The Psychology of Managerial Capital Allocation

Shir Dekel

Faculty of Science

The University of Sydney

A thesis submitted to fulfil requirements for the degree of

Doctor of Philosophy (Science)

2021

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Shir Dekel

לאמא ואבא

Acknowledgements

I am immensely grateful for my supervisors Dr Micah Goldwater, Professor Dan Lovallo, and Dr Bruce Burns. Micah was there for me from the initial struggles for an idea, to the stress of unexpected findings, to the last-minute questions about wording minutiae. Thank you for the spontaneous hallway conversations and for prioritising me when it was necessary. Dan's work and thought inspired much of the thesis experiments. Thank you for passing on valuable insights from your mentors, for challenging me, and for opening so many doors. Bruce's incisive feedback gave me confidence in the experiment development and subsequent writing. Thank you for the attention to detail and for our extended conversations.

My family were critical to the thesis even before its inception. My parents were always a model for critical and creative thought and encouraged my curiosity. In a way, the PhD really began with my backyard inventions and experiments. I am also grateful for my sisters for our Friday night jams and for allowing me to always have someone I can strive to impress.

Thank you to everyone else who supported me. To Rachel, for the love, understanding, calls, runs, activities, and donut deliveries. To Nicky, Jacob, and Mich for being my second family throughout *our* thesis. To Yoav and Josh for being my spiritual housemates and for tolerating their measly monthly quota. And to those who made an effort to read a portion of the thesis and contributed helpful insights (and to David for actually doing so).

Two years ago, my coding knowledge consisted of what I learnt at the *HTML Basics* class that I took when I was nine years old. Since then, I wrote the text of my thesis in Emacs and developed almost everything else as reproducible code. This would have not been possible without the support of the online `R` community on GitHub, Stack Exchange, and Twitter.

Chapters 4, 5, and 6 of this thesis were edited by Elite Editing, and editorial intervention was restricted to Standards D and E of the Australian Standards for Editing Practice.

Thank you all.

Preface

Navigating the document on a computer

There are links throughout the PDF document that facilitate navigation between cross-references. For most PDF viewers these links are identified by dark blue text. Otherwise, if you are using Adobe Acrobat you may instead have to look for the cursor to turn into a hand pointer (). Clicking on these links will take you to the relevant hypothesis (e.g., Hypothesis 2.1), footnote,[1] citation (e.g., Kahneman & Tversky, 1979), figure (e.g., Figure 1.1), table (e.g., Table 2.1), or section (e.g., Chapter 1) that they reference.

If you are using Preview (on a Mac) then you can subsequently return back to where you clicked on the link by pressing the key combination ⌘+[(the command key with the left square bracket). In Adobe Acrobat the key combination is ⌘+← (the command key with the left arrow key) for Mac and Alt+← (the Alt key with the left arrow key) for Windows.

The sections of the thesis can also be navigated using the linked page numbers in the Contents section below. To display the Table of Contents as a sidebar in Preview go to View > Table of Contents and then click Show at the top right of the sidebar. Alternatively, you can use the key combination ⌥+⌘+3 (the option and command keys with the number 3). To display the Table of Contents as a sidebar in Adobe Acrobat go to View > Show/Hide > Navigation panes > Bookmarks.

Appendices

Many experiments were conducted throughout the development of this thesis. Further, each experiment included multiple measures and analyses, and not all of these were directly relevant for the thesis. Therefore, the main body of the text contains the content that was deemed most important, while three appendices contained the rest of the content. These appendices contain reports of supplementary experiments, experimental materials, additional measures, data simulations, power analyses, and extra explanatory material. The appendices are organised by the

[1] Example of a footnote.

relevant empirical chapter: Appendix A for Chapter 2, Appendix B for Chapter 4, and Appendix C for Chapter 6.

Reproducibility

The thesis used R (Version 4.0.2; R Core Team, 2020)[2] for the analyses and plotting of data, and for the generation of experimental materials. `rmarkdown` (Xie et al., 2018) was used with `bookdown` (Xie, 2016) to compile the document itself. `renv` (Ushey, 2021) was used to create reproducible environments and `targets` (Landau, 2021b) was used to create a reproducible pipeline. Typesetting was done with LaTeX, based on the `oxforddown` template (Lyngs, 2019). All the components required to reproduce this document can be found through this link: https://github.com/shirdekel/phd_thesis.

References

A chapter-specific reference list is included at the end of Chapters 1, 2, 4, 6, and 7. The complete list of references is reported after the appendices.

[2]Furthermore, the following R-packages were used *afex* (Version 0.28.1; Singmann et al., 2021), *aggregation1* (Version 1.0; Dekel, 2021a), *aggregation2* (Version 1.0; Dekel, 2021b), *aggregation3* (Version 1.0; Dekel, 2021c), *aggregation4* (Version 1.0; Dekel, 2021d), *alignment1* (Version 1.0; Dekel, 2021e), *alignment2* (Version 1.0; Dekel, 2021f), *alignment3* (Version 1.0; Dekel, 2021g), *alignment4* (Version 1.0; Dekel, 2021h), *alignment5* (Version 1.0; Dekel, 2021i), *alignment6* (Version 1.0; Dekel, 2021j), *alignment7* (Version 1.0; Dekel, 2021k), *alignment8* (Version 1.0; Dekel, 2021l), *anecdotes1* (Version 1.0; Dekel, 2021m), *anecdotes2* (Version 1.0; Dekel, 2021n), *broom* (Version 0.7.7.9000; Bolker & Robinson, 2020; D. Robinson et al., 2021), *broom.mixed* (Version 0.2.6; Bolker & Robinson, 2020), *conflicted* (Version 1.0.4; Wickham, 2019a), *cowplot* (Version 1.1.1; Wilke, 2020), *devtools* (Version 2.4.1; Wickham, Hester, et al., 2021), *dflow* (Version 0.0.0.9000; McBain, 2020), *dplyr* (Version 1.0.7.9000; Wickham, François, et al., 2021), *emmeans* (Version 1.5.4.9004; Lenth, 2021), *forcats* (Version 0.5.1; Wickham, 2021a), *ggbeeswarm* (Version 0.6.0; Clarke & Sherrill-Mix, 2017), *ggplot2* (Version 3.3.4; Wickham, 2016), *janitor* (Version 2.1.0; Firke, 2021), *knitr* (Version 1.33; Xie, 2015), *lme4* (Version 1.1.27; Bates et al., 2015), *magick* (Version 2.6.0; Ooms, 2021), *magrittr* (Version 2.0.1; Bache & Wickham, 2020), *Matrix* (Version 1.3.2; Bates & Maechler, 2021), *MOTE* (Version 1.0.2; Buchanan et al., 2019), *papaja* (Version 0.1.0.9997; Aust & Barth, 2020), *patchwork* (Version 1.1.1; Pedersen, 2020), *printy* (Version 0.0.0.9003; Mahr, 2021), *purrr* (Version 0.3.4; Henry & Wickham, 2020), *pwr* (Version 1.3.0; Champely, 2020b), *rlang* (Version 0.4.11.9000; Henry & Wickham, 2021), *rmdfiltr* (Version 0.1.3; Aust, 2020), *scales* (Version 1.1.1; Wickham & Seidel, 2020), *shiR* (Version 0.0.0.9000; Dekel, 2021o), *shirthesis* (Version 0.0.0.9000; Dekel, 2021p), *snakecase* (Version 0.11.0; Grosser, 2019), *stringr* (Version 1.4.0; Wickham, 2019b), *tarchetypes* (Version 0.2.0.9000; Landau, 2021a), *tibble* (Version 3.1.2; Müller & Wickham, 2021), *tidyr* (Version 1.1.3; Wickham, 2021b), *tinylabels* (Version 0.2.1; Barth, 2021), *usethis* (Version 2.0.1; Wickham & Bryan, 2021), and *yaml* (Version 2.2.1; Stephens et al., 2020).

Ethics

The research in this thesis was approved by The University of Sydney Human Research Ethics Committee (HREC).

Project No.: 2019/056

Project Title: Business decision making

Abstract

Capital allocation decisions are critical for large organisations. Management research mainly considers such decisions from an organisational perspective, largely overlooking potential psychological influences. Therefore, this thesis investigated cognitive processes that affect capital allocation decisions. Three studies examined how participants integrated multiple kinds of cues when making their decisions. Each study presented participants with both statistical information and non-numerical semantic information. In each study, participants had the opportunity to leverage a statistical concept that arguably should be the sole basis of the decision. The first study showed participants sequential risky choices without intermittent feedback. Participants could have combined the risk across decisions to reduce the overall potential loss. However, they struggled to do this unless it was depicted visually. The second study asked participants to allocate a budget across a set of business projects. Participants could have used the variance associated with the provided forecast estimates to choose which metrics to use for the allocation. However, they only appropriately used this information when it was expressed verbally and did not when it was expressed numerically. In the third study, participants saw projects with conflicting statistical and anecdotal evidence. The anecdotes were either similar or dissimilar to the target project. Participants could have clarified the conflicting evidence by using provided information about the distribution from which the anecdote was sampled. However, they ignored this information. Despite this, participants' use of the anecdote depended on its similarity to the target project. These results show that people's capital allocation decisions are bounded by a limited understanding of certain statistical concepts, but that they are capable of more nuanced choice when properly scaffolded.

Contents

List of Figures

List of Tables

List of Abbreviations

GE . . . . General Electric.

NPV . . . . . . Net present value.

EUT . . . . . . Expected utility theory.

MBA . . . . . Master of Business Administration.

EV . . . . Expected value.

ANOVA . . . . Analysis of variance.

Executive: A man who can make quick decisions and is sometimes right.

—Elbert Hubbard (1914, p. 52)

1
Introduction

Contents

Much of modern life depends on large organisations. General Electric (GE) makes the engines that power our aircrafts, Johnson & Johnson makes our shampoo, and Google allows us to search the internet. The areas of our lives that are less affected by such private firms depend on public sector organisations such as public hospitals, schools, and police. The justification for the existence of organisations of this size is that the particular combination of individual divisions, alongside a corporate management, will lead to better performance for each of the divisions than they would have been able to generate individually. In other words, the assumption is that such organisations create a synergy—the quality of the whole will be greater than the sum of its parts.

Multi-divisional organisations are typically organised in a hierarchical structure, with a corporate management team and subsidiary divisions. Each division can be made up of several business units. For instance, some of GE's divisions include GE Aviation and GE Healthcare. Similarly, in the public sector, a hospital system may operate through multiple individual hospitals in different regions.

Such organisations therefore need to make capital allocation decisions. That is, given a limited amount of financial resources, how best to invest in the multiple divisions? Equally? Pick a winner? What metric should be used to compare across divisions? Capital allocation is a critical process to the operation and development of multi-divisional organisations.

The products and services that arise from organisations are necessarily a result of the work of many people. In GE, for instance, the factories that generate aircraft engines need to be staffed by production line workers, accountants are needed for bookkeeping, and software engineers are needed to design and maintain the production systems. Despite this, many important strategic decisions ultimately come from a very small number of people. The decisions that the CEO or other lower level executives make can have large consequences on the life of the company.

It is often assumed that a few people having a lot of decision-making power is for the best. Managers of large organisations often appear to be bold and effective decision-makers. It appears that their position of power and wealth was necessarily arrived at through high competence and rational decision-making, suggesting that the organisation is in good hands. However, there are three reasons why it may be concerning that much of an organisation's future—and by extension often many more components of the economy—depends on the decisions of a few individual. First, the role of survivorship bias in obtaining the manager's role is unclear, because the number of managers that used the same management strategy and failed is unknown. Second, decades of research has shown that people's decision-making is often fallible and that job experience does not always alleviate this fallibility. Third, managers of large organisations often face uncertain environments, which increases the likelihood of managers facing psychological biases.

1. Introduction

There are many examples of companies that suffered due to such biases. Overconfidence and confirmation bias likely played a part in Blockbuster's famous refusal of an offer to buy Netflix in 2000 (Meissner et al., 2015). Further, Roxburgh (2003) identified how Equitable Life Assurance Society unnecessarily anchored on previous interest rate performance and was unprepared when rates changed. In an example of the sunk-cost fallacy, the London Stock Exchange continued investing in an automated settlement system even when it no longer remained profitable. The Bank of England needed to step in and stop the project. Overconfidence in market entry is also a common issue, illustrated by EMI's introduction into the medical-diagnostics market with the CT scanner (Camerer & Lovallo, 1999; Horn et al., 2005). By underestimating the competition and overestimating their own capabilities they eventually incurred losses and exited the market.

One class of biases has not been well studied: capital allocation biases. While some previous work investigating these biases exists (e.g., Bardolet et al., 2011), many questions still remain unanswered. This is a rather large hole in the literature because capital allocation decisions are at the centre of executive and lower level managers' roles. When making capital allocation decisions, there are elements of the decision-making environment that can be deceiving for managers. This thesis examines how the framing of a series of business projects affects people's decisions about those projects. Specifically, the same set of projects, presented in aggregate form, is much more likely to be accepted. Further, sometimes people are distracted by extraneous semantic information, such as the relative similarity of the options.

The results of the thesis show that although people in general make sensible decisions, they fail to use critical information to inform their decisions. Specifically, information about metric variance is ignored even when other metrics are available. Further, people seem to appropriately use statistical and anecdotal information based on relevance to the situation at hand, but ignore information about the sampling of the anecdote. Not appropriately using these kinds of statistical concepts has important financial consequences, discussed below.

All the experiments in the thesis use laypeople, except for one experiment. However, past work generally shows the same biases in managers and laypeople (with some showing more bias in managers, e.g., Haigh & List, 2005). Further, upcoming studies will directly test managers to determine any potential expertise effects.

Section 1.1 will explain how the capital allocation process functions in hierarchical organisations and why it is necessary to analyse such a process with a psychological approach. Section 1.2 reviews the literature on decision-making biases and how these may apply to capital allocation decisions. Section 1.3 will then summarise the rest of the thesis chapters.

1.1 Capital Allocation in Hierarchical Organisations

The purpose of a multi-divisional organisation is to generate more value than any of the individual divisions combined. The whole should be greater than the sum of its parts. Previous work suggests that this is achieved due to factors such as reduced transaction costs (Coase, 1937; Liebeskind, 2000; Teece, 1980, 1982; Williamson, 1981), shared resources (Barney, 1991; Wernerfelt, 1984), increased competitive advantage (Porter, 1980, 1985), increased monitoring (Gertner et al., 1994), and increased synergies (Barney, 1988). The underlying logic is the same: a multi-divisional organisation will be successful if it manages its divisions using processes and resources that are shared or, better yet, are complementary.

In order to successfully manage multiple units, large organisations developed a hierarchical structure. Bower (1970) identified three levels of the typical management hierarchy: business, division, and corporate. These are equivalent to front-line (or bottom), middle, and top level managers (Noda & Bower, 1996). Early theorists suggested that the strategy for the organisation's growth is driven completely by the top managers; the rest of the organisation simply enacts their proposals. However, Mintzberg and Waters (1985) emphasised the role of an emergent strategy, in which lower level managers affect change in the organisation's strategy. Other work proposed and found evidence for an iterated process in which a strategic context may be set by top managers, but business projects advanced by lower level managers also contribute to driving the strategy of the organisation (Bower, 1970; Burgelman, 1983; Noda & Bower, 1996).

The way that capital is allocated in an organisation is very important to its growth and longevity. This process is a part of the broader process of resource allocation. A *resource* can refer to many types of assets that an organisation owns, both tangible and intangible, of which capital is only one (Wernerfelt, 1984). The capital allocation process itself is an important driver of the strategic outcomes of an organisation (Bower, 1970; Bower & Gilbert, 2005), and as a result, is an important influence on its financial performance (e.g., Arrfelt et al., 2015; Bardolet et al., 2010). Sengul et al. (2019, p. 72) describe intra-firm capital allocation as "(i) a process of determination, comparison, and selection among multiple investment alternatives, (ii) taking place across organizational levels of the firm, and (iii) influenced and constrained by the external context in which the firm is situated." In capital allocation, business-level managers typically formulate project proposals, which their division managers then evaluate. The division managers then choose the projects to send for final approval with the corporate managers. The supply of available capital is also influenced by external sources such as investors, competitors, and customers. However, this thesis focuses on the comparison and selection processes that are relevant during business project evaluation.

Managers ultimately have only limited information about the projects that they evaluate. They typically have access to descriptive information about the investment and its known properties, but also are provided with financial metrics that estimate the returns on the investment. There are many such metrics; they usually attempt to encapsulate a trade-off between predicted future gains, present losses (in the form of the capital spent to pay for the investment), and opportunity costs. Examples include net present value (NPV), internal rate of return, return on investment, cost-benefit, and pay-back period. This thesis focuses on NPV, since it is one of the most frequently used metrics for project evaluation (Graham & Harvey, 2001; Graham et al., 2015; Remer et al., 1993). NPV is the difference between the money that a project is forecasted to make and the initial investment in its development (accounting for the time value of money), as shown in Equation (1.1):

$$\mathrm{NPV} = \sum_{t=0}^{n} \frac{R_t}{(1+i)^t}, \tag{1.1}$$

where t is the time of the cash flow, i is the discount rate, R_t is the net cash flow, and n is the total number of periods. NPV is a useful metric because simply knowing that it is positive suggests that the project that it describes should be profitable. Therefore, metrics such as these have a strong influence on the decision of the manager evaluating a project.

However, there are other influences on project evaluations other than the value of the financial metrics. For instance, politics within or outside the company can lead to situations in which a decision is based on social influence or even manipulation (Garbuio & Lovallo, 2017). Such influence is not necessarily negative; it may involve qualitative feedback from, for instance, a more senior manager (Thamhain, 2014). Research has also shown that the media can have a tangible influence on managerial decision-making (Bednar et al., 2013; B. Liu & McConnell, 2013). Other sources of influence are the organisational structures and incentives that are in place both externally (Kokkinis, 2019) and internally to the organisation (Rajan et al., 2000; Ullrich & Tuttle, 2004). Such dynamics have also been the subject of economic modelling investigations (Cavagnac, 2005; Ortner et al., 2017; Reichelstein, 1997). Project proposals might also be affected by certain approval structures. For instance, managers might submit overly-optimistic project proposals if they know that the corporate team only accepts projects with a certain minimum NPV forecast.

Another potential organisational influence on capital allocation is the extent of diversification present in an organisation. A diversified organisation is one that possesses different divisions that are unrelated in some way. Penrose (1959/2009, p. 96) defined it as such:

> a firm diversifies its productive activities whenever, without entirely abandoning its old lines of product, it embarks upon the production of new products, including intermediate products, which are sufficiently different from the other products it produces to imply some significant difference in the firm's production or distribution programmes.

Previous work found that organisations that are made up of more related divisions are more successful than those that are made up of unrelated divisions (Harrison et al., 1993; Rumelt, 1974; Shelton, 1988; Wernerfelt & Montgomery, 1988). This is also true within business divisions (P. S. Davis et al., 1992). However, *more* diversified firms have also been shown to be associated with profitability (Grant & Jammine, 1988). This is usually explained by the ability for such firms to avoid risk associated with any one market. Some of the discrepancy in diversification findings has been explained to be due to the specific measures used (Lubatkin & Shrieves, 1986). It may also be because most studies used Standard Industrial Classification (SIC) codes to measure diversification (e.g., Rumelt, 1974), whereas others operationalised it using other approaches (e.g., resource-based; Harrison et al., 1993).

The advantage that related organisations have had has been explained through *synergies* (Barney, 1988). That is, an organisation with two divisions that can use their resources to better one another are better off together than separately. The 1960s saw a general rise in mergers and acquisitions from executives seeking to diversify their organisations. However, doing so simply for the sake of increasing divisions, rather than an understanding of the possible synergies, leads to the organisation actually being worth less than the sum of its parts (known as a *diversification discount*; Lang & Stulz, 1994). In fact, many organisations that acquired other businesses to diversify subsequently end up divesting them (Porter, 1987). For instance, in 2018 Australian conglomerate Wesfarmers demerged its Coles division, a successful retailer. Since then, the share price for both companies has risen by approximately 62% and 32%, respectively (Boyd, 2021).

While much of the performance of an organisation depends on influences that are external to the individual managers (e.g., organisational, political), psychological factors are often also quite consequential. For instance, on the one hand, organisational factors such as relevant support teams and approval processes may influence capital allocation depending on the extent of an organisation's extent of diversification. On the other hand, psychological factors such as ability of managers to compare between business project proposals may also impact allocation differently depending on the organisation's diversification. It is likely to be more difficult for a manager to evaluate project proposals from two dissimilar divisions that it is to evaluate those from two similar divisions. The organisational influences discussed above often assume that the manager that is making the decisions acts rationally, as per traditional economic theory. However, surveys of executives show that CEOs and CFOs often rely on non-financial factors for capital allocation decisions (Graham et al., 2015). Executives in these surveys identified manager reputation and confidence as two of the most important factors for capital allocation decisions. Further, research in psychology has shown that cognitive biases can influence such capital allocation decisions. Section 1.2 discusses such biases and the relevant implications for the thesis.

1.2 The Psychology of Capital Allocation

Managers of large organisations are generally assumed to have a superior decision-making capability compared with non-managers. However, managerial decision-making involves many of the same processes that have been shown to be affected by psychological biases in the general population (Das & Teng, 1999; McCray et al., 2002; Schwenk, 1984). Further, an organisation's success ultimately depends on strategic decisions made by top level managers (Mazzolini, 1981). Therefore, despite early work attempting to analyse such decisions using a structured organisational analysis (e.g., Mintzberg et al., 1976), it is important to understand the potential influence of psychological biases on managerial decisions. Research in the field of behavioural strategy has started to do this (Powell et al., 2011).

1. *Introduction*

Psychological research has shown that people tend to make decisions that are inconsistent with neoclassical economic theory. For instance, expected utility theory (EUT; Friedman & Savage, 1948; von Neumann et al., 1944) assumed that people have complete information when making decisions. However, both laypeople and managers of organisations are limited in the amount of information that they have and their ability to use it (Cyert et al., 1956; Simon, 1955). Such inconsistencies with economic prescription are likely to have evolutionary origins, so are sure to be adaptive in certain environments (Bettis, 2017; Gigerenzer, 2008; Haselton et al., 2009). However, there are many situations in which such inconsistency with economic theory can have bad consequences.

Research has shown many ways in which the allocation of capital in an organisation can be influenced by psychological biases. For instance, Benartzi and Thaler (2001) found that people tend to allocate their retirement fund equally between the available options, regardless of their composition. This *naive diversification* bias was also found in capital allocation for hierarchical firms (Bardolet et al., 2011). Managers allocated capital equally across the available divisions in the firm, regardless of performance. Analysis of real companies found that this behaviour is damaging to firm performance because it means that lower performing business units get subsidised by higher performing units, which are not operating at their full potential (Arrfelt et al., 2015; Bardolet et al., 2010). Subsequent studies found that business unit size also matters; capital allocation to both the smallest and largest units is disproportionate to their actual profitability levels (Bardolet et al., 2017). This was attributed to a combination of naive diversification and political power effects.

Relatedly, people tend to continue expending capital into investments that appear to be failing (Staw, 1981). This *escalating commitment* is another way that psychological biases can influence allocation in an organisation. This pattern of decision-making is likely a consequence of the sunk cost fallacy, in which people avoid "cutting their losses" even when they know that they cannot recuperate an investment (Parayre, 1995).

Managers also do not always seem to seek profit maximisation. Shapira and Shaver (2014) offered managers and Master of Business Administration (MBA) students two investments from a hypothetical firm: one with the same expected returns as the average of the firm's current investments and one with lower returns than the firm's average returns. However, both investments were profitable, so to maximise firm profits both should be chosen. Instead, participants were more likely to only choose the first investment. It seems that the firm's average returns served as an anchor, so participants did not want to reduce the firm's average returns, regardless of profitability.

The way that information is presented can also influence allocations. For instance, Yates et al. (1978) showed that people's evaluations are sensitive to the level of detail in the information provided. They found that people devalued descriptions of university courses more when they had less detail. This may be relevant for managers evaluating project proposals. A proposal might appear more attractive simply due to the level of detail in it, even if the level of detail does not correspond with the quality of each proposal.

Further, people tend to be over-confident in their decisions and forecasts. This has been shown in laypeople (E. J. Langer, 1975; Mannes & Moore, 2013; Puri & Robinson, 2007; Soll & Klayman, 2004), as well as in IT professionals (McKenzie et al., 2008) and managers (Barone-Adesi et al., 2013; Kahneman & Lovallo, 1993; Lovallo & Kahneman, 2003). This is important for higher-level managers that evaluate project proposals because the metrics that rely on forecast estimates may be biased by the over-confidence of the lower-level manager that created the proposal. Further, the higher-level manager evaluating the proposal may in turn be over-confident about its prospects due to factors that are unrelated to the underlying value. Overconfidence is also seen when considering the success of projects in hindsight (Bukszar & Connolly, 1988; Christensen-Szalanski & Willham, 1991). This means that it is less likely that managers will be able to effectively learn from both past mistakes and successes due to the potentially erroneous belief that the outcome was anticipated.

Managers often create sensitivity analyses, estimating the worst case, best case, and most likely scenario for a forecast. However, these are likely to be anchored on past experiences that further the manager's existing beliefs. In fact, prior research has shown that people are poor at constructing subjective probability distributions (e.g., Alpert & Raiffa, 1982; Schaefer & Borcherding, 1973; Tversky & Kahneman, 1974; von Holstein, 1971). Therefore, this suggests that even if the lower-level managers that construct project proposals calibrate their forecasts so that they are not over-confident, they are still likely to provide inaccurate estimates of their degree of confidence.

The above summarises many of the currently known psychological biases related to capital allocation. This thesis focuses on three essential processes within the capital allocation process: (a) risky choice, (b) the comparison between diversified businesses, and (c) the influence of prior experience. Each of these is prone to separate biases, that are also interrelated. The subsequent subsections review the literature for these processes.

1.2.1 Risky Choice

Neoclassical theories such as EUT suggest that when faced with multiple risky options people should choose the option with the highest expected value (EV), all else being equal. This means multiplying the value of each option by its probability and comparing the resulting values (first documented in Pascal, 1670/1999). For instance, imagine being presented with the following two choices:

A) a gamble that involves a 50% chance gaining $200 and a 50% chance of losing $100; or

B) gaining/losing nothing.

In option A, the EV is calculated as $200 \cdot 0.5 - 100 \cdot 0.5 = 50$. Since the EV for option A (50) is higher than the EV for option B (0), EUT would suggest that option A should be chosen.

This basic principle was extended by Bernoulli (1738/1954), who suggested that a person's subjective value of money differs depending on their current wealth. This *diminishing marginal utility* suggests that the more money a person already has, the less value acquiring more money will have for him. For example, the experience of a rich man that finds \$10 on the street is very different to the experience of a homeless man that finds \$10 (Bradley, 2013). Even though \$10 was gained in both cases, \$10 has less value to a person that already has, for example, \$1,000, than for a person that initially only has \$10. This principle is usually modelled as an power function (with a fractional exponent).

Prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) challenged EUT by suggesting that people's subjective value of money does not depend on their state of wealth—it depends on a change of wealth from a reference point. This is important because people's subjective value of money is different depending if they are gaining or losing money. Specifically, losses have a stronger psychological impact than equivalent gains. This disparity is one of the most settled and consistent findings in psychology and economics, having been well-replicated (e.g, Ruggeri et al., 2020). The fact that losses loom more than equivalent gains for the vast majority of people is referred to as *loss aversion* (Kahneman & Tversky, 1979). This finding was the primary reason that Daniel Kahneman won the Nobel Prize in Economics in 2002 (Kahneman, 2003). Loss aversion has been found with small amounts of money in experimental settings (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) and with millions of dollars in corporate settings (Koller et al., 2012; Swalm, 1966). The effect has been found in young children (Harbaugh et al., 2001), the numerous disparate cultures in which it has been tested (Weber & Hsee, 1998), and even in capuchin monkeys (Chen et al., 2006). Furthermore, a neural basis for loss aversion was identified (Tom et al., 2007). Therefore, loss aversion is clearly central to human cognition and behaviour.

The function that represents the value of a prospect describes both loss aversion and diminishing marginal utility, as shown in Equation (1.2):

$$v(x) = \begin{cases} x^{\alpha} \text{ if } x \geq 0 \\ -\lambda(-x)^{\beta} \text{ if } x < 0, \end{cases} \tag{1.2}$$

where x is the possible outcome, λ represents the loss aversion coefficient, and α and β represent the diminishing marginal utility for gains and losses, respectively.

