THE DEBIASING EFFECTS OF ALGORITHMIC ADVICE: DOES TIMING MATTER?

by

ALYSSA JOAQUIN

(Under the Direction of Elena Karahanna)

ABSTRACT

Commonly found in human decision-making, cognitive or heuristic biases are mental shortcuts that may help individuals make more efficient decisions but often result in errors. We propose that Human-AI decision systems should be designed utilizing knowledge of how to mitigate cognitive biases and thus improve human decision making. Our research examines whether algorithmic advice can mitigate human biases and how the timing of such advice influences the extent of bias mitigation. We focus on conservatism bias, a bias which causes individuals to underreact to new information, in financial decision making. The experiment compared three conditions, varying the timing of algorithmic advice. The results indicate that timing of algorithmic advice matters – individuals who received algorithmic advice prior to making a decision exhibited less conservatism bias than those who received algorithmic advice after making an initial decision and those who received none at all.

INDEX WORDS: Human-AI Decision Making, Cognitive Biases, Conservatism Bias, Algorithmic Advice, Investment Decision Making

# THE DEBIASING EFFECTS OF ALGORITHMIC ADVICE: DOES TIMING MATTER?

by

ALYSSA JOAQUIN

B.A., The University of Georgia, 2022

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2023

THE DEBIASING EFFECTS OF ALGORITHMIC ADVICE: DOES TIMING MATTER?

by

ALYSSA JOAQUIN

Major Professor: Elena Karahanna

Committee: Frederick Maier
Aaron Schecter

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2023

## ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Dr. K, for her guidance and support throughout this entire process. I could not have made it through without her kindness and patience, as well as her knowledge and expertise. I also want to thank Dr. Maier and Dr. Schecter for serving on my committee. I am grateful for their involvement and feedback.

Also, I would like to thank my friends and family for their constant support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

# CHAPTER 1

# INTRODUCTION

While research on AI has focused on detecting and mitigating AI biases in making decisions and recommendations (Orphanou et al., 2022; Balayn et al., 2021; Harris, 2020) less attention has been devoted to how algorithmic advice can mitigate human decision biases. Heuristic biases are relatively prevalent in human decision making (Tversky & Kahneman, 1974). Heuristics can help individuals reduce cognitive load and make more efficient decisions. However, while heuristics seek to reduce the complexity of decision-making tasks, they also often result in biases, lower quality decisions and suboptimal outcomes. Heuristic biases are not always consequential. Rather, their impact varies based on the importance of the decision-making task and can have serious negative consequences for certain types of decisions such as medical diagnosis and investment decisions. Despite such consequences, cognitive biases persist even for experts and experienced researchers in such professional settings across all domains (Barberis & Thaler, 2002; Kahneman et al., 1982 ). Thus, it is important to examine debiasing techniques in such contexts to mitigate cognitive biases and, thus, improve decision outcomes.

While a range of debiasing techniques have been examined, such as leveraging certain biases to nudge people to make healthier dietary decisions, smarter investment decisions, or utilize decision aids to arrive at more objective decisions (Zhang et al., 2015; Lee et al., 2011; Solomon, 2014; Zhao et al., 2015), with the prevalence of AI and algorithmic advice, there is the potential to use AI to help mitigate human decision biases. Algorithmic advice is being leveraged in a wide variety of contexts to augment human decision making. We suggest that such systems should be designed bearing in mind the cognitive biases that are prevalent in the specific context and how to mitigate these.

As such, our research objective in this thesis is *to examine whether algorithmic advice can mitigate human biases and how the timing of such advice influences the extent of bias mitigation*. Specifically, we examine (a) the impact of algorithmic advice on mitigating cognitive decision biases, and (b) how the timing of providing such advice might have different mitigating effects. Research on the timing of algorithmic advice suggests that individuals who receive advice after making an independent assessment are more likely to overweight their initial opinion, which serves as an anchor, and therefore discount the advice (Yaniv & Kleinberger, 2000). Given that the provided algorithmic advice is as good as or better than an individual's assessment, delivering the advice prior to assessment, so that they anchor on it instead, will result in greater consideration/integration of the advice and therefore, greater debiasing and accuracy.

Our research focuses on a specific type of cognitive bias that is prevalent in financial decision making: conservatism bias. Conservatism bias occurs when individuals underweight new information in the revision of their opinions (Edwards, 1968). In describing conservatism bias, Edwards compares human opinion change to Bayes' theorem, a formal rule on the updating of opinion with the addition of evidence. He demonstrates that humans update their opinions proportionally but insufficiently to Bayes's theorem. Generally, conservatism bias causes an individual to underreact to new evidence (Barberis et al., 1998). This has implications for decision making in many domains where people must incorporate evidence into their critical thinking and decision-making process. Specifically in financial decision making, evidence suggests that individuals underreact to new information or changes in the stock market leading to inaccurate judgements in evaluating potential risks and returns (Zhang et al., 2015). Evidence of conservatism bias is also provided by macro-level studies. Specifically, evidence points to the stock market underreacting to information such as dividend omissions, initiations, or an earnings report, thus exhibiting conservatism bias (Pompian, 2011; Montier, 2002).

Our study examines how algorithmic advice may mitigate conservatism bias in investment decisions. Anchoring occurs when individuals make estimates by heavily weighting their starting value or anchor, adjusting insufficiently to new evidence, resulting in final estimates that are biased towards the initial value (Tversky & Kahneman, 1974). We take the approach that, because of anchoring effects, the timing of algorithmic advice matters to the debiasing process. If an individual makes their assessment first – and based on the literature we expect that this assessment will exhibit conservatism bias - they will tend to anchor on that assessment and discount the algorithmic advice, a concept called egocentric bias. This will result in the algorithmic advice having a lesser debiasing effect. However, if the algorithmic advice is provided first, they would tend to anchor on the algorithmic advice and it will carry more weight when making their own assessment, thus mitigating conservatism bias to a greater extent. Therefore, we expect that while AI advice will have a debiasing effect, providing an individual AI advice before making their own initial assessment will have a greater debiasing effect because it will cause them to anchor on the advice, giving a better baseline, thus mitigating conservatism bias, even if they still under-adjust to new information.

To test our hypotheses, we conducted a lab experiment with three conditions: *independent* (control) where participants received no algorithmic advice, *dependent* where participants received algorithmic advice prior to making a decision, and *independent-then-revise* where participants first made an initial decision, then received algorithmic advice and made a final decision. Conservatism bias was evident in our investment decision-making task, validating our assumption of cognitive bias in this setting. Our hypothesis-testing results showed that algorithmic advice did improve accuracy, but only for the dependent condition where algorithmic advice was provided prior to making own decision. The difference between the independent and the independent-then-revise conditions was not significant. The timing of advice played a role in the debiasing effects – participants in the dependent condition experienced a greater debiasing effect than participants in the independent-then-revise condition. Participants in the independent-then-revise condition incorporated the advice into their final estimates, however, the average influence of advice was lower than expected due to a strong anchoring on their initial estimates. Thus, our research shows that accurate algorithmic advice can mitigate human biases, but the timing of such advice plays a significant role in the debiasing effects.

Our research contributes to the literature of Human-AI augmented decision making. While there is a significant amount of research that examines reducing biases in AI and algorithms, few studies exist on using artificial intelligence to remediate human biases (e.g., Lee et al., 2022; Zhang et al., 2015) and none that focuses on the timing of such advice. Specifically, prior research by Lee et al. (2022) examines whether AI advice can increase decision accuracy and reduce anchoring bias in decision-making. They found that greater acceptance of AI advice results in higher decision accuracy. However, their study indicated that decision-makers frequently exhibit egocentric bias, and therefore may not fully utilize the AI advice (Lee et al., 2022). Further, they only investigate anchoring bias, and do not explicitly discuss the timing of the AI advice itself. Research by Zhang et al. discusses remediating conservatism and loss aversion in an investment decision by providing individuals with decision aids that can help them with their calculations; however, it does not explore the use of trustworthy AI advice. Our study complements this prior work by examining both the debiasing effect of algorithmic advice and the effect of advice timing on such debiasing.

# CHAPTER 2

# LITERATURE REVIEW

Our work examines how algorithmic advice and its timing mitigate conservatism bias in an investment context. Therefore, we review literature on conservatism bias and conservatism bias in financial settings to understand the bias that we aim to mitigate. We also review literature on anchoring and advice, and the effect of timing of advice on advice taking, because they inform the theoretical arguments and mechanisms via which we expect algorithmic advice to mitigate conservatism bias.

## 2.1 Conservatism Bias

Across many fields there is much interest and research surrounding the phenomenon of how individuals update their opinions to incorporate new information. In probability and statistics, Bayes's theorem defines how to update a belief (a probability of something), given some observed evidence. Edwards (1968) states that human opinion change follows the same order of Bayes' theorem proportionally, just in a smaller amount. This insufficient updating can be described by the term *conservatism bias*, in which individuals are slow to change their prior beliefs to take in new information (Edwards, 1968), often leading to inaccurate judgements (Zhang et al., 2015). Due to the overuse of the center of the probability scale, conservatism can be observed as the underestimation of high probabilities and overestimation of low probabilities (Zhang et al., 2015).

In their research on the weighting of evidence, Griffin and Tversky (1992) argue that people focus too heavily on the strength (i.e., salience) of evidence, and not enough of the statistical weight (i.e., statistical informativeness) of evidence. One potential explanation for conservatism is that the belief updating process for new evidence is cognitively costly, especially for information in a statistical form, illustrating Griffin and Tversky's principle of under reliance on the weight of information (Hirshleifer, 2001).

