity: Evidence from Firms' Use of Twitter™." *The Accounting Review* 89 (2014): 79–112. <https://doi.org/10.2308/accr-50576>.
- BLOOMBERG. "Introducing BloombergGPT, Bloomberg's 50-Billion Parameter Large Language Model, Purpose-Built from Scratch for Finance." 2023. Available at <https://www.bloomberg.com/company/press/bloomberggpt-50-billion-parameter-llm-tuned-finance/>.
- BONNER, S. E. *Judgment and Decision Making in Accounting*. Pearson/Prentice Hall, 2008.
- CADE, N. L., CROOM, J., and KAMRATH, B. G. "News-Trade Integration: How the Combining of News and Trading Functionality on a Platform Affects Investors' Reactions to News." SSRN Working Paper (2024). <https://doi.org/10.2139/ssrn.4709548>.
- CIKUREL, D. "How Does Perceived Ease of Information Access Affect Investors' Judgments?" *Contemporary Accounting Research* n/a (2024). <https://doi.org/10.1111/1911-3846.12979>.
- CLOR-PROELL, S. M., GUGGENMOS, R. D., and RENNEKAMP, K. "Mobile Devices and Investment News Apps: The Effects of Information Release, Push Notification, and the Fear of

- Missing Out.” *The Accounting Review* 95 (2020): 95–115. <https://doi.org/10.2308/accr-52625>.
- COMMERFORD, B. P., DENNIS, S. A., JOE, J. R., and ULLA, J. W. “Man Versus Machine: Complex Estimates and Auditor Reliance on Artificial Intelligence.” *Journal of Accounting Research* 60 (2022): 171–201. <https://doi.org/10.1111/1475-679X.12407>.
- CUTTER, C. “AI Is Taking On New Work. But Change Will Be Hard—and Expensive.” WSJ. 2024. Available at <https://www.wsj.com/tech/ai/office-workers-artificial-intelligence-changes-86d8d4ab>.
- ELISEEV, E. D., and MARSH, E. J. “Understanding Why Searching the Internet Inflates Confidence in Explanatory Ability.” *Applied Cognitive Psychology* n/a (2023). <https://doi.org/10.1002/acp.4058>.
- ELLIOTT, W. B., GALE, B. T., and HOBSON, J. L. “The Joint Influence of Information Push and Value Relevance on Investor Judgments and Market Efficiency.” *Journal of Accounting Research* 60 (2022): 1049–83. <https://doi.org/10.1111/1475-679X.12400>.
- ELLIOTT, W. B., GRANT, S. M., and HOBSON, J. L. “Trader Participation in Disclosure: Implications of Interactions with Management.” *Contemporary Accounting Research* 37 (2020): 68–100. <https://doi.org/10.1111/1911-3846.12524>.
- ELLIOTT, W. B., GRANT, S. M., and HODGE, F. D. “Negative News and Investor Trust: The Role of \$Firm and #CEO Twitter Use.” *Journal of Accounting Research* 56 (2018): 1483–1519. <https://doi.org/10.1111/1475-679X.12217>.
- ELLIOTT, W. B., RENNEKAMP, K. M., and WHITE, B. J. “Does Concrete Language in Disclosures Increase Willingness to Invest?” *Review of Accounting Studies* 20 (2015): 839–65. <https://doi.org/10.1007/s11142-014-9315-6>.
- FINCHAT.IO. “How To Use FinChat: A Guide for Investors.” FinChat.Io. 2023. Available at <https://finchat.io/blog/how-to-use-finchat/>.
- FINRA. “Artificial Intelligence (AI) in the Securities Industry.” 2020. Available at <https://www.finra.org/sites/default/files/2020-06/ai-report-061020.pdf>.
- FISHER, M., GODDU, M. K., and KEIL, F. C. “Searching for Explanations: How the Internet Inflates Estimates of Internal Knowledge.” *Journal of Experimental Psychology: General* 144 (2015): 674–87. <https://doi.org/10.1037/xge0000070>.
- FISHER, M., SMILEY, A. H., and GRILLO, T. L. H. “Information without Knowledge: The Effects of Internet Search on Learning.” *Memory* 30 (2022): 375–87. <https://doi.org/10.1080/09658211.2021.1882501>.
- FLEMING, S. M. “Metacognition and Confidence: A Review and Synthesis.” *Annual Review of Psychology* 75 (2024): 241–68. <https://doi.org/10.1146/annurev-psych-022423-032425>.
- FOX, C. M. “50 Ways to Use Chat GPT-3 for Finance.” 2023. Available at <https://medium.com/chat-gpt-for-finance-and-fp-a/50-ways-to-use-chat-gpt-3-for-finance-part-1-675ae43f7b6>.
- GALLAGHER, D. “Why Microsoft Is Spreading Its AI Bets to PCs.” WSJ. 2024. Available at <https://www.wsj.com/tech/ai/why-microsoft-is-spreading-its-ai-bets-to-pcs-41e60451>.
- GRAHAM, J. R., HARVEY, C. R., and HUANG, H. “Investor Competence, Trading Frequency, and Home Bias.” *Management Science* 55 (2009): 1094–1106. <https://doi.org/10.1287/mnsc.1090.1009>.
- GRANT, S. M. “How Does Using a Mobile Device Change Investors’ Reactions to Firm Disclosures?” *Journal of Accounting Research* 58 (2020): 741–75. <https://doi.org/10.1111/1475-679X.12299>.

- GRINBLATT, M., and KELOHARJU, M. "Sensation Seeking, Overconfidence, and Trading Activity." *The Journal of Finance* 64 (2009): 549–78. <https://doi.org/10.1111/j.1540-6261.2009.01443.x>.
- HAMILTON, K. A., and YAO, M. Z. "Blurring Boundaries: Effects of Device Features on Metacognitive Evaluations." *Computers in Human Behavior* 89 (2018): 213–20. <https://doi.org/10.1016/j.chb.2018.07.044>.
- HAYES, A. F. *Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach*. Guilford Publications, 2018.
- HEATH, C., and TVERSKY, A. "Preference and Belief: Ambiguity and Competence in Choice under Uncertainty." *Journal of Risk and Uncertainty* 4 (1991): 5–28. <https://doi.org/10.1007/BF00057884>.
- HENKEL, L. A. "Point-and-Shoot Memories: The Influence of Taking Photos on Memory for a Museum Tour." *Psychological Science* 25 (2014): 396–402. <https://doi.org/10.1177/0956797613504438>.
- HODGE, F. D., KENNEDY, J. J., and MAINES, L. A. "Does Search-Facilitating Technology Improve the Transparency of Financial Reporting?" *The Accounting Review* 79 (2004): 687–703. <https://doi.org/10.2308/accr.2004.79.3.687>.
- HOFFMANN, A. O. I., and POST, T. "How Does Investor Confidence Lead to Trading? Linking Investor Return Experiences, Confidence, and Investment Beliefs." *Journal of Behavioral and Experimental Finance* 12 (2016): 65–78. <https://doi.org/10.1016/j.jbef.2016.09.003>.
- JOHNSON, G. J., BRUNER II, G. C., and KUMAR, A. "Interactivity and Its Facets Revisited: Theory and Empirical Test." *Journal of Advertising*, December (2006). <https://doi.org/10.2753/JOA0091-3367350403>.
- JOLLINEAU, S. J., and BOWEN, R. M. "A Practical Guide to Using Path Analysis: Mediation and Moderation in Accounting Research." *Journal of Financial Reporting* 8 (2023): 11–40. <https://doi.org/10.2308/JFR-2021-004>.
- KIM, A., MUHN, M., and NIKOLAEV, V. V. "Bloated Disclosures: Can ChatGPT Help Investors Process Information?" SSRN Working Paper. 2024. Available at <https://papers.ssrn.com/abstract=4425527>.
- LIBBY, R., BLOOMFIELD, R., and NELSON, M. W. "Experimental Research in Financial Accounting." *Accounting, Organizations and Society* 27 (2002): 775–810. [https://doi.org/10.1016/S0361-3682\(01\)00011-3](https://doi.org/10.1016/S0361-3682(01)00011-3).
- ODEAN, T. "Volume, Volatility, Price, and Profit When All Traders Are Above Average." *The Journal of Finance* 53 (1998): 1887–1934. <https://doi.org/10.1111/0022-1082.00078>.
- PLUTO.FI. "Introduction." 2023. Available at <https://docs.pluto.fi/>.
- PUBLIC.COM. "Meet Alpha - Your Investing Co-Pilot. Powered by GPT-4." 2023. Available at <https://public.com/alpha>.
- RAI. "JPMorgan Is Discussing Its Generative AI Projects With Regulators." Bloomberg.Com. 2023. Available at <https://www.bloomberg.com/news/articles/2023-11-09/jpmorgan-is-working-with-us-regulators-on-generative-ai-pilot-projects>.
- RENNEKAMP, K. "Processing Fluency and Investors' Reactions to Disclosure Readability." *Journal of Accounting Research* 50 (2012): 1319–54. <https://doi.org/10.1111/j.1475-679X.2012.00460.x>.
- RENNEKAMP, K., and WITZ, P. D. "Linguistic Formality and Audience Engagement: Investors' Reactions to Characteristics of Social Media Disclosures\*." *Contemporary Accounting Research* 38 (2021): 1748–81. <https://doi.org/10.1111/1911-3846.12661>.