In other words, loss aversion means that losses have more impact than equivalent gains. In fact, the impact of loss aversion can be expressed even more precisely, as a measurement of the ratio of the slopes of the curve for gains and losses. This measure tells us the average amount that losses have more impact than equivalent gains. In a sequel to the original prospect theory paper, Tversky and Kahneman (1992) measured a median coefficient (λ) of 2.25 of loss aversion. This means that people respond to losses 2.25 times more than equivalent gains. Similarly, this paper measured a median exponent (representing diminishing marginal utility, α and β) of 0.88 for both gains and losses. This means that people discount money the more of it they have by a rate of $x^{0.88}$.

Figure 1.1 shows loss aversion as the function being steeper in the domain of losses than the domain of gains. It shows diminishing marginal utility by the slight curve of the function. Equivalent changes in actual wealth from the references point (x-axis) have different impacts on the changes' subjective value (y-axis). An increase in wealth ($x = 1$) brings about an equivalent increase of value ($y = 1^{0.88} = 1$). However, a decrease in the same amount of wealth ($x = -1$) brings about a decrease in value 2.25 times the value of the equivalent gain ($y = -2.25 \cdot (-(-1))^{0.88} = -2.25$).

This research is relevant to capital allocation because the project proposals that managers evaluate invariably involve an element of risk. Therefore, managers are likely to be affected by similar effects on risk that have been shown in laypeople. However, hierarchical organisations offer an even more complex situation. Lovallo et al. (2020) found that the risk profiles of lower-level managers are lower than those of the top managers. They suggest that this may be due to lower-level managers' loss aversion to accepting projects that may jeopardise their job. However, the top

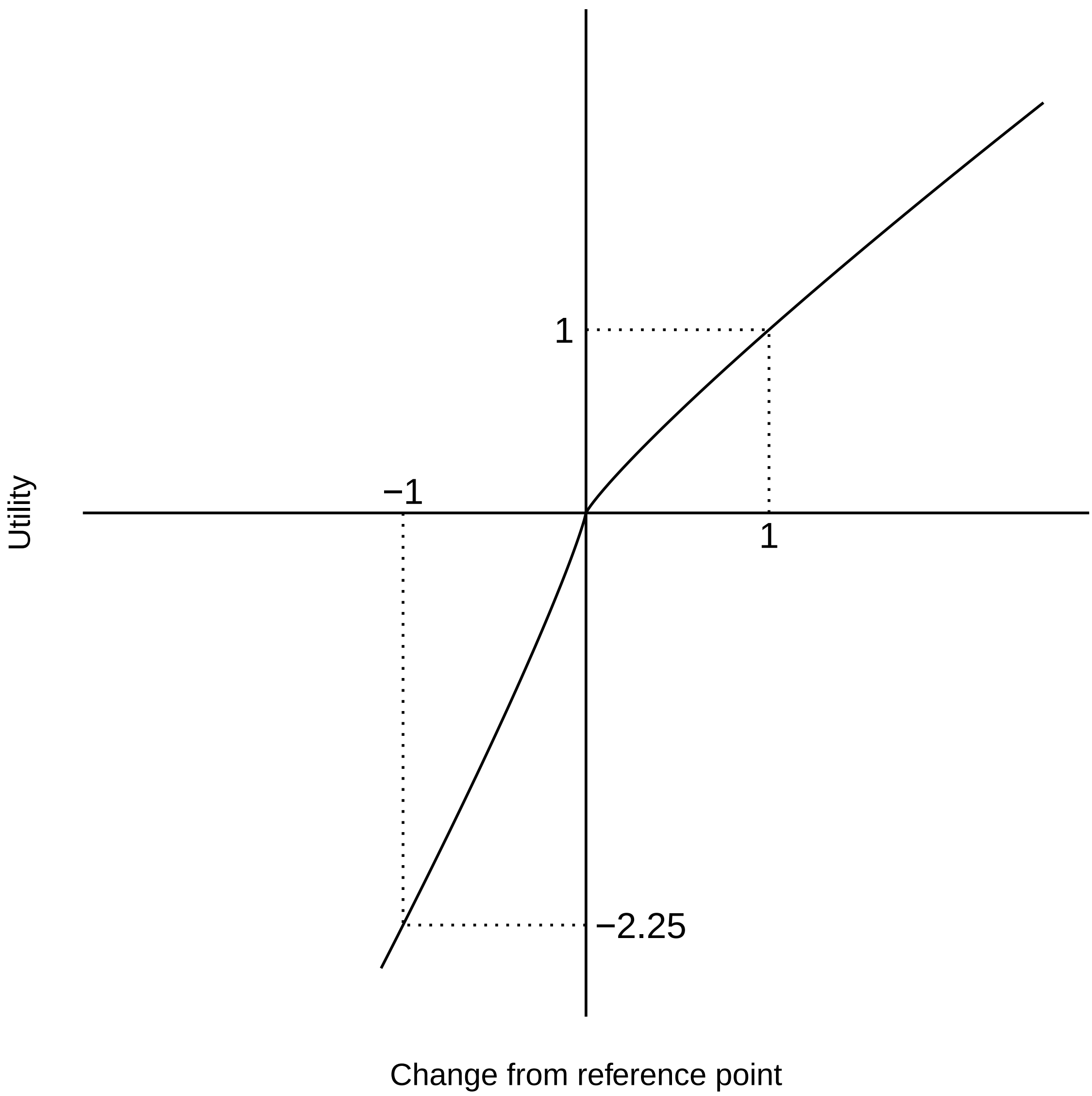


Figure 1.1: An example of the value function in prospect theory.

managers recognise that a loss in one or more business units is likely to be offset by gains in other units. Such an inconsistency in risk profiles across the levels of an hierarchical organisation fails to take advantage of the benefits of risk aggregation, which has long been understood in external markets (Markowitz, 1952).

Lovallo et al. (2020) suggested that lower-level managers' failure to aggregate risk to the degree desired by top executives is costing companies approximately a third of the total EV of new project proposals. This is an example of a negative consequence associated with ignoring statistical concepts such as risk aggregation. It is thus critical to identify ways to support risk aggregation across organisational hierarchies. The psychological literature shows that people's risk aggregation is facilitated through various choice bracketing manipulations. However, there has been no work that investigated such situations without providing participants with feedback in between decisions; this critically limits the external validity of this work because in the real world, organisations evaluate several projects before seeing the outcomes of any one decision. The experiments presented in Chapter 2 investigate the effects of choice bracketing on risk aggregation without feedback.

1.2.2 Project Similarity

When evaluating project proposals, managers are likely to be influenced by the relative similarity of the available options to each other. The extent to which this may be true is important especially since the increase firm diversification. Organisations are not only varied by the number of divisions which they possess but also by the extent of diversification. This means that managers are likely to find themselves comparing across dissimilar types of projects.

As mentioned above, there are likely many organisational and financial reasons why the extent of diversification in an organisation would impact its performance. However, the impact of psychological factors has not been investigated. Specifically, project similarity, which is an organisational factor, is likely to affect the project comparison process, which is a psychological factor. This may then have downstream consequences on firm performance through, for instance, the kinds of financial metrics that are used and how they are evaluated. Having more similar projects to compare may mean more attributes on which to evaluate, whereas a dissimilar comparison may lead to a situation in which a manager has to rely on potentially unreliable metrics.

Structure-mapping theory (Gentner, 1983; Gentner & Markman, 1997) provides a model of comparison that psychologically distinguishes similar and dissimilar allocation tasks. This framework models comparison as a process of bringing conceptual structures into alignment which, when possible, puts shared dimensions into correspondence. Alignment both highlights when two conceptual structures share dimensions, but also highlights how the two structures differ along those shared dimensions, called *alignable differences*. For example, when comparing two oil discovery projects, all the relevant processes of planning an exploration and measuring the amount of hydrocarbons in a prospect might be identical, but the specific amount measured will be different. This is the alignable difference: a difference between the two projects that is constrained within the same conceptual structure. However, when comparing between an oil field and a refinery, there will be significantly more *non-alignable differences*, because the two domains do not share component dimensions. That is, many of the processes that exist in the exploration business unit have a significantly different dimensional structure to those in the refinery business unit, such that it will be difficult to find meaningful alignments. More non-alignable differences mean that there are less opportunities to make meaningful comparisons, and so would make predicting relative project success and ranking their priority more difficult. Chapter 4 experimentally examines business project comparisons and how project alignment affects capital allocation decisions.

When evaluating projects, managers make use of financial metrics, such as NPV. However, such metrics are reliant on forecast estimates of, for instance, future cash flows. Do managers take into account such inherent variance in their decisions? This is especially important to investigate given the above discussion. In cases of non-alignable comparison managers may rely on a potentially unreliable metric. On the other hand, in an alignable comparison, managers might have the option to based their decisions on the relative reliability of different metrics. It is important to remember that all such decisions are often very consequential for the manager. That is, the project could ultimately make the company money and lead to future opportunities for the manager, or potentially cause financial harm to the company (and subsequently lead to a job loss). This is another example of the way in which ignoring certain statistical concepts—here metric variance—can have negative consequences for an organisation.

Psychological research shows that laypeople are in general quite poor at using numerical variance information (Batteux et al., 2020; Galesic & Garcia-Retamero, 2010; Konold et al., 1993; Vivalt & Coville, 2021). However, it is unclear to what extent managers would be sensitive to variance information in the metrics associated with the projects that they evaluate. On the one hand, perhaps managers' financial training will allow a consideration of such variance estimates, but this might not manifest in a situation in which managers have already been shown to be prone to biases. Chapter 4 investigates whether participants are as sensitive to verbally-instructed reliability information as they are to numerical reliability information.

1.2.3 Reasoning From Past Cases

Managers often use past events to reason and make predictions about the future (Einhorn & Hogarth, 1987). Such past events may be those that happened to the individual manager, a case from the organisation's history, or from an external source. This will especially be the case in a project evaluation scenario when a given project is hard to compare with the other projects at hand. However, managers evaluating project proposals may make inappropriate comparisons when considering the target project to other cases. For instance, people tend to limit the size of the comparison set to a small number. This is often only a handful of cases, or even one. Doing this might mean only considering potentially irrelevant surface similarity to the current situation and not aligning the underlying causal structure. Further, this might mean not considering other similar projects.

Tversky and Kahneman (1974) discussed a number of biases that may influence such processes. The availability bias is seen when people mistake the ease of retrieval of information for its frequency. Further, research on analogical retrieval showed that people are more likely to retrieve surface similar cases than those with a relational connection (Gentner et al., 1993). As such, managers are likely to recall cases that may not be sufficiently relevant to their target situation and be overly-confident about the frequency of such cases occurring. Such a focus on a particular case might then also lead to an anchoring effect, wherein other decisions might be disproportionately seen as relevant. Tversky and Kahneman (1974) also found that people are not sensitive to properties of sample size such as the greater amount of non-representative outcomes in small samples. This means that managers are even less likely to appreciate the importance of considering a large sample of cases when drawing conclusions to a target problem. Tversky and Kahneman (1974) also note an insensitivity to predictability, in which people do not take into account the reliability of the information that they have to make a prediction. This might mean that managers may struggle to ideally weigh evidence of varying degrees of reliability.

External sources that may be used to compare to a target situation include business case studies. Considering such examples of prior business decisions or events are the way that most MBA students learn about the business world. Publications such as Forbes or Harvard Business Review publicise various businesses' successes and failures and so may create an allure to use such case studies in the decision-making process. On the other hand, managers may have access to more aggregated data about their industry from, for instance, consultancy companies. How do managers use these various types of evidence in their decision-making?

Research on this topic suggests that managers tend to prefer anecdotes over statistics, unless aided (Wainberg, 2018). This is a concern because Gavetti et al. (2005) suggests that managers often make use of case studies quite poorly. The analogy literature draws a distinction between surface similarity, in which a mapping is made between easily identifiable but potentially functionally irrelevant attributes, and relational similarity, in which the underlying mechanism is considered. Are managers sensitive to the deeper causal mechanisms that underlie the anecdotes they judge? Or are they simply influenced by surface similarity? Chapter 6 investigates the extent to which people use anecdotes or aggregated data based on the relevance of the anecdote to the target project during capital allocation. It also considers whether people are sensitive to information about the distribution from which the anecdote was sampled. Ignoring this statistical concept can have negative consequences for an organisation by potentially over- or under-estimating the relevance of a past case and therefore making an ill-informed investment.

1.3 Chapter Overview

In sum, the potential consequences of a diversified hierarchical structure are that business projects will be considered one at a time, and if they are considered together, disparate project types will make comparisons hard. Considering projects one by one might mean that risk is not aggregated across projects and therefore value is lost. The difficulty to compare will lead to both potentially relying on unreliable metrics, and relying on improper anecdotal evidence. The thesis is that people often go half-way. They do not completely disregard the normative strategy, but also struggle to use statistical concepts such as aggregation, variance, and sampling.

The previous section identified three capital allocation processes that are currently under-studied and so are important to investigate further. First, the evaluation of individual project proposals may lead to managers only considering such projects one at a time, despite the opportunity of aggregating a portfolio of such projects. The choice bracketing literature suggests that there are ways of facilitating such aggregation, but does not investigate this without providing participants inter-trial feedback. Second, in situations in which managers compare multiple projects, the structural alignment literature suggests that managers in diversified firms will struggle to allocate capital, more than those in more integrated firms. Further, these managers may not be sensitive to the variance inherent in the financial metrics they rely on. Third, a difficulty to compare across existing projects may instead mean a reliance on prior case studies from personal or external experience. Research on anecdotal bias suggests that managers may rely more on such case studies than on aggregated data, but it is unclear whether their decisions will depend on anecdote relevance. Further, it is unclear if they will appropriately use information about the anecdote's sample distribution.

The rest of this thesis investigates the psychology of capital allocation decisions in three chapters that describe empirical work, two theoretical chapters, and a general discussion chapter. Chapter 2 describes two experiments that investigate the effects of choice bracketing on risk aggregation without feedback. Chapter 3 is a short theoretical chapter that discusses the difference between evaluating project proposals with inherent budget estimates and the process of allocating an existing budget top-down. Chapter 4 describes three experiments that investigate the effects of alignment and reliability type—verbal or numerical—on allocations. Chapter 5 is another short theoretical chapter that discusses the trade-offs that people make when using information to evaluate project proposal options. Chapter 6 describes two experiments that investigate the effects of anecdote similarity on the anecdotal bias. Finally, Chapter 7 discusses the theoretical and practical implications of the empirical chapters and concludes the thesis.

References

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty* (pp. 294–305). Cambridge University Press. https://doi.org/10.1017/CBO9780511809477.022

Arrfelt, M., Wiseman, R. M., McNamara, G., & Hult, G. T. M. (2015). Examining a key corporate role: The influence of capital allocation competency on business unit performance. *Strategic Management Journal*, *36*(7), 1017–1034. https://doi.org/10/gkhhc5

Bardolet, D., Brown, A., & Lovallo, D. (2017). The Effects of Relative Size, Profitability, and Growth on Corporate Capital Allocations. *Journal of Management*, *43*(8), 2469–2496. https://doi.org/10/gjvc4v

Bardolet, D., Fox, C. R., & Lovallo, D. (2011). Corporate capital allocation: A behavioral perspective. *Strategic Management Journal*, *32*(13), 1465–1483. https://doi.org/10/cn6xsb

Bardolet, D., Lovallo, D., & Rumelt, R. (2010). The hand of corporate management in capital allocations: Patterns of investment in multi- and single-business firms. *Industrial and Corporate Change*, *19*(2), 591–612. https://doi.org/10/ctxs4t

Barney, J. B. (1988). Returns to bidding firms in mergers and acquisitions: Reconsidering the relatedness hypothesis. *Strategic Management Journal*, *9*(S1), 71–78. https://doi.org/10/dmpq6x

Barney, J. B. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10/gpm

Barone-Adesi, G., Mancini, L., & Shefrin, H. (2013). *A Tale of Two Investors: Estimating Optimism and Overconfidence*. http://www.ssrn.com/abstract=2319260

Batteux, E., Bilovich, A., Johnson, S., & Tuckett, D. (2020, May 7). *Impressed by Numbers: The Extent to Which Novice Investors Favor Precise Numerical Information in a Context of Uncertainty* (SSRN Scholarly Paper No. ID 3595409). Social Science Research Network. Rochester, NY. https://doi.org/10.2139/ssrn.3595409

Bednar, M. K., Boivie, S., & Prince, N. R. (2013). Burr Under the Saddle: How Media Coverage Influences Strategic Change. *Organization Science*, *24*(3), 910–925. https://doi.org/10/gf33wh

Benartzi, S., & Thaler, R. H. (2001). Naive Diversification Strategies in Defined Contribution Saving Plans. *American Economic Review*, *91*(1), 79–98. https://doi.org/10/cfd4pb

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, *22*(1), 23–36. https://doi.org/10/cw729v (Original work published 1738)

Bettis, R. A. (2017). Organizationally Intractable Decision Problems and the Intellectual Virtues of Heuristics. *Journal of Management*, *43*(8), 2620–2637. https://doi.org/10/gf2r6w

Bower, J. L. (1970). *Managing the resource allocation process: A study of corporate planning and investment*. Division of Research, Graduate School of Business Administration, Harvard University.

Bower, J. L., & Gilbert, C. G. (Eds.). (2005). *From resource allocation to strategy.* Oxford University Press.

Boyd, T. (2021, June 3). *Wesfarmers CEO proves his worth.* Australian Financial Review. https://www.afr.com/chanticleer/wesfarmers-ceo-proves-his-worth-20210603-p57xmq

Bradley, D. (2013). Decision Theory, Philosophical Perspectives. In H. E. Pashler (Ed.), *Encyclopedia of the mind.* SAGE Publications, Inc. https://doi.org/10.4135/9781452257044.n81

Bukszar, E., & Connolly, T. (1988). Hindsight Bias and Strategic Choice: Some Problems in Learning From Experience. *Academy of Management Journal*, *31*(3), 628–641. https://doi.org/10/gjscqv

Burgelman, R. A. (1983). A Process Model of Internal Corporate Venturing in the Diversified Major Firm. *Administrative Science Quarterly*, *28*(2), 223–244. https://doi.org/10/dsvjgf

Camerer, C., & Lovallo, D. (1999). Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review*, *89*(1), 306–318. https://doi.org/10/bnkbkb

Cavagnac, M. (2005). Strategic managerial incentives under adverse selection. *Managerial and Decision Economics*, *26*(8), 499–512. https://doi.org/10/cwdbtp

Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006). How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior. *Journal of Political Economy*, *114*(3), 517–537. https://doi.org/10/fr7jw7

Christensen-Szalanski, J. J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*(1), 147–168. https://doi.org/10/fhnf22

Coase, R. H. (1937). The Nature of the Firm. *Economica*, *4*(16), 386–405. https://doi.org/10/fnm876

Cyert, R. M., Simon, H. A., & Trow, D. B. (1956). Observation of a Business Decision. *The Journal of Business*, *29*(4), 237–248. https://doi.org/10/fm8x9b

Das, T. K., & Teng, B.-S. (1999). Cognitive Biases and Strategic Decision Processes: An Integrative Perspective. *Journal of Management Studies*, *36*(6), 757–778. https://doi.org/10/c84dxp

Davis, P. S., Robinson, R. B., Pearce, J. A., & Park, S. H. (1992). Business unit relatedness and performance: A look at the pulp and paper industry. *Strategic Management Journal*, *13*(5), 349–361. https://doi.org/10/bx3wpx

Einhorn, H. J., & Hogarth, R. M. (1987). Decision making: Going forward in reverse. *Harvard Business Review*, *65*(1), 66–70. https://hbr.org/1987/01/decision-making-going-forward-in-reverse

Epictetus. (1995). *The Art of Living: The Classical Manual on Virtue, Happiness, and Effectiveness* (S. Lebell, Trans.). HarperCollins. (Original work published ca. 125)

Friedman, M., & Savage, L. J. (1948). The Utility Analysis of Choices Involving Risk. *Journal of Political Economy*, *56*(4), 279–304. https://doi.org/10.1086/256692

Galesic, M., & Garcia-Retamero, R. (2010). Statistical Numeracy for Health: A Cross-cultural Comparison With Probabilistic National Samples. *Arch Intern Med*, *170*(5), 462–468. https://doi.org/10/fmj7q3

Garbuio, M., & Lovallo, D. (2017). Does organizational politics kill company growth? *Review of International Business and Strategy*, *27*(4), 410–433. https://doi.org/10/gjscsw
Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*(8), 691–712. https://doi.org/10/b64gsr
Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10/dw52z8
Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45–56. https://doi.org/10/fm4rrb
Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The Roles of Similarity in Transfer: Separating Retrievability From Inferential Soundness. *Cognitive Psychology*, *25*(4), 524–575. https://doi.org/10/b7tzks
Gertner, R. H., Scharfstein, D. S., & Stein, J. C. (1994). Internal Versus External Capital Markets. *The Quarterly Journal of Economics*, *109*(4), 1211–1230. https://doi.org/10.2307/2118361
Gigerenzer, G. (2008). Why Heuristics Work. *Perspectives on Psychological Science*, *3*(1), 20–29. https://doi.org/10/d8h4qv
Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, *60*(2), 187–243. https://doi.org/10/fpdzrj
Graham, J. R., Harvey, C. R., & Puri, M. (2015). Capital allocation and delegation of decision-making authority within firms. *Journal of Financial Economics*, *115*(3), 449–470. https://doi.org/10/gfvz8d
Grant, R. M., & Jammine, A. P. (1988). Performance differences between the wrigley/rumelt strategic categories. *Strategic Management Journal*, *9*(4), 333–346. https://doi.org/10/fdh8xm
Haigh, M. S., & List, J. A. (2005). Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis. *The Journal of Finance*, *60*(1), 523–534. https://doi.org/10/c7jn9k
Harbaugh, W. T., Krause, K., & Vesterlund, L. (2001). Are adults better behaved than children? Age, experience, and the endowment effect. *Economics Letters*, *70*(2), 175–181. https://doi.org/10/db826t
Harrison, J. S., Hall, E. H., & Nargundkar, R. (1993). Resource Allocation as an Outcropping of Strategic Consistency: Performance Implications. *Academy of Management Journal*, *36*(5), 1026–1051. https://doi.org/10/gjscqw
Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive Rationality: An Evolutionary Perspective on Cognitive Bias. *Social Cognition*, *27*(5), 733–763. https://doi.org/10/c49fjq
Horn, J. T., Lovallo, D. P., & Viguerie, S. P. (2005). Beating the odds in market entry. *McKinsey Quarterly*, (4). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/hidden-flaws-in-strategy
Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *The American Economic Review*, *93*(5), 1449–1475. https://doi.org/10/b5gfrt

Kahneman, D., & Lovallo, D. (1993). Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking. *Management Science*, *39*(1), 17–31. https://doi.org/10/c8vntn

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10/g98

Kokkinis, A. (2019). Exploring the effects of the 'bonus cap' rule: The impact of remuneration structure on risk-taking by bank managers. *Journal of Corporate Law Studies*, *19*(1), 167–195. https://doi.org/10/gjspz9

Koller, T., Lovallo, D., & Williams, Z. (2012). Overcoming a bias against risk. *McKinsey Quarterly*, (4). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/overcoming-a-bias-against-risk

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in Students' Reasoning about Probability. *Journal for Research in Mathematics Education*, *24*(5), 392. https://doi.org/10/bq4hvm

Lang, L. H. P., & Stulz, R. M. (1994). Tobin's q, Corporate Diversification, and Firm Performance. *Journal of Political Economy*, *102*(6), 1248–1280. https://doi.org/10/dgjvg3

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32*(2), 311–328. https://doi.org/10/bhghpg

Liebeskind, J. P. (2000). Internal Capital Markets: Benefits, Costs, and Organizational Arrangements. *Organization Science*, *11*(1), 58–76. https://doi.org/10/dkmq6t

Liu, B., & McConnell, J. J. (2013). The role of the media in corporate governance: Do the media influence managers' capital allocation decisions? *Journal of Financial Economics*, *110*(1), 1–17. https://doi.org/10/gffggq

Lovallo, D., & Kahneman, D. (2003). Delusions of Success: How Optimism Undermines Executives' Decisions. *Harvard Business Review*, *81*(7).

Lovallo, D., Koller, T., Uhlaner, R., & Kahneman, D. (2020). Your Company Is Too Risk-Averse. *Harvard Business Review*, *98*(2), 104–111.

Lubatkin, M., & Shrieves, R. E. (1986). Towards Reconciliation of Market Performance Measures to Strategic Management Research. *Academy of Management Review*, *11*(3), 497–512. https://doi.org/10/ftmktj

Mannes, A. E., & Moore, D. A. (2013). A Behavioral Demonstration of Overconfidence in Judgment. *Psychological Science*, *24*(7), 1190–1197. https://doi.org/10/f443q6

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10/bhzd

Mazzolini, R. (1981). How strategic decisions are made. *Long Range Planning*, *14*(3), 85–96. https://doi.org/10/dm8852

McCray, G. E., Purvis, R. L., & McCray, C. G. (2002). Project Management under Uncertainty: The Impact of Heuristics and Biases. *Project Management Journal*, *33*(1), 49–57. https://doi.org/10/gjscsr

McKenzie, C. R., Liersch, M. J., & Yaniv, I. (2008). Overconfidence in interval estimates: What does expertise buy you? *Organizational Behavior and Human Decision Processes*, *107*(2), 179–191. https://doi.org/10/b22rw7

Meissner, P., Sibony, O., & Wulf, T. (2015). Are you ready to decide? *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/are-you-ready-to-decide

Mintzberg, H., Raisinghani, D., & Théorêt, A. (1976). The Structure of ”Unstructured” Decision Processes. *Administrative Science Quarterly*, *21*(2), 246–275. https://doi.org/10/b2t8ks

Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal*, *6*(3), 257–272. https://doi.org/10/bs2pm3

Noda, T., & Bower, J. L. (1996). Strategy making as iterated processes of resource allocation. *Strategic Management Journal*, *17*(S1), 159–192. https://doi.org/10/brg23k

Ortner, J., Velthuis, L., & Wollscheid, D. (2017). Incentive systems for risky investment decisions under unknown preferences. *Management Accounting Research*, *36*, 43–50. https://doi.org/10/gbz24d

Parayre, R. (1995). The strategic implications of sunk costs: A behavioral perspective. *Journal of Economic Behavior & Organization*, *28*(3), 417–442. https://doi.org/10/d2wg72

Pascal, B. (1999, June 10). *Pensées and other writings* (A. Levi, Ed.; H. Levi, Trans.). Oxford University Press. (Original work published 1670)

Penrose, E. (2009). *The Theory of the Growth of the Firm* (4. ed., rev. ed). Oxford University Press. (Original work published 1959)

Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press.

Porter, M. E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press.

Porter, M. E. (1987). From competitive advantage to corporate strategy. *Harvard Business Review*, *65*(3), 43–43. https://hbr.org/1987/05/from-competitive-advantage-to-corporate-strategy

Powell, T. C., Lovallo, D., & Fox, C. R. (2011). Behavioral strategy. *Strategic Management Journal*, *32*(13), 1369–1386. https://doi.org/10/d8m8pw

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, *86*(1), 71–99. https://doi.org/10/c9839j

Rajan, R., Servaes, H., & Zingales, L. (2000). The Cost of Diversity: The Diversification Discount and Inefficient Investment. *The Journal of Finance*, *55*(1), 35–80. https://doi.org/10/bdsh6h

Reichelstein, S. (1997). Investment Decisions and Managerial Performance Evaluation. *Review of Accounting Studies*, *2*(2), 157–180. https://doi.org/10/fp2zqc

Remer, D. S., Stokdyk, S. B., & Van Driel, M. (1993). Survey of project evaluation techniques currently used in industry. *International Journal of Production Economics*, *32*(1), 103–115. https://doi.org/10/bsc6bs

Roxburgh, C. (2003). Hidden flaws in strategy. *McKinsey Quarterly*, (2). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/hidden-flaws-in-strategy

Ruggeri, K., Alí, S., Berge, M. L., Bertoldo, G., Bjørndal, L. D., Cortijos-Bernabeu, A., Davison, C., Demić, E., Esteban-Serna, C., Friedemann, M., Gibson, S. P., Jarke, H., Karakasheva, R., Khorrami, P. R., Kveder, J., Andersen, T. L., Lofthus, I. S., McGill, L., Nieto, A. E., … Folke, T. (2020). Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour*, *4*(6), 622–633. https://doi.org/10/ggwrdh

Rumelt, R. P. (1974). *Strategy, Structure, and Economic Performance*. Division of Research, Graduate School of Business Administration, Harvard University.

