No Matter Your Financial Literacy: Simplicity Wins When Choosing a Fund

Zihan Gong[*]   Sebastian Müller[†]

January 9, 2025

Abstract

This study assesses the impact of GPT-4-generated fund prospectus summaries on investor comprehension and decision-making through an online experiment. A total of 305 participants were given summaries designed with varying levels of complexity, labeled as "easy" and "hard." The findings reveal that easy-to-understand summaries significantly enhance text accessibility by approximately 13% and investment willingness by 8%. This improvement is observed regardless of participants' self-reported or objectively measured financial literacy. The study also finds that individuals' self-assessed financial competence plays a more crucial role than their actual literacy in interacting with financial information and making investment decisions. These results highlight the capability of advanced language models like ChatGPT to simplify and condense complex financial disclosures, thereby potentially broadening investor participation and enhancing financial engagement.

Keywords: ChatGPT, Large Language Models, Fund Prospectus Summarization, Financial Literacy, Financial Decision Making

---

[*]Corresponding author. Technical University of Munich, TUM School of Management, Campus Heilbronn, Center for Digital Transformation, Am Bildungscampus 9, 74076 Heilbronn, Germany (zihan.gong@tum.de, Phone.: +49713126418809).

[†]Technical University of Munich, TUM School of Management, Campus Heilbronn, Center for Digital Transformation, Am Bildungscampus 9, 74076 Heilbronn, Germany (sebastian.mueller.hn@tum.de, Phone.: +49713126418806).

1   Introduction

Understanding complex financial information remains a significant challenge for individuals without a financial background, especially when making self-directed investment decisions like choosing an appropriate fund. Despite the growing emphasis on personal investment, many individuals arguably lack the expertise to interpret intricate financial documents. This knowledge gap may hinder informed decision-making. Recognizing these challenges, the SEC has taken steps to improve accessibility by mandating that companies provide key information in a more concise format through summary fund prospectuses. Since March 2009, following the Summary Prospectus Adopting Release[1], companies have been required to deliver essential information directly to investors in this simplified document, while making the full statutory prospectus available online. While the summary prospectus is designed to be concise and is written in plain English, it nonetheless incorporates standard financial terminology such as "leverage", "liabilities", "equity", "beta", etc. This use of technical jargon may still present challenges to readers unfamiliar with financial terms. Moreover, although the summary prospectus is significantly shorter than the statutory prospectus, typically comprising fewer than 10 pages, it may still be considered lengthy by investors, particularly given the vast number of funds available for selection.

Recent advancements in artificial intelligence (AI), especially Large Language Models (LLMs) like GPT-4, offer novel methods for simplifying and summarizing complex financial information. In this study, we explore how the complexity of fund prospectuses affects investors' understanding and decision-making. We use GPT-4 to generate two versions of the summary fund prospectuses with two distinct prompts: the 'easy' version, designed for individuals without a financial background, uses simple language and concise formatting to highlight key information, while the 'hard' version, crafted for experienced retail investors, employs precise technical language and emphasizes financial terminology appropriate for expert decision-making. Importantly, GPT-4's role in this study is restricted solely to the task of summarization; it does not engage in any investment decision-making.

---

[1]https://www.sec.gov/files/rules/final/2009/33-8998.pdf.

We investigate whether "easy" versus "hard" GPT-4-generated summaries improve participants' comprehension of key investment details and increase their willingness to invest. The "easy" versions are designed for individuals without a financial background, using simplified technical language and highlighting key information in a concise, accessible format. Conversely, the "hard" versions are tailored for retail investors with higher expertise and more experienced, emphasizing technical precision and financial terminology appropriate for expert decision-making. We expect that participants who read the 'easy' text will score higher on the text accessibility measure — which combines readability, understandability, and comprehensiveness — as well as on their willingness to invest. Furthermore, we hypothesize that financial literacy mitigates the challenges of reading complex text, with experienced retail investors experiencing fewer difficulties compared to laypersons when reading the "hard" version.

To test these hypotheses, we conduct an online experiment via the Prolific platform in May 2024 with a final sample of 305 participants whose primary language is English. Participants are randomly assigned to read one "easy" text and one "hard" text, with the order of reading also randomized to control for potential order effects and minimize bias. The experiment evaluates participants' perceptions of readability, understandability, comprehensiveness, and investment willingness. This experimental design facilitates both between-subjects and within-subjects analyses, allowing us to compare how different participants respond to each summarized text, as well as how individual participants' responses vary between the "easy" and "hard" versions. A statistical power analysis using G*Power [2] ensured the sample size was sufficient to achieve 95% confidence or greater. Moreover, we define two financial literacy measures, one objective measure based on the correct answers for finance-related questions and one subjective measure based on their self-evaluation. We test whether the interaction between financial literacy (both objectively measured and self-assessed) and the "easy" version of the text, as well as their individual effects, positively influences comprehensive and investment willingness scores. This approach allows us to explore if participants

---
[2]Faul et al. (2009).

with varying levels of financial expertise perceive the benefits of summarized texts differently.

Our results indicate that GPT-4-generated summaries significantly enhance both comprehension and investment willingness, particularly when the text is simplified. Participants consistently rate the easier versions of the fund prospectuses higher in terms of readability, understandability, and comprehensiveness, with combined text accessibility scores increasing by approximately 13% (or 0.54–0.65 standard deviations) compared to the harder versions. They are also 8% more likely to express willingness to invest (or 0.26–0.34 standard deviations) after reading these easier summaries. Interestingly, contrary to our hypothesis, even retail investors with higher expertise favor the "easy" version over the complex one. Both laypersons and knowledgeable retail investors, regardless of whether their financial expertise is measured objectively or through self-assessment, favor the "easy" version. This suggests that text complexity has a similar impact across different levels of financial expertise, with no substantial differences in responses based on financial literacy.

Moreover, to address potential order effects, we randomize the sequence in which participants access the text versions. Reading order has a negative and significant effect on text accessibility scores, indicating that participants rate the text they read second more favorably. This suggests that familiarity increases positive perceptions. However, the lack of significant impact of reading order on investment willingness supports that this order effect does not extend to participants' willingness to invest.

Lastly, while self-assessed financial knowledge positively influences participants' comprehension and investment willingness, it does not affect their preference for simpler text, as the interaction between self-assessed financial knowledge and the "easy" version is insignificant. Conversely, objective financial literacy does not have a positive influence on participants' comprehension or investment willingness. Furthermore, the interaction between objective financial literacy and the 'easy' version is also insignificant. In summary, simplifying complex financial information appears to benefit not just those with lower financial knowledge but also experienced retail investors.

Our study builds on prior research focusing on financial reporting complexity and its implications. Johnson et al. (2002) use the Cloze Procedure with 110 college students to demonstrate that the investment objective section of mutual fund prospectuses is not readable. They also find that students' ability to comprehend these documents correlates positively with their grade point average and the extent of their financial training, despite a tendency to overestimate their financial knowledge. Similarly, Philpot and Johnson (2007) analyze the readability of mutual fund prospectuses across 20 major U.S. fund families using Flesch scores, finding variability in clarity with risk discussions being clearer than objective/strategy sections. Beshears et al. (2009) assess the impact of the SEC's Summary Prospectus on mutual fund selection and observe no significant effect on investor choices; however, they note a reduction in the time spent on investment decisions. Recent studies like DeHaan et al. (2021) explore the "strategic obfuscation" by mutual fund managers through complex disclosures and fee structures, revealing that S&P 500 index funds use complexity to obscure high fees, leading to poor investment choices and price dispersion among similar funds. Tucker and Xia (2023) report that despite the SEC's mandate for mutual funds to use plain English, many funds fail to meet this standard. Additionally, a substantial body of literature, including studies by You and Zhang 2009, Lehavy et al. 2011, Lehavy et al. 2011, Guay et al. 2016, Kim et al. 2019, etc., focuses on the complexity of financial reporting, particularly in 10-K filings.

Since its release in November 2022, ChatGPT has demonstrated significant potential in finance. Applications of ChatGPT and other LLMs have shown promising results across a range of tasks, including classification, sentiment analysis, and summarization (Dowling and Lucey 2023; Fatouros et al. 2023; Ko and Lee 2024; Kawamura et al. 2024; Dong et al. 2023). Earlier work like Goyal et al. (2022) evaluates the impact of GPT-3 on text summarization, particularly in news domains, finding that GPT-3's prompted summaries are preferred by humans over fine-tuned models and are less prone to factual errors. Yue et al. (2023) evaluate the effectiveness of using Explainable AI (XAI) and ChatGPT to explain complex financial concepts for non-experts, demonstrating ChatGPT's potential to simplify

4

financial communication and the potential for future enhancements to improve financial literacy. Korinek (2023) discusses the capabilities of LLMs (GPT-4 and Claude 2) in the context of text summarization. Closely related to our study, Kim et al. (2024b) explore the value of LLMs in summarizing complex corporate disclosures and find that AI-generated summaries help investors by capturing essential information more effectively than the full texts, particularly influencing market reactions based on sentiment analysis. Similarly, Kim et al. (2024a) explores the application of GPT-4 in summarizing earnings call transcripts to improve decision-making in tasks such as earnings predictions and portfolio allocation, emphasizing the role of financial sophistication. Unlike our study, which highlights the universal benefits of simplified fund prospectuses for all investor groups, they demonstrate how AI tools can create tailored insights that may inadvertently widen the gap between sophisticated and less experienced investors.

Recent research has also underscored ChatGPT's role in financial advisory contexts. Niszczota and Abbas (2023) find GPT has become financially literate, increasing scores from 65% to 99% and this suggests that it is a plausible source of financial advice for laypeople. Lakkaraju et al. (2023) examine the performance of ChatGPT and other LLMs in providing personal finance advice, finding that while these models produce fluent and plausible responses, gaps in accuracy and reliability remain. Yang et al. (2023) introduce InvestLM, a financial domain-specific model, highlighting the value of fine-tuning LLMs for improved performance in generating investment-related advice. Furthermore, Fatouros et al. (2024) demonstrate that GPT-4-based frameworks like MarketSenseAI generate investment signals capable of achieving significant excess returns, showcasing the potential for LLMs in stock selection. All this research highlights ChatGPT's promising role in finance, from enhancing financial literacy and simplifying complex financial communications for non-experts to effectively summarizing corporate disclosures, providing financial advice, and supporting informed investment decisions.

The main contribution of this study lies in demonstrating the effectiveness of GPT-4 in making financial disclosures more accessible and improving investment willingness, specif-

ically through its application in generating fund prospectus summaries for various target audiences. By evaluating both simplified and complex versions of these summaries, the study highlights how text complexity influences participants' comprehension and investment decisions. The findings suggest that simplified, easy-to-understand summaries improve both text accessibility and willingness to invest, particularly among participants with lower self-assessed financial knowledge. Moreover, the easier fund prospectuses are rated more favorably by participants, leading to higher confidence and an increased likelihood of investing. This shows that when financial information is presented in a more accessible format, people are more inclined to make positive investment decisions. Furthermore, the project contributes to the understanding of how AI-generated financial summaries can bridge the gap between complex financial information and audiences with varying levels of financial expertise, offering insights into the broader potential for LLMs in financial education and communication.