### 2.1.2 Conservatism Bias in Behavioral Finance: Investment Decisions

Conservatism bias has many implications for investment decision making. For example, individuals who exhibit conservatism bias and underreact to evidence may tend to overlook information in a quarterly earnings announcement and fail to accurately adjust their valuation of shares (Barberis et al., 1998). For this scenario, Barberis et al. note that Griffin and Tversky's theory implies that individuals may insufficiently react to the isolated earnings announcement because a single earnings number appears "weakly informative" and low in strength. Similarly, an investor may exhibit conservatism when adjusting his beliefs to a dividend cut or share repurchase announcement (Barberis et al., 1998).

Hirshleifer (2001) notes that individuals are typically more prone to bias in valuing investments with limited information. However, biases can persist in a broad range of situations, and it is important to recognize that in some cases, certain biases are nearly impossible to be immune to (Hirshleifer, 2001). Some economists believe that experts, such as investment bank traders will make fewer errors due to cognitive biases (Barberis & Thaler, 2002) given their expertise and incentives. In contrast, Barberis and Thaler argue that expertise can enable cognitive biases rather than mitigate them, especially because expert predictions often receive little feedback. Prior research has examined the use of decision aids (such as an Excel spreadsheet with macros to enable calculations) for mediating cognitive biases in investment decision making (Zhang et al., 2015). However, Zhang et al. note that the results of these studies demonstrate that even experts exhibit conservatism bias. Conservatism has also been observed on a macro level, Montier (2002) describes that the stock market may underreact to key information, such as earnings reports or dividend omissions.

## 2.2 Anchoring & Advice

Previous work on anchoring and advice by Tversky and Kahneman (1974) centers around the phenomenon of adjustment and anchoring, illustrating that people's reliance on heuristics for decreasing the complexity of tasks may also lead to systematic errors. They describe situations in which individuals make an initial estimate of some specified value and then adjust this value based on additional evidence (i.e., advice), presenting a final adjusted answer (Tversky &Kahneman, 1974). Different individuals propose different initial estimates which they remain biased, despite receiving advice and undergoing an adjustment period for revising their estimate, resulting in an insufficient adjustment (Tversky & Kahneman, 1974).

## 2.3 Timing of Advice

There is considerable research on the effect of advice on individuals' decisions (e.g., Rader et al., 2015; Koehler & Beauregard, 2006; Sniezek & Buckley, 1995; Yaniv & Choshen-Hillel, 2012). Some of this work, which we review below, examines how the *timing* of advice influences decision making and decision outcomes. Most of these timing of advice studies examine the effects on decision outcomes of whether advice is provided *before or after an individual has already made a preliminary decision* (Rader et al., 2015; Koehler & Beauregard, 2006; Sniezek & Buckley, 1995; Yaniv & Choshen-Hillel, 2012 ). Research often examines the influence of advice (i.e., how much advice participants take) and the accuracy of decisions given the varied timing of advice (Rader et al., 2015; Koehler & Beauregard, 2006; Sniezek & Buckley, 1995; Yaniv & Choshen-Hillel, 2012). Rader et al. (2015) further make a distinction between anchoring studies which usually present advice of an extreme value, and advice taking studies which typically provide advice that is more central (median) to what someone would have said independently (Rader et al., 2015). Finally, research has examined the potential mediating effects of confidence in advice (Rader et al., 2015; Koehler & Beauregard, 2006; Yaniv & Choshen-Hillel, 2012). Table 1 provides a summary of these studies which we discuss below.

Building on anchoring and advice, Rader et al. (2015) explore the timing of when advice is received to determine if it changes how much people use the advice and its effects on final judgement accuracy. They conducted five studies comparing participants who formed their own opinion independently before receiving advice and then revised their estimate (*independent-then-revise*) to participants who received advice prior to making an estimate (*dependent*) and to participants who received no advice at all (*independent*). They examined the role of central advice in an attempt to better represent a normal advice taking scenario.

In the first study, they explore how estimates differ for individuals in the dependent condition versus the independent-then-revise condition. The results of Study 1 indicate that given median advice, dependent estimates were further from the advice on average than the final independent-then-revised estimates, demonstrating a "push-away" effect (Rader et al., 2015). Rader et al. suggest that when individuals in the dependent sequence are given advice, they ask a comparative question such as "is the answer higher or lower" which then pushes them in a chosen direction away from the advice. Study 2 examined whether confidence mediates the push-away effect and added an *independent* (control) condition with no advice for comparison. The results show that decision-makers who lack confidence in the advice are more likely to exhibit a push-away effect. Study 3 implemented a verbal protocol task where participants were asked to talk through their decision-making process. The data from the verbal protocol task further illustrate that individuals in the dependent condition are more likely to exhibit the push-away effect than the independent-then-revise participants.

Using a wider range of advice centrality, Study 4 examined the impact of the timing of advice on accuracy. They found that the dependent and independent-then-revise groups had similar accuracy for a wide range of advice, concluding that both groups are better than the independent estimate with no advice at all. Study 5 compared the dependent condition to the traditional anchoring paradigm, designing the dependent sequence as the standard anchoring paradigm in a new *dependent-comparative* condition. The *dependent-comparative* condition took the same form as the dependent condition but also included a comparative question "is the answer higher or lower?" in addition to the advice. The results of the *dependent-comparative* sequence were the same as the normal dependent sequence. Overall, Rader et al. found that in the context of median advice, decision-makers in the *independent-then-revise* condition tend to take more advice than those in the *dependent* condition but that in general the two groups had similar accuracy.

Many studies on advice-taking utilize similar experiments to the dependent and independent-then-revise paradigm that Rader et al. propose, though they often label the advice-timing conditions differently and often find different results. Specifically, in contrast to their own findings of a push-away effect, Rader et al. (2015) acknowledge that most anchoring research suggests that individuals in the dependent advice group would take more advice than people in the independent-then-revise group (Rader et al., 2015; Koehler & Beauregard, 2006; Sniezek & Buckley, 1995; Yaniv & Choshen-Hillel, 2012). For example, Sniezek and Buckley (1995) describe their Judge-Advisor System (JAS), which consists of a decision-maker (called judge in their study) who makes a final judgement after receiving recommendations or advice from one or more advisors. Their three advice-timing conditions include (a) independent, in which decision-makers make an initial judgement before receiving advice and then making a final judgement (corresponds to independent-then-revise), (b) cued, in which the decision-maker receives advice before making a judgement (corresponds to dependent), and (c) dependent, in which the decision-maker can only make decisions based on the advice and has no access to the question. The dependent condition attempts to restrain decision-makers from accessing internal information and is noted as being artificial because real decision-making scenarios generally give decision-makers access to the question they are being asked (Sniezek & Buckley, 1995). They also explored scenarios in which decision-makers receive conflicting advice from multiple advisors, however this is not relevant to our argument, so we only discuss their experimental findings under the no conflict condition.

In the experiment, participants completed a two-option choice task of differing levels of difficulty with content found in *Business Week*, *The Wall Street Journal*, and *The New York Times*. For example, "In the United States, 41% of all money spent on food is spent in (A) supermarkets or (B) restaurants?" (Sniezek & Buckley, 1995). Participants had to select an option, as well as a confidence interval of probability (from .5 to 1 in increments of .05). Sniezek and Buckley measured amount of advice-taking across the different conditions, observing the proportion of decision-makers who "matched" the advice. They found that individuals in the dependent condition showed a significantly larger percentage of matching than cued or independent (Sniezek & Buckley, 1995). Individuals in the cued condition also took more advice than the independent condition, however the difference was not significant. Their results suggest that the decision-maker's final choice accuracy was highest for the independent condition, followed by cued, and lowest for the dependent condition. Sniezek and Buckley attribute this to the fact that independent decision-makers experience additional information processing (Sniezek & Buckley, 1995). However, participants were randomly assigned to the role of decision-maker or advisor and therefore advisors had no more expertise than the decision-makers themselves. In the case of expert advisors, one may expect the dependent or cued group to have higher accuracy than the independent group by taking more advice.

Yaniv and Choshen-Hillel (2011) explored how individuals who form opinions prior to receiving advice are overly conservative in their belief updating, which results in suboptimal accuracy improvement. Similar to varying the timing of advice, they manipulate the timing of formation of opinions in their experiment involving calorie estimation for target foods (Yaniv and Choshen-Hillel, 2011). In the *full-view* condition, participants were able to generate initial estimates, before receiving advice and then making a final estimate (corresponding to the independent-then-revise condition). In contrast, participants in the *blindfold* condition were provided with advice but were not permitted to form initial opinions and also were not informed of the target food for which they were estimating calories. They performed three studies; however, the second and third studies manipulated the *full-view condition* so that participants did not give an initial estimate. Therefore, this is equivalent to the dependent condition.

Yaniv and Choshen-Hillel found that the blindfolded decision-makers gave estimates closer to the advice and of higher accuracy. However, they note receiving advice is beneficial to accuracy in general, which can be demonstrated through the results of *full-view* condition where decision-makers final estimates were more accurate than initial estimates. They also measured egocentric discounting by finding the percentage of participants whose final estimates did not change from their initial estimates. They found a much higher percentage of participants in the *full-view* condition than the *blindfold*. Although the *blindfold* condition allowed decision-makers to restrain their own personal opinions and take more advice, it is not a realistic condition for real-world decision making because decision-makers will typically know the questions they are being asked and will be provided information about it.

Koehler and Beauregard (2006) also explore the timing of exposure to advice. They describe the group of individuals who make their own estimate before receiving advice as "unexposed advisees" (independent-then-revise) and the group who receive advice before making their own estimate as "exposed advisees" (dependent). They distinguish between deliberate use of another person's estimate (i.e., advice) and the contaminating influence of exposure to another person's estimate, where it negatively affects an individual's ability to generate his or her own independent estimate (Koehler & Beauregard, 2006), In the case of contaminating influence, they also discuss an "illusion of confirmation" which they define as an increased confidence in the accuracy of advice.