Schaefer, R. E., & Borcherding, K. (1973). The assessment of subjective probability distributions: A training experiment. *Acta Psychologica*, *37*(2), 117–129. https://doi.org/10/dpzkfb
Schwenk, C. R. (1984). Cognitive simplification processes in strategic decision-making. *Strategic Management Journal*, *5*(2), 111–128. https://doi.org/10/b2hsgw
Sengul, M., Costa, A. A., & Gimeno, J. (2019). The Allocation of Capital within Firms. *Academy of Management Annals*, *13*(1), 43–83. https://doi.org/10/gfgvqj
Shapira, Z., & Shaver, J. M. (2014). Confounding changes in averages with marginal effects: How anchoring can destroy economic value in strategic investment assessments. *Strategic Management Journal*, *35*(10), 1414–1426. https://doi.org/10/gkhfb3
Shelton, L. M. (1988). Strategic business fits and corporate acquisition: Empirical evidence. *Strategic Management Journal*, *9*(3), 279–287. https://doi.org/10/fbhcgd
Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99. https://doi.org/10/dw3pfg
Soll, J. B., & Klayman, J. (2004). Overconfidence in Interval Estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 299–314. https://doi.org/10/dtpsvd
Staw, B. M. (1981). The Escalation of Commitment To a Course of Action. *Academy of Management Review*, *6*(4), 577–587. https://doi.org/10/fdqtjg
Swalm, R. O. (1966). Utility Theory–Insights into Risk Taking. *Harvard Business Review*, *44*(6), 123–136.
Teece, D. J. (1980). Economies of scope and the scope of the enterprise. *Journal of Economic Behavior & Organization*, *1*(3), 223–247. https://doi.org/10/fxcj3r
Teece, D. J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization*, *3*(1), 39–63. https://doi.org/10/bdwvvq
Thamhain, H. J. (2014). Assessing the Effectiveness of Quantitative and Qualitative Methods for R&D Project Proposal Evaluations. *Engineering Management Journal*, *26*(3), 3–12. https://doi.org/10/gjsm2q
Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science*, *315*(5811), 515–518. https://doi.org/10/bhtqgn
Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10/gwh
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10/cb57hk
Ullrich, M. J., & Tuttle, B. M. (2004). The Effects of Comprehensive Information Reporting Systems and Economic Incentives on Managers' Time-Planning Decisions. *Behavioral Research in Accounting*, *16*, 89–105. https://doi.org/10/b99r4n
Vivalt, E., & Coville, A. (2021, April 23). *How Do Policy-Makers Update Their Beliefs?* http://evavivalt.com/wp-content/uploads/How-Do-Policymakers-Update.pdf
von Holstein, C.-A. S. S. (1971). Two techniques for assessment of subjective probability distributions — An experimental study. *Acta Psychologica*, *35*(6), 478–494. https://doi.org/10/fgg6jn

von Neumann, J., Morgenstern, O., & Rubinstein, A. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Wainberg, J. S. (2018, January 1). Stories vs Statistics: The Impact of Anecdotal Data on Managerial Decision Making. *Advances in Accounting Behavioral Research* (pp. 127–141). Emerald Publishing Limited. https://doi.org/10.1108/S1475-148820180000021006

Weber, E. U., & Hsee, C. (1998). Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes Towards Perceived Risk. *Management Science*, *44*(9), 1205–1217. https://doi.org/10/c4rqqd

Wernerfelt, B. (1984). A Resource-Based View of the Firm. *Strategic Management Journal*, *5*(2), 171–180. https://doi.org/10/cm664p

Wernerfelt, B., & Montgomery, C. A. (1988). Tobin's q and the Importance of Focus in Firm Performance. *The American Economic Review*, *78*(1), 246–250.

Williamson, O. E. (1981). The Economics of Organization: The Transaction Cost Approach. *American Journal of Sociology*, *87*(3), 548–577. https://doi.org/10/bqzkpb

Yates, J. F., Jagacinski, C. M., & Faber, M. D. (1978). Evaluation of partially described multiattribute options. *Organizational Behavior and Human Performance*, *21*(2), 240–251. https://doi.org/10/br3dbv

Cultivate the habit of surveying and testing a prospective action before undertaking it. Before you proceed, step back and look at the big picture, lest you act rashly on raw impulse.

—Epictetus (ca. 125/1995)

2

Effect of Choice Bracketing on Risk Aggregation in Repeated-Play Gambles With no Feedback

Contents

2.1 Introduction

Investors know not to put all their eggs in one basket. Ever since work on modern portfolio theory (Markowitz, 1952), it has been clear that combining the risk of a set of individual investments reduces the overall risk of the portfolio of investments. But what about situations in which it is not clear that a set of investments fit together as a portfolio? Personal decisions such as buying a car or moving cities are typically evaluated independently, as are business decisions such as a farm investing in new cropping technology or a multi-business firm building a mine.

While these decisions are separated in time, they are often not so far apart that it is easy to learn from past outcomes (and sometimes the outcomes themselves are unclear). This is because the outcomes of large investments are often delayed. Therefore, the decision-maker cannot always use the knowledge of the returns of one investment when evaluating a subsequent investment. Any results that a farmer may identify from using a new technology will only become apparent after many seasons of use. Similarly, it will take many years for a multi-business firm to begin to estimate whether the output of a mine resulted in the expected return on investment. These are the decisions that this chapter investigates: sequences of large risky choices without immediate outcomes.

Risk aggregation is the combination of probability or variance information (or both) associated with certain outcomes for the purpose of understanding that information more comprehensively (Bjørnsen & Aven, 2019). However, the psychological literature suggests that this process may be difficult for people to use. Work on prospect theory (Kahneman & Tversky, 1979) suggests that people's evaluation of gambles does not conform to expected utility theory and is prone to framing effects. Specifically, people typically evaluate gambles one by one (Kahneman & Lovallo, 1993; Rabin & Weizsäcker, 2009; Tversky & Kahneman, 1981). Therefore, it is unlikely that people will be able to aggregate risk when they do not perceive a series of investments as a portfolio. So, what would encourage people to aggregate risk? The literature on *choice bracketing* (Read et al., 1999) shows that grouping a set of individual gambles together facilitates risk aggregation. Therefore, the current work provides two primary contributions. First, this work is the first to investigate the effect of choice bracketing on risk aggregation in independent gambles evaluated without immediate returns. Second, this work introduces novel choice bracketing manipulations.

The earlier work on risk aggregation essentially did the aggregating work for the participants. For example, experimenters provided participants with an outcome probability distribution, usually with an explicit indication to group the choices together, such as by asking for a single decision to be made on a set of identical gambles. Other work addressed the more realistic situation of a set of independent gambles. However, most of this work provided participants with the outcomes of their choices before the subsequent choice. In these paradigms participants experienced individual outcomes from the eventual outcome distribution of the gambles, meaning that aggregation was confounded with learning.

As mentioned above, in real life there is usually a significant delay between the choice a person or firm makes and the outcome of that choice, and there are likely to be several interim choices in the meantime. This is especially true for business executives, who would typically have to wait months or years before beginning to understand the consequences of their decision, and even then the outcome may be unclear. However, previous work did not investigate the effect of choice bracketing on risky choice without feedback. This is surprising, since choice bracketing is exactly the kind of process that should promote aggregation in these more realistic decisions. Therefore, this chapter investigated new ways of encouraging participants to bracket their risky choices, but with a paradigm that involves a series of independent choices without feedback. In this way, the paradigm is more isometric with real-life risky choice.

2.1.1 Multi-Play Gambles

Despite the difficulties of risk aggregation, people seem to aggregate "naively" when considering multiple gambles. Samuelson (1963) told of a colleague who rejected a gamble that involved a 50% chance of gaining \$200 and a 50% of losing \$100, despite the gamble's positive EV. That is, $200 \cdot 0.5 - 100 \cdot 0.5 = 50$. Rejection of a positive EV gamble out of fear of the possible loss is classic loss aversion. However, the same colleague said he would accept 100 plays of the same gamble. Samuelson argued that this choice is irrational.[1] Intuitively, it is clear that over the course of 100 gambles, the positive EV wins out, and a net loss of money is extremely unlikely. Samuelson's colleague was more risk averse when making a single decision about one gamble (a *single-play* gamble), than when making a single decision about multiple (in this case 100) identical gambles (a *multi-play* gamble).[2]

Wedell and Bockenholt (1994) replicated the Samuelson (1963) anecdote experimentally with a gamble involving a potential gain of \$100 and a potential loss of \$50. Participants accepted the multi-play gamble of 100 plays more than the single-play gamble. This effect has since been replicated with different outcomes and probabilities, both with hypothetical and real money. Some participants often require fewer than 10 plays of a previously rejected gamble in order to accept it (DeKay & Kim, 2005; Keren, 1991; Montgomery & Adelbratt, 1982; Redelmeier & Tversky, 1992). Other similar studies found a multi-play effect that was in the predicted direction but not significant (Barron & Erev, 2003; Benartzi & Thaler, 1999; Klos et al., 2005; T. Langer & Weber, 2001). Further, the effect is not seen when participants do not perceive gamble outcomes as fungible (DeKay, 2011; DeKay et al., 2006; DeKay & Kim, 2005) or when choice is continuous rather than discrete (Bristow, 2011).

[1] Other work suggests that it is consistent with expected utility theory, once certain assumptions are added (e.g., Aloysius, 2007; Ross, 1999). However, a normative discussion is out of the scope of the present work.

[2] This chapter uses the terminology for gamble types used in Bristow (2011), and Camilleri and Newell (2013).

However, multi-play effects are likely robust, since there is also evidence that such gambles reduce a variety of cognitive biases. These include common-ratio effects (DeKay et al., 2006; Keren, 1991; Keren & Wagenaar, 1987), preference reversals (Wedell & Böckenholt, 1990), ambiguity aversion (H.-H. Liu & Colman, 2009), and the illusion of control (Koehler et al., 1994). Participants are also more likely to use explicitly provided EVs in multi-play gambles (Li, 2003), show eye movements more congruent with an EV model than single-play gambles (Su et al., 2013), and judge multi-play gambles as riskier (Joag et al., 1990).

People prefer multi-play gambles that are displayed with an aggregated outcome distribution of those gambles than those without (Benartzi & Thaler, 1999; Coombs & Bowen, 1971; DeKay & Kim, 2005; Keren, 1991; Klos, 2013; T. Langer & Weber, 2001; Redelmeier & Tversky, 1992; Venkatraman et al., 2006; Webb & Shu, 2017). This is because these distributions present the probabilities of all the different possible outcomes, so very clearly show the rarity of a loss. Note that this does not seems to hold when returns are calculated as percentages, rather than fixed dollar amounts (Stutzer, 2013); and when participants do not perceive gamble outcomes as fungible (DeKay & Kim, 2005). However, when this effect is demonstrated, the multi-play gamble is usually set up such that its (binomial) outcome distribution shows a relatively low chance of losing any money and a very low chance of losing a lot of money. For instance, Figure 2.1 shows the outcome distribution of the Samuelson (1963) gamble played 10 times. Outcome distributions of this sort do the aggregating work for the participants, making the attractiveness of the multi-play gamble clearer. This work suggests that participants can comprehend and respond to aggregated risk, but that they struggle to compute the aggregation without external help.

2.1.2 Repeated-Play Gambles

Decisions in real life are usually sequential and rarely identical as in the multi-play paradigm (cf. Barron & Erev, 2003). That is, people tend to be confronted with individual choices whose outcomes and outcome probabilities are different

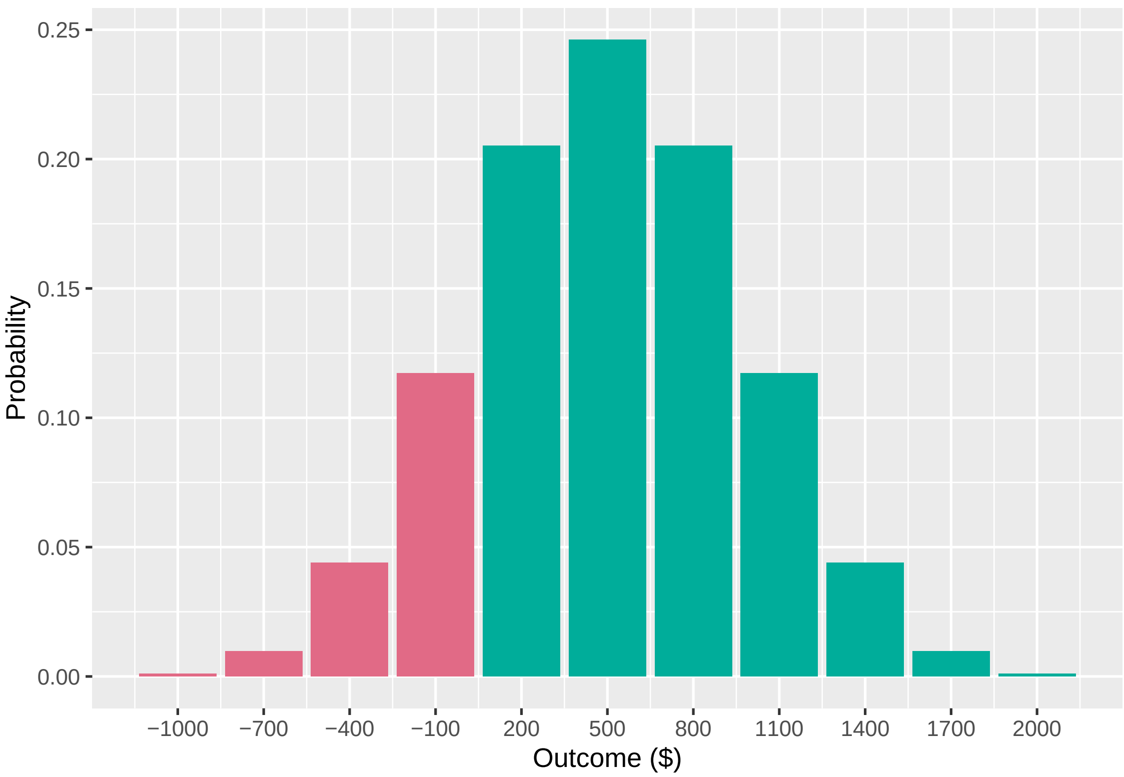


Figure 2.1: The outcome probability distribution of the Samuelson (1963) gamble (50% chance of gaining $200 and a 50% of losing $100) played 10 times. Green bars represent gains and red bars represent losses.

from one choice to another and these choices occur at different points in time. In a business setting this can be seen in decisions about whether to invest in new projects; proposals and opportunities differ widely and occur at different times. Managers are never simply asked: "here are 10 identical investments to consider; do you want all or none of them?"

In *repeated-play* (rather than multi-play) gamble paradigms, participants make decisions about a series of individual gambles. Research using this paradigm found that people are less risk averse both when outcomes for a series of gambles are evaluated less frequently and the subsequent decisions are made less frequently (Bellemare et al., 2005; Beshears et al., 2016; Gneezy & Potters, 1997; Thaler et al., 1997). People are also less risk averse (for positive EV gambles) when they receive feedback after each decision or are able to sample from the distribution of possible outcomes before making a choice (Barron & Erev, 2003; Camilleri & Newell, 2011, 2013; Hertwig et al., 2004; Jessup et al., 2008; Ludvig & Spetch, 2011; Wulff et al., 2018). Other work found that loss aversion is mitigated when people are explicitly instructed to consider the options as a part of a portfolio (Sokol-Hessner et al., 2012; Sokol-Hessner et al., 2009).

These studies are closer to real-life decisions than the multi-play gamble paradigm because they involve a set of separate gamble decisions rather than a single decision about a set of gambles. However, for the most part, the experiments used in the repeated-play gamble literature use various forms of feedback throughout the course of the experiment. That is, participants are shown the outcomes of their gambles before they make more decisions. This paradigm is known as *experience-based choice*. In *description-based choice*, on the other hand, the gamble is simply presented to the participant without any feedback, as in the multi-play gambles above. In real life, people rarely see the immediate outcomes of their risky choices, and even less so in business settings, where any return on investment often takes years to manifest.

Only a limited number of studies have used a repeated-play paradigm without feedback. For instance, Jessup et al. (2008) and Hertwig et al. (2004) investigated the effects of feedback in repeated-play gambles on the weighting of small probabilities, and had a no-feedback control condition. Other work similarly used individual description-based gambles presented sequentially (e.g., Ert & Erev, 2013; Joag et al., 1990). However, these studies did not attempt to facilitate participants' risk aggregation. Haisley et al. (2008) provided limited evidence for facilitating risk aggregation. They gave participants the opportunity to buy five (negative EV) lottery tickets, and either presented them one at a time, or together. Participants bought fewer tickets, when they considered them jointly, thereby maximising EV. However, the experimenters did not specify the outcomes and probabilities of each gamble, meaning that it is unclear if participants understood the independent lotteries as identical or non-identical. This reduces the external validity of the study, as most independent risky choice involves non-identical outcomes and probabilities. In sum, these studies were not designed to research how to facilitate risk aggregation and reduce loss aversion. The experiments in this chapter are novel because their goal is to facilitate risk aggregation without the experimental artefact of immediate feedback.

2.1.3 Choice Bracketing

Research in psychology and economics has identified ways of facilitating risk aggregation by encouraging people to group their choices. Specifically, people aggregate more when they consider the consequences of their choices together (broad bracketing) than when they consider them individually (narrow bracketing; Read et al., 1999). In multi-play gambles (especially when displayed with an outcome distribution), choices are inherently bracketed broadly because a single choice is made about multiple gambles. Similarly, studies that used repeated-play gambles facilitated risk-tolerance through what can in hindsight be considered broad bracketing. For instance, when Thaler et al. (1997) presented gamble outcomes less frequently, they allowed participants to consider longer time increments with a single evaluation.

Both the original Samuelson (1963) anecdote and its subsequent replications show that people do have an intuition for aggregation even without the risk being calculated exactly for them. This chapter tests whether that same intuition can be elicited and applied across sets of unique bets. What are the minimal conditions required to encourage aggregation? The multi-play gamble work suggests that participants can engage in a more intuitive form of aggregation when provided with the right contextual cues. Investigating the effects of more subtle cues will help shed light on the cognitive processes underlying choice bracketing. Of course, the effects of more subtle cues would not eliminate the utility of explicit financial education, but they will help the design of decision-making contexts to best align with such instruction.

One way of potentially facilitating risk aggregation is to highlight to participants the number of total options that are available to them. Sokol-Hessner et al. (2009) and Sokol-Hessner et al. (2012) reduced risk aversion using lengthy instructions that encouraged participants to "think like a trader". This meant considering all the repeated-play gambles as a portfolio, as opposed to considering them individually. However, this was quite a strong manipulation that is perhaps unrealistic in real world. A more subtle cue could involve simply making participants aware that they are going to be making a series of choices. If people possess an intuitive understanding of aggregation, as suggested above, then this kind of contextual cue will also facilitate aggregation.

In addition to simply informing participants that they will make a series of choices, making the choices more readily comparable may facilitate broad bracketing, and thus risk aggregation. Consider the inverse situation wherein a lack of comparability between choices may prevent broad bracketing, such as when an executive for a multi-business firm makes decisions across multiple distinct industries. Of course, the similarity of decision contexts does not change the maths of risk aggregation, but may well affect whether people do aggregate risk across decisions. DeKay and Kim (2005) found that multi-play effects are not seen when choices are not considered fungible. For instance, participants aggregated across dollar amounts, but not across patients in a medical decision. Therefore, people may behave similarly when considering a set of dissimilar choices if they do not consider them fungible.

There is further suggestive evidence that the similarity of a set of choices to one another will affect choice bracketing. Choices whose differences are easy to compare (alignable differences) are weighted heavier than those that are difficult to compare (Markman & Loewenstein, 2010; Markman & Medin, 1995). Increased similarity across a set of choices may both highlight the ability for those choices to be bracketed, and further facilitate risk aggregation through the comparable attributes. However, it is possible that increased similarity will facilitate risk aggregation even without a tangible benefit to the underlying calculations. That is, it is possible that simply manipulating the similarity of financially-irrelevant semantics of a choice set will make people less risk averse. If so, then this will be by virtue of an implicit risk aggregation in which the mere awareness of the possibility of a grouping of choices reduces risk aversion. It is important to investigate the effect of similarity especially because in managerial settings, executives in multi-business firms will often have to make comparisons across industries that are hard to compare. For instance, GE currently develops both analytic software products and jet engines for the military. They had been even more diversified previously, at one stage simultaneously developing home appliances and owning the NBC television network.

In addition to the similarity between choices, how choices are presented may affect how easily they are compared, and thus whether or not the multiple subsequent effects listed above would come to fruition. As mentioned above, Haisley et al. (2008) found a higher degree of EV maximisation when gambles were presented jointly, rather than separately. Similarly, Hsee et al. (1999) found that people's choices were affected by whether they viewed the attributes of the choices separately or jointly. Their *evaluability hypothesis* suggests that attributes that are difficult to evaluate will have a greater impact on joint presentation than separate presentation. Joint presentation is a form of broad bracketing because it forces a participant to view of all the components of a decision together. Participants may therefore be more likely to consider aggregating the risk involved in a set of choices when all those choices are in view. Joint presentation potentially reduces the working memory load otherwise needed to maintain that set of choices. Therefore, it is quite possible that a combination of highly similar choices, presented jointly will lead to the highest likelihood of broad bracketing, and thus risk aggregation.

Moher and Koehler (2010) replicated Gneezy and Potters (1997), but separately manipulated the number of gambles seen per trial and feedback frequency. They found that participants were less risk averse when viewing a set of three gambles per trial, than when viewing only one. However, they only found this effect with a set of identical outcomes. When outcomes were non-identical, there was no effect of presentation. However, participants were always presented with gamble outcomes for each trial, so it is unclear to what extent this influenced participants' ability to bracket broadly. In fact, when seeing gambles separately, participants were less risk averse when receiving feedback for each trial, compared to every three trials. Testing a presentation manipulation without the confound of feedback will help to clarify this effect.

2.1.4 Internal Capital Market Investment Context

Executives of large, successful firms are often viewed as fearless risk-takers who take on risky projects to generate innovation and growth. However, the available evidence suggests that executives do not view themselves that way (March & Shapira, 1987; Swalm, 1966). Executives typically evaluate multiple investments over time. Risk aggregation is sensible when investments are only partially correlated (i.e., the success of one does not influence the success of another). It is sensible to take on a set of risky investments with positive EV, where each investment has some chance of loss, because those that succeed will make up for those that failed. These benefits are well-known in stock market investment settings, thanks to Nobel laureate Harry Markowitz's work on modern portfolio theory (1952).

However, it is unclear whether the general public and even business managers use this concept, due to the extent of risk aversion in both those populations (e.g., March & Shapira, 1987; Tversky & Kahneman, 1992). In fact, executives treat risk like the rest of us; they view investments one at a time, are risk averse in the domain of gains, and are risk seeking in the domain of losses (Lovallo et al., 2020; MacCrimmon et al., 1986; Swalm, 1966). However, it is understandable why risk aggregation is foreign to most people; outside of an investment portfolio selection situation, it is unlikely for people to spontaneously group a selection of individual risky choices. Usually in life, people encounter risky choices sequentially, and so the risk of each individual choice is more salient than the aggregated risk of an arbitrary combination of choices.

Lovallo et al. (2020) showed that executives treat investments within their own company in isolation. In multi-business firms, the managers of each business unit often make the investment decisions about individual projects. Therefore, they often do not consider the scope of their decisions in the context of the entire company. For instance, Nobel laureate Richard Thaler offered 25 division managers working for the same firm a hypothetical investment that involves a 50% chance of gaining $2 million for the company and a 50% chance of losing $1 million (1999). Only three managers said they would accept the investment. However, the CEO indicated that he would have clearly preferred managers to accept all the investments. To each middle-manager, the choice represents a risk of loss for their division and potentially their job, whereas for the CEO the entire portfolio of choices represents a worthwhile risk.

This chapter investigates risky choice in the context of business project investment internal to a company because this is a real-world context where choice bracketing is important and currently under-appreciated (Lovallo et al., 2020). The participants in these experiments were taken from a population that does not have extensive managerial experience. However, in such a population a lack of risk aggregation is most likely more common, and the variables used here are readily applicable to the financial decisions that laypeople make. For instance, one of the real-world applications of the choice bracketing literature has been to use outcome distributions and increased time horizons to encourage investment in high risk, but high EV, retirement funds (e.g., Benartzi & Thaler, 1999). Otherwise, people typically prefer low risk, low EV, funds. Further, using laypeople eliminates potential differences in prior experience with the management-based decision-context. Upcoming research will focus on managers with context-specific experience to investigate the effects of that experience.

2.2 Experiment 1

Experiment 1 investigated the effect of three choice bracketing manipulations on risky choice in hypothetical capital allocation scenarios. Previous research had low ecological validity because of the use of multi-play paradigms or feedback. In this experiment, the risky choice task was a description-based repeated-play paradigm. This means that participants had to make a choice about whether to accept a number of different hypothetical investments, but were not provided with the outcome of their choices after each decision. The variables of interest were the similarity of the choices, whether the choices were presented together or separately, and whether participants were aware of the number of choices that they would be making.

The values and probabilities of the gambles were set up such that each individual gamble, as well as the aggregation of all the gambles, would be attractive to a rational agent interested in maximising EV. As such, the key dependent measure was the proportion of risky choices participants accepted.

Previous research suggests that people will be willing to make more risky choices when explicitly told to bracket their choices (Sokol-Hessner et al., 2012; Sokol-Hessner et al., 2009). Therefore, Experiment 1 tested the following hypothesis:

Hypothesis 2.1—awareness main effect. Participants that know how many projects to expect will make more risky choices than participants that are unaware.

Further, previous work suggests that joint presentation is a form of broad bracketing (e.g., Hsee et al., 1999; Moher & Koehler, 2010). Therefore, Experiment 1 tested the following hypothesis:

Hypothesis 2.2—presentation main effect. Participants will make more risky choices when seeing projects jointly than when seeing them separately.

Similarity of options has also been shown to affect the way people bracket their choices (e.g., DeKay & Kim, 2005). Therefore, Experiment 1 tested the following hypothesis:

Hypothesis 2.3—similarity main effect. Participants that see projects from the same industry will make more risky choices than participants that see projects from different industries.

Table 2.1: Experiment 1 group allocation.

Similarity	Awareness	N
High	Aware	53
High	Naive	53
Low	Aware	47
Low	Naive	45
Total		198

2.2.1 Method

2.2.1.1 Participants

One hundred and ninety-eight participants (82 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 32.52 years (*SD* = 11.42, *min.* = 18, *max.* = 69). Participants reported an average of 7.01 years (*SD* = 9.1, *min.* = 0, *max.* = 42) working in a business setting, and an average of 1.7 years (*SD* = 2.85, *min.* = 0, *max.* = 20) of business education. The mean completion time of the task was 12.04 min (*SD* = 11.29, *min.* = 3.1, *max.* = 112.4). Table 2.1 shows the allocation of participants to the different conditions.

2.2.1.2 Materials

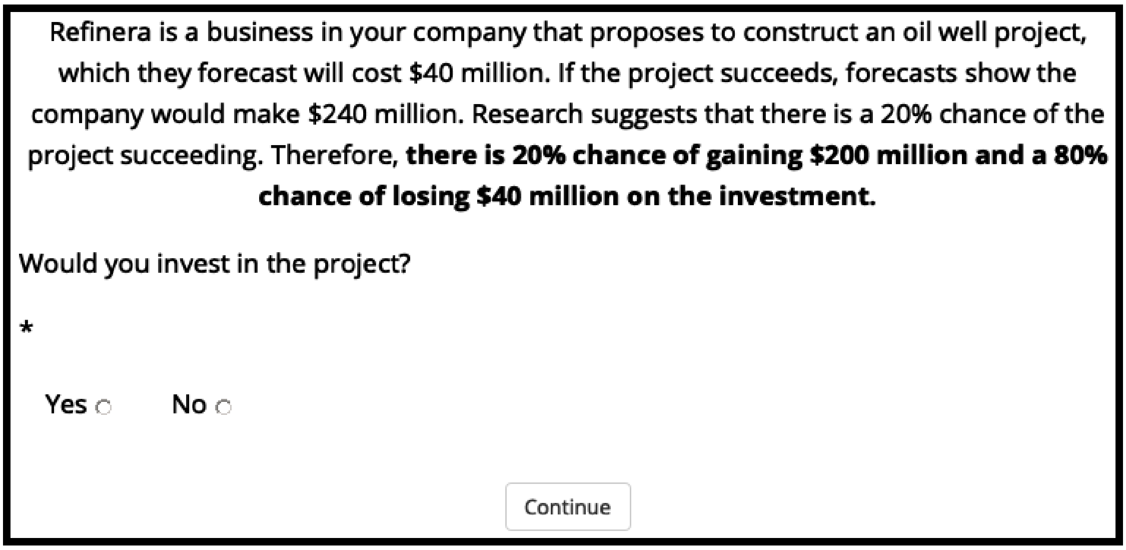
2.2.1.2.1 Instructions Participants were told to imagine that they are executives in a large company and that they will need to decide about investing in a number of hypothetical business projects. The appendix shows these instructions in Figure A.1.