- SEC. “SEC Requests Information and Comment on Broker-Dealer and Investment Adviser Digital Engagement Practices, Related Tools and Methods, and Regulatory Considerations and Potential Approaches; Information and Comments on Investment Adviser Use of Technology.” 2021. Available at <https://www.sec.gov/news/press-release/2021-167>.
- SEC. “‘Isaac Newton to AI’ Remarks before the National Press Club.” 2023a. Available at <https://www.sec.gov/news/speech/gensler-isaac-newton-ai-remarks-07-17-2023>.
- SEC. “SEC Division of Examinations Announces 2023 Priorities.” 2023b. Available at <https://www.sec.gov/news/press-release/2023-24>.
- SEC. “SEC Proposes New Requirements to Address Risks to Investors From Conflicts of Interest Associated With the Use of Predictive Data Analytics by Broker-Dealers and Investment Advisers.” 2023c. Available at <https://www.sec.gov/news/press-release/2023-140>.
- SEC. “‘AI, Finance, Movies, and the Law’ Prepared Remarks before the Yale Law School.” 2024. Available at <https://www.sec.gov/news/speech/gensler-ai-021324>.
- SEC. “SEC Roundtable on Artificial Intelligence in the Financial Industry.” 2025. Available at <https://www.sec.gov/newsroom/meetings-events/sec-roundtable-artificial-intelligence-financial-industry>.
- SHEIKH, J. “Bloomberg Uses Its Vast Data To Create New Finance AI.” 2023. Available at <https://www.forbes.com/sites/jamielsheikh/2023/04/05/the-chatgpt-of-finance-is-here-bloomberg-is-combining-ai-and-fintech/>.
- SKINNER, E. A. “A Guide to Constructs of Control.” *Journal of Personality and Social Psychology* 71 (1996): 549–70. <https://doi.org/10.1037/0022-3514.71.3.549>.
- SON, H. “Morgan Stanley Is Testing an OpenAI-Powered Chatbot for Its 16,000 Financial Advisors.” CNBC. 2023. Available at <https://www.cnbc.com/2023/03/14/morgan-stanley-testing-openai-powered-chatbot-for-its-financial-advisors.html>.
- SPARROW, B., LIU, J., and WEGNER, D. M. “Google Effects on Memory: Cognitive Consequences of Having Information at Our Fingertips.” *Science* 333 (2011): 776–78. <https://doi.org/10.1126/science.1207745>.
- STATMAN, M., THORLEY, S., and VORKINK, K. “Investor Overconfidence and Trading Volume.” *Review of Financial Studies* 19 (2006): 1531–65. <https://doi.org/10.1093/rfs/hhj032>.
- U.S. DEPARTMENT OF THE TREASURY. “U.S. Department of Treasury Releases Request for Information on Uses, Opportunities, and Risks of Artificial Intelligence in the Financial Services Sector.” 2024. Available at <https://home.treasury.gov/news/press-releases/jy2393>.
- VAN MARCKE, H., DENMAT, P. L., VERGUTS, T., and DESENDER, K. “Manipulating Prior Beliefs Causally Induces Under- and Overconfidence.” *Psychological Science* 35 (2024): 358–75. <https://doi.org/10.1177/09567976241231572>.
- WARD, A. F. “People Mistake the Internet’s Knowledge for Their Own.” *Proceedings of the National Academy of Sciences* 118 (2021): e2105061118. <https://doi.org/10.1073/pnas.2105061118>.
- WEGNER, D. M. “Transactive Memory: A Contemporary Analysis of the Group Mind.” In *Theories of Group Behavior*, edited by Brian Mullen and George R. Goethals, 185–208. Springer Series in Social Psychology. New York, NY: Springer, 1987. [https://doi.org/10.1007/978-1-4612-4634-3\\_9](https://doi.org/10.1007/978-1-4612-4634-3_9).

ZAKRZEWSKI, C. "Biden Administration Is Trying to Figure out How to Audit AI." *Washington Post*, 2023. Available at <https://www.washingtonpost.com/technology/2023/04/11/biden-commerce-department-ai-rules/>.

**FIGURE 1**  
**The Effect of Processing Tool Interactivity on Self-Assessments of Financial Information Processing Ability (Experiment 1)**