Building on these findings, the study also provides important practical and policy implications. By presenting complex financial information in a more accessible format, institutions can enhance investor confidence and encourage greater participation in financial markets. From a policy perspective, the results suggest that existing regulatory efforts, such as the SEC's plain English mandate, might not fully address all investor needs regarding disclosure accessibility. This study contributes by highlighting areas where further improvements could enhance investor comprehension and engagement. Policymakers could consider revising standards to mandate the inclusion of simplified versions of key financial documents, ensuring comprehensibility for all investors. Additionally, regulatory bodies may encourage or even mandate the use of AI-driven tools for generating simplified summaries of financial disclosures.

This paper is organized as follows. In Section 2, we describe the fund prospectuses, how we summarize them, and the online experimental setting. In Section 3, we present the empirical results and in Section 4, we conclude.

2 Data and Methodology

2.1 Selected Fund Prospectuses and Prompting

We collect a total of 60 fund summary prospectuses, consisting of 30 ETF prospectuses from State Street Global Advisors[3] and 30 mutual fund prospectuses from Fidelity[4]. The selection process is conducted randomly, without targeting specific mutual funds or ETFs. The initially downloaded prospectuses consist of approximately 5 to 8 pages, including tables and figures. For each prospectus, two distinct summarized versions are generated: an "easy" version and a "hard" version. We employ the updated GPT-4 Turbo model (gpt-4-0125-preview) for this task taking into account cost, context window, and maximum output tokens. Each prospectus is downloaded as a PDF, and we extract the content from the PDF, storing it in a variable labeled "text" for input into the prompt. The "easy" versions are produced using the following prompt:

"Your task is to generate a short, easy-to-understand summary of the summary prospectus provided in the PDF within 300 words. This summary is intended for individuals who do not have a financial background and are not familiar with financial terms. Focus exclusively on the information provided in the PDF, without referencing external sources. Explain any figures, tables, and graphs in simple terms, avoiding technical jargon, and highlight the key information essential for understanding the investment opportunity. The summary should be concise, highlighting the most important aspects for a layperson investor, and must fit within the 300-word limit. The summary prospectus is provided below, delimited by triple backticks: '''`text`'''".

The "hard" versions are generated using the following prompt:

"Your task is to generate a concise and precise summary of the summary prospectus provided in the PDF within 300 words. This summary is intended for experienced retail

---

[3]https://www.ssga.com/us/en/intermediary/fund-finder.
[4]https://www.fidelity.com/mutual-funds/fidelity-funds/overview.

investors with a strong background in finance and familiarity with financial terminology and concepts. Focus exclusively on the information provided in the PDF, without referencing external sources. Explain any figures, tables, and graphs using appropriate financial terminology, and emphasize key information critical for decision-making by sophisticated investors. The summary should accurately reflect the most important aspects for a seasoned investor or financial expert, and must fit within the 300-word limit. The summary prospectus is provided below, delimited by triple backticks: ```text```".

2.2 Differences between "easy" and "hard" Summaries

In total, we have 120 (2×60) summarized fund prospectuses. Table 1 presents the average text complexity metrics for the original documents, the summarized "easy" versions, the summarized "hard" versions, and the differences between the "easy" and "hard" versions. The table includes the most commonly used measures in financial literature for evaluating text complexity (Philpot and Johnson 2007, Loughran and McDonald 2014, Guay et al. 2016, Bonsall and Miller 2017, Du and Yu 2021). The table highlights how the complexity of the text changes between versions, with the original documents containing significantly more words and tokens than either of the summarized versions. For mutual funds, the original documents contain 5,030 words on average, while the "hard" and "easy" versions have 312 and 336 words, respectively. ETFs exhibit a similar reduction, with the original documents containing 5,361 words on average and the "hard" and "easy" versions having 345 and 371 words, respectively. The output slightly exceeds 300 words, but in the prompt, the instruction is to keep the summary within 300 words. This indicates that the prompt generally constrains the output to around 300 words while allowing for slight variations due to the complexity and content of the individual prospectuses. The token counts are much higher than the word counts for mutual funds and ETFs because tokens include subword fragments, symbols, and punctuation, which are more granular than individual words, especially in complex or formatted documents.

The Flesch Reading Ease score, which indicates how easy a text is to read (higher scores

representing easier readability), shows notable differences: For mutual funds, the original score of 50.96 drops to 41.88 in the "hard" version but increases to 57.80 in the "easy" version. Similarly, ETFs exhibit an increase in readability from the original score of 36.97 to 55.39 in the "easy" version. In contrast, other readability measures such as the Flesch-Kincaid Grade, Gunning Fog, SMOG Index, etc. follow an inverse trend, where higher scores indicate greater complexity. For these metrics, the "hard" versions exhibit higher scores, reflecting increased difficulty, while the "easy" versions show lower scores, indicating simpler texts. The scores of original documents typically fall between the two, representing an intermediate level of complexity. The differences in complexity scores between the "hard" and "easy" versions are highly statistically significant, as indicated by the t-statistics provided in Table 1. For example, the Flesch Reading Ease score for mutual funds increases by 15.92 points in the "easy" version compared to the "hard" version (t = 9.55), and for ETFs, the increase is 15.21 points (t = 8.69). Similarly, the Text Standard for mutual funds drops by 2 grades (t = 6.88), and for ETFs, it drops by 3 grades (t = 9.32), indicating simpler readability in the "easy" versions. These suggest that the prompts designed for different target audiences are effective in producing distinct levels of text complexity tailored to their respective needs. Specifically, the "hard" version, intended for readers with a strong financial background and familiarity with technical terminology, results in higher complexity and fewer words, while the "easy" version, aimed at a general audience without financial expertise, produces summaries that are simpler and more accessible. This confirms the effectiveness of our approach in adapting the content to suit varying levels of financial knowledge.

[here insert Table 1]

Both ETF and mutual fund original prospectuses consistently include all key information, including investment objectives, fees, portfolio turnover, principal investment strategies, principal investment risks, performance, portfolio management (investment adviser/portfolio manager), purchase and sale of shares, tax information, and payments to broker-dealers and other financial intermediaries. Table 2 presents the number of summarized prospectuses containing this key information and analyzes 30 ETF prospectuses and 30 mutual fund prospectuses, with each fund having both an "easy" and a "hard" summarized version.

For ETFs, all 30 "easy" and "hard" summaries include information on the fund's purpose/objective, strategies, and risks. However, while the return (fee) is included in all 30 "easy" summaries, it is only present in 26 (29) "hard" summaries. Portfolio turnover is mentioned in 1 "easy" summary and 21 "hard" summaries, and management information appears in 9 "easy" and 23 "hard" summaries. Purchase and sale information is found in 28 "easy" and 26 "hard" versions, while tax information is included in 27 "easy" and 29 "hard" summaries. For mutual funds, all 30 "easy" and "hard" versions provide information on the fund's purpose/objective, fees, and returns. However, strategy details are present in only 12 "easy" summaries compared to 22 "hard" summaries. Information on risks is found in 28 "easy" and 28 "hard" summaries. Portfolio turnover is included in 8 "easy" summaries and 28 "hard" summaries, and management information is provided in 9 "easy" and 29 "hard" summaries. Purchase and sale information is present in all 30 "easy" summaries and in 29 "hard" summaries, with tax details included in 28 "easy" and 27 "hard" summaries. Overall, the evidence in Table 2 highlights that while the majority of key information is consistently covered across both "easy" and "hard" versions, the "hard" versions generally contain more detailed information, particularly with regard to the inclusion of portfolio turnover and management information. It is important to note that in generating both "easy" and "hard" summary versions, no specific requirements regarding which information to include are given in the prompt.

[here insert Table 2]

Although the key fund variables provide essential details about investment objectives, performance potential, and risks, their consistent presence across both the "easy" and "hard" summaries indicates that these variables are not the primary drivers of participants' decisions in this study. Instead, the variation in text complexity between the two formats is the main factor shaping participants' comprehension and willingness to invest. This study focuses on how differing text formats ("easy" vs. "hard") influence participants' perceptions by altering the accessibility of the information, rather than the content itself. Since the

key information—such as returns, risks, and investment objectives—remains identical across both versions, we isolate the impact of presentation complexity on participants' understanding and investment decisions, independent of the inherent appeal or characteristics of the fund itself.

2.3   Experiment Design

To assess the effectiveness of the summarized prospectuses, we conduct an online survey via Prolific[5] in May 2024. To ensure the results are statistically significant, a statistical power analysis is performed using G*Power, as proposed by Faul et al. (2009), which indicate a minimum required sample size of 302 participants to achieve 95% confidence or greater. We recruit a total of 350 participants, limiting the sample to individuals whose primary language is English. After applying attention checks and excluding responses with missing data, the final sample consists of 305 participants.

Of the respondents, 48.2% identified as male, 50.2% as female, 0.7% as non-binary/third gender, and 0.9% declined to specify. The majority of participants (41.3%) are aged 26 to 35, followed by 34.4% aged 18 to 25. Smaller proportions are found in older age groups: 12.8% are 36 to 45, 7.2% are 46 to 55, 3.0% are 56 to 65, and 1.3% are 66 or older. Regarding education, 50.8% hold a Bachelor's degree, 32.8% have a high school diploma, 12.5% have a Master's degree, and 2.3% have a Ph.D. or higher. Additionally, 73.4% are unmarried, and 24.9% are married. In terms of employment, 49.2% are employees, 15.4% are university students, 13.8% are self-employed, and 7.9% are unemployed. Monthly income ranges from $1,000 to $2,000 for 44.6% of participants, with 18.7% earning $2,000 to $3,000, 17.4% preferring not to disclose, and smaller proportions earning more.

Our experimental design presents each participant with two summarized texts—one in an "easy" version and the other in a "hard" version. The order of presentation is randomized. After reading the first text, participants are asked to answer a series of questions. They

---

[5]https://www.prolific.com/.

then proceed to read the second text and answer the same set of questions. This approach enables a direct comparison of how participants perceive and comprehend each version of the text. Within our 305 samples, 150 participants read the "easy" version first, then the "hard" version and 155 participants read the "hard" version, then the "easy" version. The original fund prospectus is not included in the survey due to concerns that its length could negatively affect response rates (Deutskens et al. 2004, Revilla and Ochoa 2017).

The questions given to the participants are asked to evaluate a series of statements, organized into three categories: readability, understandability, and comprehensiveness. Each category includes three distinct statements, facilitating a thorough assessment of the text. Responses are recorded using a 7-point Likert scale, ranging from "strongly disagree" to "strongly agree". The scores for these categories are then calculated as the average of participants' evaluations for the three statements within each category. Additionally, participants are asked to rate their likelihood of investing in the fund based on the information provided in the summarized text, also using a 7-point scale, where 7 indicates "extremely likely to invest" and 1 indicates "extremely unlikely to invest". Following these evaluations, participants complete six basic financial literacy questions to assess their knowledge. These questions are widely used in financial literature (e.g. Lusardi and Mitchell 2011, Fernandes et al. 2014). Finally, participants provide information about their investment experience and self-assess their knowledge of investment options, choosing from four levels: basic, intermediate, advanced, or expert. They also indicate whether they have prior experience with investing, are interested in starting, or have no interest in investing. Screenshots of the online survey used for this experiment are provided in Appendix C.