2.2.1.2.2 Risky Investment Task Participants saw 10 short descriptions of business projects, and were asked whether they would invest in that project or not. Each description included the name of the hypothetical business, the amount they forecast the project to cost, the amount the project is forecast to make, and probabilities for these forecasts. The project values were selected so that the projects appeared attractive when aggregated, and unattractive when segregated (see T. Langer & Weber, 2001). These values were different for each project, but followed a set of constraints for each project's EV and the probability of any loss given the outcome distribution of all 10 projects ($P(\text{loss}_{aggregated})$). Further, there was a constraint on the gambles' loss aversion coefficient (λ), which is a measure of people's sensitivity to losses compared to gains. The constraints were:

1. EV > 0;
2. $\lambda < 2.25$; and
3. $P(\text{loss}_{aggregated}) < 0.1$.

As such, each project cannot be considered to be a loss in terms of expected value, but also would not be an easy choice for investment, because of the low λ (made to be lower than the median loss aversion coefficient calculated in Tversky & Kahneman, 1992). Further, since people are especially sensitive to loss probabilities (Kahneman & Tversky, 1979; Zeisberger, 2020), an arbitrarily low $P(\text{loss}_{aggregated})$ was chosen to make investment in the complete set of projects seem attractive. The actual probability of a loss given the outcome distribution used in the experiment was 0.09. This was calculated by summing all probabilities in the Poisson binomial distribution whose outcomes were less than zero. For comparison, $P(\text{loss}_{aggregated})$ = 0.17 for 10 plays of the Samuelson (1963) gamble. The highest probability of a loss for any single gamble ($P(\text{loss}_{single})$) was 0.80. Figure 2.2 shows an example of a description of a project in this task.

In the high similarity condition, these project descriptions were all about one type of project (in this case an oil well project) and were all from the same business. In the low similarity condition, each project was from a different industry. In the joint presentation condition, the 10 projects were all displayed on the one webpage, whereas in the separate presentation condition each was displayed on a different webpage. Participants in the aware condition saw the display shown in Figure 2.3 before their separate presentation display. Those in the naive condition simply



Refinera is a business in your company that proposes to construct an oil well project, which they forecast will cost $40 million. If the project succeeds, forecasts show the company would make $240 million. Research suggests that there is a 20% chance of the project succeeding. Therefore, **there is 20% chance of gaining $200 million and a 80% chance of losing $40 million on the investment.**

Would you invest in the project?

*

Yes ○ No ○

Continue

Figure 2.2: Example of a project choice display in Experiment 1.

You will now see 10 projects. Decide whether you would like to invest in each one.

Continue

Figure 2.3: The display seen by those in the aware condition of Experiment 1.

proceeded without this message. The financial and probability values were identical regardless of condition, and the order of each set of 10 projects was randomised.

Although the project descriptions were succinct, and the decisions in the task were made quickly, they reflect real decisions in businesses in critical ways. Companies that consider their forecast estimates probabilistically (i.e., do not simply use the most likely estimate as the only estimate) do frame their options as likelihoods of certain monetary outcomes.

2.2.1.2.3 Outcome Distribution Decision Participants were asked if they would invest in the last 10 projects they saw and were provided with a graph of the outcome probability distribution of the 10 projects. Figure A.2 shows this graph. A coding error was discovered after collecting data. This was an error in the generation of gambles, which meant that the outcome distribution decision data could not be used. Therefore, the effect of outcome distribution will not be discussed until Experiment 2, which fixed this issue. Appendix A.1.2.2 presents an analysis of these data, and describes the coding error and its implications.

2.2.1.2.4 Follow-up Gambles Participants were shown four further sets of gambles (11 total) that checked participant attention and replicated the gambles from Samuelson (1963) and Redelmeier and Tversky (1992). See Appendix A.1.1.1.3 for details.

2.2.1.3 Procedure

Participants read the instructions and completed the risky investment task, first in the separate presentation condition, and then in the joint condition. They then made the outcome distribution decision and responded to the 11 follow-up gambles.

2.2.2 Results

2.2.2.1 Project Choice

A three-way analysis of variance (ANOVA) was conducted to investigate the effects of similarity, awareness, and presentation on the proportion of participants' decision to invest in the 10 projects. As seen in Figure 2.4, participants invested more when they were told that there will be 10 projects, compared with when they were not told this, $F(1, 194) = 9.52$, $p = .002$, $\hat{\eta}_p^2 = .047$. As seen in Figure 2.5, participants invested more when viewing the projects jointly, compared with when they viewed them separately, $F(1, 194) = 28.14$, $p < .001$, $\hat{\eta}_p^2 = .127$. Although there was no main effect of similarity, $F(1, 194) = 1.63$, $p = .204$, $\hat{\eta}_p^2 = .008$, the interaction between similarity and presentation was significant, $F(1, 194) = 4.31$, $p = .039$, $\hat{\eta}_p^2 = .022$ (see Figure 2.6). Specifically, the presentation effect was stronger in the high similarity condition, $\Delta M = 0.07$, 95% CI [0.04, 0.09], $t(194) = 5.29$, $p < .001$, than in the low similarity condition, $\Delta M = 0.03$, 95% CI [0.00, 0.05], $t(194) = 2.06$, $p = .041$. These findings suggest that it is possible to facilitate risk aggregation with subtle choice bracketing manipulations.

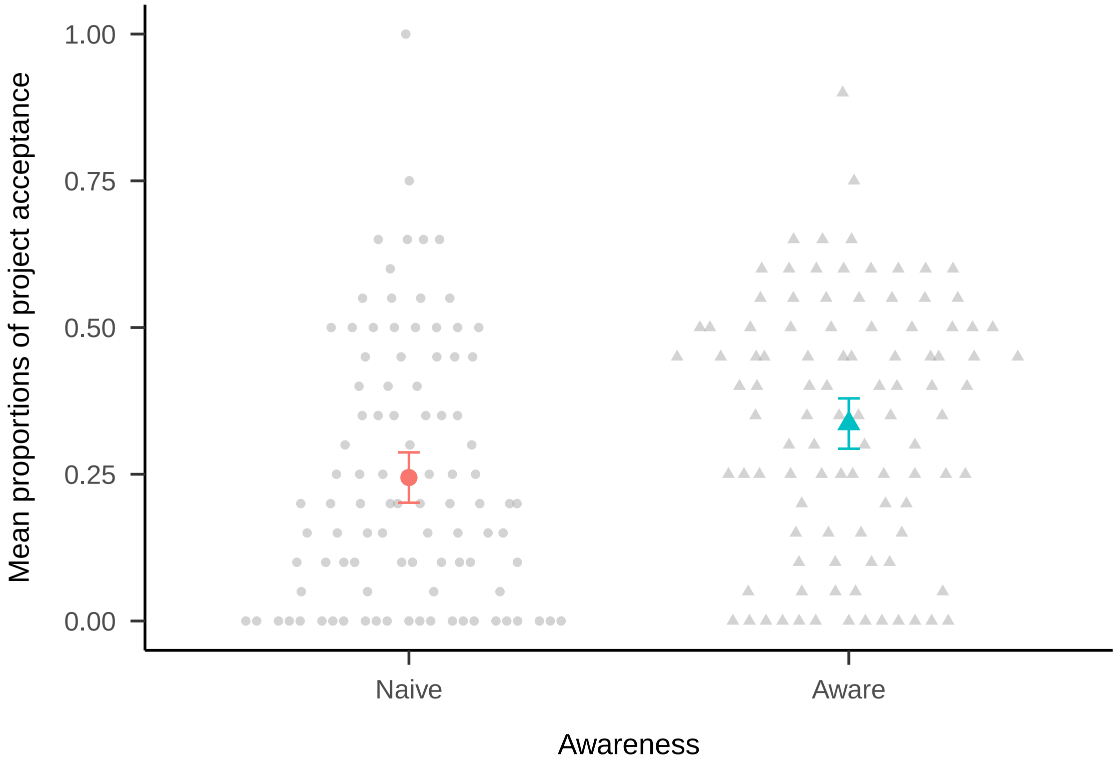


Figure 2.4: Mean proportions of decisions to invest in each set of 10 projects, by awareness condition. Error bars represent 95% confidence intervals. Raw data are plotted in the background.

2.2.2.2 Trial-by-Trial Analysis

Exploratory analyses were conducted into the possible effects of the manipulations on a trial-by trial basis. Figure A.3 shows the data for all conditions. However, the key findings are in the separate presentation. As Figure 2.7 shows, in the separate condition people are more likely to accept projects over the 10 trials, but this interacts with awareness, $b = 0.04$, 95% CI $[0.01, 0.08]$, $z = 2.32$, $p = .021$. Specifically, the relationship between choice and trial is stronger in the aware condition, $b = 0.11$, 95% CI $[0.06, 0.16]$, $z = 4.54$, $p < .001$, than in the naive condition, $b = 0.03$, 95% CI $[-0.03, 0.08]$, $z = 1.01$, $p = .311$. It seems that participants that were told the total number of projects became less risk averse as the experiment proceeded, regardless of the gamble values.

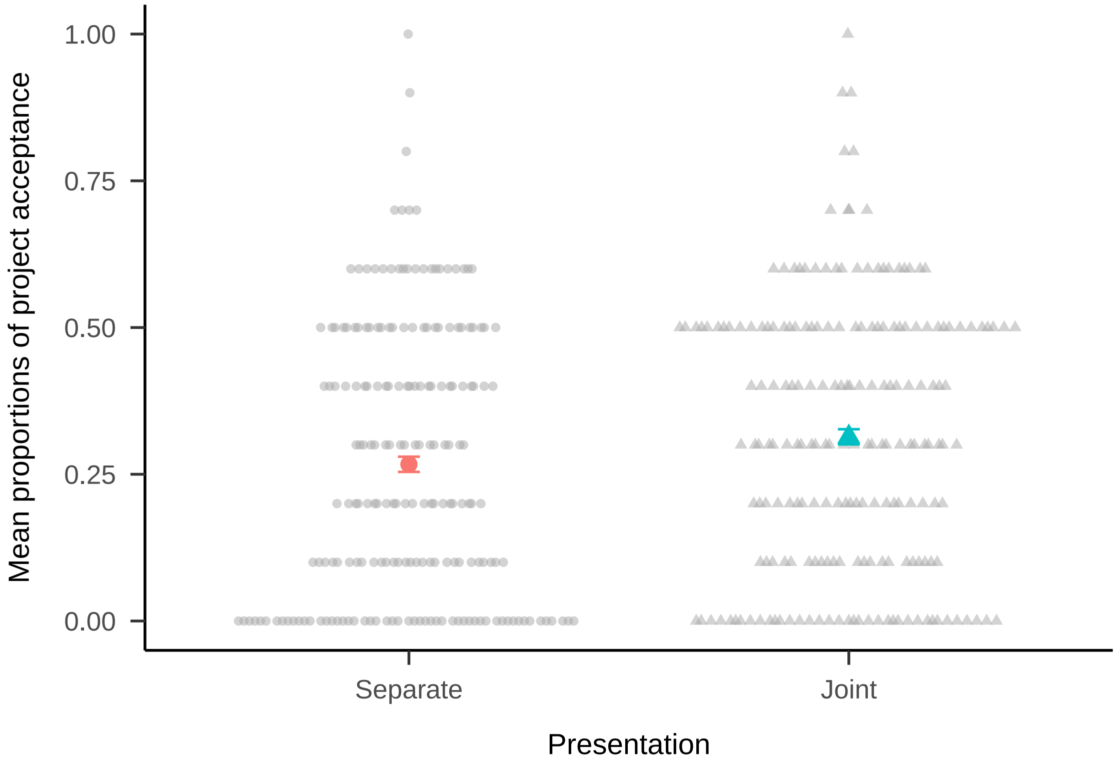


Figure 2.5: Mean proportions of decisions to invest in each set of 10 projects, by presentation condition. Error bars represent 95% confidence intervals. Here, however, the intervals are so narrow that they are sometimes obscured by the mean indicators in the plot. Raw data are plotted in the background.

2.2.3 Discussion

Experiment 1 found evidence for most of the hypotheses. Specifically, people made more risky choices when considering those choices jointly on the same page, compared to on separate pages; and when they knew how many choices were in the set. Further, the results showed an interaction between project similarity and presentation. Exploratory analyses showed that participants' risk aversion decreased as they proceeded through the trials, but only when participants were aware of the number of projects.

2.2.3.1 Presentation Effect

The presentation effect may be a result of one of two mechanisms. A mathematical aggregation explanation would mean that participants are combining the gambles into a mental representation of the probability distribution and then

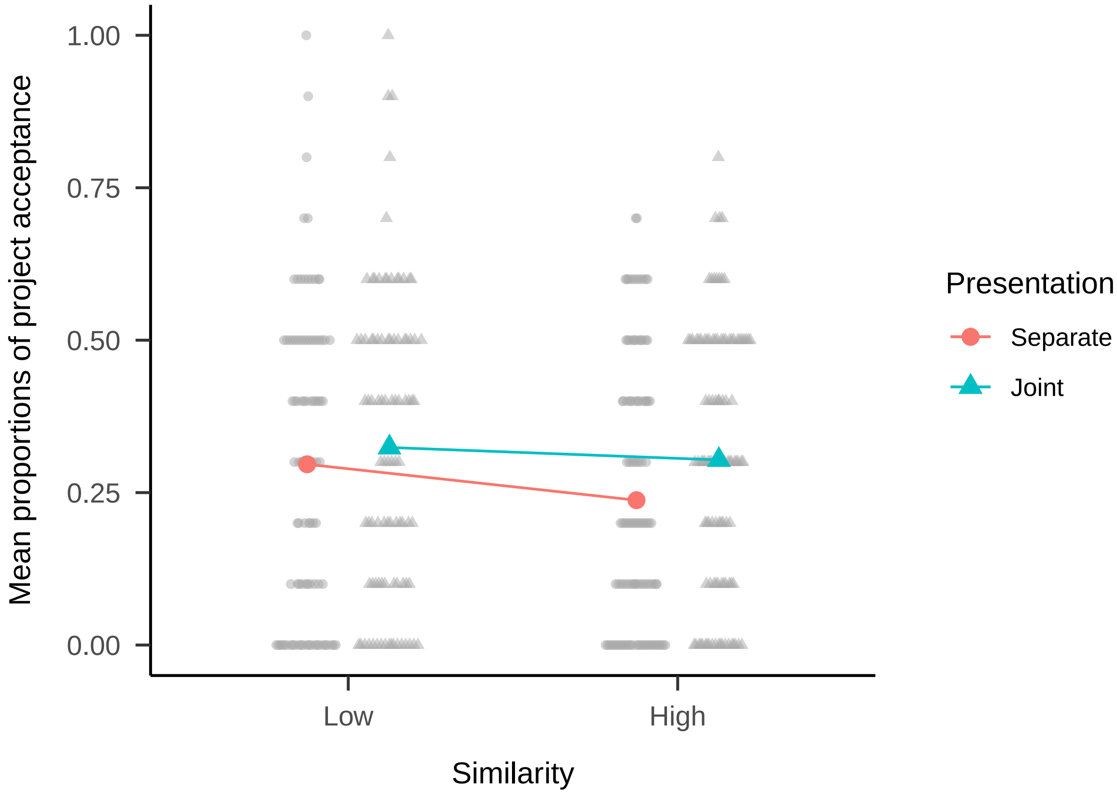


Figure 2.6: Mean proportions of decisions to invest in each set of 10 projects, by similarity and presentation conditions. In mixed factorial designs, error bars cannot be used to make inferences by "eye" across all conditions. Therefore, error bars are not included. Raw data are plotted in the background.

deciding based on the attractiveness of that distribution. A joint presentation of choices would facilitate this combination. On the other hand, people may also be using a sort of naive aggregation process when they are encouraged to group their choices together. A naive aggregation explanation would suggest that participants in the joint condition are simply more likely to realise that a few big wins could offset a few losses. Participants could have been encouraged by the joint display to consider the set of projects together. This could then lead to the conclusion that investing in a higher number of gambles might mean that the gains of some projects will pay off the losses of the other projects.

2.2.3.2 Awareness Effect

Experiment 1 found that participants that viewed the projects separately were more likely to invest in the projects as the trials went on, regardless of the actual

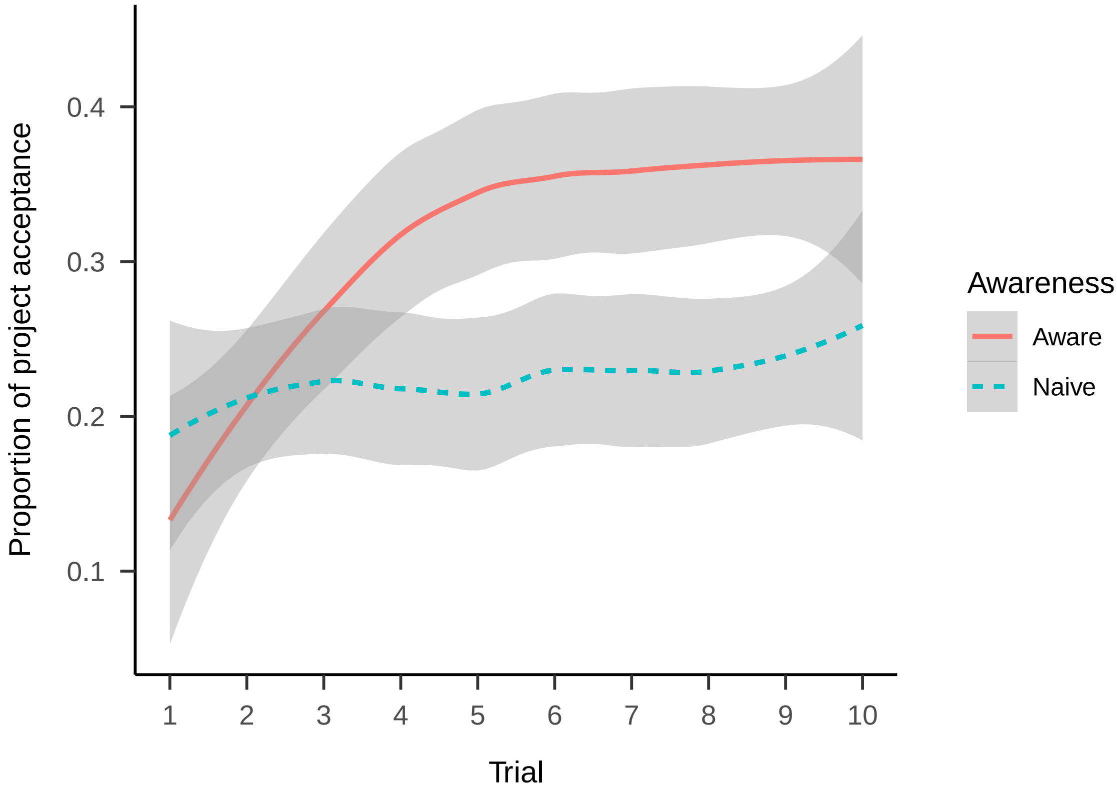


Figure 2.7: Proportion of project acceptance in the separate presentation condition, by trial and awareness conditions. LOESS method was used for smoothing over trials and the shading represents 95% confidence intervals.

gambles. Having an awareness of the total number of projects in the set could increase the likelihood that participants would naively aggregate. Specifically, knowing the number of total projects might increase the salience of the idea that the gains of some projects will offset the losses of others, because it reinforces a focus on the entire set. Another possibility is that participants had a certain aspiration level (Lopes, 1996) that they were attempting to reach. This might mean that they invested more as the task proceeded after realising that the gambles were not becoming significantly more favourable. Barron and Erev (2003, p. 219) specifically did not tell participants about the number of gambles they would experience to “avoid an ‘end of task’ effect (e.g., a change in risk attitude)”. Barron and Erev (2003) provided participants with feedback, but this should not be necessary for an aspiration level explanation since participants only need to be aware of the potential for certain gains.

This result may also be due to a Gambler's fallacy effect or the law of small numbers. This effect is characterised by people's expectation of a pattern to follow the underlying distribution of the function that generates each component. For instance, someone observing the results of a coin flip that look like HTTHTTTT might anticipate that the likelihood of "heads" is higher than that of "tails", despite the actual likelihood being 50% for either. This effect occurs in sequential decision-making, so may be relevant for the repeated-play decisions in Experiment 1. Barron and Leider (2010) found that the gambler's fallacy (in a roulette prediction task) emerges when information about past outcomes was displayed sequentially, but not when it is displayed all at once. Haisley et al. (2008) found evidence for the gambler's fallacy with a repeated-play gamble paradigm. As such, it is possible that an effect such as the Gambler's fallacy can explain the effect of the awareness manipulation. That is, participants may have thought that after a few gambles that they considered risky, the last ones were more likely to materialise. Further, this would be more likely to occur for those that knew the total number of projects, because they knew when the sequence was approaching its end.

2.2.3.3 Similarity Effect

Experiment 1 did not find a main effect of similarity in the individual choice data as predicted in Hypothesis 2.3. Instead, choice similarity interacted with the presentation condition. This interaction is harder to explain since it was not hypothesised. In fact, the results seem to suggest the opposite to what was originally expected. Initially, it was predicted that people would be less risk averse in the high similarity condition, due to the better ability to consider the isolated projects as a grouped set. Similarity was thought to act as a broad bracket, and therefore increase aggregation. That is, it was expected that seeing a set of similar projects would help participants aggregate risk when seeing them separately, more than when projects are dissimilar. Instead, project acceptance was actually numerically higher in the low than in the high similarity condition ($\Delta M = -0.06$, 95% CI $[-0.12,\ 0.00]$, $t(228.14) = -1.83$, $p = .068$) when projects were presented separately, averaging over awareness conditions.

There was no significant difference between similarity conditions regardless of presentation condition. However, allocations were significantly higher in the joint presentation condition than in the separate condition for both high and low similarity. The interaction seems to have been found due to the larger difference in the high similarity condition. Perhaps the ability to aggregate risk when projects are presented together is more made more salient when projects are similar.

Specifically, the interaction seems to be driven by the separate high similarity condition being lower, rather than by the joint high similarity being higher, as would have been expected. As such, participants could have been engaged in a naive *diversification*, rather than a naive aggregation. In "true" diversification, people would choose a set of projects that are partially (and ideally negatively) correlated, as per Markowitz (1952). However, in reality people that intend to diversify only seem diversify naively, meaning that they neglect co-variation when diversifying (e.g., Hedesstrom et al., 2006). Instead, they only seem to be looking for variety, rather than diversification in the strict sense. This *diversification bias* is also seen in product choices (Read & Loewenstein, 1995).

In Experiment 1, participants may have considered the high similarity condition as a sign that the set of projects may not be sufficiently "diversified". However, this explanation would also predict the joint presentation condition to be lower in the high similarity condition. So, perhaps when in the separate condition, participants were constantly thinking that they might be getting a different project in the next display, so rejected more projects because of the lack of diversification, but not realising that they would not be getting any other type of project. Those in the joint presentation, on the other hand, were able to see all ten projects, so already knew that there were no other projects in the set, and so were less likely to reject projects on the basis of the hope for different projects in the future.

2.2.3.4 Limitations

This experiment had two major limitations. First, proper counterbalancing was not used in the high alignment project domain, nor in the order of the within-subjects manipulation of presentation. As such, it is unclear what role these elements played in the results, especially in the presentation condition, in which participants always saw the separate condition first. Second, as mentioned above in Section 2.2.1.2.3, there was a mistake in the generation of the gamble values that meant that the individual gambles did not correspond with the distribution that participants saw. Both of these limitations were addressed in Experiment 2.

2.3 Experiment 2

Experiment 2 investigated the effect of presentation, awareness, and distribution on project choice. For the distribution manipulation, half of the sample saw an outcome probability distribution as in the previous literature (e.g., Redelmeier & Tversky, 1992; Webb & Shu, 2017) to determine their risk aversion when the gambles are explicitly aggregated. In contrast to most of the repeated-play choice literature, each choice was presented without subsequent feedback. Further, in contrast to Experiment 1, the distribution was displayed alongside each gamble, as opposed to only at the very end. This is an important manipulation because finding out whether it is effective will (a) add to the understanding of the conditions necessary for mathematical aggregation (beyond a mere intuitive sense of aggregation), and (b) suggest new ways to encourage aggregation in real-world applications.

In past work, participants were shown ordinary binomial distributions, since multi-play gambles are identical. However, there has not been an investigation of *non-identical* gamble distributions in this context. Doing this requires using a *Poisson* binomial distribution, which allows for multiple trials with different probabilities.

Further, Experiment 2 addressed potential order effects in Experiment 1 by manipulating all the main variables between-subjects. Manipulating presentation between-subjects, removes the potentially confounding factor of reduced risk aversion over time.

Experiment 2 again tested Hypotheses 2.1, and 2.2, from Experiment 1. Further, following the finding in Experiment 1 that participants in the aware condition seemed to become more risk-taking as the experiment progressed, Experiment 2 tested the following hypothesis:

Hypothesis 2.4—interaction of trial number and awareness. Participants will make more risky choices as the trials progress, but only when they are aware of the total number of projects in the set.

Further, multi-play gambles with outcome distributions have been shown to reduce risk aversion compared to multi-play gambles without distributions (e.g., Redelmeier & Tversky, 1992; Webb & Shu, 2017). Therefore, Experiment 2 tested the following hypothesis:

Hypothesis 2.5—distribution effect. Participants will make more risky choices when presented with an aggregated outcome distribution than when making the same decisions individually.

2.3.1 Method

2.3.1.1 Participants

One hundred and sixty-four participants (51 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 26.39 years ($SD = 8.63$, $min. = 18$, $max. = 72$). Participants reported an average of 2.55 years ($SD = 5.34$, $min. = 0$, $max. = 43$) working in a business setting, and an average of 1.67 years ($SD = 2.94$, $min. = 0$, $max. = 20$) of business education. The mean completion time of the task was 6.53 min ($SD = 5.15$, $min. = 1.18$, $max. = 39.93$). Table 2.2 shows the allocation of participants to the different conditions. Appendix A.2.1.1.1 describes the power analysis conducted to arrive at this sample size.

Table 2.2: Experiment 2 group allocation.

Awareness	Distribution	Presentation	N
Aware	Absent	Separate	40
Naive	Absent	Joint	41
Naive	Absent	Separate	41
Naive	Present	Separate	42
Total			164

2.3.1.2 Materials

2.3.1.2.1 Instructions Participants were shown the same instructions as in Experiment 1 (see Section 2.2.1.2.1).

2.3.1.2.2 Risky Investment Task Participants saw a similar display to the one in Experiment 1 (see Section 2.2.1.2.2), but with new gamble values, in order to fix the mistake in the Experiment 1 gamble value calculation (detailed in the appendix Section A.1.2.2).

The presentation and awareness manipulations were as in Experiment 1. However, in the distribution-present condition participants saw the outcome probability distribution of all the projects alongside the description, rather than after all the projects were seen (see Figure 2.8).

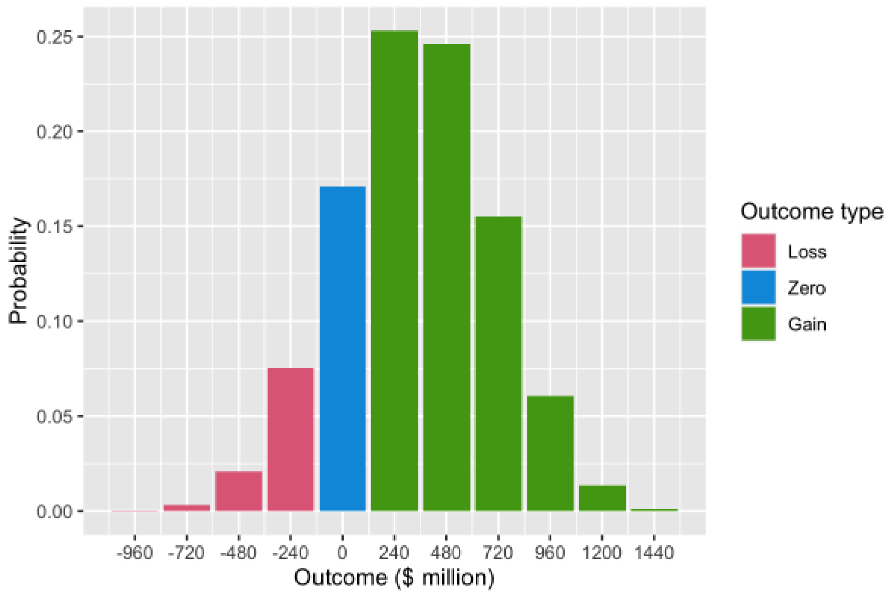
2.3.1.2.3 Follow-up Participants were asked how many projects they thought they saw, whether they were willing to accept all or none of the projects, and how many they would be willing to accept if they had to choose a number. Appendix A.2.1.2.1 shows these questions.