In their first two experiments, participants in the *unexposed* condition did not make a final estimate after seeing the advice, therefore we only focus on their third experiment. In Experiment 3, they found that exposed advisees estimates were closer to the advice than the estimates of unexposed advisees (Koehler & Beauregard, 2006). They also calculated a measure for weight of self, which attempted to quantify the weight placed on their own opinion compared to the advice. Nonetheless, despite focusing on the negative influence of the contaminating effect, Koehler and Beauregard note a potential area for future research that explores a potential benefit – if individuals tend to place a disproportionately large weight on their own judgements relative to others' opinions (conservatism bias/egocentric weighting bias), the contaminating effect might serve as a positive influence to mitigate this bias and result in judgements of higher accuracy.

Finally, Yin et al. (2020) also explored timing of algorithmic advice in clinical decision making using a sample of physicians. Though they hypothesize that physicians who make an initial diagnosis (dependent condition) prior to receiving AI advice will have lower advice taking than physicians who receive AI advice and then make a diagnosis (dependent condition), the study was research-in-progress and did not report any empirical results. We present it for completeness but given the lack of empirical evidence, we exclude it from Tables 1 and 2.

*Table 1: Summary of Studies on Timing of Advice*

| **Study** | **Objective** | **Experiment Conditions** | **Outcome Measure(s)** | **Decision Task** | **Findings** |
|---|---|---|---|---|---|
| Koehler & Beauregard (2006) Experiment 3 | Examine influence of exposure to advisor's estimate on advisee estimates and confidence | Unexposed Advisee, Exposed Advisee | Influence of Advice (Distance), Confidence Assessments | Estimating the year that a historical event occurred | Exposed advisees gave estimates closer to advisor and expressed greater confidence in accuracy of advisor's estimates than unexposed advisees |
| Rader et al. (2015) Study 1 | Examine how estimates differ for dependent and independent-then-revise sequences | 3x3 Advice timing x Advice Centrality<br><br>Advice Timing: Dependent, Independent-then-Revise<br><br>Advice Centrality: (low, high, median of independent judgements) | Influence of Advice | Estimating the age of a person in a photo | With median advice, dependent sequence judgements are further from advice than independent-then-revise.<br><br>With extreme advice, dependent sequence judgements are closer to advice than |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | independent-then-revise/ |
| Rader et al (2015). Study 2 | Explore the mechanism for the push away effect | Advice Timing: Dependent, Independent-then-Revise, Independent | Influence of Advice<br><br>Confidence in Advice | Estimating the age of a person in a photo | Confidence mediates the push-away effect |
| Rader et al (2015) Study 3 | Further explore the mechanism for the push away effect | Advice Timing: Dependent, Independent-then-Revise, Independent | Influence of Advice<br><br>Confidence in Advice<br><br>Verbal Protocol to Assess Thought Processes | Estimating the age of a person in a photo | Confidence mediates the push-away effect |
| Rader et al (2015) Study 4 | Investigate impact of advice sequence on accuracy with wider span of advice centrality | 3x3 Advice timing x Advice Centrality<br><br>Advice Timing: Dependent, Independent-then-Revise, Independent<br><br>Advice Centrality: (low, high, median of independent judgements) | Influence of Advice<br><br>Accuracy | Estimating the age of a person in a photo | Both advice sequences improve accuracy, compared to independent judgements<br><br>When advice is extreme, dependent estimates are less accurate than revised |
| Rader et al (2015) Study 5 | Explore similarities between anchoring paradigm and dependent condition, and if source | 4x3 Advice timing x Advice Centrality<br><br>Advice Timing: | Influence of Advice | Estimating the age of a person in a photo | Dependent sequence gave same results when implemented as anchoring paradigm |

| | | | | | |
|---|---|---|---|---|---|
| | of advice makes a difference | Dependent-Comparative, Dependent, Independent-then-Revise, Independent<br><br>Advice Centrality: (low, high, median of independent judgements) | | | |
| Sniezek & Buckley (1995) | Examine the cueing effect in social decision making | Advice Timing (Independent, Cued, Dependent),<br><br>Advisor Conflict (No Conflict, Conflict), | Judge Accuracy, Confidence, Overconfidence, Influence of Advice | Choice task on business events based on information from recent issues of *Business Week, The Wall Street Journal*, and *The New York Times* | Independent judges had the highest final choice accuracy and confidence, followed by Cued, then Dependent |
| Yaniv & Choshen-Hillel (2011) Study 1 | Examine if suspension of prior opinions increase accuracy | Full-View, Blindfold | Accuracy, Measures of Egocentrism, Confidence | Estimating the number of calories in a target food | Blindfold condition took more advice and gave higher accuracy than Full-View. In the Full-view condition, final estimates were more accurate than initial estimates. |

*Table 2: Timing of Advice Conditions & Findings*

| Conditions | ***Control Group (no advice)*** | ***Initial decision, then advice, then final decision*** | ***Advice then final decision*** | Advice but subjects blind to question, then final decision | Impact on Accuracy | Influence of Advice |
|---|---|---|---|---|---|---|
| Rader et al. (2015) | Independent (I) | Independent then Revise (ItR) | Dependent (D) | | (ItR=D)>I | ItR>D (Median Advice)<br><br>D>ItR (Extreme Advice) |
| Sniezek & Buckley (1995) | Initial Independent (II) | Independent (I) | Cued I | Dependent (D) | I > C > D<br><br>*(ItR>D)* | D > C (88) = I (86.5)<br><br>*(ItR=D)* |
| Koehler & Beauregard (2006) | | Unexposed Advisee (UA) | Exposed Advisee (EA) | | | EA > UA<br><br>*(D>ItR)* |
| Yaniv & Choshen-Hillel (2011) Study 1 | | Full-View (FV) | | Blindfold (B) | B > FV | B > FV |
| Yaniv & Choshen-Hillel Studies 2 & 3 | | | Modified Full-View (MFV) | Blindfold (B) | | |
| Note: Conditions in bold & italics are those of interest to the current study.<br>For ease of comparing results across studies, the results in parentheses are the study's results using the Rader et al. (2015) naming convention for conditions. | | | | | | |

Prior studies vary in the timing-of-advice conditions examined, task used, and dependent (outcome) variables (see Tables 1 & 2). There are also differences in their results.

*Timing of Advice*: Typically, studies examine: (a) a condition where the decision-maker first receives advice and then makes a judgement incorporating that advice into their own judgement (exposed, dependent, cued) and (b) the decision-maker first makes their own decision, then receives advice and revises that decision (unexposed, independent-then-revise, independent, full-view). These two conditions are often compared to a control (independent) condition where decision-makers receive no advice at all. Some studies also include conditions where the decision-maker receives the advice but not any information about the decision-making task (dependent condition in Sniezek & Buckley, blindfold condition in Yaniv & Choshen-Hillel). These conditions are less relevant to our context since they do not represent realistic decision-making scenarios.

*Decision-Making Task:* Rader et al., Koehler and Beauregard, and Yaniv and Choshen-Hillel all employed a quantitative estimation task, allowing distance measurements to compare advice and final estimates. Rader et al. used a task of estimating the age of an individual in a photograph, while Kohler and Beauregard's task was estimating the year of a historical event. Yaniv and Choshen-Hillel had their participants estimate the number of calories in a target food item.  In contrast, Sniezek and Buckley used a choice task with two options where participants had to choose one of the two options and indicate a probability from .5 to 1.00 to express confidence in the choice.

*Dependent Variables*: Dependent variables included decision accuracy (Rader et al., 2015; Sniezek & Buckley, 1995; Yaniv & Choshen-Hillel, 2011), confidence in advice (Rader et al., 2015; Koehler & Beauregard, 2006), confidence in decision (Rader et al., 2015; Sniezek & Buckley, 1995), and influence of advice (Rader et al., 2015; Koehler & Beauregard, 2006; Sniezek & Buckley, 1995; Yaniv & Choshen-Hillel, 2011).

<u>Decision Accuracy</u>: Yaniv and Choshen-Hillel measured accuracy as the mean absolute errors of participants' estimates. Rader et al. utilized two metrics for accuracy, absolute distance between the estimate and the correct answer, as well as absolute distance between the estimate

and the mean of independent estimates. Because Sniezek and Buckley utilized a choice task, accuracy was measured as the percent of items answered correctly.

Influence of Advice: For influence of advice, Rader et al. measured the absolute distance between participants' estimates and the advice, as well as the percentage of estimates that are within two and five years of the advice. Similarly, Koehler and Beauregard measured the distance (absolute deviation) of decision-maker's estimates from the advisor's estimate.

Confidence in Advice: Koehler and Beauregard measured confidence in advice by asking participants to specify the probability that the advisors' estimate was included in a fixed interval centered on the estimate. To measure confidence in both advice and participants' own estimates, Rader et al. asked participants to indicate on a scale of 1 to 7 "How accurate do you think this answer is?" and "How confident do you feel in this answer?". Sniezek and Buckley measured confidence in decision as the mean of the participants' confidence assessments.

Results for each variable vary across the different experiments. Of particular interest to our study are results related to decision accuracy (which would indicate debiasing effects) and influence of advice which would indicate the extent to which individuals anchor on their own advice or on algorithmic advice in making their final judgement.