Based on the questions, two distinct measures of financial literacy are constructed: an objective financial literacy measure (hereafter, OFI) and a subjective financial literacy measure (hereafter, SFI). The objective financial literacy measure is based on participants' responses to the six widely used financial literacy questions.(e.g. Lusardi and Mitchell 2011, Fernandes et al. 2014) The median number of correct responses is five. Participants who correctly answer five or more questions (n = 165) are classified as "financially literate", while those

who answer fewer than five questions correctly (n = 140) are classified as "financially illiterate". The subjective financial literacy measure is based on participants' self-assessment of their financial knowledge. The distribution of responses is as follows: 57% rate their knowledge as "basic", 36.1% as "intermediate", 6.9% as "advanced", and 0% as "expert". These self-assessments allow us to construct a subjective measure of financial literacy, with those rating themselves as "intermediate" or higher classified as having "high financial knowledge", and those rating themselves as "basic" classified as having "low financial knowledge". The correlation between the subjective and objective measures is 0.35. This rather low positive correlation is not surprising as it is well-documented that individuals are prone to overconfidence. Previous literature has used survey data to highlight the discrepancy between actual ability and perceived ability (e.g., Bhandari and Deaves 2006, Asaad 2015).

The median time to complete the survey is approximately 9.1 minutes, with an average completion time of 8.6 minutes. Participants spend an average of 1.4 minutes on the first essay and 1.3 minutes answering the related questions. For the second essay, participants spend an average of 1.5 minutes, with an additional 1.3 minutes on the corresponding questions. A comparison of the time spent on the two texts reveals no significant differences. Compensation for participating in the survey averages £9/hour. In our experimental setup, the primary objective is to assess participants' comprehension and preferences rather than their actual financial behavior or investment decisions. As such, incentive-compatible pay is less critical, since the study focuses on their evaluations of the provided fund prospectus summaries, relying on participants' intrinsic motivation to respond thoughtfully.

3   Data Analysis

3.1   Survey Data Analysis - T-Test

We start our analysis by within-subjects and between-subjects t-tests on the collected survey data. Given that participants respond to a 7-point Likert scale across three categories

(readability, understandability, and comprehensiveness), each comprising three statements, we calculate the mean score for each category. The text accessibility score is then derived as the average of the scores for readability, understandability, and comprehensiveness. For investment willingness, the score is based on a single statement, not an average. Higher scores indicate greater readability, enhanced understandability, improved comprehensiveness, and increased investment willingness.

Table 3 displays the results of within-subjects paired t-tests comparing participants' scores for various dimensions across the "easy" and "hard" versions of the texts, as well as between their first and second readings. The table has five key dimensions: Readability, understandability, comprehensiveness, text accessibility score, and investment willingness. The first part of the table shows that, on average, participants rate the easy text higher across all dimensions, no matter whether they read the easy text for the first time or for the second time. For readability, the easy text had a significantly higher score (5.51) compared to the hard text (4.51), with a statistically significant difference of 1.00 (t-statistic: 10.00). Similar trends are observed for understandability, where the easy text score of 5.54 compared to 4.61 for the hard text, with a difference of 0.93 (t-statistic 10.22). The comprehensiveness scores also follow this trend, with the easy text scoring 5.44 and the hard text scoring 4.95, yielding a significant difference of 0.49 (t-statistic 7.19). For the text accessibility score, which is an average of the readability, understandability, and comprehensiveness measures, the easy text is rated 5.50 compared to 4.69 for the hard text, with a difference of 0.81 (t-statistic 10.15). Investment willingness also shows a significant difference, with participants more likely to invest after reading the easy text (average score of 4.38) compared to the hard text (average score of 3.89), yielding a significant difference of 0.49 (t-statistic 5.25).

The second part of the table compares participants' scores between the first and second readings, regardless of whether the text is easy or hard. The results indicate a slight increase in scores on the second reading across all dimensions, though the magnitude of these differences is less pronounced. For example, the readability score increases from 4.80 to 5.22 after the second reading, with a difference of 0.42 (t-statistic: 3.66). Similarly, understandability

and text accessibility scores show a slight but statistically significant increase in the second reading. These results suggest that participants, now more familiar with the text, rate the second reading slightly higher. This increased familiarity may enhance comprehension, resulting in slightly elevated investment willingness scores; however, no significant difference is observed between the two readings concerning investment willingness.

[here insert Table 3]

Table 4 shows the results of within-subjects paired t-tests comparing participants' scores for five dimensions—based on the order in which they read the "easy" and "hard" versions of the texts. Participants are divided into two groups: 150 participants read the easy text first followed by the hard text, while 155 participants read the hard text first followed by the easy text. The first row of the table shows the average scores for participants who first read the easy text, while the second row shows the scores for those who first read the hard text.

For participants who read the easy text first and then the hard text, the results indicate that the easy text receives significantly higher scores across all dimensions. For readability, the easy text scores 5.14 compared to 4.54 for the hard text, with a significant difference of 0.60 (t-statistic: 4.40). Similar patterns are observed for understandability (5.28 vs. 4.58, with a difference of 0.70, t-statistic: 5.24), comprehensiveness (5.22 vs. 4.83, difference of 0.39, t-statistic: 3.69), and text accessibility score (5.21 vs. 4.65, difference of 0.56, t-statistic: 4.81). Investment willingness also shows a significant difference, with participants more likely to invest after reading the easy text (average score of 4.15) compared to the hard text (average score of 3.78), with a difference of 0.37 (t-statistic: 3.03).

For participants who read the hard text first and then the easy text, the results again show that the easy text is rated more favorably across all dimensions, but with slightly larger differences. For readability, the easy text scores 5.88 compared to 4.48 for the hard text, yielding a significant difference of 1.40 (t-statistic: 10.02). The differences for under-

15

standability (5.80 vs. 4.65, a difference of 1.15, t-statistic: 9.47), comprehensiveness (5.65 vs. 5.06, a difference of 0.59, t-statistic: 6.77), and text accessibility score (5.78 vs. 4.73, a difference of 1.05, t-statistic: 9.91) are also highly significant. Investment willingness is also significantly higher after reading the easy text (average score of 4.59) compared to the hard text (average score of 3.99), with a difference of 0.60 (t-statistic: 4.32).

In summary, results shown in Table 4 suggest that, regardless of reading order, the easy text consistently receives higher scores across all dimensions. Notably, participants who read the hard text first tend to rate the easy text more favorably, with larger differences in all dimensions compared to those who read the easy text first. One possible explanation for this is the contrast effect: after reading a harder, more complex text, the easier text may appear significantly clearer and more understandable by comparison. The difficulty of the first text could heighten participants' appreciation for the simplicity and clarity of the easy text, leading to more favorable ratings. For example, one similar finding is reported by Hartzmark and Shue (2018), who find that investors mistakenly perceive today's earnings news as more impressive if yesterday's earnings surprise was negative, and less impressive if yesterday's surprise was positive.

[here insert Table 4]

We then focus on the analysis of between-subjects. Table 5 presents the results of independent two-sample t-tests comparing participants' scores for five categories, based on their first and second readings of the easy and hard texts. The sample is divided by reading order, with comparisons made between the easy and hard text for both the first reading and the second reading, regardless of which text version participants read first. The number of participants who first read the easy text is 150 and the hard text is 155, and for the second reading, the number of participants who read the hard text is 155 and the easy text is 150.

For the first reading, participants rate the easy text significantly higher in readability (5.14 vs. 4.48), understandability (5.28 vs. 4.65), and text accessibility score (5.21 vs. 4.73),

with significant differences of 0.66, 0.63, and 0.48, respectively. However, there is no significant difference in comprehensiveness (5.22 vs. 5.06) or investment willingness (4.15 vs. 3.99), with both differences of 0.16. For the second reading, participants again rate the easy text higher across all dimensions. Readability (5.88 vs. 4.54), understandability (5.80 vs. 4.58), comprehensiveness (5.65 vs. 4.83), text accessibility score (5.78 vs. 4.65), and investment willingness (4.59 vs. 3.78) all show significant differences, with larger gaps compared to the first readings. The t-statistics for these differences range from 4.41 to 9.05, indicating highly significant results.

Overall, this table demonstrates that participants consistently rate the easy text more favorably across all dimensions, which is similar to what we find in the within-subjects analysis (Table 3 and Table 4). The differences between the easy and hard texts are more pronounced during the second reading, likely because participants, already familiar with the content, perceive the easy text as even clearer and more engaging than the hard text. This familiarity effect enhances their comprehension and increases their investment willingness, as reflected in the significantly larger differences observed in readability, understandability, and investment willingness during subsequent readings.

[here insert Table 5]

In addition to the "easy" and "hard" versions, we also include two different types of fund prospectuses: mutual fund prospectus and ETF prospectus. Table 6 presents the results of independent two-sample t-tests comparing participants' scores for ETF versus mutual fund summarized texts across five dimensions. During the first reading, there are no statistically significant differences between the ETF and mutual fund summarzied texts across any of the dimensions. During the second reading, there are still no significant differences in most dimensions, but investment willingness shows a marginally significant difference. Participants are slightly more inclined to invest after reading the mutual fund text compared to the ETF text (4.38 vs. 4.02), with a difference of 0.36 (t-statistic: 1.90). Understandability shows a small but marginally significant difference (0.28, t-statistic: 1.74) at the 90% confidence

level. Overall, this table suggests minimal differences between how participants perceive ETF and mutual fund summarized texts across readability, understandability, comprehensiveness, and text accessibility score. However, investment willingness slightly favors mutual fund texts during the second reading. Hence, the differences are marginal, and for further analysis, we ignore the influence of different types of prospectus.

[here insert Table 6]

3.2 Pooled Cross-Sectional Regression Analysis

In this section, we conduct several regression analyses. For the pooled regression analysis, both evaluations are included as separate observations, allowing for a comprehensive analysis of all text evaluations. Consequently, the total number of observations in this analysis is 610, reflecting the two evaluations from each participant.

Specifically, we conduct the following regressions to investigate how text complexity (*Easy*), *PageOrder*, and participant characteristics (*OFI*, *SFI*, and *Demo*) influence text accessibility scores and investment willingness. In the following analysis, we include only the text accessibility scores, as they represent the average of the readability, understandability, and comprehensiveness scores. This aggregate measure captures all relevant aspects of text evaluation, eliminating the need to analyze these dimensions separately. To ensure reliable results, we calculate t-statistics with standard errors grouped by text, combining responses to the same text, whether in its easy or hard version, into one group. Our focus is more on the interaction terms. If β_5 or β_6 is positive, it means that the combined effect of being financially literate and reading an easy text is greater than the sum of the individual effects of financial literacy and text complexity alone. In other words, being financially literate enhances the benefit of reading an easier text. If they are negative, it indicates that the combined effect of financial literacy and text complexity is less than the sum of their individual effects, implying financial literacy may reduce the benefit of reading an easier text.