2.3.1.3 Procedure

Participants read the instructions and completed the risky investment task in their respective conditions. After seeing the individual projects, participants were then asked the three follow-up questions.

Below is the probability distribution of final outcomes if all projects were chosen.

The numbers on the x-axis (labelled 'Outcome') represent the final amounts of money possible if you chose to invest in all the projects. The numbers on the y-axis (labelled 'Probability') represent the likelihoods of each of the possible outcomes. Negative final outcomes (losses) are shown in red, positive final outcomes (gains) are shown in green, and a final outcome of zero (no loss or gain) is shown in blue.



Indicate below whether you would invest in the following:

Refinera is a business in your company that proposes to construct an oil well project, which they forecast will cost $40 million. If the project succeeds, forecasts show the company would make $240 million. Research suggests that there is a 30% chance of the project succeeding. Therefore, **there is 30% chance of gaining $200 million and a 70% chance of losing $40 million on the investment.***

Yes ○

No ○

Continue

Figure 2.8: An example of a display seen by those in the separate distribution-present condition of Experiment 2.

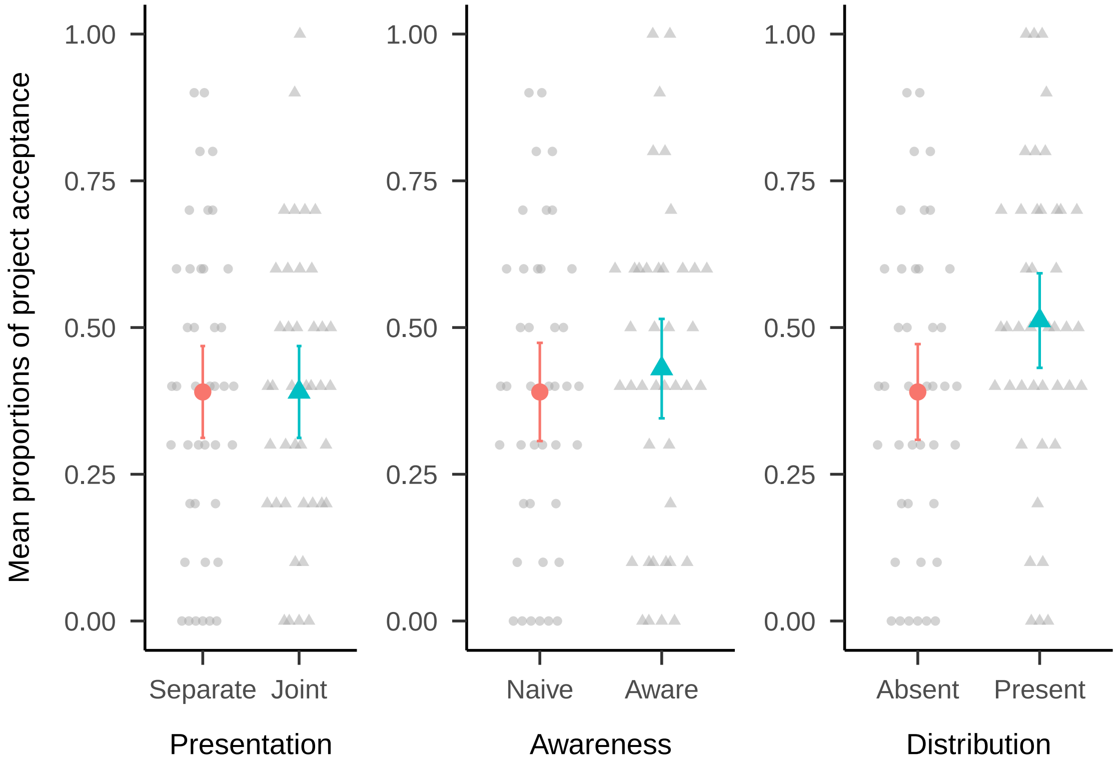


Figure 2.9: Mean proportion of project acceptance for the presentation, awareness, and distribution effects. The condition on the left of each effect is the reference condition (separate presentation, naive awareness, distribution absent). As such, it is identical for the three effects. Error bars represent 95% confidence intervals. Raw data are plotted in the background.

2.3.2 Results

2.3.2.1 Project Investment

The project investment data were analysed as proportions of choice per participant, as in Experiment 1. Each experimental condition was compared to the same control condition (separate presentation, naive awareness, and distribution absent). Figure 2.9 shows these data. The difference between presentation conditions was not significant, $F(1, 80) = 0.00$, $p > .999$, $\hat{\eta}_p^2 = .000$. Similarly, the difference between awareness conditions was not significant, $F(1, 79) = 0.44$, $p = .508$, $\hat{\eta}_p^2 = .006$. However, those that that saw a distribution chose to invest significantly more (51.19%) than those that did not see a distribution (39.02%), $F(1, 81) = 4.46$, $p = .038$, $\hat{\eta}_p^2 = .052$.

Further, as Figure 2.10 shows, it doesn't seem as if the previous awareness

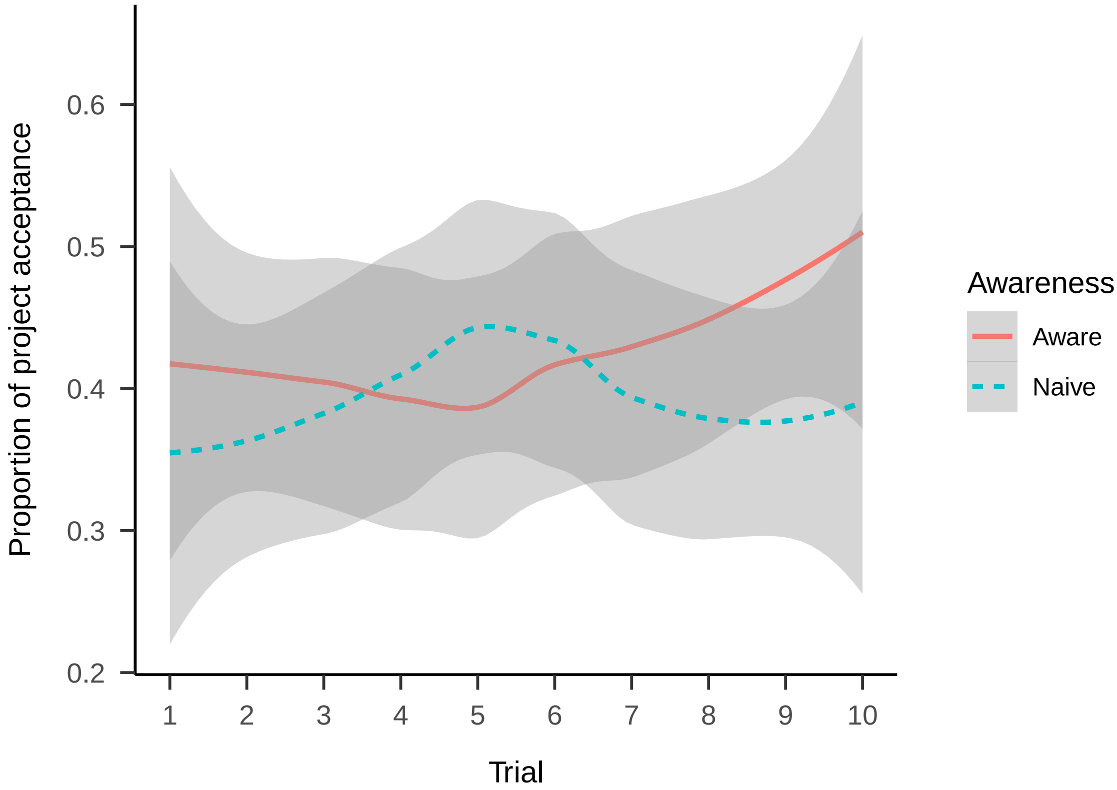


Figure 2.10: Mean project acceptance for separate presentation, distribution absent condition, by awareness and trial. LOESS method was used for smoothing over trials and the shading represents 95% confidence intervals.

by trial effect was replicated.

2.3.2.2 Follow-up

The portfolio choice data from both the number and binary questions were congruent with the above, finding that those in the distribution condition were more likely to invest (see Appendix A.2.2).

2.3.3 Discussion

Experiment 2 found support for Hypothesis 2.5. Seeing an outcome distribution of a business project portfolio had a strong effect on participants' decision-making. Participants indicated that they would invest in more projects and were more likely to indicate that they would invest in the entire portfolio. However, the awareness and presentation effects found in Experiment 1 (see Section 2.2.2) did not replicate.

These findings provide evidence for choice bracketing. That is, people do seem to be primarily considering gambles one at a time. Further, these findings suggest that that the main bottleneck for appropriately aggregating a set of gambles is a computational one. That is, people simply cannot mentally combine the outcomes and probabilities in a way that sufficiently approximates the outcome distribution display.

The lack of replication of the awareness and presentation effects provides evidence against a naive aggregation account of the distribution effect. Specifically this suggests that the distribution effect is a result of a lack of ability to mathematically combine risk, rather than naive aggregation. If some of the bottleneck was attributable to a lack of realisation that the individual gambles could be grouped together, then the effects from Experiment 1 should have replicated. Instead it seems that even when people have an opportunity to consider an entire set of risky choices together (and consider that the gains may outweigh the losses), they do not do this.

In Experiment 2, all the gambles came from the same domain. This was done to attempt to replicate the relevant effects from Experiment 1. However, there could have been something about that particular domain that led to the lack of replication. A follow-up experiment addressed this issue by presenting participants with 20 gambles from 10 different industries and still did not replicate the awareness effect (see Appendix A.4).

2.4 General Discussion

When making one decision about a series of risky choices, it is clear that people have an intuitive sense of the advantages of risk aggregation (e.g., Samuelson, 1963). However, because risky choices are typically made one at a time in the real world, this chapter aimed to identify whether (and how) this intuition could be leveraged in this more realistic scenario. Overall, there was little evidence that subtle cues could tap into this intuitive advantage of risk aggregation, and clear visualisations of outcome distributions were needed to assist people's risk aggregation. This suggests that the act of deciding can create a strong cognitive barrier to treating a series of decisions as if they were one. However, as elaborated below, the success of the outcome distribution for overcoming this cognitive barrier in the current paradigm is a novel and important finding.

This chapter found that some choice bracketing facilitated risk aggregation in description-based repeated-play gambles. This paradigm has never been a target of research. Early work on risk aggregation involved multi-play gambles, which treated gambles as simultaneous and identical. However, most risky choice outside the lab involves considering multiple choices independently, as in repeated-play paradigms. Most repeated-play paradigms have involved providing participants with feedback, or allowing them to sample from outcome distributions. Large real-life investments are different, as their outcomes are not eventuated immediately (and do not allow for distribution sampling). The limited prior work using description-based repeated-play gambles did not consider the effect of choice bracketing on risk aggregation. As such, the paradigm used in this chapter allowed for the investigation of choice bracketing in a way that is more isomorphic with real-life prescriptions.

Experiment 1 found evidence for the effects of similarity, presentation, and awareness of the number of projects. Experiment 2 found evidence for the effect of an outcome distribution but did not replicate the presentation and awareness effects. Subsequent follow up experiments (reported in Appendices A.3 and A.4) again tested the similarity and awareness effects. These experiments found evidence for naive diversification (an advantage for low similarity) when considering all projects once and did not replicate the trial-by-trial interaction from Experiment 1.

Therefore, in addition to the novelty of the paradigm itself, this chapter found that choice bracketing facilitates risk aggregation, if aided by the aggregated distribution. As per Hypothesis 2.5, Experiment 2 found that showing a distribution of outcome probabilities without inter-trial feedback reduced risk aversion. Further, there was mixed evidence for Hypothesis 2.3, such that people were less risk averse when the set of projects they saw were dissimilar, but only when offered them as a portfolio (see Appendix A.3). There was only minimal evidence for Hypotheses 2.1 and 2.2, suggesting that viewing projects together and an awareness of the number of projects are not sufficient to encourage aggregation. Altogether, it seems that subtle contextual cues are often not sufficient to encourage risk aggregation and that people need risk to be is aggregated for them explicitly in order to understand the benefits of aggregation.

2.4.1 Theoretical Implications

The finding that participants are less risk averse when provided with an aggregated outcome distribution is congruent with previous work (e.g., Redelmeier & Tversky, 1992). However, when distributions have been previously used, gambles were identical—as in multi-play paradigms—and used immediate feedback for repeated-play paradigms (e.g., Benartzi & Thaler, 1999). As mentioned previously, both these paradigms have limited ecological validity because usually people are faced with non-identical sequential choices and do not receive immediate feedback. This work is the first to provide evidence for this aggregation effect with non-identical gambles without feedback.

The other choice bracketing findings that showed little success with aiding aggregation are less congruent with previous research. Sokol-Hessner et al. (2009) and Sokol-Hessner et al. (2012) found that encouraging participants to make decisions akin to a professional investor increased the amount of risky choices they made. The results showed that a subtler manipulation—whether or not participants were aware of the number of choices to be made—is not sufficient to encourage aggregation. Hsee et al. (1999) found that useful, but hard-to-interpret, attributes were used more when the options were presented jointly, rather than separately. In the case of these experiments, the “hard to interpret” element of the decision set was the risk of the projects. Contrary to Hsee et al. (1999), it seems that risk was not always accounted for more when projects were presented jointly, rather than separately. More study is needed to understand whether the effects that were seen in Experiment 1 but not replicated in the subsequent experiments are due to statistical chance or unexplored elements of the experiment.

Research on the effect of option similarity on choice (e.g., Markman & Medin, 1995) suggests that alignable differences are more important than non-alignable differences. Further, the effects of multi-play gambles and outcome distributions on risk aggregation are only seen when participants perceive the options as fungible (e.g., DeKay & Kim, 2005). As such, it was predicted that a set of investments that involve the same type of investment would be seen as more similar, and therefore be considered as fungible. Hypothesis 2.3 predicted that this would facilitate a broad bracketing, and therefore more risk aggregation.

Instead, the results showed that choice similarity did not affect individual project allocations. However, when participants were given an all-or-nothing choice for the entire set of projects, those that viewed dissimilar projects were more likely to take the entire set projects than those that viewed similar projects. This is different from the initial hypothesis, however, it may still suggest an effect of choice bracketing. That is, this effect was only found when participants were asked about the entire portfolio of projects, rather than when they had a chance to make a choice about each project. The way that the question was framed may have acted to broadly bracket the choices by forcing the choice.

A diversified portfolio is one whose investments are uncorrelated or negatively correlated. According to portfolio theory (Markowitz, 1952), a diversified portfolio is preferred to one that is not diversified, because it reduces the probability of a loss. When some investments have losses, others will have gains—the root of “don’t keep all your eggs in one basket.” Typically, questions of gamble aggregation assume that each gamble is independent. That is, the gambles are uncorrelated. As such, aggregation of a portfolio already assumes that the portfolio is somewhat diversified (or at least that the gambles aren’t perfectly correlated).

In the case of the similarity effect, the choice bracketing did not seem to encourage aggregation, but instead appears to have encouraged a naive diversification

(Hedesstrom et al., 2006; Read & Loewenstein, 1995). It could not have been actual diversification, because the projects did not contain correlational information. Rather, participants could have been more eager to accept the project portfolio due to the higher variability between projects (due to the similarity manipulation).

This finding suggests that there may be trade-off between aggregation and diversification. The literature shows that people prefer multi-play gambles to single-play gambles. However, participants in this chapter were more likely to aggregate diverse repeated-play gambles to similar repeated-play gambles when these were bracketed broadly. Therefore, people are likely to still need choice bracketing. That is, diverse repeated-play gambles that are not bracketed are simply individual single-play gambles.

One way to test this explanation is by using identical gambles. This chapter used unique gambles to increase ecological validity. However, the above explanation would predict that participants prefer non-identical repeated-play gambles to identical repeated-play gambles when these are bracketed. However, when these gambles are not presented as a portfolio, it is likely that the identical gambles would be preferred overall because the non-identical gambles would be represented as individual single-play gambles.

It is also possible that similarity effects were not seen because the sequence of gambles itself led to naive aggregation for all conditions. One way that this could be tested is by interweaving other tasks in-between the gambles to break them up. Then similarity may play a role by allowing bracketing across otherwise distinct gambles. Multiple sets of gambles can be interweaved with similarity alone creating the potential sets. The prediction is that without similarity the gambles would not be aggregated.

2.4.1.1 How Does Choice Bracketing Facilitate Aggregation?

Much of the literature (e.g., Benartzi & Thaler, 1999) is not clear about why choice bracketing occurs. Some explain the effect of bracketing on aggregation using risk aversion (e.g., Read et al., 1999), while others refer to the increased weighting of potential losses (Webb & Shu, 2017).

Decision-from-experience *sampling* studies explain the underweighting of rare events (as opposed to the overweighting that occurs with decisions-from-description) by sampling bias and recency effects (e.g., Hertwig et al., 2004; Wulff et al., 2018). That is, they explain that people are less risk averse for positive EV gambles because when they sample from the distribution they only sample a small amount (usually approximately 20 times) so they do not experience rare events very often. Also, the latter half of the sequence of sampling is significantly more predictive than the former (recency effect). Some decision-from-experience *feedback* studies explain this effect by "choice inertia" (Camilleri & Newell, 2011). That is, "the tendency to repeat the last choice, irrespective of the obtained outcome" (p. 383). However, there is not much more elaboration beyond this. Repeated-play gambles show more underweighting than multi-play gambles. This is said to be due to a "reliance on a very small set of samples" (Camilleri & Newell, 2013, p. 64). However, this explanation does not account for repeated-play effects independently.

The experiments in this chapter shed some light about the mechanisms behind why choice bracketing may affect risk aggregation in repeated-play gambles without feedback. Two explanations were proposed: participants may realise that some gains will offset the losses, or they may need explicit aggregation. Not finding evidence for the subtle choice bracketing manipulations suggests that people do not intuitively consider that the gains of their choices may offset the potential losses. Perhaps the possibility of recouped losses would become more salient when other participants are explicitly told of this possibility, as in Sokol-Hessner et al. (2009). Their explicit instruction manipulation is introduced above as appearing unrealistically strong, but the results of this chapter suggest that people do need very explicit scaffolding in order to use risk aggregation.

2.4.2 Practical Implications

This research implies some prescriptions for capital allocation decision-making. For instance, even if managers implement processes that encourage a joint evaluation of projects, this may be insufficient to encourage aggregation. Projects need to very explicitly be considered as individual components in a portfolio in order to facilitate better risk aggregation. Some companies are already implementing processes that make this more explicit (Lovallo et al., 2020). This is especially important for those that would still have to evaluate projects separately. Further, this work shows the importance of being explicit about the forecasted probabilities of project success. Doing this is necessary for the aggregation process. Even more ideal would be to forecast project success using an entire probability distribution for the different possible outcomes. However, research shows that people struggle to construct such distributions (e.g., Alpert & Raiffa, 1982; Schaefer & Borcherding, 1973; Tversky & Kahneman, 1974; von Holstein, 1971) and Chapter 4 shows that people struggle to use such variance information when making allocation decisions. Regardless, the benefits of risk aggregation can be used even if forecast information is limited (e.g., only a point estimate and a probability) and only one project is being considered. Specifically, a proposed project can be seen in a larger context by aggregating it with projects from the immediate past.

Interestingly, participants were less risk averse about a portfolio of projects when industries differed, compared to when they were all from the same industry. Simply manipulating the similarity of financially-irrelevant semantics of a set of choices affected participants' risk aversion. This has implications for managerial settings. Executives in multi-business firms often have to make capital allocation decisions that involve comparing dissimilar projects. How can an oil well exploration project be appropriately compared to an oil refinery? Or to a microchip project? Chapter 4 suggests that evaluating dissimilar business projects is more difficult to comparing similar projects. The current work suggests that managers may actually be *less* likely to realise the benefits of aggregation when they are in a less diversified company. As such, managers should complement an understanding of aggregation with that of diversification. This might help to avoid being biased by a lack of variety of projects despite a potentially high level of diversification.

2.4.3 Future Research

The main novelty of the experiments in this chapter comes from increasing ecological validity of risky choice problems by removing inter-trial feedback. Future work should test even more realistic scenarios. Such studies should involve managers, ideally in multi-business firms. Investigating whether the choice bracketing findings from these experiments replicates in a sample of managers will help to determine whether these results could be applied to real-world managerial decision-making. This is especially important since Haigh and List (2005) found that professional traders show more myopic loss aversion than students. Further, the similarity, awareness, and presentation manipulations should be tested with managers since it is possible that they have a greater sense of naive aggregation and are therefore more likely to be more amenable to such manipulations. The addition of extra payment for better performance on the task might also assist in making the task more isomorphic with real-world managerial decisions. Further, in the present experiments, participants viewed the projects all in the space of one session. However, this is not completely isomorphic to real life, where managers make many other decisions that are unrelated to the large risky investments at their companies. Future research should test participants over a longer period of time (as in Beshears et al., 2016) in order to see whether the effects of the manipulations replicate in a more realistic environment.

References

Aloysius, J. A. (2007). Decision making in the short and long run: Repeated gambles and rationality. *British Journal of Mathematical and Statistical Psychology*, *60*(1), 61–69. https://doi.org/10/czgkh3

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty* (pp. 294–305). Cambridge University Press. https://doi.org/10.1017/CBO9780511809477.022

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*(3), 215–233. https://doi.org/10/d3jsr8

Barron, G., & Leider, S. (2010). The role of experience in the Gambler's Fallacy. *Journal of Behavioral Decision Making*, *23*(1), 117–129. https://doi.org/10/d3p92r

Bellemare, C., Krause, M., Kröger, S., & Zhang, C. (2005). Myopic loss aversion: Information feedback vs. investment flexibility. *Economics Letters*, *87*(3), 319–324. https://doi.org/10/dfjq7n

Benartzi, S., & Thaler, R. H. (1999). Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments. *Management Science*, *45*(3), 364–381. https://doi.org/10/cjhp6t

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2016). Does Aggregated Returns Disclosure Increase Portfolio Risk Taking? *The Review of Financial Studies*, *30*(6), 1971–2005. https://doi.org/10/gjscs7

Bjørnsen, K., & Aven, T. (2019). Risk aggregation: What does it really mean? *Reliability Engineering & System Safety*, *191*, 106524. https://doi.org/10/gjscst

Bristow, R. E. (2011). *There is more to Life than Expected Values: Results of Free Distributions of Multiple-Play Gambles* (Masters thesis). Miami University. Oxford, Ohio. https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=miami1304352729

Camilleri, A. R., & Newell, B. R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review*, *18*(2), 377–384. https://doi.org/10/cz55xk

Camilleri, A. R., & Newell, B. R. (2013). The long and short of it: Closing the description-experience "gap" by taking the long-run view. *Cognition*, *126*(1), 54–71. https://doi.org/10/f4gq3w

Coombs, C. H., & Bowen, J. N. (1971). A test of VE-theories of risk and the effect of the central limit theorem. *Acta Psychologica*, *35*(1), 15–28. https://doi.org/10/dm5gbv

DeKay, M. L. (2011). Are Medical Outcomes Fungible? A Survey of Voters, Medical Administrators, and Physicians. *Medical Decision Making*, *31*(2), 338–353. https://doi.org/10/b539tb

DeKay, M. L., Hershey, J. C., Spranca, M. D., Ubel, P. A., & Asch, D. A. (2006). Are medical treatments for individuals and groups like single-play and multiple-play gambles? *Judgment and Decision Making*, *1*(2), 12. http://journal.sjdm.org/jdm06133.pdf

DeKay, M. L., & Kim, T. G. (2005). When things don't add up: The role of perceived fungibility in repeated-play decisions. *Psychological Science*, *16*(9), 667–672. https://doi.org/10/ddgt5v

Ert, E., & Erev, I. (2013). On the descriptive value of loss aversion in decisions under risk: Six clarifications. *Judgment and Decision Making*, *8*(3), 22. http://journal.sjdm.org/12/12712/jdm12712.pdf

Gneezy, U., & Potters, J. (1997). An Experiment on Risk Taking and Evaluation Periods. *The Quarterly Journal of Economics*, *112*(2), 631–645. https://doi.org/10/bpkbhz

Haigh, M. S., & List, J. A. (2005). Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis. *The Journal of Finance*, *60*(1), 523–534. https://doi.org/10/c7jn9k

Haisley, E., Mostafa, R., & Loewenstein, G. (2008). Myopic risk-seeking: The impact of narrow decision bracketing on lottery play. *Journal of Risk and Uncertainty*, *37*(1), 57–75. https://doi.org/10/czj8x7

Hedesstrom, T. M., Svedsater, H., & Garling, T. (2006). Covariation neglect among novice investors. *Journal of Experimental Psychology*, *12*(3), 155–165. https://doi.org/10/ftmd77

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), 534–539. https://doi.org/10/b274n8

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, *125*(5), 576–590. https://doi.org/10/dm3fvc

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback Produces Divergence From Prospect Theory in Descriptive Choice. *Psychological Science*, *19*(10), 1015–1022. https://doi.org/10/bgb3qs

Joag, S. G., Mowen, J. C., & Gentry, J. W. (1990). Risk perception in a simulated industrial purchasing task: The Effects of single versus multi-play decisions. *Journal of Behavioral Decision Making*, *3*(2), 91–108. https://doi.org/10/czwknv

Kahneman, D., & Lovallo, D. (1993). Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking. *Management Science*, *39*(1), 17–31. https://doi.org/10/c8vntn

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10/g98

Keren, G. (1991). Additional tests of utility theory under unique and repeated conditions. *Journal of Behavioral Decision Making*, *4*(4), 297–304. https://doi.org/10/bqqkt4

Keren, G., & Wagenaar, W. A. (1987). Violation of utility theory in unique and repeated gambles. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 387. https://doi.org/10/dkr96j

Klos, A. (2013). Myopic loss aversion: Potential causes of replication failures. *Judgment and Decision Making*, *8*(5), 13. http://journal.sjdm.org/12/121229/jdm121229.pdf

Klos, A., Weber, E. U., & Weber, M. (2005). Investment Decisions and Time Horizon: Risk Perception and Risk Behavior in Repeated Gambles. *Management Science*, *51*(12), 1777–1790. https://doi.org/10/bbrvhd

Koehler, J. J., Gibbs, B. J., & Hogarth, R. M. (1994). Shattering the illusion of control: Multi-shot versus single-shot gambles. *Journal of Behavioral Decision Making*, *7*(3), 183–191. https://doi.org/10/fwmwjs

Langer, T., & Weber, M. (2001). Prospect Theory, Mental Accounting, and Differences in Aggregated and Segregated Evaluation of Lottery Portfolios. *Management Science*, *47*(5), 716–733. https://doi.org/10/fcfk69

Li, S. (2003). The role of Expected Value illustrated in decision-making under risk: Single-play vs multiple-play. *Journal of Risk Research*, *6*(2), 113–124. https://doi.org/10/cz6phv

Liu, H.-H., & Colman, A. M. (2009). Ambiguity aversion in the long run: Repeated decisions under risk and uncertainty. *Journal of Economic Psychology*, *30*(3), 277–284. https://doi.org/10/d5p9kw

Lopes, L. L. (1996). When Time Is of the Essence: Averaging, Aspiration, and the Short Run. *Organizational Behavior and Human Decision Processes*, *65*(3), 179–189. https://doi.org/10/fdtw45

Lovallo, D., Koller, T., Uhlaner, R., & Kahneman, D. (2020). Your Company Is Too Risk-Averse. *Harvard Business Review*, *98*(2), 104–111.

Ludvig, E. A., & Spetch, M. L. (2011). Of Black Swans and Tossed Coins: Is the Description-Experience Gap in Risky Choice Limited to Rare Events? *PLoS ONE*, *6*(6), e20262. https://doi.org/10/ds5q2k

MacCrimmon, K. R., Wehrung, D. A., & Stanbury, W. T. (1986). *Taking risks: The management of uncertainty*. Free Press.