In terms of *accuracy*, Sniezek and Buckley's results showed that decision-makers who made an initial decision before receiving advice (independent-then-revise) had higher accuracy than those who received advice prior to making any decision (dependent). Rader et al. expected to find similar results to Sniezek and Buckley, however they found that these two groups had around the same amount of accuracy, both surpassing individuals who receive no advice at all. Sniezek and Buckley propose that participants in the independent-then-revise group perform more information processing than the dependent group resulting in the difference in accuracy for the two groups. In general, prior literature suggests that the independent-then-revise condition seems to have equivalent or higher accuracy than the dependent condition. However, we posit this to the differing quality of advice in these studies and thus for accurate advice, we are arguing the opposite, that the dependent condition will have greater accuracy than the independent-then-revise.

For *influence of advice*, Koehler and Beauregard's results indicate that decision-makers who see advice prior to making a judgement (dependent condition) gave final estimates closer to the advisor's estimate, compared to individuals who received advice after first making an independent judgement (independent-then-revise), while Rader et al. found that for median advice, decision-makers in the dependent condition gave estimates further from advice than individuals in the independent-then-revise condition. Sniezek and Buckley observed that the influence of advice for the two groups took was around the same. One possible reason for the difference in results is explained below by the variable *confidence in advice*.

Studies have also measured *confidence in advice* as a dependent variable. Rader et al. found that confidence in advice mediates the push-away effect, where participants in the dependent condition move away from the advice. Koehler and Beauregard found that the more advice a participant took, the more confidence the participant had in the advice. This could potentially explain the difference in findings for *influence of advice* – participants in Koehler and Beauregard's experiment expressed greater confidence in the advice and thus it had a greater influence on their estimates. While Rader et al. found that participants who lack confidence in the advice are more likely to exhibit the push-away effect, explaining the smaller influence of advice on their estimates. Yaniv and Choshen-Hillel found that although participants in the blindfold condition were more accurate on average, they had less confidence than the full-view condition. Though confidence is not a dependent variable of interest, we plan to measure and control for it in our study.

In general, these studies indicate that the findings for accuracy and amount of advice taking are mixed. Consistent with Koehler and Beauregard, we argue that decision-makers in the dependent condition will take more advice compared to decision-makers in the independent-then-revise condition. Although this contrasts the results of Sniezek and Buckley, and Rader et al., we are focusing on scenarios where decision-makers seek advice from experts. We provide our hypotheses and rationale for these hypotheses next.

# CHAPTER 3

# HYPOTHESES

Prior literature suggests, and empirically demonstrates, that individuals making investment decisions exhibit conservatism bias (Barberis et al., 1998; Barberis & Thaler, 2002; Hirshleifer, 2001; Zhang et al., 2015). We posit that unbiased (i.e., accurate) algorithmic advice would help mitigate this bias. In general, decision-makers who receive algorithmic advice, regardless of the timing, will incorporate the advice into their decision making and, therefore, given that the advice is accurate, they will adjust their judgements closer to an unbiased (accurate) estimate (Rader et al., 1995).

Prior research has shown that individuals typically adjust their estimates around 20-30% towards the advice (Harvey & Fischer, 1997), consistent with Yaniv and Kleinberger's (2000) findings that individuals place a weight on their own estimates of around 70-80%. Although final estimates of decision-makers in the dependent (advice first) sequence are found to be closer to the advice, final estimates of individuals in the independent-then-revise sequence still reflect a deliberate attempt to incorporate the advice into this estimate (Koehler & Beauregard, 2006). Regardless of the magnitude of shift toward advice, decision-makers have shown that they update their beliefs based on advice, resulting in final decisions closer to the advice. Given that this advice is unbiased, incorporation of advice results in less biased decisions. Thus, we suggest the following hypothesis:

*Hypothesis 1: Decisions with algorithmic advice (i.e., dependent and independent-then-revise sequences) will exhibit less conservatism bias (will be more accurate) than decisions without algorithmic advice (i.e., independent condition).*

Even though we expect algorithmic advice to debias conservatism bias to some extent, we expect that the timing of advice would influence the magnitude of the debiasing effect. Specifically, we expect that in the independent-then-revise condition, participants will tend to anchor on their initial estimate and discount the algorithmic advice, resulting in final estimates further from the AI advisor's advice. Therefore, the final decisions of participants in the independent-then-revise sequence will be further from the algorithmic advice than those of participants in the dependent sequence. Literature on anchoring and adjustment has shown that people anchor on initial estimates, insufficiently adjusting to new evidence (Tversky & Kahneman, 1974). We would expect that decision-makers in the dependent condition who are given the advice first will tend to similarly anchor on the AI advice, and therefore will rely on the AI advice to a greater extent. Given that the AI advice does not suffer from conservatism bias, we would expect that in dependent condition AI advice will have a greater debiasing effect. Therefore, we suggest the following hypothesis:

*Hypothesis 2: Algorithmic advice in the dependent sequence will have a greater debiasing effect (i.e., will lead to more accurate decisions) than algorithmic advice in the independent-then-revise sequence.*

# CHAPTER 4

# RESEARCH METHODOLOGY

## 4.1 Research Design

We tested our hypotheses about (a) the impact of algorithmic advice on mitigating conservatism bias; and (b) the impact of the timing of such advice on bias mitigation, using a lab experiment with three conditions: *independent* (control) where no algorithmic advice is provided; *independent-then-revise* where participants make an independent decision, then receive algorithmic advice and then make their final decision; and *dependent* where participants receive algorithmic advice and then receive the task information to make their decision.

*Materials*

We adapted the experimental materials by Zhang et al. (2015) who conducted a lab experiment to examine how to mitigate conservatism bias and loss aversion through the use of decision support systems. We use the same experimental material and task as in the Zhang et al. study to examine conservatism bias but our study (a) provides AI advice rather than a decision-tool (excel with macros) to make the decision; and (b) varies the timing of algorithmic advice. Materials include a set of randomly generated (multivariate normal distribution) company profiles with ratings from 1-5 for four features: competitor strength, leadership ability, market condition, and proprietary technology (see Appendix A). Features were used to quantify company success likelihood through two equations listed below – odds ratio and success likelihood. Company profiles were categorized into nine groups based on their likelihood of success. This was done by rounding each company's success likelihoods to the nearest decile, as in Zhang et al. (2015). "That is, the profiles were grouped into nine deciles from 10% to 90% at 10% intervals. Profiles that fell in [0, .05] and [.95, 1] were discarded because they only covered 5% range. The probability judgment stage and the decision-making stage then randomly drew stimuli from the nine decile groups" (Zhang et al. 2015, p. 2213).

$$Odds\ Ratio = -2.4 + 0.4L + 0.6P + 0.2M - 0.4C$$

$$Success\ Likelihood = \frac{e^{OddsRatio}}{1 + e^{OddsRatio}}$$

## 4.2 Participants

Participants were Amazon Mechanical Turk workers. To guarantee quality responses, we used only AMT workers with a HIT approval rate above 97% (i.e., workers of the AMT marketplace who have a high percentage of completed tasks that are approved by requesters). To ensure statistical power according to Statistics Kingdom's sample size calculator (https://www.statskingdom.com/sample_size_regression.html), to detect medium size effects of .25, with a statistical test power of .80, we needed a sample size of 158 participants, with 53 individuals in each group. Participants were randomly assigned to one of the three experimental conditions: dependent, independent-then-revise, and independent (control).

## 4.3 Task & Procedure

The experiment was conducted online using Qualtrics. Before the start of the experiment, each participant provided informed consent for participation in the experiment (see Appendix B). In addition, participants completed a pre-task survey to measure (a) individual differences in processing information and (b) their knowledge of and interest in investing and evaluating startups. The individual differences in processing information constructs included "Need for Cognition" (Yang & Smith, 2009) and "Need for Cognitive Closure" (Federico et al., 2007). Participants were presented with the task scenario and were required to answer a few questions correctly in order to indicate understanding of the task and be granted access into the experiment. In addition, participants completed a post-task survey to measure confidence in the answers they provided and in the advice that they received (see Table 3).

*Table 3: Pre-Task & Post-Task Items*

| **Pre-Task Survey** | |
|---|---|
| Need for Cognition<br>Source: Yang and Smith (2009) | For each statement below, please indicate the extent that you agree with the statement (1 = strongly disagree, 5= strongly agree). Please answer the questions thoughtfully |
| | I don't like to do a lot of thinking |
| | I try to avoid situations that require thinking in depth about something. |
| | I prefer to do something that challenges my thinking abilities rather than something that requires little thought. |
| | I prefer complex to simple problems. |
| | Thinking hard and for a long time about something gives me little satisfaction. |
| Need for cognitive closure<br>Source: Federico et al. (2007) | For each statement below, please indicate the extent that you agree with the statement (1 = strongly disagree, 5 = strongly agree). Please answer the questions thoughtfully |
| | I get very upset when things around me aren't in their place |
| | Generally, I avoid participating in discussions on ambiguous and controversial problems. |
| | I prefer to be with people who have the same ideas and tastes as myself. |
| | I feel uncomfortable when I do not manage to a give a quick response to problems that I face. |
| | Any solution to a problem is better than remaining in a state of uncertainty. |
| | I prefer activities where it is always clear what is to be done and how it needs to be done. |
| | I prefer things to which I am used to those I do not know and cannot predict. |
| Attention Checks | On this statement, click "somewhat agree". |
| | On this statement, click "strongly disagree". |
| Knowledge on evaluating startup companies' success likelihood. | Please answer the following questions about your financial knowledge and interest. |
| | How do you rate your knowledge about evaluating the financial success of a company?<br>1 = none at all, 5 = a great deal |

| | |
|---|---|
| | How do you rate your knowledge of evaluating the success of a startup?<br>1 = none at all, 5 = a great deal |
| | Have you ever invested in startup companies?<br>1=yes, 2=no |
| Interest in investing in startup companies. | How do you rate your interest in investing in startup companies?<br>1 = none at all, 5 = a great deal |
| **Post-Task Survey** | |
| Confidence in advice Source: Rader et al. (2015), See et al. (2011) | How accurate do you think the AI advice was?<br>1 = not at all, 7 = extremely |
| | How confident are you in the AI advice that has been provided?<br>1 = not at all, 7 = extremely |
| | How reliable do you think the AI advisor was?1 = not at all, 7 = extremely |
| Confidence in their own estimate | How accurate do you think your answers were?<br>1 = not at all, 7 = extremely |
| | How confident are you in the answers that you provided?<br>1= not at all, 7 = extremely |

Each participant performed the role of a venture capitalist evaluating a startup company's probability of success, given company profiles and advice from an AI advisor (except for the control group). To provide a common base for the expertise of the AI advisor, the participants were told that the AI advisor has an 80% accuracy. The experiment consisted of one training block and one test block. The training block consisted of 18 randomly ordered trials balanced across the nine intervals from 10% to 90% of success likelihood, and the test block consisted of nine randomly ordered trials balanced across the same nine intervals (see Appendices C and D). To prevent participants from identifying a pattern among the nine balanced test block trials, we randomly inserted four trials which were later thrown out.