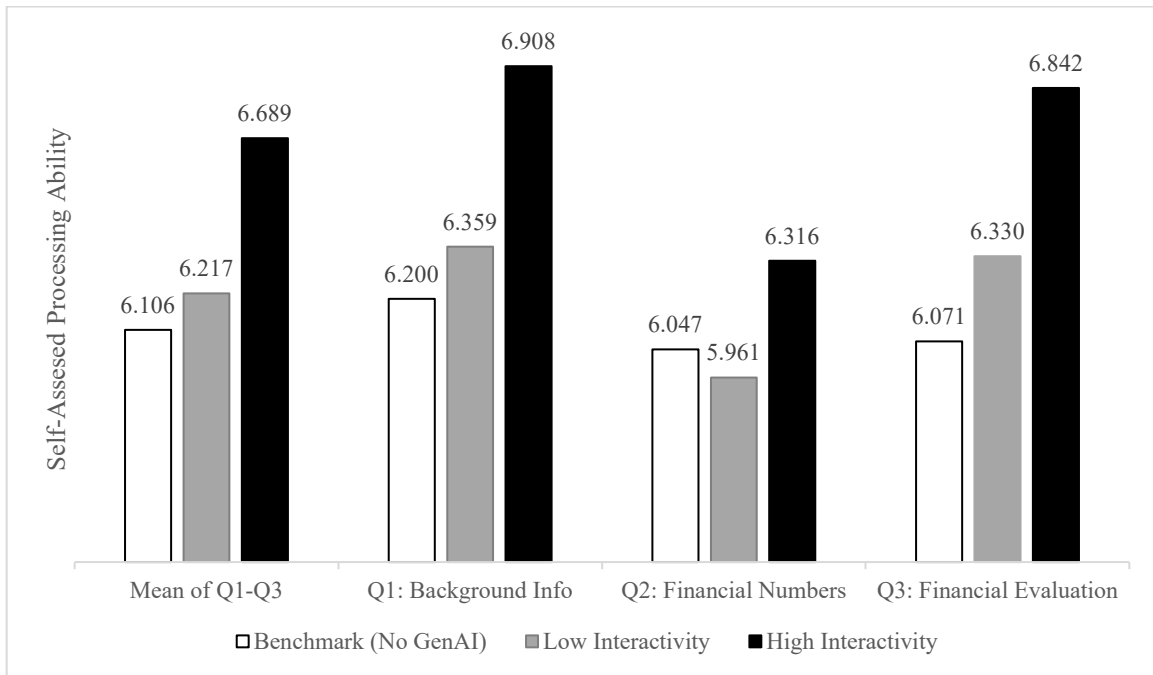
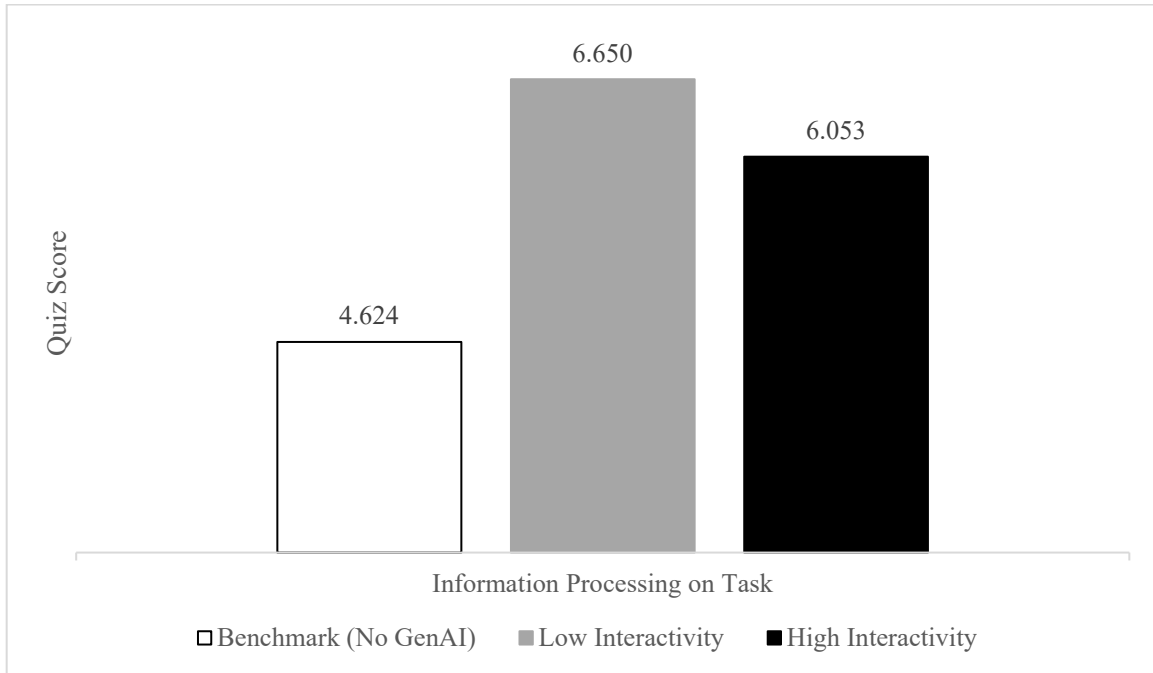


Figure 1 depicts observed mean values for *Self-Assessed Processing Ability* in Experiment 1. For each of three sub-topics, participants see a set of sample questions (see Appendix A, Panel C) and answer “How well could you answer detailed questions about [topic] like these?” on a scale from 1 (very poorly) to 11 (very well). I use the average of participants’ judgments to the three sub-topics to form my primary dependent variable of *Self-Assessed Processing Ability*. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I also include a benchmark condition in which participants answer questions without FinGPT. Table 1 presents descriptive statistics and planned comparisons for *Self-Assessed Processing Ability*.

**FIGURE 2**  
**The Effect of Processing Tool Interactivity on Financial Information Processing**  
**(Experiment 1)**



---

Figure 2 depicts observed mean values for *Information Processing* in Experiment 1. I measure *Information Processing* using their performance on a ten-question quiz designed to align with the ten questions they were tasked with answering in the study. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I also include a benchmark condition in which participants answer questions without FinGPT. Table 2 presents descriptive statistics and planned comparisons for *Information Processing*.

**FIGURE 3**  
**The Effect of Processing Tool Interactivity on Self-Assessments of Financial Information Processing Ability (Experiment 2)**

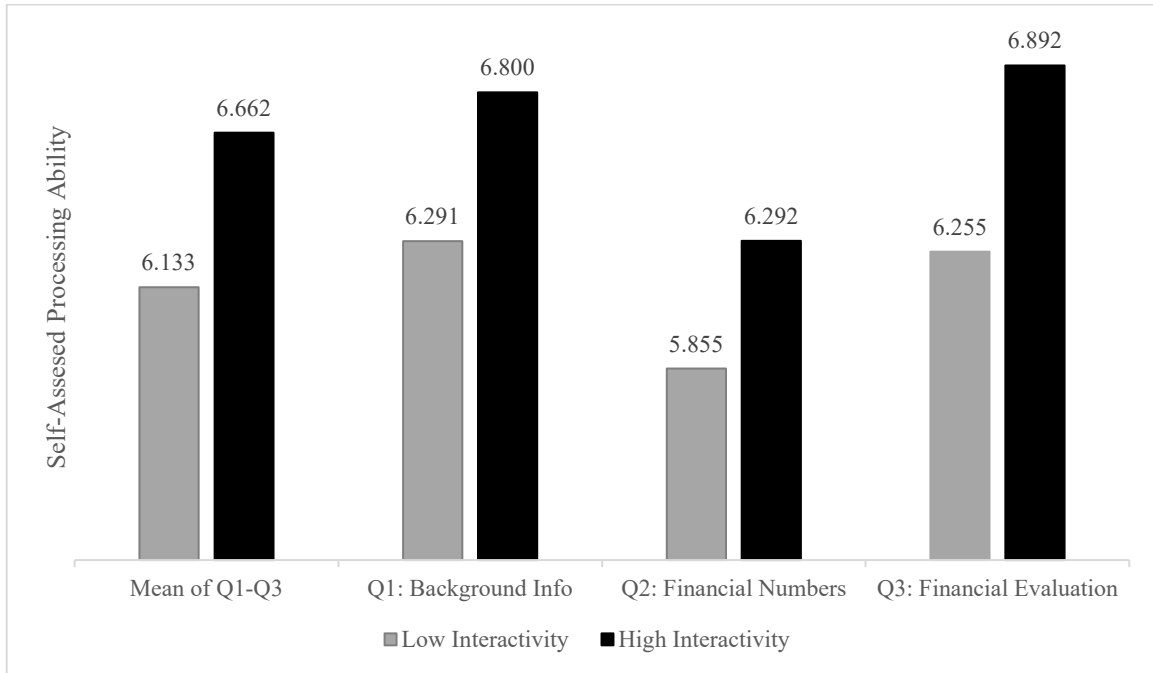
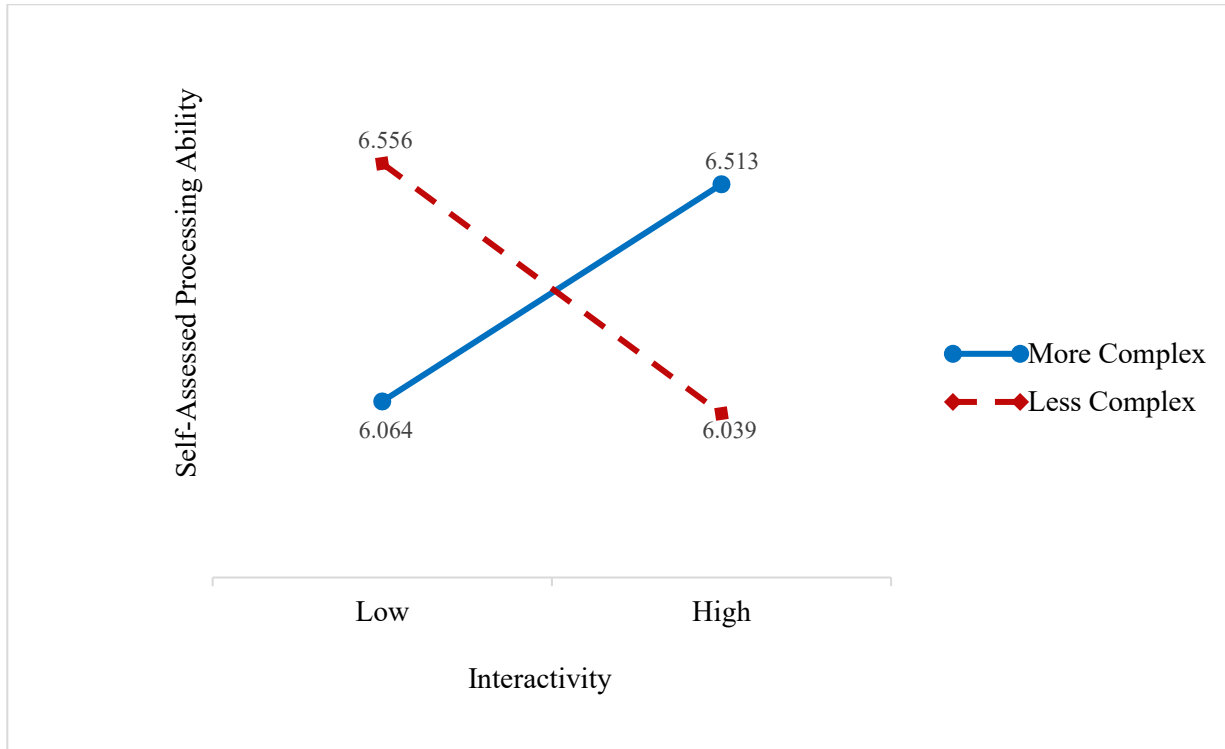