18

$$TestAccessibility_i = \alpha + \beta_1 Easy_i + \beta_2 OFI_i + \beta_3 SFI_i + \beta_4 PageOrder_i + \tag{1}$$

$$\beta_5 Easy_i * OFI_i + \beta_6 Easy_i * SFI_i + \beta_7 demo_i + \epsilon \tag{2}$$

$$InvestmentWillingness_i = \alpha + \beta_1 Easy_i + \beta_2 OFI_i + \beta_3 SFI_i + \beta_4 PageOrder_i + \tag{3}$$

$$\beta_5 Easy_i * OFI_i + \beta_6 Easy_i * SFIi + \beta_7 demo_i + \tag{4}$$

$$\beta_8 FundCharacteristics_i + \epsilon \tag{5}$$

Here, OFI_i indicates whether a participant is objectively financially literate (1 for financially literate, 0 otherwise); SFI_i is based on participants' self-assessment, with 0 representing "basic" knowledge and 1 representing "intermediate" or "advanced" knowledge. $Easy_i$ takes the value of 1 if the participant read the easy text and 0 if they read the hard text. $PageOrder_i$ equals 1 if the text appears first and 0 otherwise. We also include interaction terms between OFI_i and $Easy_i$, as well as between SFI_i and $Easy_i$. Demographic information ($demo_i$), which includes gender, age, education, income, and employment status, is included to control for other influences. $FundCharacteristics_i$, which are assumed to influence only investment willingness and not participants' understanding of the summarized texts, are also included. These characteristics comprise total annual fund operating expenses, 1-year average annual total returns (before tax), 5-year average annual total returns (before tax), and 10-year average annual total returns (before tax). The disturbance term ϵ_i is assumed to have a mean of zero, capturing random variations unexplained by the model. T-statistics are computed using standard errors clustered by the same funds to ensure reliable inference.

Table 7 presents the results of the regression examining factors that influence participants' text accessibility scores, which range from 1 to 7, with higher scores indicating stronger clarity of the summarized prospectus. The variable $Easy$ consistently shows a strong and positive impact on text accessibility scores, with a coefficient of 0.81 across all models, indicating

that participants rate the easier text significantly higher. SFI also exhibits a significant positive effect, with coefficients ranging between 0.51 and 0.56, suggesting that participants with higher self-assessed financial knowledge tend to provide higher scores. In contrast, the objective measure OFI remains insignificant, showing no notable effect on text accessibility. The variable $PageOrder$ has a consistently negative and significant coefficient, ranging from -0.24 to -0.26, suggesting that participants' second reading results in higher accessibility compared to the first text they read. This finding may reflect initial unfamiliarity with the task or material, leading to lower scores on the first reading compared to subsequent ones. Controlling for demographic information, captured by $Demo$, does not change the significance or direction of the key variables. The t-statistics are calculated using standard errors clustered by the same funds. Overall, the results demonstrate that participants' text accessibility scores are significantly influenced by text simplicity ($Easy$) and self-assessed financial knowledge (SFI), while other variables like OFI and demographic factors have minimal impact.

[here insert Table 7]

Building on this, the analysis presented in Table 8 further investigates whether the effects of text simplicity ($Easy$) and financial literacy (OFI and SFI) interact to influence text accessibility scores. The variable $Easy$ remains a strong predictor, with coefficients ranging from 0.68 to 0.81, showing that participants consistently rate the easier text more favorably. Similarly, SFI has a significant positive effect, with coefficients between 0.53 and 0.61, highlighting that individuals with higher self-assessed financial knowledge tend to perceive texts as more accessible. The interaction terms between OFI and $Easy$, as well as SFI and $Easy$, are insignificant, suggesting that financial literacy—whether objectively measured or self-assessed—does not amplify or diminish the benefits of easier texts. $PageOrder$, on the other hand, consistently exhibits a negative and significant effect, with coefficients ranging from -0.24 to -0.25, indicating that participants rate the first text they read lower, potentially due to initial unfamiliarity with the task. Demographic controls ($Demo$) do not alter the results, affirming the robustness of these findings. This analysis demonstrates that text

20

simplicity ($Easy$) and self-assessed financial knowledge (SFI) are the key factors influencing text accessibility scores.

[here insert Table 8]

Table 9 presents the results of a pooled regression analysis examining the impact of various factors on participants' investment willingness, ranging from 1 to 7, where higher scores indicate a stronger likelihood to invest. The variable $Easy$ shows a consistent and significant positive effect on investment willingness across all models, with coefficients around 0.49. This indicates that participants who read easier texts are more inclined to invest. SFI also exhibits a strong and significant positive impact, with coefficients between 0.66 and 0.75, suggesting that participants with higher self-assessed financial knowledge are more likely to express a willingness to invest. In contrast, OFI displays a smaller and only marginally significant effect in some models, indicating that participants' objective financial literacy has a limited influence on their investment willingness. The variable $PageOrder$ has a negative but insignificant coefficient, suggesting that the sequence in which participants read the texts does not substantially affect their willingness to invest. Controlling for demographic factors ($Demo$) and fund characteristics does not alter the significance or direction of these results, underscoring their robustness. The t-statistics are calculated using standard errors clustered by the same funds, ensuring reliable inference. Overall, the results highlight that text simplicity ($Easy$) and self-assessed financial knowledge (SFI) are key drivers of investment willingness, while objective financial literacy (OFI), page order, demographic factors, and fund characteristics have minimal impact.

[here insert Table 9]

Table 10 extends the analysis by examining whether the effects of text simplicity ($Easy$) on investment willingness are moderated by participants' financial literacy, using interaction terms between $Easy$ and both OFI and SFI. The results confirm that $Easy$ continues to have a significant positive impact on investment willingness across all models, with coefficients ranging from 0.42 to 0.56. This reaffirms that participants exposed to easier texts are

21

more likely to express a willingness to invest. Similarly, SFI remains a strong and significant predictor, with coefficients ranging from 0.76 to 0.87, indicating that participants with higher self-assessed financial knowledge are more inclined to invest. The interaction terms between $Easy$ and OFI ($OFI*Easy$) and between $Easy$ and SFI ($SFI*Easy$) are found to be statistically insignificant across all models. This suggests that the relationship between text simplicity and investment willingness does not depend on participants' financial literacy, whether measured objectively or subjectively. In other words, the benefit of reading an easier text applies uniformly to participants regardless of their financial literacy level. As in previous models, $PageOrder$ exhibits a negative but insignificant effect, indicating that the order of text reading does not significantly impact participants' investment willingness. Demographic factors ($Demo$) and fund characteristics are included as controls and do not alter the key findings. The t-statistics are calculated using standard errors clustered by the same funds, ensuring reliable inference. Overall, this analysis highlights that text simplicity ($Easy$) and self-assessed financial knowledge (SFI) are primary drivers of investment willingness, with little evidence of moderating effects from financial literacy.

[here insert Table 10]

Taking together, text simplicity ($Easy$) has a stronger effect on text accessibility scores (coefficients ranging from 0.68 to 0.81) than on investment willingness (coefficients ranging from 0.42 to 0.56), indicating that easier texts are more impactful in improving clarity than driving investment decisions. Self-assessed financial knowledge (SFI) is consistently significant for both outcomes, with larger coefficients for investment willingness (0.66 to 0.87), suggesting it plays a key role in shaping decisions. In contrast, objective financial literacy (OFI) has no significant effect on text accessibility and only a marginal influence on investment willingness. Interaction terms between $Easy$ and financial literacy measures (OFI and SFI) are insignificant, showing that the benefits of simplicity apply uniformly across participants. $PageOrder$ negatively affects text accessibility but has no significant impact on investment willingness. Demographic controls and fund characteristics do not alter these results.

22

3.3   Robustness

This section provides robustness checks by dividing the sample based on reading order and financial literacy, replicating the regressions from the previous section. The dependent variables are text accessibility scores and investment willingness (1 to 7), with $Easy$, OFI, SFI, $PageOrder$, and interaction terms as independent variables. Control variables include demographics, fund fees, and returns. See Appendix A for the tables.

In Table A.1, we present the regression results on text accessibility scores, showing the effect of reading order. Panel A (first reading) shows that $Easy$ has a positive and significant effect, while OFI is not significant and SFI has a positive effect. Interaction terms ($OFIEasy$ and $SFIEasy$) are insignificant. In Panel B (second reading), the effect of $Easy$ strengthens, with SFI remaining significant and other variables unchanged. Table A.2 shows results for investment willingness. In Panel A (first reading), $Easy$ and OFI are not significant, but SFI has a strong positive effect. In Panel B (second reading), $Easy$ becomes significant, while SFI remains significant, and other variables remain unchanged.

Table A.3 presents the regression results for text accessibility scores, divided by financial literacy (OFI) and self-assessed financial knowledge (SFI). The $Easy$ variable consistently shows a significant positive effect, stronger among financially literate participants ($OFI = 1$), while $PageOrder$ negatively influences scores, indicating higher ratings for the second text. Similarly, Table A.4 shows that $Easy$ significantly improves investment willingness across all groups, with notable effects among financially literate participants and those with low self-assessed knowledge. $PageOrder$ has no significant impact on investment willingness, and including demographic and fund control variables does not change the results, confirming their robustness.

Overall, these robustness checks confirm that text simplicity enhances comprehension and investment willingness, particularly during repeated readings. The impact of financial

literacy and reading order is less pronounced, with text simplicity being the key driver of both higher ratings and investment willingness.

4   Conclusions

Using an online experiment, this study highlights the advantages of using GPT-4 to generate shorter, simplified fund prospectuses that enhance comprehension and investment willingness among participants. Text complexity plays a crucial role: Easier, more readable texts significantly improve participants' understanding, as evidenced by higher readability, understandability, and comprehensibility scores, and boost confidence in making investment decisions. These benefits are consistent across all participant groups, including both laypersons and experienced investors, regardless of their financial literacy or self-assessed knowledge. This underscores the potential of AI-generated content to democratize access to complex financial information, making it more accessible to a broad audience.

While reading order does not affect investment willingness, participants tend to rate the second text they read higher in terms of readability and comprehensibility, likely due to increased familiarity with the content and evaluation process. This suggests that repeated exposure to financial information, regardless of complexity, could lead to more favorable evaluations but does not necessarily increase willingness to invest.

The findings confirm GPT-4's suitability for generating summaries that simplify complex financial disclosures, improving engagement and decision-making across diverse audiences. Future research could explore the real-world impact of AI-generated summaries when integrated into digital advisory platforms or robo-advisors. Investigating the long-term effects of exposure to simplified financial content on financial literacy and decision-making skills could provide further insights. Additionally, assessing the applicability of AI-generated summaries in other financial domains, such as insurance or retirement planning, would broaden the scope of these findings.

From a policy and practical perspective, these results suggest that financial institutions and advisors may leverage AI tools like GPT-4 to create accessible and user-friendly disclosures, fostering greater client engagement and trust. Even regulators, such as the SEC, might consider mandating simplified financial summaries to ensure equal access to actionable information. However, simplified summaries must meet disclosure standards and avoid misrepresentation. Pilot programs under regulatory oversight could evaluate AI-generated disclosures' accuracy and usability.