*Training Block:* Within each trial of the training block, participants observed a company profile and were instructed to select one of three options – fail, not sure, succeed – to indicate their estimate of the company's success outcome (see Appendix C). Then participants were shown the actual outcome of the company, success or fail, and the actual success likelihood percentage. The purpose of the training block was for participants to gain familiarity with the relationship between company profile ratings and success likelihood.

*Testing Block:* Within each trial of the testing block, participants again were displayed a company profile, but this time they were instructed to indicate their estimate of the company's success likelihood by selecting the probability of success from one of 11 options ranging from 0% to 100% by intervals of 10%. In addition to the company profile, the dependent and independent-then-revise groups were also presented with AI advice, while the independent (control) group only had access to the company profile when making their estimate (see Appendix D). In the dependent condition, participants received the algorithmic advice and the company profile at the same time (see Appendix D) and were asked to provide an estimate of the company's success likelihood. In the independent-then-revise condition, participants were provided with the company information and asked to provide an estimate of the company's success likelihood (see Appendix D). They then received the algorithmic advice, along with a reminder of their initial estimate, and were asked to provide a final estimate of the company's success likelihood.

## 4.4 Measures

*Dependent variable*: Our main dependent variable is decision accuracy which serves as a proxy for debiasing. We measured *decision accuracy* by calculating the absolute distance between participants' estimates and the AI advice which has 100% accuracy. In addition, we assessed *influence of advice* in the independent-then-revise condition by measuring the amount of change between participants' initial answers and their final answers. We were unable to measure *influence of advice* in the dependent and independent (control) groups because both groups only provide a final estimate per question. Because AI advice is our measure of *accuracy*, we cannot measure i*nfluence of advice* by directly comparing estimates to the AI advice.

*Evidence of conservatism bias*: To provide evidence of conservatism bias, we also assessed the *presence of conservatism*, which can be manifested as the underestimation of high probabilities and overestimation of low probabilities, across the control group using the procedure in Zhang et al. (2015). Specifically, we plotted the estimated success likelihoods of participants in the independent (control) condition as a function of the actual success likelihoods across the nine intervals. For actual success likelihoods higher than 60%, we expect to see participants' estimated success likelihoods generally lower than the actual success likelihood. For actual success likelihoods below 40%, we expect to participants' estimates to be higher in general than the actual success likelihood.

*Control variables:* We measure several control variables such as knowledge of task, interest in task, need for cognition, need for cognitive closure, confidence in advice, and confidence in their own estimates (see Table 3).

CHAPTER 5

# DATA ANALYSIS & RESULTS

*Sample*

Data were collected from 198 MTurks, of these we discarded 15 because they failed the attention checks or because they provided the same probability estimate of success across all companies. As a result, our sample consists of 183 individuals – 64 in the dependent group, 60 in the independent-then-revise group, and 59 in the independent (control) group.

Table 4 presents descriptive statistics for our control variables. To derive these, we averaged all items for the same construct. All items were measured on a scale 1-5 (strongly disagree to strongly agree), except for one item in 'knowledge on evaluating startup companies' success likelihood' which was measured on a binary scale of 0/1.

*Table 4: Descriptive Statistics*

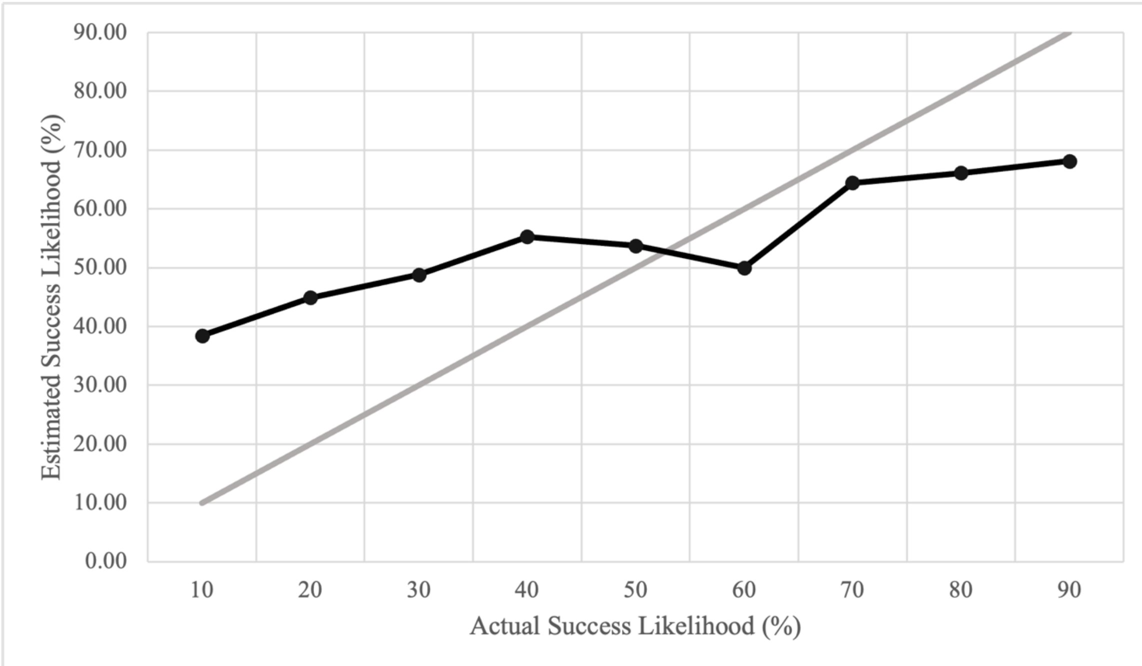
| | **Overall** | | **Independent (Control) Group** | | **Dependent Group** | | **Independent-then-Revise Group** | |
|---|---|---|---|---|---|---|---|---|
| | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* | *Mean* | *SD* |
| **Need for Cognition** | 3.61 | 0.71 | 3.61 | 0.64 | 3.54 | 0.8 | 3.67 | 0.68 |
| **Need for Cognitive Closure** | 3.63 | 0.66 | 3.62 | 0.65 | 3.61 | 0.72 | 3.67 | 0.60 |
| **Knowledge on evaluating startup companies' success likelihood** | 2.93 | 0.50 | 2.79 | 0.52 | 2.99 | 0.50 | 2.99 | 0.47 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Interest in investing in startup companies** | 3.75 | 0.98 | 3.56 | 1.07 | 3.75 | 0.98 | 3.93 | 0.86 |
| **Confidence in AI Advice** | 3.90 | 0.66 | N/A | N/A | 3.92 | 0.71 | 3.88 | 0.61 |
| **Confidence in Answer** | 3.92 | 0.68 | 3.81 | 0.80 | 4.00 | 0.64 | 3.92 | 0.60 |

*Test of Assumption of Conservatism Bias*

Our premise is that individuals making financial decisions exhibit conservatism bias. We tested this assumption in our independent (control) group where participants did not receive any algorithmic advice. Our experimental results indicate that participants in the independent (control) group overestimated low probabilities and underestimated high probabilities, thus providing evidence of conservatism bias in their decision-making (Figure 1). Similar to Zhang et al. (2015), for actual values below 40%, participants' estimated success likelihoods were higher than the actual success likelihoods, and for actual values above 60%, participants' estimated success likelihoods were lower on average than the actual success likelihoods. As Zhang et al. (2015) described, the proximity of participants' estimates to the actual values suggests that participants learned to predict the different levels of success likelihood. We calculated the Pearson correlation coefficient and found that the estimated success likelihoods were significantly correlated with the actual success likelihoods, $r(7) = .9377, p < .001$. Additionally, the R-squared was 0.879 and the RMSE was 3.319.

*Figure 1: Evidence of Conservatism Bias*




Having validated our assumption of conservatism bias, we proceed to test our hypotheses.

*Hypothesis Testing*

To assess our hypotheses, we ran an ANOVA followed by planned contrasts to determine if there were significant effects on accuracy of a) receiving algorithmic advice and b) varying the timing of such advice.