Figure 3 depicts observed mean values for *Self-Assessed Processing Ability* in Experiment 2. For each of three sub-topics, participants see a set of sample questions (see Appendix A, Panel C) and answer “How well could you answer detailed questions about [topic] like these?” on a scale from 1 (very poorly) to 11 (very well). I use the average of participants’ judgments to the three sub-topics to form my primary dependent variable of *Self-Assessed Processing Ability*. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, response speed, and nonverbal communication—i.e., the four facets of processing tool interactivity. See Appendix B for details. Table 3 presents descriptive statistics and planned comparisons for *Self-Assessed Processing Ability*.

**FIGURE 4**  
**The Effect of Processing Tool Interactivity and Task Complexity on Self-Assessments of Financial Information Processing Ability (Experiment 3)**



---

Figure 4 depicts observed mean values for *Self-Assessed Processing Ability* in Experiment 3. For each of three sub-topics, participants see a set of sample questions (see Appendix A, Panel C) and answer “How well could you answer detailed questions about [topic] like these?” on a scale from 1 (very poorly) to 11 (very well). I use the average of participants’ judgments to the three sub-topics to form my primary dependent variable of *Self-Assessed Processing Ability*. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I manipulate task complexity by varying the number of processing steps required to answer the questions. Table 4 presents descriptive statistics and planned comparisons for *Self-Assessed Processing Ability*.

**FIGURE 5**  
**The Effect of Processing Tool Interactivity and Task Complexity on Self-Assessed Ability to Evaluate an Investment (Experiment 3)**

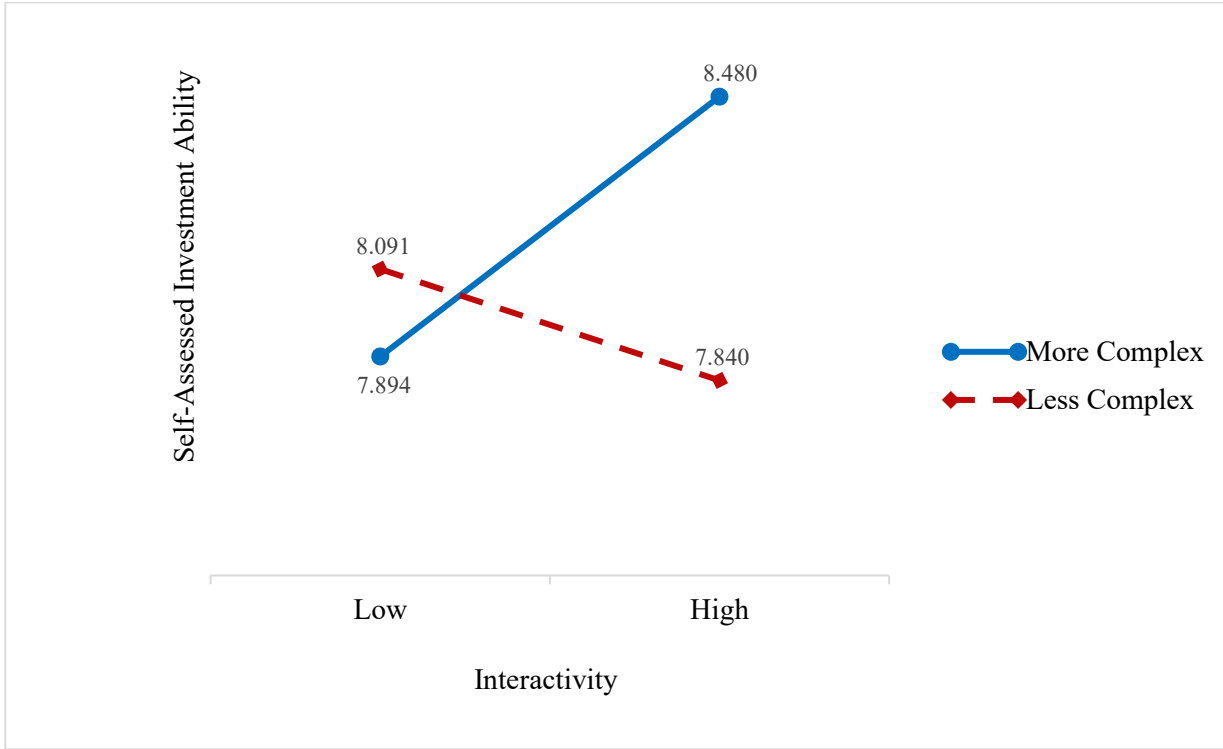


Figure 5 depicts observed mean values for *Self-Assessed Investment Ability* in Experiment 3. I measure *Self-Assessed Investment Ability* by asking participants “How comfortable do you feel about your ability to evaluate FitWear as a potential investment?” Participants respond on an 11-point scale with endpoints 1 (*very uncomfortable*) and 11 (*very comfortable*). I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I manipulate task complexity by varying the number of processing steps required to answer the questions. Table 5 presents descriptive statistics and planned comparisons for *Self-Assessed Investment Ability*.