March, J. G., & Shapira, Z. (1987). Managerial Perspectives on Risk and Risk Taking. *Management Science*, *33*(11), 1404–1418. https://doi.org/10/ft2phq

Markman, A. B., & Loewenstein, J. (2010). Structural comparison and consumer choice. *Journal of Consumer Psychology*, *20*(2), 126–137. https://doi.org/10/d7b49c

Markman, A. B., & Medin, D. L. (1995). Similarity and Alignment in Choice. *Organizational Behavior and Human Decision Processes*, *63*(2), 117–130. https://doi.org/10/c8z7r9

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10/bhzd

Moher, E., & Koehler, D. J. (2010). Bracketing effects on risk tolerance: Generalizability and underlying mechanisms. *Judgment and Decision Making*, *5*(5), 8. http://journal.sjdm.org/10/10422/jdm10422.pdf

Montgomery, H., & Adelbratt, T. (1982). Gambling decisions and information about expected value. *Organizational Behavior and Human Performance*, *29*(1), 39–57. https://doi.org/10/cvgjp4

Rabin, M., & Weizsäcker, G. (2009). Narrow Bracketing and Dominated Choices. *American Economic Review*, *99*(4), 1508–1543. https://doi.org/10/fk4rmz

Read, D., & Loewenstein, G. (1995). Diversification bias: Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied*, *1*(1), 34. https://doi.org/10/fcgvrw

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice Bracketing. *Journal of Risk and Uncertainty*, *19*(1), 171–197. https://doi.org/10/dh3rmv

Redelmeier, D. A., & Tversky, A. (1992). On the Framing of Multiple Prospects. *Psychological Science*, *3*(3), 191–193. https://doi.org/10/ctw2k6

Ross, S. A. (1999). Adding Risks: Samuelson's Fallacy of Large Numbers Revisited. *The Journal of Financial and Quantitative Analysis*, *34*(3), 323–339. https://doi.org/10/bj6r8r

Samuelson, P. A. (1963). Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia*, *57*(98), 108–113. https://www.casact.org/sites/default/files/database/forum_94sforum_94sf049.pdf

Schaefer, R. E., & Borcherding, K. (1973). The assessment of subjective probability distributions: A training experiment. *Acta Psychologica*, *37*(2), 117–129. https://doi.org/10/dpzkfb

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2012). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience*, *8*(3), 341–350. https://doi.org/10/fx5cn6

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*(13), 5035–5040. https://doi.org/10/fhdrcw

Stutzer, M. (2013). Misperceptions of long-term investment performance Insights from an experiment. *The Journal of Behavioral Finance & Economics*, *3*(1), 1–20. http://www.aobf.org/wp-content/uploads/2020/06/1-Stutzer-and-Grant.pdf

Su, Y., Rao, L.-L., Sun, H.-Y., Du, X.-L., Li, X., & Li, S. (2013). Is making a risky choice based on a weighting and adding process? An eye-tracking investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1765–1780. https://doi.org/10/gjscr2

Swalm, R. O. (1966). Utility Theory–Insights into Risk Taking. *Harvard Business Review*, *44*(6), 123–136.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral decision making*, *12*(3), 183–206. https://doi.org/10/d4njp3

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *The Quarterly Journal of Economics*, *112*(2), 647–661. https://doi.org/10/fcf346

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10/gwh

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10/fj3z3r

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10/cb57hk

Venkatraman, S., Aloysius, J. A., & Davis, F. D. (2006). Multiple prospect framing and decision behavior: The mediational roles of perceived riskiness and perceived ambiguity. *Organizational Behavior and Human Decision Processes*, *101*(1), 59–73. https://doi.org/10/dszh5v

von Holstein, C.-A. S. S. (1971). Two techniques for assessment of subjective probability distributions — An experimental study. *Acta Psychologica*, *35*(6), 478–494. https://doi.org/10/fgg6jn

Webb, E. C., & Shu, S. B. (2017). Is broad bracketing always better? How broad decision framing leads to more optimal preferences over repeated gambles.

Judgment and Decision Making, *12*(4), 382. http://journal.sjdm.org/17/17317/jdm17317.pdf

Wedell, D. H., & Bockenholt, U. (1994). Contemplating Single versus Multiple Encounters of a Risky Prospect. *The American Journal of Psychology*, *107*(4), 499. https://doi.org/10/b4fs2p

Wedell, D. H., & Böckenholt, U. (1990). Moderation of preference reversals in the long run. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 429–438. https://doi.org/10/bmn8hf

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*(2), 140–176. https://doi.org/10/gc2s4c

Zeisberger, S. (2020, May 15). *Do People Care About Loss Probabilities?* https://ssrn.com/abstract=2169394

3

Joint Evaluation of Multiple Projects

Chapter 2 found that people struggle to aggregate risk even when provided with choice bracketing cues that could have built on an intuitive sense of how aggregation reduces risk. The finding that people are more likely to accept many gambles at once (e.g., Samuelson, 1963; Wedell & Bockenholt, 1994), even without any aids to calculate risk, suggests that people can gain an intuition for the benefits of aggregation. Yet, in the current work, people instead considered projects one at a time and only leveraged the benefits of aggregation when given an explicit visualisation of what it entails.

This shows that it is important to change organisational policy to encourage considering business projects jointly. Doing this means that the risk can be concurrently aggregated. In real-life capital allocation scenarios, when managers evaluate projects sequentially, an aggregated distribution can also be presented using any number of projects that were considered in the recent past. This means that a strategy of project risk aggregation can be implemented at any stage in an organisation's lifespan. Relatively new ventures can implement these recommendations by waiting until a certain number of project proposals have been accrued before aggregating.

Considering projects jointly is also useful for accountability purposes. The usual incentive structure in organisations that judges each project outcome independently is likely to punish risk-taking due to its potential negative consequences and not due to the information that was available at the time of evaluation. Framing a set of projects as a portfolio means that any subsequent success or failure of one project can be traced back to the entire batch, and the performance of the whole portfolio can be evaluated.

Business projects might not always be either accepted or rejected, as in Chapter 2. Instead, top-level managers might ask for project proposals from lower-level managers, and then allocate funds from the available budget. An organisation might also have a initial “culling” phase, and a subsequent ranking phase. When initially considering a set of projects, some might be rejected according to certain rules. For instance, an NPV might not meet a certain minimum cut-off. The remaining projects in the set can then be ranked in order of priority and receive an allocation of capital from the budget.

A few potential problems arise at the point that projects are considered jointly for ranking and allocation. For instance, it might not be easy to compare between the projects in the set. As discussed in Chapter 1, diversification of business units has become very popular in large organisations. Therefore, most hierarchical organisations are likely to face difficult comparisons when deciding on how to rank and allocate capital to projects that originated in different divisions. A non-hierarchical organisation that develops one type of product may be able to simply compare across any number of intrinsic project attributes, whereas a diversified organisation is likely to have to rely on more abstract financial metrics, such as NPV. Such metrics are “abstract” because they can be applied in almost any domain.

For instance, when comparing across two oil well projects, there can be both attributes intrinsic to the project, such as the amount of hydrocarbons that are extracted per hour, and also the more abstract financial metrics. There is a potential interaction between the ease that managers have to compare across the projects and the kinds of measures that are used to make the comparison. Two similar projects, such as two oil wells can be evaluated using litres of hydrocarbons extracted per hour, whereas an oil refinery cannot. In the case that two dissimilar projects are compared, managers can use financial metrics to compare across domains. This can lead to comparable accuracy as long as the abstract metrics are as reliable as the intrinsic project features.

A concern that arises out of a reliance on such metrics is that underlying variance is not taken into account. Forecast estimates such as NPV rely on many assumptions and contain much inherent uncertainty, so managers that use them should be cautious about over-relying on them. Chapter 4 tests people's sensitivity to forecast estimate variance information. That is, will people use NPV more when the variance information suggests that it is a reliable measure, than when the information suggests that it is unreliable?

Chapter 2 manipulated project presentation and found no significant difference between when projects were considered jointly or separately. This was explained by the bounds on people's ability to intuitively aggregate. However, it was unclear what components of the projects people focused on both because they were not explicitly manipulated and because the task involved a binary choice (accept or reject). A relative allocation measure for multiple projects with systematically varied attributes would allow to determine the influence of those different attributes. Therefore, Chapter 4 considers the situation in which people are already presented with choices together and asked to evaluate the projects by allocating a hypothetical budget.

Further, Chapter 4 identifies the factors that affect people's decisions independently from the potential risk of losing hypothetical money, which is a large reason for the effects in the previous chapter. Risk aversion is accounted for by making it clear that no losses are possible. This is achieved by using only positive NPVs, which implies that the project is not forecasted to lose money.

Chapter 4 also manipulates how easy the project attributes are to compare. This helps identify the ways that decision-making in a diversified organisation may be different to that of a more integrated organisation. Chapter 2 manipulated similarity by either showing a set of projects from the same industry or a set from different industries. This was meant to simulate an integrated and diversified firm, respectively. This manipulation was not as strong because there were no project attributes that could be aligned or not. That is, there was nothing actually non-alignable. This may explain the equivocal similarity effect. In Chapter 4, alignability is more fully manipulated by having project attributes be critical to the evaluation. These project features are shown explicitly so that the difficulty of the comparison is more obvious.

It is not possible to compare apples and oranges. But it is possible to compare apples and oranges in terms of some specific attribute—to say that apples deliver twice as many calories per dollar or that oranges deliver twice as many vitamin C units per dollar.

—C. L. Robinson (1944, p. 13)

4

Project Similarity Bias and Variance Neglect in Forecast Metric Evaluation

Contents

4.1 Introduction

One of the most important tasks faced by executives is the allocation of capital within their companies. This requires the ranking of projects by importance and predicted success, and allocating the limited capital accordingly (not unlike a scientific funding agency). Ranking of projects necessitates comparing them across a number of dimensions. For example, the executive of an oil company may have received multiple oil exploration proposals. Determining what makes one oil exploration project better than another is relatively simple. However, consider a different scenario in which the executive must allocate capital between an oil exploration project and an oil refinery project. The dimensions of oil refinery projects that distinguish superior from inferior projects may be totally different from those of oil exploration projects. Consider a funding agency having to decide between two cognitive scientists or between a cognitive scientist and a physicist in awarding a fellowship. What makes a physics proposal better for the field of physics than a cognitive science proposal for cognitive science?

Structure-mapping theory (Gentner, 1983; Gentner & Markman, 1997) provides a model of comparison that psychologically distinguishes these two kinds of allocation tasks. This framework models comparison as a process of mapping and alignment of the shared dimensions of two conceptual structures. This mapping process reveals the shared dimensions of the two structures as well as the differences in those shared dimensions (known as *alignable differences*). For example, when comparing two oil exploration projects, the process for measuring the quantity of hydrocarbons in a prospective oil field may be identical, but the specific quantities measured will differ. This is known as an alignable difference; that is, the difference constrained within the same dimension. However, when comparing an oil field and a refinery, there will be a significantly higher number of *non-alignable differences*, because the two domains do not share component dimensions. That is, the dimensional structure of processes in the exploration project will be substantially different from that of processes in the refinery project, making it difficult to find meaningful alignments. With a higher number of non-alignable differences, there are fewer opportunities to make meaningful comparisons, leading to greater difficulty in predicting project success and ranking projects. This chapter experimentally examined project comparisons and how such comparisons may affect capital allocation decisions. The working hypothesis is that projects that have a higher number of alignable differences will lead to more precise and informed project predictions and rankings compared with projects with non-alignable differences.

However, what happens when a task demands that two domains be aligned but they are too disparate to align? Experimentally, this territory is somewhat uncharted. It is expected that, when required, people will grasp at any piece of information available and attempt to abstract and infer that which is reasonable to ease the alignment. This occurs frequently in business settings. Because corporate enterprises continue to embrace diversification strategies in their investments, they must constantly make capital allocation decisions involving highly disparate domains. To overcome these difficult comparisons, executives rely on various financial measures that, in theory, may be applied to any project or business proposal. These financial measures work well to ease the burden of difficult comparisons because they ignore the complexities of individual projects and focus solely on financial information such as total cost and projected profits. Therefore, projects that are difficult to compare may be evaluated more easily by comparing individual numerical measures.

The most common financial measure that is used by executives in order to value business project proposals is NPV (Graham & Harvey, 2001; Graham et al., 2015; Remer et al., 1993). NPV is the difference between the forecasted revenue of a project and the initial investment in its development (accounting for the time value of money), as shown in Equation (1.1). NPV is commonly used in decisions about capital allocation and investment. The basic rule is that if a project has a positive NPV, it is financially viable, and if it has a negative NPV, it is not. However, the use of NPV has been criticised, by both academics and practitioners (Fox, 2008; Willigers et al., 2017). The main criticism is that there can be underlying sources of variance in NPV that are not reflected in the final measure, which is expressed as a single numerical value. For instance, NPV is dependent on the projected cash inflows for each year of the project. However, financial forecasting is frequently inaccurate and prone to optimism bias (Lovallo & Kahneman, 2003;

Puri & Robinson, 2007). Therefore, there is bound to be variation in the reliability of NPV measures as a function of the forecasting error in the cash flow calculations. Project duration and the discount rate are further sources of variance that may be hidden by the single numerical value of NPV.

The secondary goal of this research is to investigate the extent to which people are sensitive to variance information (from financial forecasting) when making capital allocation decisions. This consideration is especially important in the capital allocation situations illustrated above, when executives need to compare projects with disparate domains and must, therefore, rely on NPV. This matters because the NPV of different domains may have different underlying forecasting error, potentially compromising the utility of using NPV as the basis of comparison. Do executives sufficiently account for the inherent sources of variance in the measure on which they rely so heavily? Research shows that people are effective at extracting variance information when exposed to numerical sequences (Rosenbaum et al., 2020). However, they struggle to use variance information when it is represented numerically (Batteux et al., 2020; Galesic & Garcia-Retamero, 2010; Konold et al., 1993; Vivalt & Coville, 2021).

4.1.1 Experiment Summary

Experiment 1 investigated the effect of project alignment on the decision-making of naive participants asked to allocate capital to a set of fictional projects. Naive participants were assumed to have no requisite knowledge about NPV reliability; thus, NPV reliability level was manipulated by directly telling participants whether or not the given NPV was reliable. For this experiment, it was predicted that when projects are alignable, participants who are told NPV is reliable would use it in their decision-making, while participants who are told that NPV was unreliable would not use it in their decision-making. However, when projects are not alignable, it was predicted that participants would use NPV, regardless of the stated NPV reliability level.

Experiment 2 investigated the decision-making of management students in a similar situation to Experiment 1. The main difference was that instead of telling participants whether or not the NPV was reliable, the level of *numerical* NPV reliability—that is, the width of the numerical range around the average NPV—was manipulated. Similar to Experiment 1, it was predicted that participants would rely more on NPV in non-alignable projects than in alignable projects. However, it was predicted that numerical reliability level would have no effect because there is little evidence that people are sensitive to variance information when it is shown numerically.

Experiment 3 also tested the effects of project alignment and reliability level in a non-business population but manipulated both verbal and numerical reliability to enable a direct comparison. The term *reliability level* is used to describe the manipulation of whether NPV was expressed as a reliable measure or not, while *reliability type* is used to describe the manipulation of whether reliability was expressed verbally or numerically. Experiment 3 predicted a reliability level effect for the verbal reliability condition but not the numerical reliability condition. Further, this experiment used project descriptions with clearer profitability indicators and added a larger selection of business industries.

4.2 Experiment 1

Experiment 1 investigated the effects of project alignment and explicit NPV reliability information on capital allocation decisions. The structural alignment literature suggests that people place more weight on alignable differences than they do on non-alignable differences. It was expected that participants would rely more on NPV than on other product attributes in their decision-making because NPV may be applied to every product. However, this effect should vary with participants' perceived NPV reliability level. That is, if other project dimensions are alignable, the use of NPV may depend on its reliability. However, it was predicted that in projects with low alignment, there will be a greater reliance on

NPV as the sole alignable difference, regardless of its stated reliability. These effects were measured by considering the linear relationship between NPV and the money allocated to each project. Critically, the NPV and intrinsic features of each project shown to participants were inversely related. Therefore, a positive NPV trend will indicate a heavier reliance on NPV, whereas a negative trend will indicate a heavier reliance on the intrinsic project features. First, Experiment 1 tested the following omnibus hypothesis:

Hypothesis 4.1—overall effect. The alignment × reliability level × NPV interaction is significant.

Initially, specific effects were tested by excluding the no NPV condition (in which participants were not given NPV information). Given the difficulty of comparing dissimilar projects, participants were expected to rely more heavily on NPV when project attributes are not alignable compared with when they are alignable. Therefore, Experiment 1 tested the following hypothesis:

Hypothesis 4.2—alignment effect. The linear NPV trend will be stronger for projects with low alignment than for projects with high alignment.

Participants' budget allocations were expected to depend on the provided NPV reliability information. However, this is more likely when there are multiple aligned metrics from which to choose compared with when NPV only is alignable. Therefore, Experiment 1 tested the following hypothesis:

Hypothesis 4.3—the NPV reliability level effect depends on alignment. The NPV × reliability level interaction will be stronger in the high alignment than in low alignment.

Specifically, when projects are similar, it is expected that participants will rely more on NPV if they are told that NPV is reliable (leading to a positive NPV trend) but more on the intrinsic features of projects if they are told that NPV is unreliable (leading to a negative NPV trend). However, when projects are dissimilar, it is expected that participants will rely solely on NPV, regardless of what they are told about its reliability. Therefore, Experiment 1 tested the following hypotheses:

Hypothesis 4.4—NPV reliability level in high alignment. When projects have high alignment, the NPV trend will be stronger when NPV reliability is high compared with when NPV reliability is low.

Hypothesis 4.5—NPV reliability level in low alignment. When projects have low alignment, the NPV trend will not differ significantly between the two reliability level conditions.

A no NPV condition was used to gain a better understanding of participants' baseline response to materials when they had no information about NPV. The extent of participants' reliance on NPV was determined by comparing this no NPV condition to the conditions in which NPV was present. When projects are similar, this condition was expected to be equivalent to the low NPV reliability condition because in this condition participants should disregard NPV. When projects are dissimilar, this condition was expected to show the average participant value judgements of the project descriptions, because they only had the intrinsic project features for their evaluations. This was expected to result in a flat NPV trend. Therefore, Experiment 1 tested the following hypotheses:

Hypothesis 4.6—effect of NPV information for projects with high alignment. For projects with high alignment, the positive NPV trend will be stronger for projects with high NPV reliability compared with projects with no NPV information.

Hypothesis 4.7—effect of NPV information for projects with low alignment. For projects with low alignment, the positive NPV trend will be stronger for projects with both low and high NPV reliability compared with projects with no NPV information.

Table 4.1: Experiment 1 group allocation.

Project alignment	Reliability level of net present value (NPV)	N
High	High	26
High	Low	17
High	No NPV	17
Low	High	21
Low	Low	16
Low	No NPV	21
Total		118

4.2.1 Method

4.2.1.1 Participants

One hundred and eighteen participants (55 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 29.42 years (SD = 9.25, $min.$ = 18, $max.$ = 73). Table 4.1 shows the allocation of participants to the different conditions. NPV was varied within subjects.

4.2.1.2 Materials

4.2.1.2.1 Instructions Participants, who did not necessarily have business experience, were first shown an instructions page with information about the task and NPV. These instructions also informed participants about whether NPV as a financial measure was reliable or unreliable for the specific project. Participants in the low NPV reliability level conditions were told that NPV was an unreliable metric, while those in the high NPV reliability level conditions were told that NPV was a reliable metric. Instructions provided to participants in the no NPV condition did not include an explanation of NPV or its reliability. Critically, participants were asked to invest in products with a high objective value (because a higher-quality product is not always better in the consumer goods market). Given that participants may not use this instruction when directly viewing the projects, Experiment 3 used projects whose attributes inherently expressed their quality. Appendix B.1.1.1.1 shows the instructions used in Experiment 1.

	Project 1	Project 2	Project 3	Project 4	Project 5
Product	Laptop	Laptop	Laptop	Laptop	Laptop
RAM (GB)	4	8	32	2	16
Hard drive (GB)	500	750	2000	250	1000
Resolution (px)	900	1080	1440	768	1200
Processor (Ghz)	2.4	3.2	3.8	1.6	3.6
NPV ($)	663	495	70	887	252

Figure 4.1: An example of a high alignment display in Experiment 1.

4.2.1.2.2 Project Display Participants were provided with a set of fictional business projects to which they were asked to allocate capital. Alignment manipulation was reinforced through visual presentation. Projects with high alignment were displayed in a table listing their various attributes (see Figure 4.1). In this group, each project involved the same product type with consistent concrete attributes. The table format was more appropriate for the high alignment condition because all dimensions were shared. In contrast, projects with low alignment were presented as paragraphs describing their relevant attributes (see Figure 4.2). In this group, each project was a different product with concrete attributes specific to that product. In both alignment conditions, each project description included an NPV. Critically, the values of the concrete attributes were always in conflict with the NPV. For instance, Project 4 always had the lowest value for each concrete attribute but always had the highest NPV. This meant that participants' allocations could be used as a proxy for their degree of dependence on NPV.

Presentation style was potentially a confounding factor. This was addressed in Experiment 3 by using the table format for both alignment conditions.

4.2.1.2.3 Allocation Participants completed a capital allocation task (see Figure 4.3) adapted from Bardolet et al. (2011) in which they were asked to allocate

PROJECT 1

This project is about developing a new shampoo. It will have a 400 mL capacity, and the patented Dandruff Reduction Factor was 17 at testing. The fragrance was optimally effective for 3 metres, and the Safety Authority gave it a 81% safety rating. The NPV is estimated to be $685.

PROJECT 2

This project is about developing a new laundry machine. The machine will have a 12-star energy rating and an 8L capacity. The maximum speed rate is 900 rpm and it will have six different cycle programs. The NPV is estimated to be $500.

PROJECT 3

This project is about developing a new mountain bike. It will have a tensile strength of 910 megapascals, and a suspension for travel of 200mm. It will have a 12-speed cassette and is guaranteed for at least three tours. The NPV is estimated to be $81.

PROJECT 4

This project is about developing a new laptop computer. It will have 2GB of RAM and a hard drive with 250GB capacity. The resolution will be 768px, and the processor speed will be 1.6 Ghz. The NPV is estimated to be $894.

PROJECT 5

This project is about developing a new backpack. It will have eight separate compartments for different types of storage, and a total capacity of 30L. The company will offer a four-year warranty, and the material is an 800-denier nylon fabric. The NPV is estimated to be $251.

Figure 4.2: An example of a low alignment display in Experiment 1.

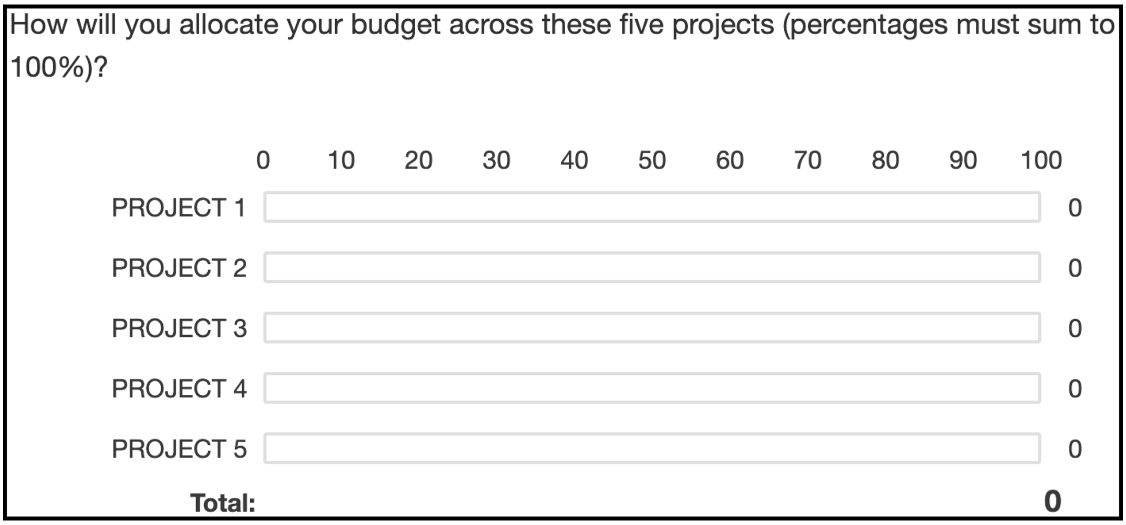


Figure 4.3: The allocation task.

a hypothetical yearly budget across the given five projects.

4.2.1.2.4 Additional Measures Other measures apart from allocation were included. The stimuli for and analyses of these measures are reported in Appendices B.1.1.1 and B.1.2, respectively. Specifically, participants were asked to forecast the future returns of the projects (see Figure B.4), rank the projects (see Figure B.5), indicate their confidence in their decisions (see Figure B.6), and justify their decisions (see Figure B.7).

4.2.1.3 Procedure

After reading the relevant instruction page, participants allocated to the low alignment conditions completed the forecasting task directly beneath each project display. For the high alignment conditions, this was done directly beneath all projects. Participants were then asked to rank the projects and subsequently answer the allocation, confidence, and justification questions.

4.2.2 Results

A mixed factorial ANOVA was conducted to investigate the effects of project alignment and NPV reliability level on participants' budget allocations. As shown in Figure 4.4, the alignment × NPV reliability level × NPV interaction was significant, $F(6.57, 367.76) = 2.18$, $p = .039$, $\hat{\eta}_p^2 = .038$. The analyses excluding the no NPV condition showed the expected results. The NPV trend averaged across both reliability level conditions was stronger for the low alignment conditions than for the high alignment conditions, $M = 61.70$, 95% CI [33.02, 90.37], $t(76) = 4.29$, $p < .001$. This shows that people relied more on NPV when projects were dissimilar than when they were similar.

Further, the NPV × NPV reliability level interaction was stronger in the high alignment conditions than in the low alignment conditions, $M = 67.81$, 95% CI [10.47, 125.16], $t(76) = 2.36$, $p = .021$. Specifically, in the high alignment conditions, the NPV trend was stronger in the high NPV reliability condition than in the low NPV reliability condition, $M = -63.47$, 95% CI [−100.00, −26.94], $t(112) = -3.44$, $p = .001$. In the low alignment conditions, there was no significant difference between the two reliability conditions, $M = 4.35$, 95% CI [−34.52, 43.21], $t(112) = 0.22$, $p = .825$. This shows that participants only used the NPV reliability information in their allocation decisions when projects were similar, not when they were dissimilar.

The comparison with the no NPV condition revealed the expected pattern. For the high alignment group, the linear NPV trend was significantly weaker in the no NPV condition than in the high NPV reliability condition, $M = 75.70$, 95% CI [39.17, 112.24], $t(112) = 4.11$, $p < .001$, but not the low NPV reliability condition, $M = 12.24$, 95% CI [−27.94, 52.41], $t(112) = 0.60$, $p = .547$. However, in the low alignment group, the linear NPV trend was significantly weaker for the no NPV condition compared with both the low NPV reliability condition, $M = 64.63$, 95% CI [25.76, 103.50], $t(112) = 3.29$, $p = .001$, and the high NPV reliability condition, $M = 60.29$, 95% CI [24.14, 96.43], $t(112) = 3.30$, $p = .001$.

The mean ranking, confidence, and forecast data were all largely congruent with the allocation findings (see Appendix B.1.2). The results also show that the forecasts of those in the low alignment condition had higher standard deviations

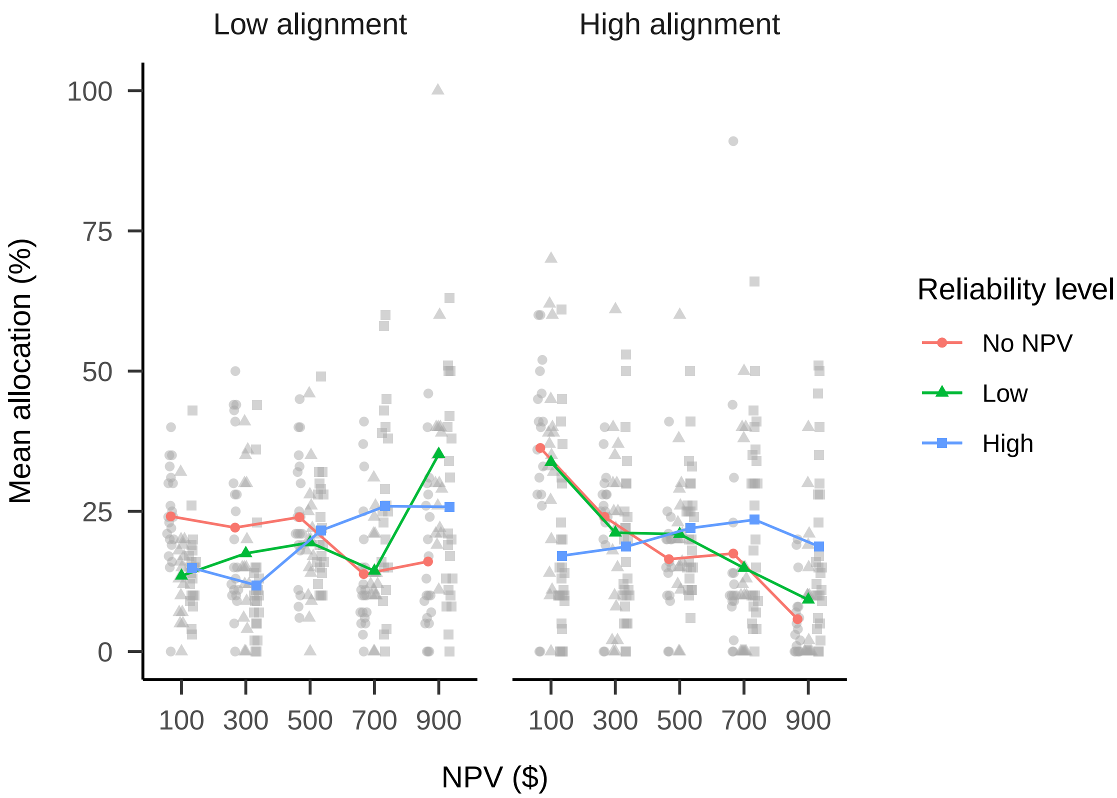


Figure 4.4: Mean allocation across NPV, by project alignment and reliability level conditions. In mixed factorial designs, error bars cannot be used to make inferences by "eye" across all conditions. Therefore, error bars are not included. Raw data are plotted in the background. When interpreting this figure, consider the linear trends in NPV.

than those in the high alignment condition (see Appendix B.1.2.4). However, this was not replicated in subsequent experiments (see Appendices B.5.2.2 and B.6.2.2).