References

Asaad, C. T. (2015). Financial literacy and financial behavior: Assessing knowledge and confidence. *Financial Services Review*, 24(2):101–117.

Beshears, J., Choi, J. J., Laibson, D., Madrian, B. C., et al. (2009). *How does simplified disclosure affect individuals' mutual fund choices?* Number w14859. National Bureau of Economic Research.

Bhandari, G. and Deaves, R. (2006). The demographics of overconfidence. *The Journal of Behavioral Finance*, 7(1):5–11.

Bonsall, S. B. and Miller, B. P. (2017). The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies*, 22:608–643.

DeHaan, E., Song, Y., Xie, C., and Zhu, C. (2021). Obfuscation in mutual funds. *Journal of Accounting and Economics*, 72(2-3):101429.

Deutskens, E., De Ruyter, K., Wetzels, M., and Oosterveld, P. (2004). Response rate and response quality of internet-based surveys: An experimental study. *Marketing letters*, 15:21–36.

Dong, M. M., Stratopoulos, T. C., and Wang, V. X. (2023). A scoping review of chatgpt research in accounting and finance. *Theophanis C. and Wang, Victor Xiaoqi, A Scoping Review of ChatGPT Research in Accounting and Finance (December 30, 2023)*.

Dowling, M. and Lucey, B. (2023). Chatgpt for (finance) research: The bananarama conjecture. *Finance Research Letters*, 53:103662.

Du, S. and Yu, K. (2021). Do corporate social responsibility reports convey value relevant information? evidence from report readability and tone. *Journal of business ethics*, 172:253–274.

Fatouros, G., Metaxas, K., Soldatos, J., and Kyriazis, D. (2024). Can large language models beat wall street? unveiling the potential of ai in stock selection. *arXiv preprint*

arXiv:2401.03737.

Fatouros, G., Soldatos, J., Kouroumali, K., Makridis, G., and Kyriazis, D. (2023). Transforming sentiment analysis in the financial domain with chatgpt. *Machine Learning with Applications*, 14:100508.

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using g*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4):1149–1160.

Fernandes, D., Lynch Jr, J. G., and Netemeyer, R. G. (2014). Financial literacy, financial education, and downstream financial behaviors. *Management science*, 60(8):1861–1883.

Goyal, T., Li, J. J., and Durrett, G. (2022). News summarization and evaluation in the era of gpt-3. arxiv. *arXiv preprint arXiv:2209.12356*.

Guay, W., Samuels, D., and Taylor, D. (2016). Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2-3):234–269.

Hartzmark, S. M. and Shue, K. (2018). A tough act to follow: Contrast effects in financial markets. *The Journal of Finance*, 73(4):1567–1613.

Johnson, D. T., Bauerly, R. J., and Waggle, D. (2002). Are mutual fund prospectuses written in plain english? *Managerial Finance*, 28(7):43–55.

Kawamura, K., Li, Z., Lin, C.-K., and Mcdanel, B. (2024). Revelata at the finllm challenge task: Improving financial text summarization by restricted prompt engineering and fine-tuning. In *Proceedings of the Eighth Financial Technology and Natural Language Processing and the 1st Agent AI for Scenario Planning*, pages 146–152.

Kim, A., Kim, D. S., Muhn, M., Nikolaev, V. V., and So, E. C. (2024a). Ai, investment decisions, and inequality. *Chicago Booth Accounting Research Center Research Paper, Fama-Miller Working Paper, MIT Sloan Research Paper*.

Kim, A., Muhn, M., and Nikolaev, V. V. (2024b). Bloated disclosures: can chatgpt help investors process information? *Chicago Booth Research Paper*, (23-07):2023–59.

Kim, C., Wang, K., and Zhang, L. (2019). Readability of 10-k reports and stock price crash risk. *Contemporary accounting research*, 36(2):1184–1216.

Ko, H. and Lee, J. (2024). Can chatgpt improve investment decisions? from a portfolio management perspective. *Finance Research Letters*, 64:105433.

Korinek, A. (2023). Generative ai for economic research: Use cases and implications for economists. *Journal of Economic Literature*, 61(4):1281–1317.

Lakkaraju, K., Vuruma, S. K. R., Pallagani, V., Muppasani, B., and Srivastava, B. (2023). Can llms be good financial advisors?: An initial study in personal decision making for optimized outcomes. *arXiv preprint arXiv:2307.07422*.

Lehavy, R., Li, F., and Merkley, K. (2011). The effect of annual report readability on analyst following and the properties of their earnings forecasts. *The accounting review*, 86(3):1087–1115.

Loughran, T. and McDonald, B. (2014). Measuring readability in financial disclosures. *the Journal of Finance*, 69(4):1643–1671.

Lusardi, A. and Mitchell, O. S. (2011). Financial literacy around the world: an overview. *Journal of pension economics & finance*, 10(4):497–508.

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

Philpot, J. and Johnson, D. T. (2007). Mutual fund performance and fund prospectus clarity. *Journal of Financial Services Marketing*, 11:211–216.

Revilla, M. and Ochoa, C. (2017). Ideal and maximum length for a web survey. *International Journal of Market Research*, 59(5).

Tucker, A. M. and Xia, Y. (2023). Promise & peril of plain english: Mutual fund disclosure readability. *Harv. Bus. L. Rev.*, 13:59.

Yang, Y., Tang, Y., and Tam, K. Y. (2023). Investlm: A large language model for investment using financial domain instruction tuning. *arXiv preprint arXiv:2309.13064.*

You, H. and Zhang, X.-j. (2009). Financial reporting complexity and investor underreaction to 10-k information. *Review of Accounting studies*, 14:559–586.

Yue, T., Au, D., Au, C. C., and Iu, K. Y. (2023). Democratizing financial knowledge with chatgpt by openai: Unleashing the power of technology. *Available at SSRN 4346152.*

Table 1: Summarized Text Complexity Scores

This table presents the number of words, the number of tokens, reading time and various complexity scores for mutual fund and exchange-traded fund (ETF) prospectuses. Each prospectus is available in three versions: original, hard, and easy. The "Diff" column shows the difference between the hard and easy versions. For text complexity measures, higher scores typically indicate greater complexity: except higher Flesch Reading Ease scores indicate easier readability, higher scores in Flesch Kincaid Grade, Gunning Fog, SMOG Index, Automated Readability Index, Coleman Liau Index, Linsear Write Formula, Dale Chall Readability Score, Text Standard, Spache Readability, and Mcalpine Eflaw indicate more complex text. Except for the Flesch Reading Ease score, positive differences indicate higher scores in the hard versions, suggesting more complexity, while negative differences indicate higher scores in the easy versions, suggesting simpler text. Statistical significance of the differences between hard and easy versions is indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. And t-statistics are in parentheses.

Score	Mutual Fund				Exchange Traded Fund			
	Original	Hard	Easy	Diff	Original	Hard	Easy	Diff
Number of Words	5,030	312	336	-24*** (-3.69)	5,361	345	371	-26*** (-4.08)
Number of Tokens	152,448	429	434	-5 (-0.53)	102,209	457	459	-2 (-0.17)
Flesch Reading Ease	50.96	41.88	57.80	-15.92*** (-9.95)	36.97	40.72	55.93	-15.21*** (-8.60)
Flesch Kincaid Grade	11.03	11.55	9.51	2.04*** (6.81)	14.62	12.21	10.17	2.04*** (6.26)
Gunning Fog	11.02	13.93	11.81	2.12*** (6.39)	13.45	13.76	12.11	1.65*** (5.41)
Smog Index	14.26	14.58	13.03	1.55*** (5.75)	16.15	14.77	13.09	1.68*** (6.34)
Automated Readability Index	13.32	13.83	11.47	2.36*** (6.34)	16.70	14.45	11.81	2.64*** (7.26)
Coleman Liau Index	11.95	14.12	11.01	3.11*** (10.40)	12.82	14.00	10.82	3.18*** (10.70)
Linsear Write Formula	15.46	11.84	11.36	0.48 (0.60)	17.53	13.63	10.63	3.00*** (3.74)
Dale Chall Readability Score	8.42	11.40	9.50	1.90*** (13.70)	7.74	10.85	9.39	1.46*** (12.94)
Text Standard	11th Grade	13th Grade	10th Grade	3*** (6.02)	16th Grade	14th Grade	11th Grade	3*** (6.42)
Spache Readability	4.84	5.81	5.09	0.72*** (6.58)	5.68	5.92	5.30	0.62*** (6.06)
Mcalpine Eflaw	27.15	22.32	23.82	-1.5** (-2.30)	36.32	25.29	26.40	-1.11 (-1.41)
Reading Time	136.12	26.10	25.27	0.83* (1.57)	329.11	28.60	27.41	1.19** (2.18)

Table 2: Key Information in Summarized Prospectus

This table shows the number of prospectuses with the following key information. The key information includes the fund's investment objectives, fees and expenses, principal investment strategies, principal risks, portfolio turnover, fund performance, portfolio management, purchase and sale information, and tax information. We have 30 ETF prospectuses and 30 mutual fund prospectuses. For each type of prospectus, there are two versions: one is easy and one is hard.

Number of Text	ETF		Mutual Fund	
	Easy	Hard	Easy	Hard
Purpose/Objective	30	30	30	30
Fee	30	29	30	30
Return	30	26	30	30
Turnover	1	21	8	28
Strategy	30	30	12	22
Risk	30	30	28	28
Management	9	23	9	29
Purchase and Sale Information	28	26	30	29
Tax	27	29	28	27

Table 3: Test Full-Sample Difference within Subject

This table presents paired t-test results comparing the differences in participants' scores for easy versus hard text and between their first and second readings within the sample (N = 305). The scores are across five dimensions: readability, understandability, comprehensiveness, text accessibility, and investment willingness. The text accessibility score is the average of the readability, understandability, and comprehensiveness scores. The readability score, understandability score, comprehensive score, and text accessibility score range from 1 to 7, where 1 indicates "strongly disagree" and 7 indicates "strongly agree". The investment willingness ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". The "Easy" and "Hard" columns display the average scores for participants who read easy and hard texts, respectively. The "Diff" column shows the score differences, with t-statistics in parentheses. Additionally, the table compares scores from the first and second readings, with the "First" and "Second" columns showing the average scores for each reading, and the "Diff" column indicating the differences, with corresponding t-statistics in parentheses. Significance levels are denoted as *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Score	readability			understandability			comprehensiveness			text accessibility			investment		
	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff
Average	5.51	4.51	1.00***	5.54	4.61	0.93***	5.44	4.95	0.49***	5.50	4.69	0.81***	4.38	3.89	0.49***
			(10.00)			(10.22)			(7.19)			(10.15)			(5.25)
	First	Second	Diff	First	Second	Diff	First	Second	Diff	First	Second	Diff	First	Second	Diff
Average	4.80	5.22	-0.42***	4.96	5.20	-0.24**	5.14	5.25	-0.11	4.97	5.22	-0.25***	4.07	4.19	-0.12
			(-3.66)			(-2.34)			(-1.49)			(-2.81)			(-1.28)