**FIGURE 6**  
**The Effect of Processing Tool Interactivity and Task Complexity on Willingness to Invest**  
**(Experiment 3)**

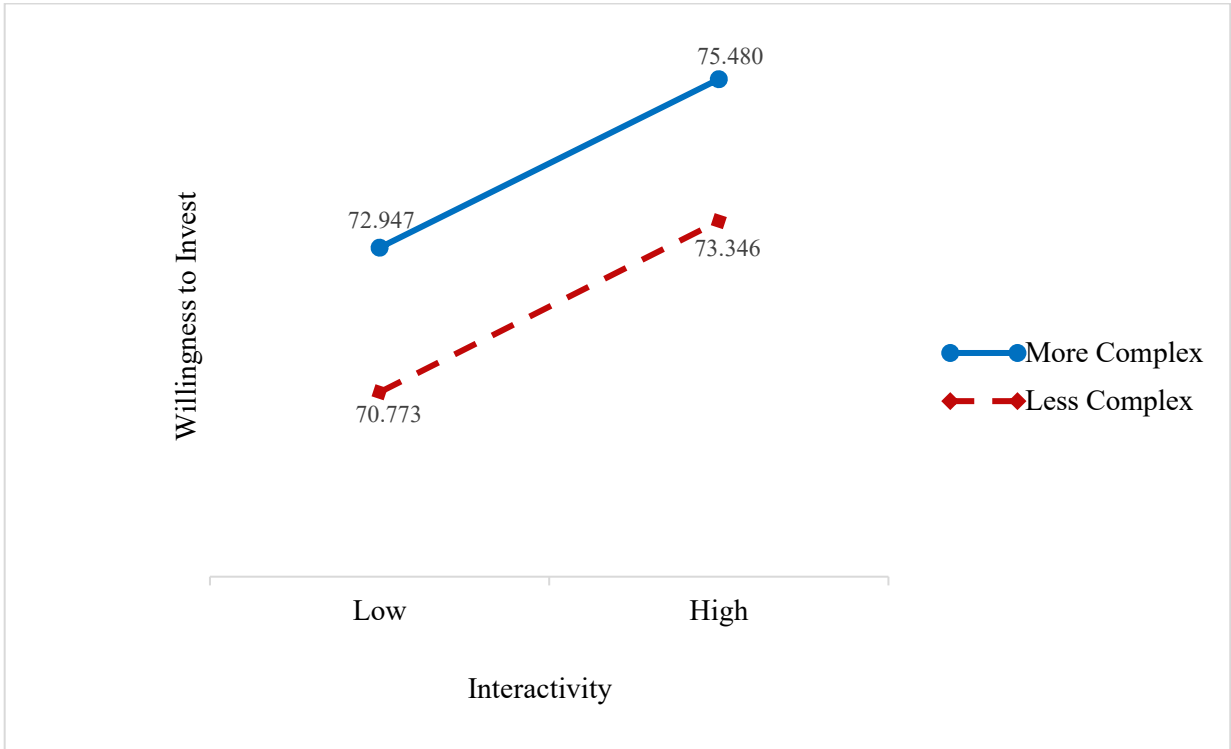
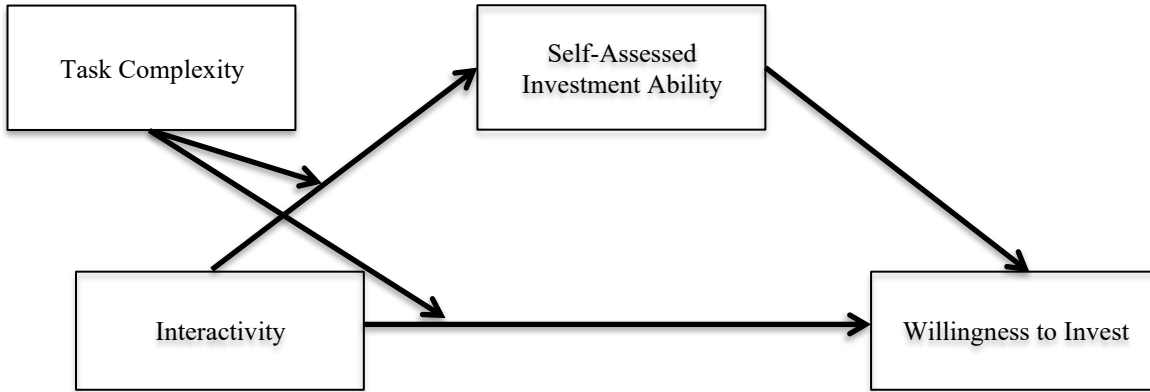


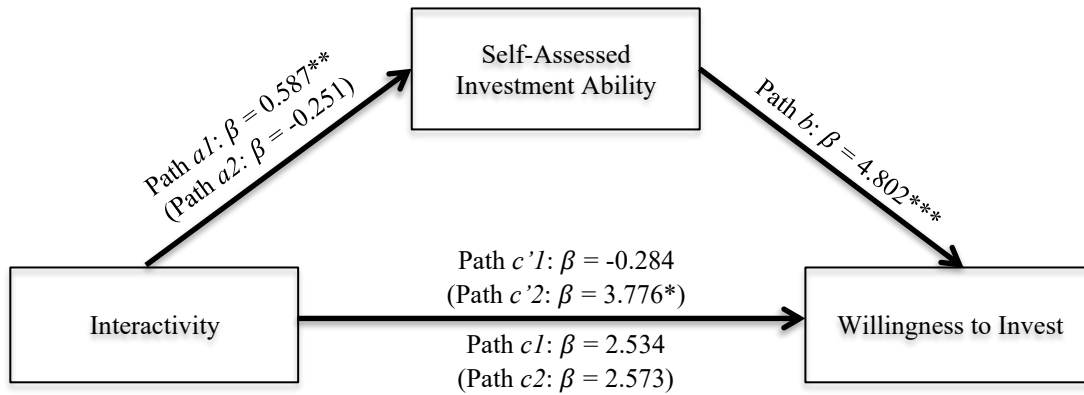
Figure 6 depicts observed mean values for *Willingness to Invest* in Experiment 3. Participants rate (1) the attractiveness of an investment in the firm’s stock and (2) the likelihood that they would invest in the firm’s stock on a sliding scale from 0 (“Not at all attractive/likely”) to 100 (“Very attractive/likely”). *Willingness to Invest* is an average of these measures. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I manipulate task complexity by varying the number of processing steps required to answer the questions. Table 6 presents descriptive statistics and planned comparisons for *Willingness to Invest*.

**FIGURE 7**  
**The Interactive Indirect Effects of Processing Tool Interactivity and Task Complexity on Willingness to Invest via Self-Assessments of Investment Ability (Experiment 3)**

**Panel A: Theorized Model (Hayes Model 8)**



**Panel B: Observed Model for More Complex (Less Complex) Condition**



Indirect Effect – More Complex ( $a1*b$ ): (0.563, 4.881)  
 Indirect Effect – Less Complex ( $a2*b$ ): (-3.384, 0.908)

Index of Moderated Mediation (0.851, 7.236)

Figure 7 graphically depicts my path model for Experiment 3. The model uses *Interactivity* as the predictor, *Self-Assessed Investment Ability* as the mediating variable, *Task Complexity* as the moderator, and *Willingness to Invest* as the dependent variable. Path *a1* (*a2*) reflects the direct effect of *Interactivity* on *Self-Assessed Investment Ability* within the more complex (less complex) condition. Path *b* reflects the association between *Self-Assessed Investment Ability* and *Willingness to Invest*. To test for indirect effects, I construct 90 percent two-tailed confidence intervals for the product of paths *a* and *b*. I use 1,000 bootstrapped resamples of data with replacement (Hayes [2018]). A statistically significant indirect effect requires that zero not appear within the confidence interval. Path *c'1* (*c'2*) reflects the direct effect of *Interactivity* on *Willingness to Invest* after accounting for the indirect effects within the more complex (less complex) condition, while path *c1* (*c2*) reflects the total effect (indirect effects plus direct effects). Table 5 presents descriptive statistics for *Self-Assessed Investment Ability* and Table 6 presents descriptive statistics for *Willingness to Invest*.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