## 5.1 Hypothesis 1 Results

Our first hypothesis posited that decisions with algorithmic advice (i.e., participants in the dependent and independent-then-revise conditions) would have higher accuracy than decisions without algorithmic advice (i.e., participants in the independent (control) condition). Table 5 describes the mean and variance for accuracy for each condition. Because accuracy for each estimate was measured as the distance from the advice, lower distance indicated higher accuracy. The dependent condition had the highest accuracy, followed by independent-then-revise, with the independent (control) condition having the lowest accuracy. The ANOVA, as described in Table 6, revealed that the between group difference was significant, with F-statistic = 10.3646 > F-critical = 3.05, p =5.4899e-05. To determine which mean differences were significant, we ran two analyses with planned contrasts, adjusting the p-values using the Bonferroni method to avoid increased Type I error rates (Table 7). The first analysis compared the independent (control) condition to both the dependent and independent-then-revise conditions combined (H1), and the dependent condition to the independent-then-revise (H2). The second analysis compared the independent (control) condition to the dependent and independent-then-revise conditions separately (H1), and again, the dependent condition to the independent-then-revise condition (H2).

To assess our first hypothesis, we focus on the results of the planned contrasts comparing the independent (control) group to the dependent and independent-then-revise groups (Table 8). In the first analysis, the difference between the independent group and the dependent and independent-then-revise groups (combined) was significant with a p-value of 0.00380. In the second analysis, the planned contrasts revealed that the difference between the independent (control) condition and the dependent condition was significant, with an adjusted p-value of 6.31e-05. However, the difference between the independent (control) condition and the independent-then-revise condition was not found to be significant, with an adjusted p-value of 0.77008. Thus, the findings of our first analysis support our first hypothesis, that decisions with algorithmic advice (dependent and independent-then-revise) are more accurate than decisions without algorithmic advice (independent). However, the findings of our second analysis indicate that participants in the dependent group made more accurate decisions than participants in the independent (control) group, but there is not enough evidence to suggest that decisions of independent-then-revise group have higher accuracy than those of the independent (control) group.

*Table 5: Mean and Variance of Accuracy by Condition*

| **Condition** | **Mean Distance from Advice (Accuracy)** | **Variance** |
|---|---|---|
| Independent (Control) | 20.6779661 | 57.6848425 |
| Dependent | 12.7256944 | 118.202222 |
| Independent-then-Revise | 18.5740741 | 127.352654 |

*Table 6: ANOVA Results*

| | **df** | **Sum of Squares** | **Mean Squared** | **F** | **P-value** | **F crit** |
|---|---|---|---|---|---|---|
| **Between Group** | 2 | 2108.197 | 1054.098 | 10.3646 | 5.4899e-05 | 3.05 |
| **Residuals/Within Group** | 180 | 18306.27 | 101.70 | | | |

*Table 7: Planned Contrasts*

| | **Planned Contrasts** | **Independent/Control (I)** | **Dependent (D)** | **Independent-then-Revise (ITR)** |
|---|---|---|---|---|
| **Analysis 1** | *I vs. (D vs. ITR)* | -1 | .5 | .5 |
| | *D vs. ITR* | 0 | -1 | 1 |
| **Analysis 2** | *I vs. D* | -1 | 1 | 0 |
| | *I vs. ITR* | -1 | 0 | 1 |
| | *D vs. ITR* | 0 | -1 | 1 |

*Table 8: Planned Contrasts Results*

| | **Planned Contrasts** | **Estimate** | **Std. Error** | **t value** | **Pr(>\|t\|)** |
|---|---|---|---|---|---|
| **Analysis 1** | *I vs.(D vs. ITR)* | -5.028 | 1.595 | -3.512 | 0.00380 ** |
| | *D vs. ITR* | 5.848 | 1.812 | 3.227 | 0.00297 ** |
| **Analysis 2** | *I vs. D* | -7.952 | 1.820 | -4.369 | 6.31e-05 *** |
| | *I vs. ITR* | -2.104 | 1.849 | -1.138 | 0.77008 |
| | *D vs. ITR* | 5.848 | 1.812 | 3.227 | 0.00446 ** |

## 5.2 Hypothesis 2 Results

Our second hypothesis suggested that algorithmic advice in the dependent condition will lead to more accurate decisions than algorithmic advice in the independent-then-revise condition. To assess this hypothesis, we look at the planned contrasts comparing the dependent and the independent-then-revise conditions (Table 7 and Table 8). In the first analysis, the difference between the two conditions was significant, with an adjusted p-value of 0.00297. In the second analysis, the difference between the dependent and independent-then-revise conditions was also found to be significant, with an adjusted p-value of 0.00446. Thus, our results of both analyses support our hypothesis that the timing of algorithmic advice does indeed matter. Participants who received algorithmic advice *prior* to making an initial decision had greater accuracy than participants who made an initial decision, then received algorithmic advice and made a final revised decision.

*ANCOVA Analysis*

We also ran an ANCOVA to account for effects of the control variables measured through the pre-test and post-test surveys. The results of the ANCOVA indicated that none of these covariates were significant, though interest in investing in startups approached significance (Table 9).

*Table 9: ANCOVA Results*

| | **Sum of Squares** | **df** | **f value** | **Pr(>\|f\|)** |
|---|---|---|---|---|
| **Condition** | 770.0 | 1 | 6.3143 | 0.01335 * |
| **Need for Cognition** | 15.0 | 1 | 0.1228 | 0.72667 |
| **Need for Cognitive Closure** | 2.5 | 1 | 0.0206 | 0.88618 |

| | | | | |
|---|---|---|---|---|
| **Knowledge on evaluating startup companies' success likelihood.** | 25.3 | 1 | 0.2072 | 0.64979 |
| **Interest in investing in startup companies** | 457.6 | 1 | 3.7531 | 0.05514 |
| **Confidence in Answer** | 27.2 | 1 | 0.2230 | 0.63765 |
| **Confidence in AI Advice** | 182.0 | 1 | 1.4927 | 0.22427 |
| **Residuals** | 14144.8 | 116 | | |

*Additional Analyses*

We ran a number of additional analyses. First, we tested whether participants in the three conditions differed in their confidence in their own estimates. The ANOVA results suggest there was no significant difference between groups, with p = 0.318 (Table 10).

*Table 10: ANOVA Results for Confidence in Own Estimates*

| | **df** | **Sum of Squares** | **Mean Squared** | **F** | **P-value** |
|---|---|---|---|---|---|
| **Group** | 2 | 1.08 | 0.5378 | 0.154 | 0.318 |
| **Residuals** | 180 | 83.86 | 0.4659 | | |

Then we tested whether participants in the algorithmic advice conditions differed in their confidence in the AI advice. The ANOVA results suggest there was no significant difference between groups, with p = 0.694 (Table 11).

*Table 11: ANOVA Results for Confidence in AI Advice*

| | **df** | **Sum of Squares** | **Mean Squared** | **F** | **P-value** |
|---|---|---|---|---|---|
| **Group** | 1 | 0.07 | 0.0680 | 0.155 | 0.694 |
| **Residuals** | 122 | 53.42 | 0.4379 | | |

Finally, although we cannot measure the influence of advice for all three conditions, we could calculate influence of advice in the independent-then-revise condition by taking the difference between one's pre-advice initial estimate and post-advice final estimate. The average influence of advice for participants in the independent-then-revise condition was found to be around 12.72%.

# CHAPTER 6

# IMPLICATIONS & CONTRIBUTIONS

Given the limited amount of research on the use of algorithmic advice for mitigating human decision biases, our research aimed to examine whether algorithmic advice can mitigate cognitive biases and how the timing of such advice impacts the extent of bias mitigation. Specifically, we examined how algorithmic advice may mitigate conservatism bias in investment decisions. The results indicate that conservatism bias is present in the investment decision-making task. We found that decisions where algorithmic advice is given prior to an initial decision exhibit less conservatism bias than decisions without algorithmic advice and decisions where an initial decision is made prior to receiving.

Our research shows that algorithmic advice did improve accuracy, however, the results for the independent-then-revise group were not significant. We found that the timing of algorithmic advice played a role in the debiasing effects – the dependent group experienced a greater debiasing effect than both the independent (control) group and the independent-then-revise group. Although participants in the independent-then-revise did incorporate the advice into their revised estimates, there was strong anchoring on the initial own estimate and the magnitude of adjustment was not sufficient to result in a significantly higher accuracy for the independent-then-revise than control groups.

Our research contributes to the literature of Human-AI augmented decision-making and has implications for the design of AI-augmented decision systems. It suggests that, when known human decision biases are present, accurate algorithmic advice provided *before* a human decision-maker makes their own decision can have a debiasing effect. Designing AI-augmented decision systems with human decision biases in mind can help mitigate the effects of such biases in decision-making. Although it is highly unlikely that algorithmic advice will be able to fully mitigate cognitive biases, it is worthwhile to design systems that attempt to minimize their effects in situations where decision-makers commonly exhibit certain biases.

Our findings on the timing of advice have broader practical implications beyond debiasing. They suggest that for algorithmic advice to have a more significant influence on decision making, the advice should be provided early before an individual forms their own assessment. Whether this is desirable will depend on the quality of algorithmic advice vis-a-vis human expertise, algorithmic biases versus human biases, and whether it is desirable to nudge the decision-maker or user towards the algorithmic advice or not.

It is important to acknowledge that while the advice we provided was accurate, algorithmic advice in many applications may not be unbiased, but rather may be biased or inaccurate. Cognitive biases can be reflected in the data used to train algorithms, as well as the decision to use a certain algorithm over another. The presence of cognitive biases in algorithmic advice introduces many new questions. For example, if the algorithmic bias reflects the same bias as that of the human decision-maker, which is possible given how machine learning algorithms are trained, would receiving algorithmic advice further amplify the bias leading to worse decisions? Does the level of cognitive bias present in algorithmic advice vary based on the nature of the algorithm (e.g., machine learning versus rule-based)? More research is needed to examine the effect of algorithmic advice on cognitive biases under these different circumstances.