Table 4: Test Easy-Hard Difference based on Page Order within Subject

This table shows paired t-test results comparing the differences in participants' scores for easy versus hard text and the order in which they read the texts. The scores are across five dimensions: readability, understandability, comprehensiveness, text accessibility, and investment willingness. The text accessibility score is the average of the readability, understandability, and comprehensiveness scores. The readability score, understandability score, comprehensive score, and text accessibility score range from 1 to 7, where 1 indicates "strongly disagree" and 7 indicates "strongly agree". The investment willingness ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". Participants are divided into two groups: 150 participants first read easy text and then hard text, while 155 participants first read hard text and then easy text. The first row of the table shows the average scores for readability, understandability, comprehensiveness, text accessibility, and investment willingness for participants who first read the easy text followed by the hard text. The second row shows the average scores for participants who first read the hard text followed by the easy text. The "Diff" column indicates the differences, with corresponding t-statistics in parentheses. Significance levels are indicated as follows: $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Score	readability			understandability			comprehensiveness			ovetext accessibilityrall			investment		
	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff
Average	5.14	4.54	0.60*** (4.40)	5.28	4.58	0.70*** (5.24)	5.22	4.83	0.39*** (3.69)	5.21	4.65	0.56*** (4.81)	4.15	3.78	0.37*** (3.03)
	Hard	Easy	Diff	Hard	Easy	Diff	Hard	Easy	Diff	Hard	Easy	Diff	Hard	Easy	Diff
Average	4.48	5.88	-1.40*** (-10.02)	4.65	5.80	-1.15*** (-9.47)	5.06	5.65	-0.59*** (-6.77)	4.73	5.78	-1.05*** (-9.91)	3.99	4.59	-0.60*** (-4.32)

Table 5: Test Easy-Hard Difference between Subject

This table presents independent two-sample t-test results comparing the differences in participants' scores based on the order of reading easy and hard texts. The scores are across five dimensions: readability, understandability, comprehensiveness, text accessibility, and investment willingness. The text accessibility score is the average of the readability, understandability, and comprehensiveness scores. The readability score, understandability score, comprehensive score, and text accessibility score range from 1 to 7, where 1 indicates "strongly disagree" and 7 indicates "strongly agree". The investment willingness ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". The sample is divided by page order, with the first row showing scores for participants who first read easy text (N = 150) and hard text (N = 155), and the second row showing scores for those who second read hard text (N = 155) and easy text (N = 150). The "Diff" column indicates the differences, with corresponding t-statistics in parentheses. Significance levels are denoted as $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Score	readability			understandability			comprehensiveness			text accessibility			investment		
first time	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff
Average	5.14	4.48	0.66*** (3.83)	5.28	4.65	0.63*** (4.12)	5.22	5.06	0.16 (1.26)	5.21	4.73	0.48*** (3.50)	4.15	3.99	0.16 (0.91)
second time	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff	Easy	Hard	Diff
Average	5.88	4.54	1.34*** (9.05)	5.80	4.58	1.22*** (8.35)	5.65	4.83	0.82*** (6.55)	5.78	4.65	1.13*** (8.60)	4.59	3.78	0.81*** (4.41)

Table 6: Test ETF-Mutual Fund Difference between Subject

This table presents independent two-sample t-test results comparing the differences in participants' scores for ETF versus mutual fund texts. The scores are across five dimensions: readability, understandability, comprehensiveness, text accessibility, and investment willingness. The text accessibility score is the average of the readability, understandability, and comprehensiveness scores. The readability score, understandability score, comprehensive score, and text accessibility score range from 1 to 7, where 1 indicates "strongly disagree" and 7 indicates "strongly agree". The investment willingness ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". For the first reading, 153 participants read the ETF text and 152 participants read the mutual fund text. For the second reading, 157 participants read the ETF text and 148 participants read the mutual fund text. The "Diff" column indicates the differences, with corresponding t-statistics in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Score	readability			understandability			comprehensiveness			text accessibility			investment		
first time	ETF	MF	Diff	ETF	MF	Diff	ETF	MF	Diff	ETF	MF	Diff	ETF	MF	Diff
Average	4.74	4.87	-0.13	4.88	5.03	-0.15	5.06	5.21	-0.15	4.89	5.04	-0.15	4.08	4.06	0.02
			(-0.74)			(-0.97)			(-1.16)			(-1.03)			(0.11)
second time	ETF	MF	Diff	ETF	MF	Diff	ETF	MF	Diff	ETF	MF	Diff	ETF	MF	Diff
Average	5.10	5.34	-0.24	5.06	5.34	-0.28*	5.20	5.30	-0.10	5.12	5.33	-0.21	4.02	4.38	-0.36*
			(-1.40)			(-1.74)			(-0.68)			(-1.38)			(-1.90)

Table 7: Pooled Regression for Text Accessibility Scores without Interaction Terms

This table reports the regression coefficients. The dependent variable is the text accessibility score, which ranges from 1 to 7, where 1 indicates 'strongly disagree' and 7 indicates 'strongly agree.' The text accessibility score represents the average of the readability, understandability, and comprehensiveness scores. OFI is the objective financial literacy, a binary variable. if it is equal to 1, it means the participant is financially literate, otherwise, it is 0. SFI is based on participants' self-assessment, with 0 representing "basic" knowledge and 1 representing "intermediate" or "advanced" knowledge. $Easy$ is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. $PageOrder$ is a binary variable, where 1 means this is the text that the participants first read and 0 means the participants' second read. $Demo$ indicates demographic information, including gender, age, education, marriage, income, and employment. The sample size is 610. T-statistics are computed using standard errors clustered by the same funds and are reported in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

	Text Accessibility Scores										
Easy	0.81***				0.81***	0.81***	0.81***	0.81***	0.81***	0.81***	0.81***
	(8.60)				(8.66)	(8.79)	(8.55)	(8.71)	(8.07)	(8.62)	(8.95)
OFI		0.10			0.10		0.10		-0.10	-0.10	-0.08
		(0.96)			(1.11)		(1.12)		(-1.10)	(-1.13)	(-0.81)
SFI			0.53***			0.53***		0.53***	0.56***	0.56***	0.51***
			(6.14)			(6.01)		(5.91)	(6.25)	(6.20)	(5.17)
PageOrder				-0.26**			-0.24**	-0.24**		-0.24**	-0.24**
				(-2.37)			(-2.26)	(-2.21)		(-2.21)	(-2.30)
Demo											controlled

Table 8: Pooled Regression for Text Accessibility Scores with Interaction Terms

This table reports the regression coefficients. The dependent variable is the text accessibility score, which ranges from 1 to 7, where 1 indicates 'strongly disagree' and 7 indicates 'strongly agree.' The text accessibility score represents the average of the readability, understandability, and comprehensiveness scores. OFI is the objective financial literacy, a binary variable. if it is equal to 1, it means the participant is financially literate, otherwise, it is 0. SFI is based on participants' self-assessment, with 0 representing "basic" knowledge and 1 representing "intermediate" or "advanced" knowledge. $Easy$ is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. $PageOrder$ is a binary variable, where 1 means this is the text that the participants first read and 0 means the participants' second read. $Demo$ indicates demographic information, including gender, age, education, marriage, income, and employment. The sample size is 610. T-statistics are computed using standard errors clustered by the same funds and are reported in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

	Text Accessibility Scores						
$Easy$	0.70***	0.81***	0.68***	0.81***	0.72***	0.70***	0.70***
	(5.03)	(6.11)	(4.86)	(6.06)	(4.59)	(4.49)	(4.33)
OFI	-0.01		-0.02		-0.21	-0.23	-0.21
	(-0.05)		(-0.13)		(-1.46)	(-1.60)	(-1.43)
SFI		0.53***		0.53***	0.60***	0.61***	0.56***
		(3.64)		(3.54)	(4.20)	(4.13)	(3.56)
$OFI * Easy$	0.20		0.23		0.23	0.26	0.26
	(1.00)		(1.14)		(1.16)	(1.35)	(1.36)
$SFI * Easy$		0.01		-0.00	-0.07	-0.09	-0.09
		(0.04)		(-0.01)	(-0.36)	(-0.47)	(-0.46)
$PageOrder$			-0.25**	-0.24**		-0.25**	-0.25**
			(-2.36)	(-2.20)		(-2.31)	(-2.41)
$Demo$							controlled

Table 9: Pooled Regression for Investment Willingness without Interaction Terms

This table reports the regression coefficients. The dependent variable is the Investment Willingness, which ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". OFI is the objective financial literacy, a binary variable. if it is equal to 1, it means the participant is financially literate, otherwise, it is 0. SFI is the self-assessment of financial knowledge from the survey participant. If it is equal to 1, it means the participants rate their level of knowledge as intermediate or advanced; if it is 0, they rate the knowledge as basic. $Easy$ is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. $PageOrder$ is a binary variable, where 1 means this is the text that the participants first read and 0 means the participants' second read. $Demo$ represents demographic information, including gender, age, education, marriage, and employment. $FundCharacteristics$ contains the following information: Total annual Fund operating expenses, 1 year Average Annual Total Returns (before tax), 5 year Average Annual Total Returns (before tax) and 10 year Average Annual Total Returns (before tax). The sample size is 610. T-statistics are computed using standard errors clustered by the same funds and are reported in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

	Investment Willingness Scores											
Easy	0.49***				0.49***	0.49***	0.49***	0.49***	0.49***	0.49***	0.49***	0.48***
	(3.62)				(3.65)	(3.69)	(3.66)	(3.70)	(3.68)	(3.69)	(3.87)	(3.81)
OFI		0.26*			0.26**		0.26**		-0.00	-0.00	0.07	0.07
		(1.92)			(2.02)		(2.01)		(-0.02)	(-0.02)	(0.53)	(0.51)
SFI			0.75***			0.75***		0.75***	0.75***	0.75***	0.66***	0.66***
			(6.21)			(6.16)		(6.11)	(6.23)	(6.20)	(4.69)	(4.63)
PageOrder				-0.12			-0.12	-0.12		-0.12	-0.12	-0.06
				(-0.93)			(-0.89)	(-0.88)		(-0.88)	(-0.90)	(-0.51)
Demo											controlled	controlled
FundCharacteristics												controlled

Table 10: Pooled Regression for Investment Willingness with Interaction Terms

This table reports the regression coefficients. The dependent variable is the Investment Willingness, which ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". OFI is the objective financial literacy, a binary variable. if it is equal to 1, it means the participant is financially literate, otherwise, it is 0. SFI is the self-assessment of financial knowledge from the survey participant. If it is equal to 1, it means the participants rate their level of knowledge as intermediate or advanced; if it is 0, they rate the knowledge as basic. $Easy$ is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. $PageOrder$ is a binary variable, where 1 means this is the text that the participants first read and 0 means the participants' second read. $Demo$ represents demographic information, including gender, age, education, marriage, and employment. $FundCharacteristics$ contains the following information: Total annual Fund operating expenses, 1 year Average Annual Total Returns (before tax), 5 year Average Annual Total Returns (before tax) and 10 year Average Annual Total Returns (before tax). The sample size is 610. T-statistics are computed using standard errors clustered by the same funds and are reported in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

	Investment Willingness Scores							
Easy	0.43**	0.56***	0.42**	0.56***	0.49**	0.48**	0.48**	0.48**
	(2.24)	(2.94)	(2.19)	(2.93)	(2.21)	(2.17)	(2.32)	(2.28)
OFI	0.20		0.19		-0.10	-0.11	-0.04	-0.01
	(1.07)		(1.03)		(-0.55)	(-0.59)	(-0.18)	(-0.06)
SFI		0.83***		0.83***	0.87***	0.87***	0.79***	0.76***
		(4.05)		(4.02)	(4.25)	(4.23)	(3.88)	(3.72)
OFI * Easy	0.12		0.13		0.20	0.21	0.21	0.15
	(0.48)		(0.53)		(0.79)	(0.84)	(0.86)	(0.62)
SFI * Easy		-0.17		-0.17	-0.24	-0.25	-0.25	-0.18
		(-0.58)		(-0.60)	(-0.79)	(-0.82)	(-0.88)	(-0.66)
PageOrder			-0.12	-0.12		-0.12	-0.12	-0.07
			(-0.91)	(-0.90)		(-0.94)	(-0.96)	(-0.56)
Demo							controlled	controlled
FundCharacteristics								controlled

Appendix A

This appendix presents the tables for robustness checks, illustrating the regression results across various analyses. These include tests for the impact of reading order and financial literacy on text accessibility and investment willingness. The tables provide detailed coefficients, significance levels, and control variables for each model.