**TABLE 1**  
**Test of H1: The Effect of Processing Tool Interactivity on Self-Assessments of Financial Information Processing Ability (Experiment 1)**

<b>Panel A: Descriptive Statistics —Mean [Standard Deviation]</b>			
Question	Condition		
	Benchmark (No GenAI)	Low Interactivity	High Interactivity
Q1: Background Info	6.200 [2.293]	6.359 [2.090]	6.908 [2.060]
Q2: Financial Numbers	6.047 [2.464]	5.961 [2.240]	6.316 [2.499]
Q3: Financial Evaluation	6.071 [2.240]	6.330 [2.041]	6.842 [1.960]
Average Self-Assessment	6.106 [1.952] <i>n</i> = 85	6.217 [1.780] <i>n</i> = 103	6.689 [1.651] <i>n</i> = 76

<b>Panel B: Planned Comparisons</b>			
Source of Variation	<i>df</i>	<i>t</i> -Stat	<i>p</i> -Value
High Interactivity > Low Interactivity	177	1.807	0.036 <sup>†</sup>
High Interactivity > Benchmark	159	2.032	0.022 <sup>†</sup>
Low Interactivity vs. Benchmark	186	0.407	0.684

Table 1 presents descriptive statistics and planned comparisons for *Self-Assessed Processing Ability* in Experiment 1. For each of three sub-topics, participants see a set of sample questions (see Appendix A, Panel C) and answer “How well could you answer detailed questions about [topic] like these?” on a scale from 1 (very poorly) to 11 (very well). I use the average of participants’ judgments to the three sub-topics to form my primary dependent variable of participants’ *Self-Assessed Processing Ability*. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details.. I also include a +1 benchmark condition in which participants answer questions without FinGPT. Figure 1 depicts observed mean values for *Self-Assessed Processing Ability*.

<sup>†</sup> *p*-value is one-tailed given theory predicts a directional relation.

**TABLE 2**  
**The Effect of Processing Tool Interactivity on Financial Information Processing**  
**(Experiment 1)**

<b>Panel A: Descriptive Statistics – Mean [Standard Deviation]</b>				
Question	Benchmark (No GenAI)	Condition		
		Low Interactivity	High Interactivity	
<i>First Company</i>				
Q1	0.412	0.825	0.684	
Q2	0.376	0.650	0.447	
Q3	0.682	0.845	0.789	
Q4	0.294	0.359	0.382	
Q5	0.341	0.369	0.342	
<i>Second Company</i>				
Q1	0.518	0.806	0.750	
Q2	0.047	0.825	0.645	
Q3	0.871	0.932	0.921	
Q4	0.671	0.670	0.711	
Q5	0.412	0.369	0.382	
Total	4.624 [1.34] <i>n</i> = 85	6.650 [1.564] <i>n</i> = 103	6.053 [1.624] <i>n</i> = 76	
<b>Panel B: Planned Comparisons</b>				
Source of Variation		<i>df</i>	<i>t</i> -Stat	<i>p</i> -Value
High Interactivity vs. Low Interactivity		177	2.487	0.014
High Interactivity vs. Benchmark		159	6.121	< 0.001
Low Interactivity vs. Benchmark		186	9.440	< 0.001

Table 2 presents descriptive statistics and planned comparisons for *Information Processing* in Experiment 1. I measure *Information Processing* using their performance on a ten-question quiz designed to align with the ten questions they were tasked with answering in the study. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I also include a +1 benchmark condition in which participants answer questions without FinGPT. Figure 2 depicts observed mean values for *Information Processing*.

**TABLE 3**  
**The Effect of Processing Tool Interactivity on Self-Assessments of Financial Information Processing Ability (Experiment 2)**

<b>Panel A: Descriptive Statistics —Mean [Standard Deviation]</b>			
Question	Condition		
	Low Interactivity	High Interactivity	
Q1: Background Info	6.291 [1.629]	6.800 [1.914]	
Q2: Financial Numbers	5.855 [2.360]	6.292 [2.517]	
Q3: Financial Evaluation	6.255 [2.154]	6.892 [1.937]	
Average Self-Assessment	6.133 [1.697] <i>n</i> = 55	6.662 [1.825] <i>n</i> = 65	
<b>Panel B: Planned Comparisons</b>			
Source of Variation	<i>df</i>	<i>t</i> -Stat	<i>p</i> -Value
High Interactivity > Low Interactivity	118	1.641	0.052 <sup>†</sup>

Table 3 presents descriptive statistics and planned comparisons for *Self-Assessed Processing Ability* in Experiment 2. For each of three sub-topics, participants see a set of sample questions (see Appendix A, Panel C) and answer “How well could you answer detailed questions about [topic] like these?” on a scale from 1 (very poorly) to 11 (very well). I use the average of participants’ judgments to the three sub-topics to form my primary dependent variable of participants’ *Self-Assessed Processing Ability*. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, response speed, and nonverbal communication—i.e., the four facets of processing tool interactivity. See Appendix B for details. Figure 3 depicts observed mean values for *Self-Assessed Processing Ability*.

<sup>†</sup> *p*-value is one-tailed given theory predicts a directional relation.

**TABLE 4**  
**Test of H2a: The Effect of Processing Tool Interactivity and Task Complexity on Investors’ Self-Assessments of Financial Information Processing Ability (Experiment 3)**

<b>Panel A: Descriptive Statistics – Means [Standard Deviations]</b>					
GenAI Interactivity					
Task Complexity	Low	High	Row Means		
More Complex	6.064 [2.334] n = 94	6.513 [2.529] n = 102	6.298 [2.442] n = 196		
Less Complex	6.556 [2.143] n = 99	6.039 [2.589] n = 94	6.304 [2.379] n = 193		
Column Means	6.316 [2.246] n = 193	6.286 [2.563] n = 196	6.301 [2.408] n = 389		
<b>Panel B: Two-Way Analysis of Variance (ANOVA) Results</b>					
Source of Variation	SS	df	MS	F-Stat	p-Value
Interactivity	0.110	1	0.110	0.019	0.890
Complexity	0.008	1	0.008	0.001	0.971
Interactivity × Complexity	22.650	1	22.650	3.917	0.049
Error	2226.401	385	5.783		
<b>Panel C: Planned Comparisons</b>					
Source of Variation		df	t-Stat	p-Value	
Effect of Interactivity given More Complex		194	1.289	0.099 <sup>†</sup>	
Effect of Interactivity given Less Complex		191	1.513	0.132	

Table 4 presents descriptive statistics and planned comparisons for *Self-Assessed Processing Ability* in Experiment 3. For each of three sub-topics, participants see a set of sample questions (see Appendix A, Panel C) and answer “How well could you answer detailed questions about [topic] like these?” on a scale from 1 (very poorly) to 11 (very well). I use the average of participants’ judgments to the three sub-topics to form my primary dependent variable of *Self-Assessed Processing Ability*. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I manipulate task complexity by varying the number of processing steps required to answer the questions. Figure 4 depicts observed mean values for *Self-Assessed Processing Ability*.

<sup>†</sup> p-value is one-tailed given theory predicts a directional relation.