Furthermore, we examined one specific cognitive bias – conservatism bias. Given the plethora of biases that are present in different decision-making contexts, future research can examine the effect of algorithmic advice, and its timing, on mitigating other cognitive biases.

Finally, the results of the study should be interpreted bearing in mind the possibility that participants in the experiment may not give their full effort to the task, especially given that it is an online experiment. To minimize this possibility, we included several attention checks throughout the experiment (and dropped participants who failed the attention checks) and offered a reward based on accuracy to incentivize maximum effort. Nonetheless, it is important to replicate the study in a real-life decision-making context to assess the generalizability of the results.

In conclusion, while many studies have focused on AI-biases, this study focuses on how AI can overcome human biases. We show that accurate algorithmic advice has a significant debiasing effect, but that the timing of such advice is material. We hope that the results of the study can generate additional research to understand how to design Human-AI collaborative decision-making systems that can help humans make better - less biased - decisions.

.

REFERENCES


Balayn, A., Lofi, C., & Houben, G. J. (2021). Managing bias and unfairness in data for decision support: a survey of machine learning and data engineering approaches to identify and mitigate bias and unfairness within data management and analytics systems. *The VLDB Journal*, *30*(5), 739-768.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, *49*(3), 307-343.

Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, *1*, 1053-1128.

Edwards, W. (1968). Conservatism in human information processing. *Formal representation of human judgment*.

Federico, C. M.; Jost, J. T.; Pierro, A.; Kruglanski, A. W. (2007): The Need for Closure and Political Attitudes: Final Report for the ANES Pilot. ANES Pilot Study Report, No. nes01190, from http://www.electionstudies.org/resources/papers/Pilot2006/nes011904.pdf, letzter Zugriff 11.05.2010.

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive psychology*, *24*(3), 411-435.

G. Harris, C. (2020, April). Mitigating cognitive biases in machine learning algorithms for decision making. In *Companion Proceedings of the Web Conference 2020* (pp. 775-781).

Harvey, N., & Fischer, I. (1997). Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organizational behavior and human decision processes*, *70*(2), 117-133.

Hirshleifer, D. (2001). Investor psychology and asset pricing. *The journal of Finance*, *56*(4), 1533-1597.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge university press.

Koehler, D. J., & Beauregard, T. A. (2006). Illusion of confirmation from exposure to another's hypothesis. *Journal of Behavioral Decision Making*, *19*(1), 61-78.

Lee, M. K., Kiesler, S., and Forlizzi, J. Mining behavioral economics to design persuasive technology for healthy choices. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI '11 (2011), 325–334.

Lee, K., Woo, H. G., Cho, W., & De Jong, S. (2022, January). When Can AI Reduce Individuals' Anchoring Bias and Enhance Decision Accuracy? Evidence from Multiple Longitudinal Experiments. In *Proceedings of the 55th Hawaii International Conference on System Sciences*.

Montier, J., & Strategy, G. E. (2002). Applied behavioural finance: Insights into irrational minds and market. *Unpublished manuscript*.

Orphanou, K., Otterbacher, J., Kleanthous, S., Batsuren, K., Giunchiglia, F., Bogina, V., ... & Kuflik, T. (2022). Mitigating Bias in Algorithmic Systems—A Fish-eye View. *ACM Computing Surveys*, *55*(5), 1-37.

Pompian, M. M. (2011). *Behavioral finance and wealth management: how to build investment strategies that account for investor biases*. John Wiley & Sons.

Rader, C. A., Soll, J. B., & Larrick, R. P. (2015). Pushing away from representative advice: Advice taking, anchoring, and adjustment. *Organizational Behavior and Human Decision Processes*, *130*, 26-43.

See, K. E., Morrison, E. W., Rothman, N. B., & Soll, J. B. (2011). The detrimental effects of power on confidence, advice taking, and accuracy. *Organizational behavior and human decision processes*, *116*(2), 272-285.

Sniezek, J. A., & Buckley, T. (1995). Cueing and cognitive conflict in judge-advisor decision making. *Organizational behavior and human decision processes*, *62*(2), 159-174.

Solomon, J. Customization bias in decision support systems. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI ’14 (2014), 3065–3074.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *science*, *185*(4157), 1124-1131.

Yang, X., & Smith, R. E. (2009). Beyond attention effects: Modeling the persuasive and emotional effects of advertising creativity. *Marketing Science*, *28*(5), 935-949.

Yaniv, I., & Choshen-Hillel, S. (2012). Exploiting the wisdom of others to make better decisions: Suspending judgment reduces egocentrism and increases accuracy. *Journal of Behavioral Decision Making*, *25*(5), 427-434.

Yaniv, I., & Kleinberger, E. (2000). Advice taking in decision making: Egocentric discounting and reputation formation. *Organizational behavior and human decision processes*, *83*(2), 260-281.

Yin, J., Ngiam, K. Y., & Teo, H. H. (2020). Work design in healthcare artificial intelligence applications: The role of advice provision timing.

Zhang, Y., Bellamy, R. K., & Kellogg, W. A. (2015, April). Designing information for remediating cognitive biases in decision-making. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 2211-2220).

Zhao, J. C., Fu, W.-T., Zhang, H., Zhao, S., and Duh, H. B.-L. To risk or not to risk? improving financial risk-taking of older adults by online social information. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work and Social Computing*, CSCW ’15 (2015), to appear.

APPENDIX A: SAMPLE COMPANIES

| | |
|---|---|
| Leadership Ability | ★★★☆☆ |
| Proprietary Technology | ★★★★☆ |
| Market Condition | ★★☆☆☆ |
| Competitor Strength | ★☆☆☆☆ |

| **Rounded Decile** | **Interval** | **Leadership Ability** | **Proprietary Technology** | **Market Condition** | **Competitor Strength** | **Actual Success Likelihood** |
|---|---|---|---|---|---|---|
| 10% | [.05, .1499] | 2 | 1 | 2 | 4 | .09975 |
| 20% | [.15, .2499] | 1 | 2 | 4 | 3 | .231475 |
| 30% | [.25, .3499] | 3 | 1 | 3 | 2 | .310026 |
| 40% | [.35, .4499] | 2 | 2 | 4 | 2 | .401312 |
| 50% | [.45, .5499] | 2 | 3 | 5 | 3 | .50 |
| 60% | [.55, .6499] | 5 | 4 | 2 | 5 | .598688 |
| 70% | [.65, .7499] | 4 | 3 | 1 | 1 | .689974 |
| 80% | [.75, .8499] | 4 | 4 | 2 | 1 | .832018 |
| 90% | [.85, .9499] | 4 | 5 | 5 | 2 | .916827 |

APPENDIX B: INFORMED CONSENT FORM

You are being asked to take part in a research study. The information provided here will help you decide if you want to be in the study. Please ask the researcher(s) below if there is anything that is not clear or if you need more information.


| | |
|---|---|
| **Investigator** | Alyssa Joaquin (Under the Direction of Dr. Elena Karahanna) |
| | Institute for Artificial Intelligence |
| | University of Georgia |
| | Agj04689@uga.edu |


Purpose: The purpose of this research is to examine how people assess whether a startup company is likely to succeed. This assessment can inform whether or not to invest in the startup.

Procedure: If you choose to be in the study, you will complete an online survey in which you evaluate the success likelihood of multiple companies, given a company profile. The expected duration of your participation in the experiment is about one hour.

Right to Withdraw: Your involvement in this study is voluntary, and you may choose not to participate or to stop at any time.

Risks & Benefits: There is no foreseeable risk or discomfort associated with our experiment, other than a loss of time. Collection of data and survey responses using the internet involves the same risk that a person would encounter in everyday use of the internet, such as fatigue or breach of confidentiality. While the researchers have taken every reasonable step to protect your confidentiality, there is always the possibility of interception or hacking of the data by third parties that is not under the control of the research team. There will be no costs for participating. The benefits to participation include a base compensation of $3.75 if you complete the experiment and up to $1.30 in bonus for accuracy of decisions. The findings from this project will provide information on how to best help human decision makers in making financial decisions.

Privacy & Confidentiality: We will take steps to protect your privacy, but there is a small risk that your information could be accidentally disclosed to people not connected to the research. To reduce this risk, your Mechanical Turk Worker ID will be used to distribute the payment to you, but we will not store your worker ID with your survey responses. Please be aware that your Mturk Worker ID can potentially be linked to information about you on your Amazon Public Profile page, however, we will not access any personally identifying information from your Amazon Public Profile.

Use of your information for future research: No information collected from this study will be used or distributed for future research.

If you have any questions, suggestions, or concerns regarding this study or you want to withdraw from the study please contact Alyssa Joaquin at 301-803-0793, agj04689@uga.edu. If you have any complaints or questions about your rights as a research volunteer, contact the IRB at 706-542-3199 or by email at IRB@uga.edu.

If you agree to participate in this research study, please click “I Agree” and begin the survey.

## APPENDIX C: TRAINING BLOCK DESCRIPTION AND QUESTIONS

This experiment is designed as an investment game. You will play the role of an investor who will evaluate whether or not to invest in certain companies. To assess whether to invest or not, your task is to estimate the likelihood of success of the company based on four criteria: leadership ability, proprietary technology, market condition, and competitor strength. Higher ratings (more stars) in leadership, proprietary technology, and market condition suggest higher likelihood of success, whereas a higher rating in competitor strength suggests lower likelihood of success.

For the first part of the experiment, you will participate in a training session where you will be provided the company profile and you will be asked to predict whether the company will succeed or fail. After you provide your prediction, you will be provided with the actual (correct) outcome. This will help you calibrate your assessment of the company's probability of success based on these four criteria. We will provide you with 18 company profiles as training.