Table A.1: Regression for Text Accessibility Scores based on Page Order

This table reports the regression coefficients. The dependent variable is the text accessibility score, which ranges from 1 to 7, where 1 indicates "strongly disagree" and 7 indicates "strongly agree". OFI is the objective financial literacy, a binary variable. if it is equal to 1, it means the participant is financially literate, otherwise, it is 0. SFI is the self-assessment of financial knowledge from the survey participant. If it is equal to 1, it means the participOants rate their level of knowledge as intermediate or advanced; if it is 0, they rate the knowledge as basic. $Easy$ is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. $Demo$ represents demographic information, including gender, age, education, marriage, and employment. The sample size is 305. T-values are reported in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

	Text Accessibility Scores						
Panel A: First Time							
$Easy$	0.48***			0.29	0.44**	0.16	0.33*
	(3.50)			(1.42)	(2.45)	(0.75)	(1.80)
OFI		0.10		-0.10		-0.15	
		(0.68)		(-0.52)		(-0.73)	
SFI			0.53***		0.47**		0.42**
			(3.82)		(2.50)		(2.10)
$OFI*Easy$				0.35		0.41	
				(1.27)		(1.43)	
$SFI*Easy$					0.13		0.16
					(0.48)		(0.56)
O							
$Demo$						controlled	controlled
Panel B: Second Time							
$Easy$	1.13***			1.08***	1.18***	1.05***	1.16***
	(8.60)			(5.56)	(6.92)	(5.08)	(6.43)
OFI		0.10		0.10		0.06	
		(0.66)		(0.53)		(0.29)	
SFI			0.53***		0.58***		0.50**
			(3.66)		(3.11)		(2.52)
$OFI*Easy$				0.11		0.15	
				(0.40)		(0.54)	
$SFI*Easy$					-0.13		-0.12
					(-0.51)		(-0.42)O
$Demo$						controlled	controlled

41

Table A.2: Regression for Investment Willingness based on Page Order

The dependent variable is the investment willingness, which ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". OFI is the objective financial literacy, a binary variable. if it is equal to 1, it means the participant is financially literate, otherwise, it is 0. SFI is the self-assessment of financial knowledge from the survey participant. If it is equal to 1, it means the participants rate their level of knowledge as intermediate or advanced; if it is 0, they rate the knowledge as basic. $Easy$ is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. $Demo$ represents demographic information, including gender, age, education, marriage, and employment. $FundCharacteristics$ contains the following information: Total annual Fund operating expenses, 1 year Average Annual Total Returns (before tax), 5 year Average Annual Total Returns (before tax) and 10 year Average Annual Total Returns (before tax). The sample size is 305. T-values are reported in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

				Investment Willingness					
Panel A: First Time									
Easy	0.17			-0.09	0.08	-0.16	0.05	-0.17	-0.01
	(0.91)			(-0.33)	(0.34)	(-0.58)	(0.22)	(-0.61)	(-0.05)
OFI		0.33*		0.10		0.11		0.14	
		(1.79)		(0.41)		(0.41)		(0.53)	
SFI			0.79***		0.68***		0.64**		0.62**
			(4.42)		(2.71)		(2.41)		(2.32)
OFI ∗ Easy				0.44		0.48		0.40	
				(1.20)		(1.26)		(1.03)	
SFI ∗ Easy					0.23		0.18		0.26
					(0.66)		(0.48)		(0.69)
Demo						controlled	controlled	controlled	controlled
FundCharacteristics								controlled	controlled
Panel B: Second Time									
Easy	0.81***			0.92***	1.05***	0.87***	0.97***	0.99***	1.03***
	(4.41)			(3.37)	(4.38)	(3.07)	(3.95)	(3.44)	(4.16)
OFI		0.19		0.32		0.33		0.37	
		(0.98)		(1.20)		(1.21)		(1.30)	
SFI			0.70***		0.98***		0.87***		0.84***
			(3.75)		(3.77)		(3.23)		(3.01)
OFI ∗ Easy				-0.18		-0.15		-0.27	
				(-0.47)		(-0.39)		(-0.69)	
SFI ∗ Easy					-0.57		-0.50		-0.51
					(-1.57)		(-1.33)		(-1.33)
Demo						controlled	controlled	controlled	controlled
FundCharacteristics								controlled	controlled

Table A.3: Regression for Scores based on Financial Literacy

The dependent variable is the text accessibility score, which ranges from 1 to 7, where 1 indicates "strongly disagree" and 7 indicates "strongly agree". *Easy* is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. *PageOrder* is a binary variable, where 1 means this is the text that the participants first read and 0 means the participants' second read. *Demo* represents demographic information, including gender, age, education, marriage, income, and employment. The sample size is 330 in Panel A, 280 in Panel B, 262 in Panel C, and 348 in Panel D. T-values are reported in parentheses. Significance levels are indicated as follows: $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

	Text Accessibility Scores			
Panel A: Financial Literate				
Easy	0.90***		0.91***	0.91***
	(6.90)		(7.00)	(7.35)
PageOrder		-0.26*	-0.28**	-0.28**
		(-1.84)	(-2.18)	(-2.29)
Demo				controlled
Panel B: Financial Illiterate				
Easy	0.70***		0.68***	0.68***
	(4.92)		(4.82)	(4.80)
PageOrder		-0.26*	-0.21	-0.21
		(-1.74)	(-1.46)	(-1.45)
Demo				controlled
Panel C: High Financial Knowledge				
Easy	0.81***		0.84***	0.80***
	(6.19)		(6.14)	(6.18)
PageOrder		-0.26*	-0.23*	-0.23*
		(-1.84)	(-1.73)	(-1.74)
Demo				controlled
Panel D: Low Financial Knowledge				
Easy	0.81***		0.81***	0.81***
	(6.12)		(6.14)	(6.23)
PageOrder		-0.26*	-0.26*	-0.26**
		(-1.85)	(-1.95)	(-1.98)
Demo				controlled

Table A.4: Regression for Investment Willingness based on Financial Literacy

The dependent variable is the investment willingness, which ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest. OFI is the financial literacy, a binary variable. if it is equal to 1, it means the participant is financially literate, otherwise, it is 0. $Easy$ is a binary variable, if it is equal to 1, it means that participants read an easy text, and if it is equal to 0, the participants read a hard text. $PageOrder$ is a binary variable, where 1 means this is the text that the participants first read and 0 means the participants' second read. $Demo$ represents demographic information, including gender, age, education, marriage, and employment. $FundCharacteristics$ contains the following information: Total annual Fund operating expenses, 1 year Average Annual Total Returns (before tax), 5 year Average Annual Total Returns (before tax) and 10 year Average Annual Total Returns (before tax). The sample size is 330 in Panel A, 280 in Panel B, 262 in Panel C, and 348 in Panel D. T-values are reported in parentheses. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

	Investment Willingness			
Panel A: Financial Literate				
$Easy$	0.55***		0.55***	0.55***	0.50***
	(2.99)		(3.00)	(3.11)	(2.76)
$PageOrder$		-0.06	-0.08	-0.08	-0.02
		(-0.33)	(-0.42)	(-0.44)	(-0.11)
$Demo$				controlled	controlled
$FundCharacteristics$					controlled
Panel B: Financial Illiterate				
$Easy$	0.43**		0.42**	0.42**	0.43**
	(2.32)		(2.25)	(2.31)	(2.35)
$PageOrder$		-0.20	-0.17	-0.17	-0.11
		(-1.08)	(-0.92)	(-0.95)	(-0.60)
$Demo$				controlled	controlled
$FundCharacteristics$					controlled
Panel C: High Financial Knowledge				
$Easy$	0.40**		0.39**	0.39**	0.38**
	(2.01)		(1.99)	(2.07)	(1.96)
$PageOrder$		-0.08	-0.06	-0.06	-0.01
		(-0.38)	(-0.31)	(-0.32)	(-0.03)
$Demo$				controlled	controlled
$FundCharacteristics$					controlled
Panel D: Low Financial Knowledge				
$Easy$	0.56**		0.56***	0.56***	0.54***
	(3.40)		(3.40)	(3.44)	(3.27)
$PageOrder$		-0.16	-0.16	-0.16	-0.13
		(-0.96)	(-0.97)	(-0.98)	(-0.80)
$Demo$				controlled	controlled
$FundCharacteristics$					controlled

Appendix B

This appendix provides detailed regression analyses and supplementary materials to support the findings presented in the main text. It includes regression tables examining the relationship between text complexity measures and dependent variables such as readability, understandability, comprehensiveness, and investment willingness.

Table B.1 presents the regression results examining the relationship between various text complexity measures and five dependent variables: readability, understandability, comprehensiveness, text accessibility score, and investment willingness. The Number of Words shows a small but significant positive relationship with all dependent variables, suggesting that longer texts are generally perceived more favorably. In contrast, the Number of Tokens does not exhibit a significant effect on any of the dependent variables. For Flesch Reading Ease, the relationship is positive across all dependent variables, consistent with our t-test analysis. This indicates that as the text becomes easier to read, participants demonstrate better understanding, readability, and comprehensiveness scores, and show greater willingness to invest. Conversely, other complexity measures such as the Flesch-Kincaid Grade, Gunning Fog, and SMOG Index have negative coefficients, implying that more difficult texts result in lower readability, understandability, comprehensiveness, and investment willingness scores, which also corroborates our main findings. Finally, Reading Time does not have a significant influence on any of the dependent variables.

[here insert Table B.1]

Next, we examine how the complexity of the fund prospectus affects investors' willingness to invest in the fund. Table B.2 presents the regression coefficients, showing that all three independent variables—readability, understandability, and comprehensiveness—have a positive and significant effect on investment willingness, regardless of whether demographic information is controlled for. These findings suggest that as the texts become clearer and more comprehensive, participants' willingness to invest increases.