**TABLE 5**  
**Test of H2b: The Effect of Processing Tool Interactivity and Task Complexity on Self-Assessed Ability to Evaluate an Investment (Experiment 3)**

<b>Panel A: Descriptive Statistics – Means [Standard Deviations]</b>					
GenAI Interactivity					
Task Complexity	Low	High	Row Means		
More Complex	7.894	8.480	8.199		
	[2.207]	[1.627]	[1.944]		
	n = 94	n = 102	n = 196		
Less Complex	8.091	7.840	7.969		
	[1.813]	[1.964]	[1.887]		
	n = 99	n = 94	n = 193		
Column Means	7.995	8.173	8.085		
	[2.012]	[1.820]	[1.917]		
	n = 193	n = 196	n = 389		
<b>Panel B: Two-Way Analysis of Variance (ANOVA) Results</b>					
Source of Variation	SS	df	MS	F-Stat	p-Value
Interactivity	2.746	1	2.746	0.755	0.386
Complexity	4.758	1	4.758	1.307	0.254
Interactivity × Complexity	17.022	1	17.022	4.677	0.031
Error	1401.185	385	3.639		
<b>Panel C: Planned Comparisons</b>					
Source of Variation		df	t-Stat	p-Value	
Effect of Interactivity given More Complex		194	2.130	0.017 <sup>†</sup>	
Effect of Interactivity given Less Complex		191	0.921	0.358	

Table 5 presents descriptive statistics and planned comparisons for *Self-Assessed Investment Ability* in Experiment 3. I measure *Self-Assessed Investment Ability* by asking participants “How comfortable do you feel about your ability to evaluate FitWear as a potential investment?” Participants respond on an 11-point scale with endpoints 1 (*very uncomfortable*) and 11 (*very comfortable*). I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details. I manipulate task complexity by varying the number of processing steps required to answer the questions. Figure 5 depicts observed mean values for *Self-Assessed Investment Ability*.

<sup>†</sup> *p*-value is one-tailed given theory predicts a directional relation.

**TABLE 6**  
**Test of H2c: The Effect of Processing Tool Interactivity and Task Complexity on**  
**Willingness to Invest (Experiment 3)**

<b>Panel A: Descriptive Statistics – Means [Standard Deviations]</b>					
GenAI Interactivity					
Task Complexity	Low	High	Row Means		
More Complex	72.947 [17.037] n = 94	75.480 [15.429] n = 102	74.265 [16.228] n = 196		
Less Complex	70.773 [17.808] n = 99	73.346 [16.385] n = 94	72.026 [17.134] n = 193		
Column Means	71.832 [17.426] n = 193	74.457 [15.890] n = 196	73.154 [16.700] n = 389		
<b>Panel B: Two-Way Analysis of Variance (ANOVA) Results</b>					
Source of Variation	SS	df	MS	F-Stat	p-Value
Interactivity	633.230	1	633.230	2.277	0.132
Complexity	450.812	1	450.812	1.621	0.204
Interactivity × Complexity	0.038	1	0.038	0.001	0.991
Error	107087.344	385	278.149		

Table 6 presents descriptive statistics and planned comparisons for *Willingness to Invest* in Experiment 3. Participants rate (1) the attractiveness of an investment in the firm’s stock and (2) the likelihood that they would invest in the firm’s stock on a sliding scale from 0 (“Not at all attractive/likely”) to 100 (“Very attractive/likely”). *Willingness to Invest* is an average of these measures. I manipulate processing tool interactivity by varying FinGPT’s reciprocity, responsiveness, and nonverbal communication—i.e., three of the four facets of processing tool interactivity. See Appendix B for details.. I manipulate task complexity by varying the number of processing steps required to answer the questions. Figure 6 depicts observed mean values for *Willingness to Invest*.

**APPENDIX A**  
**Experimental Excerpts**

**Panel A: Earnings Release Questions (Experiments 1 and 2)**

Company 1 (Sole Savvy)

1. *How are Sole Savvy's different sales channels trending and what does this suggest for the future of the company?*
2. *What are the company's plans for share repurchases and/or dividends and what do these plans indicate?*
3. *How did Sole Savvy's SG&A percentage compare to last year and what does this indicate?*
4. *How has the company adjusted its capital expenditure strategy in the current year and what might this suggest?*
5. *Were there any material events that took place after the fiscal year-end? If so, how might these events influence Fiscal 2024 performance?*

Company 2 (FitWear)

1. *What are FitWear's key risk factors and how well are they mitigating those risks?*
2. *Which geographic segments contributed most to FitWear's overall performance, and what does this indicate?*
3. *How is FitWear's store count trending and what does this mean for the company?*
4. *What are FitWear's asset and liability levels and how healthy is the company in terms of liquidity?*
5. *How promising does FitWear's forecast look for the upcoming fiscal year?*

## Panel B: Manipulations

### Benchmark

Answer the five questions about Sole Savvy listed below to the best of your ability.

Questions to answer:

1. How are Sole Savvy's different sales channels trending and what does this suggest for the future of the company?
2. What are the company's plans for share repurchases and/or dividends and what do these plans indicate?
3. How did Sole Savvy's SG&A percentage compare to last year and what does this indicate?
4. How has the company adjusted its capital expenditure strategy in the current year and what might this suggest?
5. Were there any material events that took place after the fiscal year-end? If so, how might these events influence Fiscal 2024 performance?

Submit Answers

### Low Interactivity (Experiment 1 and 3)

Use FinGPT to answer the five questions about Sole Savvy listed below.

Questions to answer:

1. How are Sole Savvy's different sales channels trending and what does this suggest for the future of the company?
2. What are the company's plans for share repurchases and/or dividends and what do these plans indicate?
3. How did Sole Savvy's SG&A percentage compare to last year and what does this indicate?
4. How has the company adjusted its capital expenditure strategy in the current year and what might this suggest?
5. Were there any material events that took place after the fiscal year-end? If so, how might these events influence Fiscal 2024 performance?

Submit Answers

**FinGPT**

**How are Sole Savvy's different sales channels trending and what does this suggest for the future of the company?**

In 2023, Sole Savvy's in-person sales grew by 2.7% to \$7,900 million, while online sales surged by 14.6% to \$2,120 million compared to the prior year. This suggests that the company's online sales, though making up a smaller percentage of total sales, are growing at a faster rate than in-person sales. The trend indicates a shift in consumer behavior towards online shopping, highlighting the potential for the company's future growth to be increasingly driven by its online sales channel.

**What are the company's plans for share repurchases and/or dividends and what do these plans indicate?**

In September 2023, Sole Savvy's Board of Directors approved a share repurchase program, allowing the company to buy back up to \$1.6 billion worth of its common stock in the future. Additionally, the Board of Directors announced a new quarterly dividend of \$0.42 per share. These actions indicate that Sole Savvy is focused on returning value to its shareholders. The share repurchase program signals the company's confidence in its financial position and future prospects, while the dividend payment reflects its commitment to providing a return to shareholders.

**How did Sole Savvy's SG&A percentage compare to last year and what does this indicate?**

In 2023, Sole Savvy's SG&A percentage decreased from 24.6% to 24.4% compared to the previous year. This reduction implies that the company managed to control and optimize its expenses in relation to its total sales, suggesting improved cost efficiency. Sole Savvy attributes this improved efficiency to increased labor productivity and reduced impacts from COVID-19-related factors compared to the prior year.

**How has the company adjusted its capital expenditure strategy in the current year and what might this suggest?**

In 2023, Sole Savvy increased its capital expenditures to \$270 million, up from \$200 million in the previous year. The increase in capital expenditures reflects new investments in digital infrastructure and upgraded supply chain systems. This suggests the company is committed to adapting to changing consumer preferences. Specifically, the focus on digital infrastructure supports the growth in online sales, indicating the company's strategic emphasis on the online sales channel, which is growing at a faster rate than in-person sales.

**Were there any material events that took place after the fiscal year-end? If so, how might these events influence Fiscal 2024 performance?**

On January 10, 2024, Sole Savvy entered into a partnership agreement with innoFT Tech, a technology company specializing in wearable fitness and health products. This strategic alliance aims to merge the worlds of fashion and wellness. The partnership may influence Fiscal 2024 performance by attracting a new segment of health-conscious consumers and driving sales through the integration of fitness and wellness into the retail experience. Additionally, the partnership may strengthen Sole Savvy's position in the market and differentiate its offerings.