4.2.3 Discussion

Experiment 1 found evidence for the effect of project alignment on laypeople's decision-making in capital allocation scenarios. Specifically, when projects were comparable, participants used NPV when they were told that it was reliable, but did not when they were told that it was unreliable. However, they used NPV regardless of its reliability when it was the only shared dimension across products.

Experiment 1 manipulated *verbal* NPV reliability. That is, participants were explicitly told whether NPV was considered to be a reliable metric or not. However, in the real-world the reliability of a metric is more commonly expressed in numerical form, such as a range around an estimate. Experiment 2 attempted to replicate the alignment effects, while manipulating the *numerical* NPV reliability associated with each project, rather than the verbal reliability as used in Experiment 1. Further, people with sufficient experience with financial theory and analysis may be able to successfully draw inferences from such information. Therefore, Experiment 2 used a sample of students enrolled in a Master of Management degree, instead of the laypeople used in Experiment 1.

4.3 Experiment 2

Experiment 2 investigated the effects of project alignment and numerically-expressed NPV reliability information on capital allocation decisions. In Experiment 1, the information about NPV reliability level was communicated explicitly (e.g., "NPV is unreliable"). However, in Experiment 2, only the actual NPV information itself was communicated without the conclusion about its reliability. Specifically, participants were given a range of predicted values (akin to a confidence interval). Therefore, while Experiment 1 manipulated *verbal* NPV reliability, Experiment 2 manipulated *numerical* NPV reliability. Further, Experiment 2 included participants with more business experience. This experiment tested whether the previous findings of an alignment effect would be replicated using participants with more business experience. The experiment also tested whether this population is sensitive to variance in forecasts.

Hypothesis 4.2 was again tested to investigate the alignment effect in the new sample. However, the other hypotheses tested in Experiment 1 were not retested because Experiment 2 manipulated numerical reliability and did not include a no NPV condition. Research has shown that people are poor at reasoning with numerical variance information (Batteux et al., 2020; Galesic & Garcia-Retamero, 2010; Konold et al., 1993; Vivalt & Coville, 2021). Therefore, Experiment 2 tested the following hypothesis:

Hypothesis 4.8—no effect of numerical reliability. The NPV $\times$ reliability level interaction is not significant in either alignment condition.

Experiment 2 also investigated the ability to quickly change participants' understanding, if they did not initially use numerical NPV reliability in their allocations. Therefore, participants were presented with a short lecture on the importance of paying attention to variance in financial decision-making. However, this lecture was not sufficient to inform participants' use of numerical reliability (see Appendix B.2). Further, Experiment 2 investigated whether participants would be over-confident in their understanding of NPV (as in Long et al., 2018). These results are also reported in Appendix B.2 because they are not highly relevant to this chapter.

4.3.1 Method

4.3.1.1 Participants

Fifty-four participants (28 female) were recruited from a Master of Management degree at an Australian university. Age information was not recorded. Both the reliability level (low and high) and project alignment (low and high) conditions were presented to subjects, and the order of presentation was counterbalanced.

4.3.1.2 Materials

4.3.1.2.1 Instructions Participants were shown similar instructions to those used in Experiment 1 (see Section 4.2.1.2.1). However, they were given more NPV information (including discount rate and initial investment). Appendix B.2.1.1.1 shows the full instructions.

4.3.1.2.2 NPV Test Participants were asked to complete a short, simple test to check their understanding of NPV (see Appendix B.2.1.1.2).

4.3.1.2.3 Project Display As shown in Figures 4.5 and 4.6, projects were displayed as they were in Experiment 1. However, a second set of projects with different product types and descriptions was added to enable within-subjects manipulation. Along with the single numerical NPV, participants were provided with the forecasted cash flow ranges used to calculate the NPV. In the low NPV reliability condition, ranges were ±85% around the mean (e.g., \$150–\$1,850 if forecast mean was \$1,000); while in the high NPV reliability condition, ranges were ±5% around the mean (e.g., \$950–\$1,050 if the forecast mean was \$1,000). A wide range indicated that the measure had low reliability, while a narrow range indicated that the measure had high reliability. Participants were told to treat each set of projects independently.

4.3.1.2.4 NPV Knowledge Ratings Participants were asked to rate their confidence in knowledge of NPV at multiple points in the experiment. Appendix B.2.1.1.3 shows an example of this display.

4.3.1.2.5 Variance Lecture Participants were given a short lecture on the importance of paying attention to variance information in an attempt to increase their use of numerical reliability information in their allocations (see Appendix B.2.1.1.4 for more details and the lecture slides).

4.3.1.3 Procedure

Participants were provided with the instructions and an explanation of NPV before completing a simple test to demonstrate their understanding of NPV. They then completed four counterbalanced capital allocation trials (one for each condition combination) before viewing a brief presentation on the importance of paying attention to variance in financial decision-making. Participants then repeated two of the trials that they had completed earlier. They were shown the allocation values they had provided earlier and were given the opportunity to change them. Participants rated their knowledge of NPV before and after completing the NPV test and then rated it again after completing the four project displays. They were then asked to rate their knowledge of NPV before and after the variance presentation.

4.3.2 Results

A within-subjects factorial ANOVA was conducted to investigate the effects of NPV, project alignment, and numerical NPV reliability on participants' project

PROJECT 1

This project is about developing a new pair of wireless earphones. They will have a frequency response of 16-40,000Hz and a sensitivity of 90 db/mW. The battery life is 8 hours and the pair will isolate noise up to 35 dB. The range of the cash inflow for the first year is $861-$10,619. The NPV is $227.27.

PROJECT 2

This project is about developing a new wrist watch. It will be water resistant up to 50m and will have one extra timing feature. The hardness of the glass face is rated 4 on the Moh scale and the strap is 10% leather. The range of the cash inflow for the first year is $966-$11,914. The NPV is $881.82.

PROJECT 3

This project is about developing a new treadmill. It will 12 training programs for different interests and abilities and 10 speed levels. It will also have two small compartments for storage and three adjustable inclination levels. The range of the cash inflow for the first year is $832.50-$10,267.50. The NPV is $81.82.

PROJECT 4

This project is about developing a new couch. It will have a guarantee for 10 years and a lightfastness level of 5. The cover's ability to resist abrasion has been tested to handle 50,000 cycles and it has a softness rating of 70%. The range of the cash inflow for the first year is $906-$11,174. The NPV is $490.91.

PROJECT 5

This project is about developing a new bottle of perfume. It will have a 100 mL capacity, and the scent will be able to last for 12 hours. The fragrance concentration will be 30%, and there will be two layers of notes to the scent. The range of the cash inflow for the first year is $925.50-$11,414.50. The NPV is $654.55.

Figure 4.5: An example of a low alignment, low reliability display in Experiment 2.

	Project 1	Project 2	Project 3	Project 4	Project 5
Product	Laptop	Laptop	Laptop	Laptop	Laptop
RAM (GB)	4	8	32	2	16
Hard drive (GB)	500	750	2000	250	1000
Resolution (px)	900	1080	1440	768	1200
Processor (Ghz)	2.4	3.2	3.8	1.6	3.6
Cash inflow range for Year 1 ($)	$5,890-$6,510	$5,738-$6,342	$5,244-$5,796	$6,137-$6,783	$5,538.50-$6,121.50
NPV ($)	$636.36	$490.91	$18.18	$872.73	$300.00

Figure 4.6: An example of a high alignment, high reliability display in Experiment 2.

allocations (see Figure 4.7). The alignment $\times$ NPV reliability level $\times$ NPV interaction was significant, $F(2.81, 148.75) = 3.95$, $p = .011$, $\hat{\eta}_p^2 = .069$. However, this appeared to be driven by the difference between alignment conditions in the interaction between the quadratic NPV trend and NPV reliability level, $\Delta M = -42.28$, 95% CI $[-76.96,\ -7.59]$, $t(53) = -3.14$, $p = .011$, even after applying a Šidák correction. The same interaction but using a the linear NPV trend was not significant, $\Delta M = -6.13$, 95% CI $[-31.50,\ 19.25]$, $t(53) = -0.62$, $p = .954$. Further, the linear NPV trend did not differ between the reliability level conditions in either the low alignment condition, $\Delta M = -3.19$, 95% CI $[-18.77,\ 12.40]$, $t(53) = -0.41$, $p = .683$ or the high alignment condition, $\Delta M = 2.94$, 95% CI $[-12.63,\ 18.52]$, $t(53) = 0.38$, $p = .706$. However, averaged across reliability level, the linear NPV trend was stronger in the low alignment condition than in the high alignment condition, $\Delta M = 28.19$, 95% CI $[5.57,\ 50.81]$, $t(53) = 2.50$, $p = .016$. This suggests that participants relied more on NPV when projects were dissimilar compared with when they were similar.

The ranking data were congruent with these results, while the confidence data

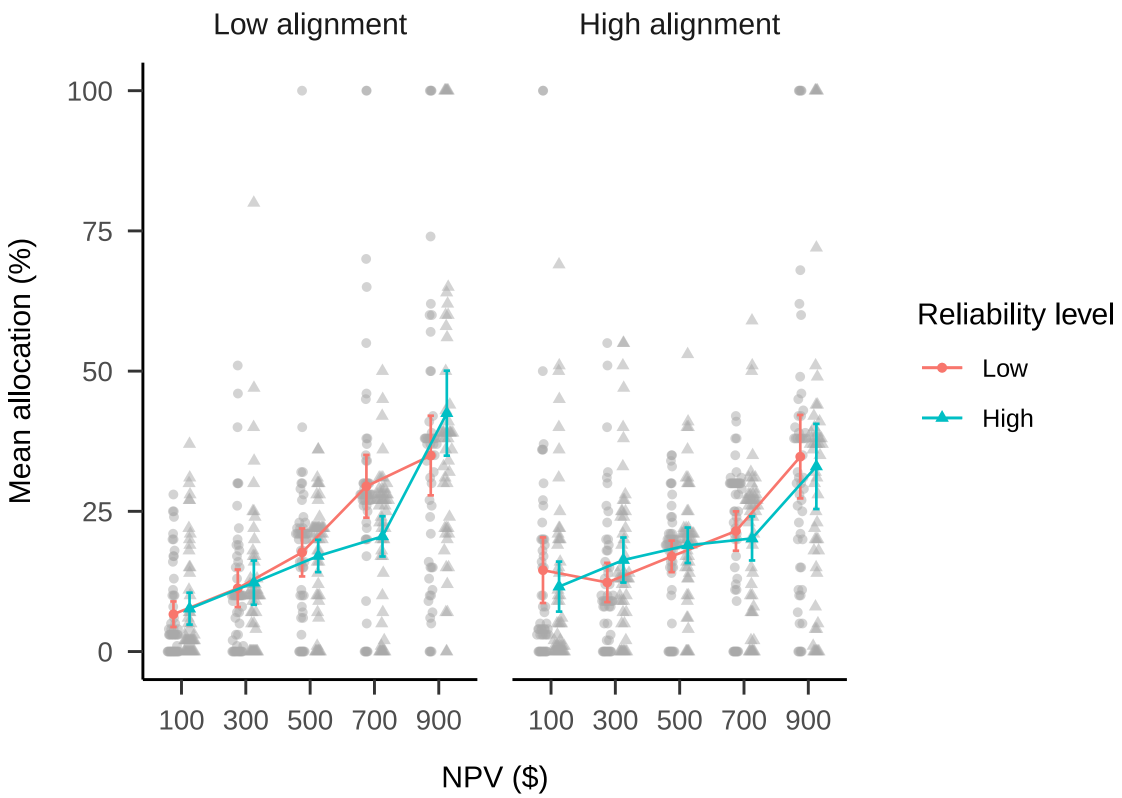


Figure 4.7: Mean allocation across NPV, by project alignment and reliability level conditions. Error bars represent 95% confidence intervals, calculated from the within-subjects standard errors using the method from Cousineau and O'Brien (2014). Raw data are plotted in the background.

were less so. Further, the findings on over-confidence from Long et al. (2018, Study 1) were not replicated with NPV knowledge, and the variance lecture did not facilitate participants' use of numerical reliability information. These analyses are reported in Appendix B.2.2.

4.3.3 Discussion

Based on participants with real-world business experience, Experiment 2 replicated the alignment effect found in Experiment 1. That is, participants relied more on NPV when faced with a set of dissimilar projects than when faced with similar projects, supporting Hypothesis 4.2. Experiment 2 also found evidence for Hypothesis 4.8, with no significant differences between the numerical reliability conditions. While Experiment 2 did not replicate the interaction found in Experiment 1, it should be emphasised that these are two different effects. In

Experiment 1, participants were explicitly told whether the NPV measure was reliable, while in Experiment 2, they were provided with variance information that merely implied NPV reliability. Thus, the results of Experiment 2 show that business students were affected by the comparability of projects but not by numerical NPV reliability information. Specifically, participants appeared to focus only on the NPV itself for a specific project, not on the underlying noisiness of the measure.

The participants in Experiment 2 seemed to rely on NPV more than those in Experiment 1. This was seen by the steeper linear trends in Experiment 2. This discrepancy may be due to the difference in domain experience and exposure to financial metrics in formal study. However, the extra explanation and testing of NPV for the management students may have also increased its salience. In sum, the Experiment 2 sample showed clearer trends of NPV reliance, but importantly was still affected by similarity even when it was manipulated within-subjects.

Experiment 1 tested NPV reliability expressed verbally, while Experiment 2 tested NPV reliability expressed numerically. However, the difference in findings was confounded by the different populations that were sampled. Further, in both experiments, the business projects consisted of a limited number of domains. It is unclear to what extent these specific domains influenced the results. These projects were centred around consumer products, which were chosen to be more easily accessible to participants without business experience. However, the individual features of a project do not necessarily indicate its profitability. For instance, a laptop with a low storage capacity may be more profitable than one with a high storage capacity because of consumer goods markets. Experiment 3 addressed these limitations.

Another limitation of Experiments 1 and 2 was the potential confounding effect of presentation style. The two alignment conditions differed in the number of alignable differences, but also in the way that the information was presented. The information in the low alignment condition was presented as paragraphs, while the information in the high alignment condition was presented as a table. While it is likely that these data types would be presented in this way in the business setting, it is important to rule out that this difference did not unnecessarily increase task difficulty. Therefore, Experiment 3 attempted to replicate this effect while controlling for presentation style

4.4 Experiment 3

Experiment 3 investigated the effects of project alignment, NPV, NPV reliability type and NPV reliability level on participants' budget allocations. Experiment 1 manipulated NPV reliability level using verbal prompts. That is, participants were explicitly told whether or not NPV was reliable for a certain project. Experiment 2 investigated whether people were able to extract the same reliability information using numerical prompts. That is, participants were provided with NPVs with either wide or narrow ranges, indicating either low or high reliability, respectively. Moreover, given that laypeople were sampled for Experiment 1, and Master of Management students were sampled for Experiment 2, it was not possible to compare the two reliability types (verbal and numerical) without ruling out the potential confounding effect of population type. Thus, similar to Experiments 1 and 2, Experiment 3 manipulated project alignment, NPV and NPV reliability level but also added reliability type. Further, presentation style was a possible confounding factor in the previous experiments. That is, projects in the high alignment condition were always displayed in a table, while projects in the low alignment condition were always displayed as paragraphs. This possible confounder was excluded in Experiment 3 by using the same presentation style for both alignment conditions.

In Experiment 3, the expected results for the verbal reliability condition replicated those of Experiment 1. The numerical reliability condition may replicate the findings of Experiment 2. However, a pilot experiment (detailed in Appendix B.8) found no significant differences between numerical reliability conditions. Appendix B.3 shows a simulation of the hypothesised effects, with the numerical reliability effects

Table 4.2: Experiment 3 group allocation.

Project alignment	Reliability type	N
High	Explicit	112
High	Implicit	112
Low	Explicit	112
Low	Implicit	112
Total		448

based on the findings of the pilot experiment. Therefore, Experiment 3 retested Hypotheses 4.1, 4.2, 4.3, 4.4, and 4.5 for the verbal reliability condition, but was agnostic between whether the numerical reliability condition will look more like the pattern found in the pilot experiment or the pattern found in Experiment 2.

4.4.1 Method

4.4.1.1 Participants

Four hundred and forty-eight participants (176 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 41.65 years (*SD* = 10.3, *min.* = 29, *max.* = 78). Participants reported an average of 6.94 years (*SD* = 8.23, *min.* = 0, *max.* = 43) working in a business setting, and an average of 3.73 years (*SD* = 6.27, *min.* = 0, *max.* = 45) of business education. The mean completion time of the task was 11.35 min (*SD* = 10.79, *min.* = 1.92, *max.* = 183.7). Table 4.2 shows the allocation of participants to the different conditions. The two reliability level conditions (low and high) were presented within subjects and the order of their presentation was randomised. Similar to the previous experiments, NPV varied within subjects. Therefore, each participant saw two separate project displays. Appendix B.3.1.1.1 describes the power analysis conducted to arrive at the sample size.

4.4.1.2 Materials

4.4.1.2.1 Instructions Participants were given instructions similar to those in the previous experiments, with an added explanation about the NPV reliability information for each reliability type (see Appendix B.3.1.2.1). Further, they completed a test of basic NPV understanding. Further, they completed a test on basic NPV understanding, which also functioned as an attention check.

4.4.1.2.2 Project Display The project displays were similar to those used in the previous experiments. However, participants were given the same presentation style for both alignment conditions. Each display had a table describing the projects in the set, including ranking and allocation inputs. Project details were presented as bullet points within the relevant cells of the table. Figure 4.8 shows an example of a low alignment, low verbal reliability display; and Figure 4.9 shows an example of a high alignment, high numerical reliability display.

Three elements were counterbalanced: (a) the association between reliability level and project set (two variations), (b) the association between business name and NPV (five latin square variations), and (c) project variation (five variations per alignment condition). When counterbalancing for the high alignment group, projects varied by project type (e.g., whether the five projects all described oil wells or microchips). When counterbalancing for the low alignment group, projects varied by their intrinsic features (e.g., whether the oil well project in the set indicated a probability of finding oil of 96% or 90%). Table column order and project display order were both randomised.

4.4.1.2.3 Interstitial Page Prior to each project being displayed, participants were shown an interstitial page, which was used to (a) introduce the next display, and (b) check the participant's attention (given that no input was required, participants could easily skip the page without reading the text). See Appendix B.3.1.2.2 for an example.

Carefully read through the project descriptions below and then do the following: 1. Rank the projects between 1 (highest) and 5 (lowest) in order of investment priority in the relevant "Project Ranking" row input; and 2. Allocate each project a percentage (a number between 0 and 100) of the total budget in the relevant "Project Allocation" row input.

Relevant information	Project 1	Project 2	Project 3	Project 4	Project 5
Project ranking	Ranking:	Ranking:	Ranking:	Ranking:	Ranking:
Project allocation (%)	Allocation:	Allocation:	Allocation:	Allocation:	Allocation:
Business name	Pressbloom	Cweb	Pharmacore	Erectic	Railmont
Project type	national newspaper	software	pharmaceutical	high-rise construction	railway
Predicted project features	• Newspapers printed: 50,000 a day • Number of weekly advertisers: 80 • Ink that is not discarded due to impurities: 5,000L a day	• Code written: 1,000 lines a day • Security rating: 60% • Number of potential customers in first year: 3 million	• Pills pressed: 300,000 an hour • Shelf life: 20 months • Probability of symptom reduction after a week: 90%	• High-rises built: 8 a year • Probability that the builders complete construction within a month of the due date: 70% • Number of tenant expressions of interest: 100	• Railway lines built: 5 a decade • Number of seats filled by paying customers at peak hour: 2,000 • Time before the train carriages will need to be serviced: 12 years
NPV ($)	501 million. (In this industry, NPV is an **unreliable** predictor of a project's profits.)	611 million. (In this industry, NPV is an **unreliable** predictor of a project's profits.)	722 million. (In this industry, NPV is an **unreliable** predictor of a project's profits.)	806 million. (In this industry, NPV is an **unreliable** predictor of a project's profits.)	416 million. (In this industry, NPV is an **unreliable** predictor of a project's profits.)

Continue

Figure 4.8: An example of a low alignment, low verbal reliability display in Experiment 3.

Carefully read through the project descriptions below and then do the following: 1. Rank the projects between 1 (highest) and 5 (lowest) in order of investment priority in the relevant "Project Ranking" row input; and 2. Allocate each project a percentage (a number between 0 and 100) of the total budget in the relevant "Project Allocation" row input.

Relevant information	Project 1	Project 2	Project 3	Project 4	Project 5
Project ranking	Ranking:	Ranking:	Ranking:	Ranking:	Ranking:
Project allocation (%)	Allocation:	Allocation:	Allocation:	Allocation:	Allocation:
Business name	Liquid Pipeline	Enfuel	Petroyield	Refinera	Oilpier
Project type	oil well	oil well	oil well	oil well	oil well
Predicted project features	• Oil extracted: 3,400L an hour • Time the machinery lasts before requiring maintenance: 11 years • Probability of finding oil: 96%	• Oil extracted: 2,000L an hour • Time the machinery lasts before requiring maintenance: 7 years • Probability of finding oil: 90%	• Oil extracted: 3,870L an hour • Time the machinery lasts before requiring maintenance: 13 years • Probability of finding oil: 99%	• Oil extracted: 2,470L an hour • Time the machinery lasts before requiring maintenance: 8 years • Probability of finding oil: 92%	• Oil extracted: 2,940L an hour • Time the machinery lasts before requiring maintenance: 10 years • Probability of finding oil: 94%
NPV ($)	494-546 million. (Midpoint: 520.)	792-876 million. (Midpoint: 834.)	409-453 million. (Midpoint: 431.)	697-771 million. (Midpoint: 734.)	598-662 million. (Midpoint: 630.)

Continue

Figure 4.9: An example of a high alignment, high numerical reliability display in Experiment 3.

4.4.2 Results

A mixed factorial ANOVA was conducted to investigate the effects of NPV, project alignment, NPV reliability level, and NPV reliability type on participants' project allocations (see Figure 4.10 for the main results and Appendix B.3.2.1 for the remainder of the hypothesised allocation effects). The four-way interaction (alignment × reliability level × NPV × reliability type) was not significant, $F(3.20, 1, 420.19) = 0.71$, $p = .555$, $\hat{\eta}_p^2 = .002$. Regardless, the primary hypotheses were supported.

4.4.2.1 Verbal Reliability

The three-way interaction (alignment × reliability level × NPV amount) in the verbal reliability condition was not significant, $\Delta M = 13.42$, 95% CI $[-1.27,\ 28.11]$, $t(444) = 1.80$, $p = .073$. This is because NPV reliability level interacted with NPV in both alignment conditions. This is a different pattern from Experiment 1 where there was no effect of NPV reliability level in the low alignment condition. In the high alignment condition, the interaction between the linear NPV trend and NPV reliability level was significant, $\Delta M = -36.63$, 95% CI $[-47.02,\ -26.25]$, $t(444) = -6.93$, $p < .001$. Specifically, the trend was stronger for the high reliability condition, $\Delta M = 27.26$, 95% CI $[17.69,\ 36.83]$, $t(444) = 5.60$, $p < .001$, compared with the low reliability condition, $\Delta M = -9.38$, 95% CI $[-18.86,\ 0.11]$, $t(444) = -1.94$, $p = .053$. This shows that, similar to Experiment 1, participants' allocations depended on verbally expressed NPV reliability. In low alignment, there was also an interaction between the linear NPV trend and NPV reliability level, $\Delta M = -23.21$, 95% CI $[-33.60,\ -12.83]$, $t(444) = -4.39$, $p < .001$. This suggests that allocations also depended on verbal reliability in the low alignment condition.

However, another aspect of the data suggests a greater use of NPV in the low alignment condition. The linear NPV trend was stronger in the low alignment condition than in the high alignment condition when averaged over reliability level, $\Delta M = 28.97$, 95% CI $[17.68,\ 40.26]$, $t(444) = 5.04$, $p < .001$. This suggests that when NPV reliability was expressed verbally, similar to Experiment 1, participants relied more on NPV when projects were dissimilar than when they were similar.

Overall, participants used NPV less when it was described as less reliable in both high and low alignment conditions, and further, used NPV more when projects were less alignable regardless of how reliable NPV was described to be.

4.4.2.2 Numerical Reliability

The numerical reliability data were analysed differently to the verbal reliability data because the effects of interest here were the alignment and reliability level effects. The linear NPV trend was stronger in the low alignment condition, averaged over reliability level (with Bonferroni adjustment), $\Delta M = 15.19$, 95% CI [0.78, 29.60], $t(444) = 2.64$, $p = .034$. This pattern was the same as that found for the verbal reliability condition above and in Experiment 2. Further, the linear NPV trend was not significantly different between the reliability level conditions for both the low alignment condition, $\Delta M = 1.64$, 95% CI [−11.61, 14.90], $t(444) = 0.31$, $p > .999$, and high alignment condition, $\Delta M = -1.21$, 95% CI [−14.46, 12.05], $t(444) = -0.23$, $p > .999$. This indicates that participants did not use numerical NPV reliability to inform their allocations.

Similar to the verbal reliability condition, the use of NPV was stronger in the low alignment condition than it was in the high alignment condition. However, unlike the verbal reliability condition, allocations did not depend on numerical reliability in either the low or the high alignment condition. In the verbal reliability condition, allocations depended on NPV reliability level in both alignment conditions.

4.4.3 Discussion

Hypotheses 4.1, 4.2, 4.3, and 4.4 were supported in the verbal reliability condition. This shows that, while overall participants preferred to use NPV as a proxy for project quality in their allocations, they still used verbal reliability information. Specifically, when projects were similar, participants used NPV when they were told that it was reliable, and used alternative metrics when told that it was not reliable. However, in Experiment 3, no support was found for Hypothesis 4.5. It was expected that participants in the low alignment condition would use NPV regardless of the reliability level conditions, as in Experiment 1. Rather, they used NPV less when told that it was unreliable. However, they primarily used NPV overall, as shown by the positive NPV trend in both reliability level conditions.