Table B.1: Regression for Text Complexity Measures

This table presents regression coefficients with t-statistics in parentheses. The dependent variables include readability scores, understandability scores, comprehensiveness scores, text accessibility scores (the average of the previous three scores, ranging from 1 to 7, where 1 indicates "strongly disagree" and 7 indicates "strongly agree"), and investment willingness (ranging from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest"). The independent variables are text complexity metrics. Significance levels are indicated as follows: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

	Readability	Understandability	Comprehensiveness	Text Accessibility	Investment
Number of Words	0.01***	0.01***	0.00**	0.01***	0.01***
	(3.65)	(3.30)	(2.16)	(3.34)	(2.63)
Number of Tokens	0.00	0.00	0.01	0.00	0.00**
	(1.19)	(0.98)	(0.60)	(1.02)	(1.99)
Flesch Reading Ease	0.04***	0.04***	0.02***	0.03***	0.02***
	(6.79)	(6.80)	(4.34)	(6.57)	(2.78)
Flesch Kincaid Grade	-0.22***	-0.20***	-0.11***	-0.18***	-0.10**
	(-5.98)	(-5.97)	(-3.86)	(-5.79)	(-2.34)
Gunning Fog	-0.23***	-0.20***	-0.11***	-0.18***	-0.12***
	(-6.18)	(-5.69)	(-3.75)	(-5.73)	(-2.84)
Smog Index	-0.26***	-0.23***	-0.12***	-0.20***	-0.11**
	(-5.66)	(-5.42)	(-3.44)	(-5.32)	(-2.25)
Automated Readability Index	-0.19***	-0.17***	-0.10***	-0.15***	-0.08**
	(-6.08)	(-5.98)	(-4.01)	(-5.88)	(-2.27)
Coleman Liau Index	-0.21***	-0.20***	-0.11***	-0.17***	-0.10***
	(-7.11)	(-7.03)	(-4.69)	(-6.90)	(-2.98)
Linsear Write Formula	-0.03	-0.02	-0.01	-0.02	0.00
	(-1.50)	(-1.00)	(-0.42)	(-1.10)	(0.23)
Dale Chall Readability Score	-0.47***	-0.42***	-0.23***	-0.37***	-0.25***
	(-8.12)	(-7.79)	(-4.97)	(-7.67)	(-3.83)
Text Standard	-0.18***	-0.15***	-0.09***	-0.14***	-0.08***
	(-6.63)	(-6.05)	(-4.07)	(-6.14)	(-2.71)
Spache Readability	-0.73***	-0.65***	-0.35***	-0.58***	-0.35***
	(-6.65)	(-6.34)	(-4.06)	(-6.26)	(-2.83)
Mcalpine Eflaw	0.01	0.01	0.01	0.01	0.01
	(0.58)	(0.74)	(0.45)	(0.64)	(0.68)
Reading Time	-0.03	-0.04	-0.02	-0.03	0.02
	(-1.27)	(-1.61)	(-1.11)	(-1.44)	(0.60)

Table B.2: Impact on Investment Willingness

This table presents regression coefficients with t-statistics in parentheses. The dependent variable is the Investment Willingness, which ranges from 1 to 7, where 1 indicates "extremely unlikely to invest" and 7 indicates "extremely likely to invest". The independent variables are the scores of readability, understandability, and comprehensiveness. *Demo* refers to demographic information, which is included as a control variable. The sample size is 610. Significance levels are indicated as follows: $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

	Investment Willingness	
readability	0.26***	0.22***
	(3.60)	(3.05)
understandability	0.17**	0.16*
	(1.96)	(1.88)
comprehensiveness	0.39***	0.44***
	(5.39)	(5.92)
Demo		Controlled

Appendix C

In this appendix, we include screenshots from the experiment design to provide a visual summary of the materials and methods used.

SPDR Portfolio S&P 500 Value ETF

The SPDR® Portfolio S&P 500® Value ETF (SPYV) aims to mirror the performance, before fees and expenses, of the S&P 500 Value Index, which tracks U.S. equity securities of large capitalizations with value characteristics. The ETF employs a sampling strategy, meaning it may not hold all securities in the index but selects a subset to match the index's risk and return profile. The fund's management fees are notably low at 0.04%, with no distribution and service (12b-1) fees, reflecting a total annual operating expense of 0.04%.

The ETF's portfolio turnover rate for the last fiscal year was 27%, indicating moderate trading activity which could influence transaction costs and tax implications in taxable accounts. The fund predominantly invests in securities comprising the index, dedicating at least 80% of its assets to this purpose, and may also invest in cash equivalents or use futures contracts for index performance alignment and cash flow management.

The S&P 500 Value Index focuses on large-cap U.S. equities listed on major exchanges, selecting stocks based on book value to price ratio, earnings to price ratio, and sales to price ratio. As of August 31, 2023, the index included 403 stocks, with significant allocations in healthcare and technology sectors as of July 31, 2023.

Investing in the ETF involves risks such as market volatility, equity investing risks, value stock risks, and large-capitalization securities risks. The fund's performance can fluctuate due to various factors, including economic conditions and market changes. The ETF's past performance, including annual total returns and quarterly highs and lows, provides insight into its volatility and potential for investment loss.

The ETF is managed by SSGA Funds Management, Inc., with Karl Schneider and Mark Krivitsky as the primary portfolio managers. Shares of the ETF are traded on the NYSE Arca and other platforms, with prices subject to market fluctuations, potentially trading at a premium or discount to the net asset value (NAV).

For tax purposes, the fund's distributions may be taxed as ordinary income, qualified dividend income, or capital gains, with specific tax implications for investments through tax-advantaged arrangements. The ETF may also involve payments to brokers or financial intermediaries for sales and promotional activities, potentially creating a conflict of interest.

Figure C.1: Summarized "hard" Fund Prospectus

Fidelity Stock Selector Large Cap Value Fund

The Fidelity® Stock Selector Large Cap Value Fund (FSLVX) is designed for investors looking to grow their money over time by investing in large companies that are considered to be undervalued. These companies are typically larger in size, similar to those found in the Russell 1000® Index or the S&P 500® Index. The fund aims to pick stocks that are priced lower than what the fund managers believe they are worth based on various factors like the company's earnings, cash flow, and growth potential.

Investors don't pay any direct fees when they buy or sell shares of the fund, but there are annual operating expenses. These expenses amount to 0.83% of your investment each year. This means if you invest $10,000, you would pay about $83 in fees annually. The management fee, which is part of these expenses, can vary up to 0.20% higher or lower than 0.60%, depending on how well the fund performs compared to the Russell 1000® Value Index.

The fund's portfolio turnover rate was 67% last year, indicating it replaced about two-thirds of its investments. This turnover can lead to higher transaction costs and taxes, especially if you're investing through a taxable account.

Investing in this fund comes with risks, including stock market volatility, the potential for investments in foreign markets to be more unstable, and the chance that "value" stocks may not rebound as expected. The fund's performance can fluctuate, as shown by its yearly returns ranging from a high of 32.24% in 2013 to a low of -9.18% in 2018.

You can buy or sell shares of the fund through Fidelity's website, by phone, or by mail, with no minimum purchase required. Keep in mind, any earnings from the fund are subject to federal income tax, unless you're investing through a tax-advantaged account like a 401(k) or IRA.

Lastly, the fund may pay financial intermediaries for sales and services, which could influence their recommendation to choose this fund over others. Always ask your investment professional for more information if you're concerned about potential conflicts of interest.

Figure C.2: Summarized "easy" Fund Prospectus

2. Please indicate how much you agree or disagree with the following statement:

	strongly disagree	disagree	somewhat disagree	neither agree or disagree	somewhat agree	agree	strong agree
The language used in the content is clear and easy to understand.	○	○	○	○	○	○	○
The sentence structure makes it easy for me to follow the information.	○	○	○	○	○	○	○
I swim across the Atlantic Ocean to get to work every day. This is an attention check.	○	○	○	○	○	○	○
The text uses straightforward words and simple sentences, making it easy to comprehend.	○	○	○	○	○	○	○

3. Please indicate how much you agree or disagree with the following statement:

	strongly disagree	disagree	somewhat disagree	neither agree or disagree	somewhat agree	agree	strong agree
The information presented is logically organized and flows well.	○	○	○	○	○	○	○
I can easily grasp the main concepts presented in the content.	○	○	○	○	○	○	○
The content avoids unnecessary complexity and is straightforward.	○	○	○	○	○	○	○

Figure C.3: Survey Questions on Readability and Understandability with Embedded Attention Check

4. Please indicate how much you agree or disagree with the following statement:

	strongly disagree	disagree	somewhat disagree	neither agree or disagree	somewhat agree	agree	strong agree
The content covers all the necessary information on the topic.	○	○	○	○	○	○	○
Important aspects of the topic are adequately addressed and easily accessible.	○	○	○	○	○	○	○
The material is comprehensive, meeting expectations for a thorough understanding.	○	○	○	○	○	○	○

5. Please rate your likelihood on a scale from 1 to 7

	Extremely Unlikely	Unlikely	Somewhat Unlikely	Neutral	Somewhat Likely	Likely	Extremely Likely
Based on the information provided about the investment fund, how likely are you to invest in this fund?	○	○	○	○	○	○	○

Figure C.4: Survey Questions on Comprehensiveness and Investment Willingness

16. How would you rate your level of knowledge about various investment options?

○ Basic
○ Intermediate
○ Advanced
○ Expert

17. Have you had any prior experience with investing?

○ Yes, some experience
○ No, but I'm interested in starting
○ No, and I have no interest in investing

Figure C.5: Survey Questions on Investment Knowledge and Experience

Appendix D

The financial literacy questions presented above are taken from the work of Lusardi and Mitchell (2011) and Fernandes et al. (2014), which are well-regarded in the field of financial education research.

1. Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow?

   - More than $102
   - Exactly $102
   - Less than $102
   - Do not know
   - Refuse to answer

2. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account?

   - More than today
   - Exactly the same
   - Less than today
   - Do not know
   - Refuse to answer

3. Please tell me whether this statement is true or false: "Buying a single company's stock usually provides a safer return than a stock mutual fund."

   - True
   - False

- Do not know

- Refuse to answer

4. Considering a long time period (for example, 10 or 20 years), which asset described below normally gives the highest return?

    - Savings accounts

    - Stocks

    - Bonds

    - Do not know

    - Refuse to answer

5. Normally, which asset described below displays the highest fluctuations over time?

    - Savings accounts

    - Stocks

    - Bonds

    - Do not know

    - Refuse to answer

6. When an investor spreads his money among different assets, does the risk of losing a lot of money:

    - Increase

    - Decrease

    - Stay the same

    - Do not know

    - Refuse to answer

Declaration

We acknowledge the use of ChatGPT[https://chat.openai.com/] and Grammarly in helping us review our writing at the final stage.

We used the following prompt: "Is this paragraph in line with academic writing? If not, what is the revised version?".

We reviewed the feedback generated by ChatGPT and Grammarly critically and, based on this, revised our writing using the output or persists of it.