### High Interactivity (Video example)

Use FinGPT to answer the five questions about Sole Savvy listed below.

Questions to answer:

1. How are Sole Savvy's different sales channels trending and what does this suggest for the future of the company?
2. What are the company's plans for share repurchases and/or dividends and what do these plans indicate?
3. How did Sole Savvy's SG&A percentage compare to last year and what does this indicate?
4. How has the company adjusted its capital expenditure strategy in the current year and what might this suggest?
5. Were there any material events that took place after the fiscal year-end? If so, how might these events influence Fiscal 2024 performance?

Submit Answers

**FinGPT**

**How are Sole Savvy's different sales channels trending and what does this suggest for the future of the company?**

In 2023, Sole Savvy's in-person sales grew by 2.7% to \$7,900 million, while online sales increased by 14.6% to \$2,120 million. This indicates that the company's online sales, though a smaller percentage of total sales, are growing at a faster rate than in-person sales. The trend suggests a shift in consumer behavior towards online shopping. Consequently, the future growth of Sole Savvy may increasingly be driven by its online sales channel.

**What are the company's plans for share repurchases and/or dividends and what do these plans indicate?**

Sole Savvy's Board of Directors approved a share repurchase program allowing the company to buy back up to \$1.6 billion worth of its common stock in the future. They also announced a quarterly dividend of \$0.42 per share. These plans indicate that Sole Savvy is focused on returning value to its shareholders. The share repurchase program signals the company's confidence in its financial position and future prospects, while the dividend payment reflects its commitment to providing a return to shareholders.

**How did Sole Savvy's SG&A percentage compare to last year and what does this indicate?**

In 2023, Sole Savvy's SG&A percentage decreased from 24.6% to 24.4% compared to the previous year. This reduction suggests that the company managed to control and optimize its expenses relative to its total sales, implying improved cost efficiency. The improved SG&A percentage arises from increased labor productivity and reduced impacts from COVID-19-related factors compared to the prior year.

How has the company adjusted its capital expenditure strategy in the current year and what might this suggest?

## **Panel C: Sample Questions Listed for Each Topic of Self-Assessment**

### Topic 1: Company Background

1. What are the key success factors for the company?
2. In what areas is the company investing money?
3. What are the company's biggest threats or obstacles?

### Topic 2: Financial Numbers

1. What is the company's gross margin and profit margin?
2. Did the company have any significant one-time or extraordinary expenses?
3. How did the company's cash balance change and what were the sources of cash flow?

### Topic 3: Financial Performance

1. How would you assess the company's overall financial performance?
2. In what areas has the company's performance improved or deteriorated?
3. How are the company's growth prospects and outlook?

### Panel D: Post-Experimental Quiz Questions (Experiment 1 only)

1. Sole Savvy's sales results suggest the company has \_\_\_\_\_
  - a. Slowing demand in certain international markets
  - b. Declining sales from in-person channels
  - c. Increasing demand for footwear relative to clothing
  - d. Shifting consumer preference for online vs. in-person shopping
2. Sole Savvy announced a quarterly dividend of \$ \_\_\_\_\_ per share
  - a. 0.42
  - b. 0.16
  - c. 0.27
  - d. 0.05
3. Sole Savvy's SG&A percentage suggests the company's cost efficiency \_\_\_\_\_ in 2023 relative to 2022
  - a. Improved
  - b. Worsened
  - c. Did not change
4. Sole Savvy's increase in capital expenditures was driven by investing in \_\_\_\_\_
  - a. Digital infrastructure and supply chain systems
  - b. Share repurchases and dividends
  - c. New stores and renovations
  - d. Acquisitions of other companies
5. Sole Savvy announced they are acquiring InnoFit, a company that specializes in \_\_\_\_\_
  - a. Lightweight athletic footwear
  - b. Fitness-based data analytics
  - c. Wearable fitness and health devices
  - d. Athletic-friendly material production
6. FitWear mitigates the risk of economic uncertainty by \_\_\_\_\_
  - a. Hedging risk using investment strategies
  - b. Diversifying across geographies and product lines
  - c. Maintaining an advanced economic forecasting system
  - d. Building consumer loyalty
7. FitWear's \_\_\_\_\_ segment contributed **most** to 2023 performance
  - a. Asia
  - b. South America
  - c. Europe
  - d. North America
8. FitWear's **total** store count \_\_\_\_\_ during the current year
  - a. Did not change
  - b. Decreased
  - c. Increased
9. FitWear's overall liquidity position \_\_\_\_\_ during the current year
  - a. Strengthened
  - b. Did not change
  - c. Weakened
10. FitWear provided a forecast for \_\_\_\_\_ different metrics
  - a. 3
  - b. 2
  - c. 1
  - d. 0

## **Panel E: Earnings Release Questions (Experiment 3)**

### More Complex Conditions

1. *How are FitWear's different sales channels trending and what does this suggest for the future of the company?*
2. *Which geographic segments contributed most to FitWear's overall performance, and what does this indicate?*
3. *How has the company adjusted its capital expenditure strategy in the current year and what might this suggest?*
4. *What are FitWear's asset and liability levels and how healthy is the company in terms of liquidity?*
5. *How promising does FitWear's forecast look for the upcoming fiscal year?*

### Less Complex Conditions

1. *How much was FitWear's store sales revenue?*
2. *How much was FitWear's North America sales revenue?*
3. *How much were FitWear's capital expenditures?*
4. *How much does FitWear have in assets?*
5. *How much does FitWear forecast for earnings per share (EPS)?*

## APPENDIX B

### Operationalization of the Four Facets of Processing Tool Interactivity

Facet	Conceptual Definition	Operationalization: Experiments 1 & 3	Operationalization: Experiment 2
Reciprocity	The level of iterative, two-way communication between the user and the tool.	<p>High: Participants actively submit questions to FinGPT and receive responses in a back-and-forth chat.</p> <p>Low: Participants do not submit questions, but instead view static pre-generated FinGPT output. No back-and-forth between participants and FinGPT.</p>	<p>High: All participant questions and FinGPT responses displayed in back-and-forth chat.</p> <p>Low: Only one participant question and FinGPT response shown at a time in FinGPT window.</p> <p><i>Note: Unlike E1 and E3, all participants in E2 actively submit questions to FinGPT.</i></p>
Responsiveness	The degree to which the tool's output is a function of the user's input.	<p>High: Participants' questions directly determine the responses FinGPT provides.</p> <p>Low: Participant does not provide input and cannot influence FinGPT output.</p>	<p>High: FinGPT provides participant-requested information only with direct response.</p> <p>Low: FinGPT provides participant-requested information alongside other irrelevant information (i.e., answers to other questions).*</p>
Response Speed	The immediacy with which the tool generates output.	<p>High: FinGPT provides responses almost immediately after participants submit questions.</p> <p>Low: Not applicable because there is no active back-and-forth (static text shown immediately).</p>	<p>High: FinGPT provides responses almost immediately after participants submit questions.</p> <p>Low: FinGPT provides responses after a 30-second delay.</p>
Nonverbal Communication	The extent to which the tool uses supplementary channels, such as visual or auditory cues, to communicate output.	<p>High: FinGPT types out responses word-by-word in real-time.</p> <p>Low: No real-time typing; all text shown at once.</p>	<p>High: FinGPT types out responses word-by-word in real-time.</p> <p>Low: No real-time typing; FinGPT responses appear without animation.</p>

\*Similar to Johnson et al. (2006), the manipulation of responsiveness in Experiment 2 also influences reciprocity by varying the amount of back-and-forth interaction needed to extract all information from FinGPT.