For the second part of the experiment, you will be provided with 13 profiles, and you will be asked to estimate the **actual probability of success**. You will be rewarded points for accuracy of your responses which will translate to additional compensation.

The training session will start next. Remember, for this session you will need to estimate whether a company will succeed or fail based on the company profile that includes four criteria.

<table>
<tr>
<td>Question 1</td>
<td>Given the provided company profile, indicate whether you think this company will fail or succeed.
<table>
<tr><td>Leadership Ability</td><td>★★☆☆☆</td></tr>
<tr><td>Proprietary Technology</td><td>★☆☆☆☆</td></tr>
<tr><td>Market Condition</td><td>★★☆☆☆</td></tr>
</table>
</td>
</tr>
</table>

<table>
<tr><td></td><td><table><tr><td>Competitor Strength</td><td>★★★★☆</td></tr></table>a. Fail<br>b. Not Sure<br>c. Succeed</td></tr>
<tr><td>Actual Binary Outcome for Question 1</td><td>This company will likely fail (estimated probability of success is 9.9%).</td></tr>
<tr><td>Question 2</td><td>Given the provided company profile, indicate whether you think this company will fail or succeed.<table><tr><td>Leadership Ability</td><td>★★★★☆</td></tr><tr><td>Proprietary Technology</td><td>★★★★☆</td></tr><tr><td>Market Condition</td><td>★★☆☆☆</td></tr><tr><td>Competitor Strength</td><td>★☆☆☆☆</td></tr></table>a. Fail<br>b. Not Sure<br>c. Succeed</td></tr>
<tr><td>Actual Binary Outcome for Question 2</td><td>This company will likely succeed (estimated probability of success is 83.2%).</td></tr>
</table>

## APPENDIX D: TESTING BLOCK QUESTIONNAIRE

**Control Condition**

Now, you will be provided with 13 profiles and this time you will have to estimate the **actual probability of success** for each company. You will be rewarded points for accuracy of your responses* which will translate to additional compensation.

As in the training block, higher ratings (more stars) in leadership, proprietary technology, and market condition suggest higher likelihood of success, whereas a higher rating in competitor strength suggests lower likelihood of success.

<table>
<tr><th colspan="2">Independent (Control) Condition</th></tr>
<tr><td>Question 1</td><td>Given the provided company profile, indicate your estimation of the company’s success likelihood (probability) by selecting one of the percentages below.
<table>
<tr><td>Leadership Ability</td><td>★★★☆☆</td></tr>
<tr><td>Proprietary Technology</td><td>★★★★☆</td></tr>
<tr><td>Market Condition</td><td>★★★☆☆</td></tr>
<tr><td>Competitor Strength</td><td>★★☆☆☆</td></tr>
</table>
a. 0%<br>
b. 10%<br>
c. 20%<br>
d. 30%<br>
e. 40%<br>
f. 50%<br>
g. 60%<br>
h. 70%<br>
i. 80%<br>
j. 90%<br>
k. 100%</td></tr>
<tr><td>Question 2</td><td>Given the provided company profile, indicate your estimation of the company’s success likelihood (probability) by selecting one of the percentages below.
<table>
<tr><td>Leadership Ability</td><td>★☆☆☆☆</td></tr>
</table></td></tr>
</table>

<table>
<tr><td></td><td><table><tr><td>Proprietary Technology</td><td>★★★☆☆</td></tr><tr><td>Market Condition</td><td>★★★☆☆</td></tr><tr><td>Competitor Strength</td><td>★★☆☆☆</td></tr></table>
a. 0%<br>b. 10%<br>c. 20%<br>d. 30%<br>e. 40%<br>f. 50%<br>g. 60%<br>h. 70%<br>i. 80%<br>j. 90%<br>k. 100%</td></tr>
<tr><td>…</td><td>…</td></tr>
</table>

**Dependent Conditions**

Now, you will be provided with 13 profiles along with advice from an AI advisor. The AI advisor has an accuracy level of approximately 80%. This time your task is to estimate the **actual probability of success** for each company. You will be rewarded points for accuracy of your responses which will translate to additional compensation.

As in the training block, higher ratings (more stars) in leadership, proprietary technology, and market condition suggest higher likelihood of success, whereas a higher rating in competitor strength suggests lower likelihood of success.

<table>
<tr><th colspan="2">Dependent Condition</th></tr>
<tr><td>Question 1 (advice included)</td><td>Given the provided company profile and AI advice, indicate your estimation of the company's success likelihood (probability) by selecting one of the percentages below.
<table><tr><td>Leadership Ability</td><td>★★★☆☆</td></tr><tr><td>Proprietary Technology</td><td>★★★★☆</td></tr><tr><td>Market Condition</td><td>★★★☆☆</td></tr><tr><td>Competitor Strength</td><td>★★☆☆☆</td></tr></table>
Your AI advisor suggests the success likelihood 70%<br>a. 0%</td></tr>
</table>

<table>
<tr><td></td><td>b. 10%<br>c. 20%<br>d. 30%<br>e. 40%<br>f. 50%<br>g. 60%<br>h. 70%<br>i. 80%<br>j. 90%<br>k. 100%</td></tr>
<tr><td>Question 2 (advice included)</td><td>Given the provided company profile and AI advice, indicate your estimation of the company’s success likelihood (probability) by selecting one of the percentages below.
<table>
<tr><td>Leadership Ability</td><td>★☆☆☆☆</td></tr>
<tr><td>Proprietary Technology</td><td>★★★☆☆</td></tr>
<tr><td>Market Condition</td><td>★★★☆☆</td></tr>
<tr><td>Competitor Strength</td><td>★★★☆☆</td></tr>
</table>
Your AI advisor suggests the success likelihood 30%<br>a. 0%<br>b. 10%<br>c. 20%<br>d. 30%<br>e. 40%<br>f. 50%<br>g. 60%<br>h. 70%<br>i. 80%<br>j. 90%<br>k. 100%</td></tr>
<tr><td>…</td><td>…</td></tr>
</table>

**Independent-then-Revise Condition**

Now, you will be provided with 13 profiles and this time you will have to estimate the **actual probability of success** for each company. Once you provide your estimate, you will be provided with advice from an AI advisor, and you will be asked to enter your final estimate of probability of success. The AI advisor has an accuracy level of approximately 80%. You will be rewarded points for accuracy of your responses* which will translate to additional compensation.

As in the training block, higher ratings (more stars) in leadership, proprietary technology, and market condition suggest higher likelihood of success, whereas a higher rating in competitor strength suggests lower likelihood of success.

<table>
<tr><th colspan="2">Independent-then-Revise Condition</th></tr>
<tr><td>Question 1</td><td>Given the provided company profile, indicate your estimation of the company’s success likelihood (probability) by selecting one of the percentages below.
<table>
<tr><td>Leadership Ability</td><td>★★★☆☆</td></tr>
<tr><td>Proprietary Technology</td><td>★★★★☆</td></tr>
<tr><td>Market Condition</td><td>★★★☆☆</td></tr>
<tr><td>Competitor Strength</td><td>★★☆☆☆</td></tr>
</table>
a. 0%<br>b. 10%<br>c. 20%<br>d. 30%<br>e. 40%<br>f. 50%<br>g. 60%<br>h. 70%<br>i. 80%<br>j. 90%<br>k. 100%</td></tr>
<tr><td>Question 1: Revising with AI Advice</td><td>You now have the option to revise your previous estimation based on your AI advisor’s suggestion. You previously selected [selected choice]. Your AI advisor suggests the success likelihood 70%. Please select your final estimation.
<table>
<tr><td>Leadership Ability</td><td>★★★☆☆</td></tr>
<tr><td>Proprietary Technology</td><td>★★★★☆</td></tr>
<tr><td>Market Condition</td><td>★★★☆☆</td></tr>
<tr><td>Competitor Strength</td><td>★★☆☆☆</td></tr>
</table>
a. 0%<br>b. 10%<br>c. 20%<br>d. 30%<br>e. 40%<br>f. 50%<br>g. 60%<br>h. 70%<br>i. 80%</td></tr>
</table>

<table>
<tr><td></td><td>j. 90%<br>k. 100%</td></tr>
<tr><td>Question 2</td><td>Given the provided company profile, indicate your estimation of the company’s success likelihood (probability) by selecting one of the intervals.
<table>
<tr><td>Leadership Ability</td><td>★☆☆☆☆</td></tr>
<tr><td>Proprietary Technology</td><td>★★★☆☆</td></tr>
<tr><td>Market Condition</td><td>★★★☆☆</td></tr>
<tr><td>Competitor Strength</td><td>★★★☆☆</td></tr>
</table>
a. 0%<br>b. 10%<br>c. 20%<br>d. 30%<br>e. 40%<br>f. 50%<br>g. 60%<br>h. 70%<br>i. 80%<br>j. 90%<br>k. 100%</td></tr>
<tr><td>Question 2: Revising with AI Advice</td><td>You now have the option to revise your previous estimation based on your AI advisor’s suggestion. You previously selected [selected choice]. Your AI advisor suggests the success likelihood 30%. Please select your final estimation.
<table>
<tr><td>Leadership Ability</td><td>★☆☆☆☆</td></tr>
<tr><td>Proprietary Technology</td><td>★★★☆☆</td></tr>
<tr><td>Market Condition</td><td>★★★☆☆</td></tr>
<tr><td>Competitor Strength</td><td>★★★☆☆</td></tr>
</table>
a. 0%<br>b. 10%<br>c. 20%<br>d. 30%<br>e. 40%<br>f. 50%<br>g. 60%<br>h. 70%<br>i. 80%<br>j. 90%<br>k. 100%</td></tr>
<tr><td>…</td><td>…</td></tr>
</table>