Further, Experiment 3 replicated the finding of Experiment 2 for the numerical reliability condition. Specifically, participants relied more on NPV when projects were dissimilar but, critically, did not use numerical range information to influence their allocations. A pilot study (documented in Appendix B.8) replicated the

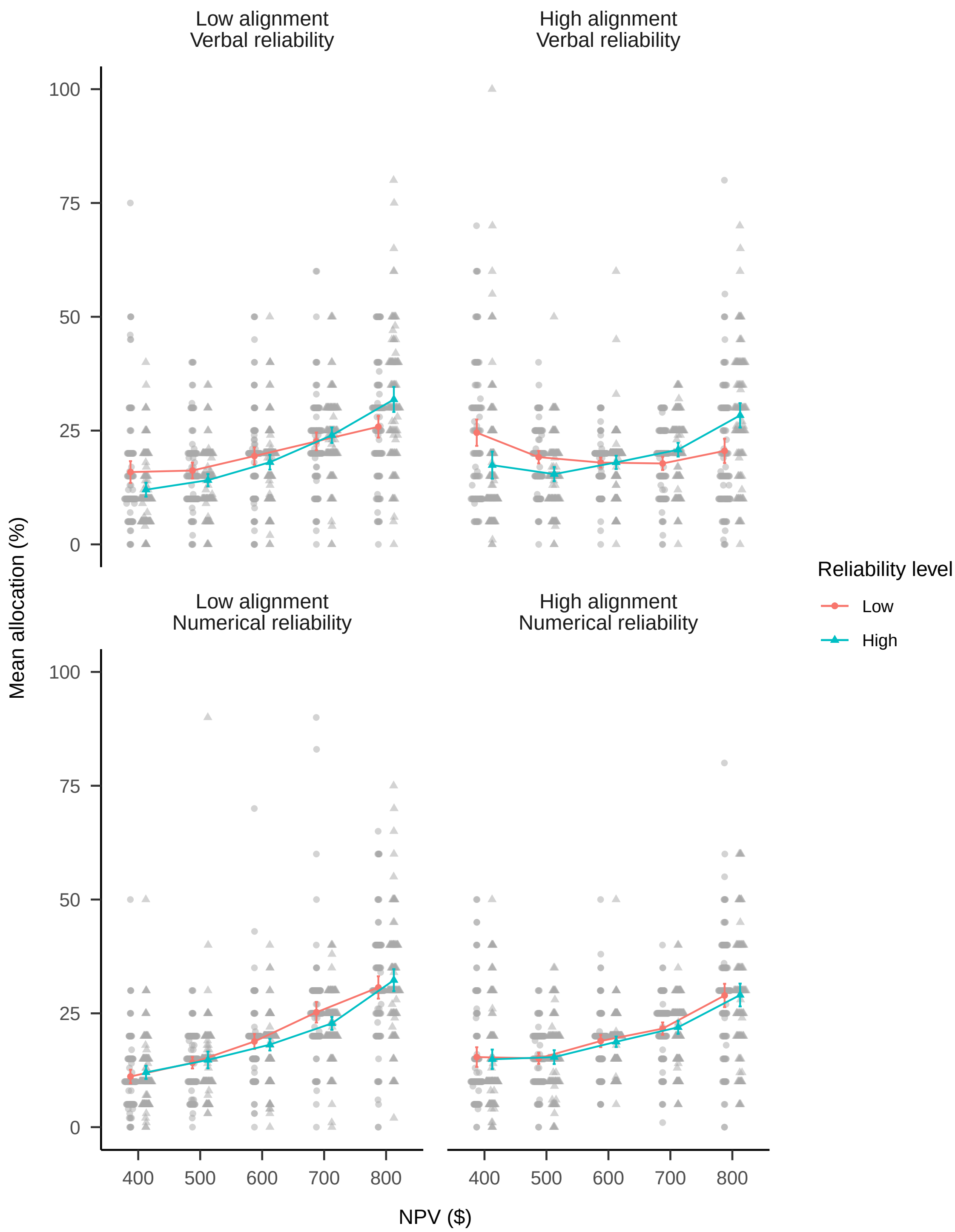


Figure 4.10: Mean allocation across NPV, by alignment, reliability level, and reliability type conditions. Error bars represent 95% confidence intervals, calculated from the within-subjects standard errors using the method from Cousineau and O'Brien (2014). Raw data are plotted in the background.

results of Experiment 1 in the verbal reliability condition, but did not replicate the results of Experiment 2 in the numerical reliability condition. That is, when faced with numerical ranges as the NPV reliability information, participants did not even use the midpoint in their decisions. The results of Experiment 3 suggest that the finding in the pilot experiment may have been spurious or due to an unexplored component of the experimental design, but this can only be deteremined with future research.

4.5 General Discussion

Across three experiments there were two main findings: (a) NPV is used more when options are difficult to compare in the low alignment conditions; and (b) people do not consider numerical variance information, despite this being important to the reliability of the NPV forecasts. This pattern with numerical reliability information contrasted with the frequent use of verbal indicators of reliability level. This numerical variance neglect is surprising, since other work showed that people can readily extract variance information when experiencing numerical sequences (Rosenbaum et al., 2020). Both the verbal and numerical effects were consistent for both naive and experienced participants, indicating their persistence. People make use of metrics with alignable differences when required to compare disparate options. However, they do not always use alternative metrics, even when they are available.

Experiment 1 found that participants did not use NPV in their allocation decisions when they were told that it was unreliable but did use it when told it was reliable. Experiment 2 found that participants with some business experience relied more on NPV for capital allocation when the rest of the information was non-alignable compared with when it was alignable. However, they did not take into account numerical reliability information when making these decisions. Experiment 3 found further evidence of these effects within one experimental design.

Alignable differences have been shown to be important into decision-making in many settings (Markman & Loewenstein, 2010; Markman & Medin, 1995). The experiments presented in this chapter are novel in terms of the effects of project alignment on capital allocation. Further, these experiments considered the extent to which the reliability of an alignable measure (NPV) affects the way in which it is used. This depended on the availability of other alignable differences in the set of choices. If other alignable differences were available, then participants were willing to reduce their use of a reportedly unreliable alignable measure (or use it when told that it was reliable). However, when no other alignable differences were available, then the alignable, albeit unreliable, measure was more likely to be used. This was found in both Experiments 1 and 3, as well as in a pilot study to a lesser extent (reported in Appendix B.4).

Financial measures such as NPV are useful because of their alignability. That is, they may serve as an alignable difference, regardless of the inherent similarities between a set of projects. Psychologically, these measures are useful because they allow for relevant inferences (Lassaline, 1996) and because they offer an abstraction of concrete details (Doumas & Hummel, 2013). However, the structural alignment account does not directly speak to real-world implications when there is a need for non-alignable comparisons. NPV is a type of abstraction that facilitates the comparison of different aspects of a company. For instance, the use of NPV may facilitate the comparison of an oil field project with a refinery project. However, this increased alignment could actually hide important information because it does not consider the finer complexities inherent in each business unit. The forecasts used to calculate NPV for each business unit are based on different indicators, and there are likely to be differences between each unit's estimates. Thus, one can imagine a continuum of comparisons in which the usefulness of comparison increases with the level of alignability but depends on the level of abstraction that is required to achieve the alignment.

The finding that participants, even those with some business experience, did not sufficiently consider variance information is surprising but understandable. It is surprising because financial decision-making largely depends on the consideration of different sources of variance (e.g., risk, volatility, and uncertainty). At the same time, it is understandable because research from psychology and statistics education shows that statistics students and people in general have a poor ability to draw statistical inferences (e.g., Galesic & Garcia-Retamero, 2010; Konold et al., 1993). Future research should investigate the conditions under which individuals' sensitivity to variance information may be facilitated. For instance, it is unclear whether it is merely salience that is lacking, meaning that visual aids could be useful, or whether a further explicit explanation of statistical inference is necessary. The findings of a pilot experiment suggest that participants struggle to use numerical NPV reliability information, even when given explicit instructions (see Appendix B.7).

A possible limitation of these experiments is the use of NPV as the only financial metric. In the business world, there are many metrics that serve similar functions and are used as tools to deal with non-alignable options. Therefore, future research should attempt to replicate the current findings using different financial measures.

Future research should also investigate the boundary conditions of the reliability type effect. That is, people appear to respond to explicit reliability information but not to variance information that only implies reliability. Future research should attempt to identify the minimal variance information that participants need to understand the relevant implications for reliability. Participants may simply not notice the variance information or assume that it is irrelevant. For instance, future research could test participants in a condition in which the variance information is more salient.

References

Bardolet, D., Fox, C. R., & Lovallo, D. (2011). Corporate capital allocation: A behavioral perspective. *Strategic Management Journal*, *32*(13), 1465–1483. https://doi.org/10/cn6xsb

Batteux, E., Bilovich, A., Johnson, S., & Tuckett, D. (2020, May 7). *Impressed by Numbers: The Extent to Which Novice Investors Favor Precise Numerical Information in a Context of Uncertainty* (SSRN Scholarly Paper No. ID 3595409). Social Science Research Network. Rochester, NY. https://doi.org/10.2139/ssrn.3595409

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, *46*(4), 1149–1151. https://doi.org/10/f6vdsw

Doumas, L. A. A., & Hummel, J. E. (2013). Comparison and Mapping Facilitate Relation Discovery and Predication (O. Sporns, Ed.). *PLoS ONE*, *8*(6), 1–8. https://doi.org/10/gjscsn

Fox, R. (2008–April 3). A brief critical history of NPV. *BAA Conference*, 16. http://usir.salford.ac.uk/id/eprint/9291/2/NPV_paper.pdf

Galesic, M., & Garcia-Retamero, R. (2010). Statistical Numeracy for Health: A Cross-cultural Comparison With Probabilistic National Samples. *Arch Intern Med*, *170*(5), 462–468. https://doi.org/10/fmj7q3

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10/dw52z8

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45–56. https://doi.org/10/fm4rrb

Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, *60*(2), 187–243. https://doi.org/10/fpdzrj

Graham, J. R., Harvey, C. R., & Puri, M. (2015). Capital allocation and delegation of decision-making authority within firms. *Journal of Financial Economics*, *115*(3), 449–470. https://doi.org/10/gfvz8d

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in Students' Reasoning about Probability. *Journal for Research in Mathematics Education*, *24*(5), 392. https://doi.org/10/bq4hvm

Lassaline, M. E. (1996). Structural alignment in induction and similarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 754–770. https://doi.org/10/fq9fww

Long, A. R., Fernbach, P. M., & De Langhe, B. (2018). Circle of Incompetence: Sense of Understanding as an Improper Guide to Investment Risk. *Journal of Marketing Research*, *55*(4), 474–488. https://doi.org/10/gjscr7

Lovallo, D., & Kahneman, D. (2003). Delusions of Success: How Optimism Undermines Executives' Decisions. *Harvard Business Review*, *81*(7).

Markman, A. B., & Loewenstein, J. (2010). Structural comparison and consumer choice. *Journal of Consumer Psychology*, *20*(2), 126–137. https://doi.org/10/d7b49c

Markman, A. B., & Medin, D. L. (1995). Similarity and Alignment in Choice. *Organizational Behavior and Human Decision Processes*, *63*(2), 117–130. https://doi.org/10/c8z7r9

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, *86*(1), 71–99. https://doi.org/10/c9839j

Remer, D. S., Stokdyk, S. B., & Van Driel, M. (1993). Survey of project evaluation techniques currently used in industry. *International Journal of Production Economics*, *32*(1), 103–115. https://doi.org/10/bsc6bs

Rosenbaum, D. M., Glickman, M., & Usher, M. (2020, July 30). *Extracting summary statistics of rapid numerical sequences* (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/6scav

Vivalt, E., & Coville, A. (2021, April 23). *How Do Policy-Makers Update Their Beliefs?* http://evavivalt.com/wp-content/uploads/How-Do-Policymakers-Update.pdf

Willigers, B. J. A., Jones, B., & Bratvold, R. B. (2017). The Net-Present-Value Paradox: Criticized by Many, Applied by All. *SPE Economics & Management*, *9*(04), 090–102. https://doi.org/10/gjscsx

5

Seeking Alignment in Past Cases

Chapter 4 found that people do not sufficiently perceive the importance of numerical variance information in capital allocation. This is important when business projects are dissimilar because people may fail to pay attention to the differing variance underlying NPV across different domains. However, there are also implications for high alignment scenarios. When projects are alignable, managers are likely to be able to use abstract metrics as well as intrinsic project features. Managers may use a metric such as NPV, the variance of which may suggest a lack of reliability, despite being able to use intrinsic project features. Therefore, they may miss the opportunity to use different and potentially more reliable measures.

Therefore, the evaluation of a non-alignable set of projects has many potential pitfalls. This situation is likely to occur in most hierarchical organisations, especially those that are highly diversified. As discussed in Chapter 3, a solution for managers who fail to aggregate the risk of multiple projects may be for them to concurrently evaluate projects as a portfolio. However, the solution to the evaluation of dissimilar projects in diversified organisations is likely to involve significantly more difficult structural changes in the organisation. For instance, this may mean divesting certain divisions of the organisation, as GE has done in the last few years (Scott, 2018).

Other solutions are also possible. For instance, organisations may develop a more normative use of metrics and take into account underlying uncertainties. However, this change may require substantially more statistical reasoning abilities than should be expected of managers without better decision-making guidelines. Another solution for managers is to seek evidence from similar projects from outside of the organisation. This may be useful because a diversified organisation may not have enough points of reference for a project proposal. It would also mean that substantial organisational restructuring such as divestment or training managers in statistical reasoning would not be required.

Evidence from similar projects may come in the form of an individual case study from another organisation or a research report that describes a statistical result. Case studies are especially important in managerial decision-making because they are used extensively in business school teaching materials. Therefore, managers are likely to seek case studies that may be used to inform their decisions. However, do they believe that a single case study is more useful than statistical data? The literature on anecdotal bias suggests that they might. Chapter 6 considers the influence of anecdotes on project allocation when they conflict with statistical evidence.

Previous work shows that people often do not give evidence appropriate weighting in their decisions (Griffin & Tversky, 1992). Statistical and anecdotal evidence often conflict because statistical estimates commonly refer to the mean value of a distribution, while individual cases may be sampled from either tail of the distribution. This comparison may produce conflicting information, especially if the distribution is skewed; therefore, it is important to appropriately weigh their influence when making a decision. In the same way that intrinsic project features conflicted with the abstract financial metrics in Chapter 4, anecdotal evidence conflicts with financial metrics of the target project in Chapter 6.

Chapter 6 considers how people deal with such conflicting information. That is, do they focus on one metric or use a trade-off? In the previous chapter, participants did not appear to predominantly use any one specific cue. The fact that those in the low alignment condition relied more on NPV compared with those in the high alignment condition means that the latter were still referring to intrinsic project features to some extent. Specifically, the influence of different measures may have been integrated in a type of trade-off. However, there was no clear way of determining this because the allocation measure was aggregated in the analysis. In Chapter 6, however, conditions are set up so that it is possible to determine whether participants were using anecdotes exclusively, partially, or not at all.

We like stories, we like to summarize, and we like to simplify

—Nassim Nicholas Taleb (2007, p. 63)

6

Anecdotal Bias in Capital Allocation Depends on Anecdote Similarity

Contents

6.1 Introduction

A good story is often more persuasive than data. While usually harmless in daily settings, poor judgement arising from a bias towards anecdotal evidence can lead to large-scale negative consequences. Perhaps the most prominent example of such an error in judgement is the belief that a vaccine causes a certain disorder based on isolated stories, despite contradictory scientific evidence. An analogous error exists in settings such as managerial decision-making. In business, managers use analogies, known as *case studies*, as a part of their strategic decision-making. Case studies are examples of previous situations considered similar by the decision-maker and are used to draw inferences about a target problem. Case studies are known as *anecdotes* when comparing them with aggregated data.

Many businesses use case studies to inform their decisions but often struggle to use them successfully (Gavetti & Rivkin, 2005). This may be attributable to the prominence of companies that are either highly successful or highly unsuccessful. That is, people are often uninterested in average outcomes but are captivated by both positive and negative extreme outcomes. The increased salience of an anecdote may increase its influence over that of useful statistical data. Further, increased anecdotal salience may also shift attention away from structural similarities in favour of more surface similarities. Both of these issues may explain the unsuccessful use of case studies.

The first consideration when using a case study is its merit relative to available aggregated statistical data. That is, if the case study is a single data point in a set of other relevant cases, then using the statistical properties of the larger sample is more inferentially informative than using a single case from within the sample (unlike perhaps when the single case is somehow the most relevant example from the sample). Despite the utility and availability of large sample data, research has shown that people often prefer anecdotal evidence over statistical data (Freling et al., 2020; Jaramillo et al., 2019; Reinard, 1988; Shen et al., 2015).

However, if this larger sample is not available (or is ignored), then the second consideration when using a case study is the extent of its similarity to the target problem. Research on the psychology of similarity judgements distinguishes between surface and relational similarity (Gentner, 1983). The consensus of this research is that the more conceptual structures that two cases share, the more useful they are in decision-making (Lassaline, 1996; Markman & Medin, 1995). Therefore, case studies that are similar to a target problem on a merely surface level are less useful than those that are related through a shared conceptual structure.

Previous research has considered the role of similarity and analogical reasoning in business-related decision-making (e.g., Gavetti et al., 2005). Others have investigated the influence of anecdotes in capital allocation decisions and the impact of anecdote similarity on their persuasiveness (summarised below). However, it is unclear to what extent an anecdote's similarity to the target problem will affect its influence on capital allocation decisions. Further, it is unclear whether people will be sensitive to information about the distribution from which the anecdote was sampled.

6.1.1 Anecdotal Bias

Anecdotal bias refers to the influence of anecdotal evidence over statistical evidence on people's beliefs. Journalists, for instance, are well aware of the power of anecdotes. An analysis of approximately 29,000 New York Times editorials showed a reliance on anecdotes to drive arguments (Al Khatib et al., 2017). While some studies have concluded that statistics are more persuasive than anecdotes (e.g., Allen & Preiss, 1997; Hoeken, 2001; Hornikx, 2005) and others provided more cautious conclusions (Winterbottom et al., 2008), a number have found evidence for anecdotal bias (e.g., Jaramillo et al., 2019; Ratcliff & Sun, 2020; Reinard, 1988; Reinhart, 2006; Shen et al., 2015). Zebregs et al. (2015) suggest that this disparity in findings might be attributable to statistics affecting beliefs and attitudes, and anecdotes affecting intention. A more recent meta-analysis of 61 studies found that, overall, statistical evidence is more persuasive than anecdotal evidence (Freling et al., 2020). However, even if statistical evidence is more persuasive overall, anecdotes that add no additional information to co-presented statistics may still influence people's judgement (Jaramillo et al., 2019). Further, the meta-analysis found that people tend to prefer anecdotal evidence over statistical data when the stakes are more emotional, medical, or relevant to the decision-maker. In business, decisions are clearly relevant to the decision-maker.

6.1.2 Anecdotal Bias in Business

It is important to investigate anecdotal bias in business because of its implications for managers' use of case studies. There are many cases of managers successfully using analogies from anecdotal cases but also of failures to analogise correctly (Gavetti et al., 2005; Gavetti & Rivkin, 2005). There is very little research on anecdotal bias in business, but the existing work finds clear evidence of the effect. In fact, the recent meta-analysis by Freling et al. (2020) included the work of Wainberg et al. (2013) as one such paper.

Wainberg et al. (2013) gave a sample of managers and other professionals a choice between two audit firms, which varied in terms of their audit deficiencies for various clients. The experiment was designed in such a way that the statistical evidence favoured one firm, while the anecdotal evidence favoured the other firm. Participants were allocated to one of five conditions. Participants in the *anecdotes only* condition were given anecdotal examples of firm deficiencies, while those in the *anecdotes & statistics* condition were given the same anecdotal examples as well as the number of clients and deficiencies found. However, participants were not explicitly provided with the proportions of these values. Participants in the *statistics only* condition were given this proportions information as well as the number of clients without deficiencies but no detailed examples of deficiencies. The *anecdotes & enhanced statistics* condition included both anecdotes and the information in the statistics only condition. The terminology here is confusing because nothing about the way the statistics are presented to the participants is "enhanced" beyond how they are presented in the statistics only condition. However, the *anecdotes & enhanced statistics—judgment orientation* condition emphasised the importance of proportions and keeping absolute numbers in context.

Wainberg et al. (2013) measured the percentage of participants who chose firms favoured by the statistical data, finding evidence of anecdotal bias. Participants in the anecdotes only and anecdotes & statistics conditions equally chose the firm favoured by statistical data. However, participants in the anecdotes & enhanced statistics condition were less likely to choose this firm compared with those in the statistics only condition, even when the underlying proportions were made explicit. This shows evidence of anecdotal bias because participants ignored contradictory statistical data. The lack of difference between the anecdotes & statistics condition and the anecdotes only condition implies that the anecdotal bias effect was "complete". That is, the presented statistics did not play a role in influencing participants' choice of firm. A "partial" effect would have occurred if more participants in the anecdotes & statistics condition had chosen the firm compared with participants in the anecdotes only condition. This would have meant that statistics played at least some role in influencing choice.

The other important finding in this work is that anecdotal bias was reduced by highlighting relevant statistical features and providing an explanation of statistical inference. This is important because it suggests that potential psychological biases can be reduced with a reframing of provided information and an explanation of relevant statistical concepts.

Wainberg (2018) conducted a similar study to that of Wainberg et al. (2013) but with a capital budgeting task as opposed to a binary choice. Participants had to choose between three production line machines for a mid-sized company that prints circuit boards. The statistical data suggested that Machine A was better than Machine B, and Machine B was better than Machine C. Participants were given only statistical information or statistical information along with an anecdote. The anecdote was in the form of an email from a colleague who recommended against Machine A (the best option). Similar to Wainberg et al. (2013), participants were assigned to *anecdote & statistics* and *statistics only* conditions. In the *judgement* *orientation I* and *judgement orientation II* conditions, participants were told to "think like a scientist" and received either a short or a long explanation, respectively, of the importance of statistical inference.

Wainberg (2018) found evidence for anecdotal bias. Including a contradictory anecdote alongside statistical evidence (the anecdote & statistics condition) reduced the proportion of participants who chose Machine A. The study also found that the addition of instructions that emphasised scientific thinking reduced this bias. Unlike Wainberg et al. (2013), Wainberg (2018) could not determine whether the anecdotal bias was a complete or partial because there was no anecdote only condition. Further, neither work considered the effect of the anecdote's similarity to the target problem.

6.1.3 Effect of Similarity

Arguably, the extent of one's reliance on an anecdote should depend on its similarity to the target problem. Previous work has examined the importance of weighting previous cases according to their similarity to the present situation (Gilboa & Schmeidler, 1995; Lovallo et al., 2012). For instance, consider a medical treatment with contradictory statistical and anecdotal evidence; that is, a large-scale aggregated study has found that the treatment has 99% efficacy, while someone reports on social media that they became sick as a side-effect of the treatment. While the decision to use the treatment should be informed more by the aggregated data than by the anecdotal data, an individual may have reason to be concerned if the person who became sick was their identical twin. Therefore, the inference that the individual may also need to be cautious about the treatment arises from a specific causal model based on the shared genetics of the two cases.

There have been mixed results regarding the effect of anecdote similarity on the extent of anecdotal bias. Hoeken and Hustinx (2009, Study 3) found evidence for the effect of similarity on anecdotal bias for a variety of claims. As well as manipulating whether participants received a claim supported by anecdotal or statistical evidence, they manipulated whether the anecdotal evidence was similar or dissimilar to the claim that it was supporting. They found that similar anecdotes were more persuasive than dissimilar anecdotes. Using a student sample, Hoeken (2001) did not find evidence for the effect of similarity to a local government proposal. Similarly, Hornikx (2018) considered the effect of similarity on anecdotal bias in local government policy decision-making. The researchers did not find an effect for similarity or for anecdotes. However, they measured persuasiveness, and it may be that requiring participants to make more concrete decisions will create a more realistic scenario.

Apart from the need to determine the effect of similarity on the anecdotal bias effect, it is important to clarify how such an effect might work. Research on analogical reasoning has distinguished between simple surface similarity and deeper relational similarity (Gentner, 1983). As mentioned above, one's use of an anecdote should depend on the extent to which it is associated by an underlying causal mechanism or mere surface similarity. Imagine a manager of a multi-divisional company deciding on the allocation of capital between an oil well project and a technology project. Would hearing of a recent failed oil well project at another company influence the manager's allocation decision? If so, would it influence the manager's decision because the anecdote has similarities to the target oil well project (surface similarity)? Or would the manager seek out the underlying reason for the failure of the other company's oil well project to identify whether it is relevant to the target oil project (relational similarity)? The experiments presented in this chapter investigated whether the anecdotal bias effect arose from causal inductive reasoning or merely the surface similarity with the target project.

6.1.4 Experiment Summary

Experiment 1 investigated whether anecdotal bias in a capital allocation paradigm depends on anecdote similarity. Further, it tested whether providing additional statistical information encourages participants to consider the statistics over the anecdote. Experiment 1 used a negative anecdote, which is an example of an unsuccessful case. This kind of anecdote has been shown to produce anecdotal bias in both medical (Jaramillo et al., 2019) and business (Wainberg, 2018) decision-making. However, Jaramillo et al. (2019) found less bias in positive anecdotes, which are examples of successful cases, and Wainberg (2018) did not consider these at all. Therefore, Experiment 2 attempted to replicate the effect of similarity on anecdotal bias using a positive anecdote. Further, Experiment 2 provided participants with information about the sample distribution of the anecdote, whereas Experiment 1 did not. This allowed for an informal test of whether people are sensitive to such information.

6.2 Experiment 1

Experiment 1 investigated the effects of anecdote similarity and bias on capital allocation. Participants were assigned to the same conditions as those in Wainberg (2018) except that an anecdote only condition was included and the judgement orientation I condition was excluded. They were then asked to allocate a hypothetical budget between two business projects. Participants were also presented with a case study that was either similar or dissimilar to the target project (but still from the same industry). Further, for the conditions in which statistical evidence was provided, participants were presented with aggregated information about the success of similar projects in the form of NPV as well as a reliability measure. One project was clearly better than the other in terms of the statistical data, but the anecdotal evidence suggested the opposite.

Previous research has found that people are persuaded more by negative anecdotes than by positive statistical data in capital allocation scenarios (Wainberg, 2018). While studies have shown that similar anecdotes are more persuasive than dissimilar anecdotes (Hoeken & Hustinx, 2009, Study 3), it is unclear how the anecdotal bias effect may depend on anecdote similarity. Thus, the main question is whether anecdotal bias will be greater when the anecdote is similar to the target project compared with when it is dissimilar. The target project was supported by the statistics but was inconsistent with the anecdotes. Further, Experiment 1 only used negative anecdotes. Therefore, the experiment would show evidence of anecdotal bias if participants assigned to the statistics only condition allocated more money to the target project compared with those in the anecdote & statistics condition. Therefore, Experiment 1 tested the following hypothesis:

Hypothesis 6.1—anecdotal bias depends on the similarity of negative anecdotes. Budget allocations to the target project will be higher when statistics only are presented compared with when statistics are accompanied by an anecdote with high similarity to the target project. In addition, budget allocations will not be affected by anecdotes with low similarity. That is, the statistics only condition will not differ from the low-similarity anecdote & statistics condition.

Experiment 1 predicted that that the anecdotal bias effect will be complete, as in Wainberg et al. (2013). Specifically, the participants presented with the high-similarity anecdote along with the statistics would not use any statistical information. Testing the high similarity condition will provide an equivalent test to that of Wainberg et al. (2013). Therefore, Experiment 1 tested the following:

Hypothesis 6.2—effect of statistics for negative anecdotes. Participants in the high-similarity anecdote & statistics condition (without the enhanced statistics explanation) and those in the high-similarity anecdote only condition will allocate capital equally to the target project.

Participants with additional information on the importance of scientific thinking and statistical data may be less affected by anecdotes. Wainberg (2018) termed this the *judgement orientation II* condition, while in this experiment it is termed the *anecdote & enhanced statistics condition*. Unlike the anecdote & enhanced statistics condition in Wainberg et al. (2013), the statistical information here is actually "enhanced" because of the accompanying text. Experiment 1 tested whether the effect of additional information on anecdotal bias would be replicated in a capital allocation scenario. Therefore, Experiment 1 tested the following hypothesis:

Table 6.1: Experiment 1 group allocation.

Evidence type	Project alignment	N
Anecdote & enhanced statistics	High	41
Anecdote & enhanced statistics	Low	41
Anecdote & statistics	High	41
Anecdote & statistics	Low	40
Anecdote only	High	41
Anecdote only	Low	40
Statistics only	NA	40
Total		284

Hypothesis 6.3—effect of enhanced statistics for negative anecdotes. Participants in the high-similarity anecdote & enhanced statistics condition will allocate more capital to the target project than those in the high-similarity anecdote & statistics condition.

6.2.1 Method

6.2.1.1 Participants

Two hundred and eighty-four participants (197 female) were recruited from a cohort of psychology undergraduates at The University of Sydney. Participants were compensated with course credit. The average age was 20.84 years (SD = 4.93, *min.* = 18, *max.* = 58). Participants reported an average of 1.68 years (SD = 3.63, *min.* = 0, *max.* = 32) working in a business setting, and an average of 0.81 years (SD = 1.57, *min.* = 0, *max.* = 12) of business education. The mean completion time of the task was 22.24 min (SD = 97.45, *min.* = 1.67, *max.* = 1,101.48). Table 6.1 shows the allocation of participants to the different conditions. Appendix C.1.1.1.1 describes the power analysis conducted to arrive at this sample size.

6.2.1.2 Materials

6.2.1.2.1 Instructions All participants were first shown general instructions explaining the task. Subsequent instructions shown to participants depended on their experimental condition. Those in the anecdote only condition were told that they would be shown a case study of a failed project and an analysis of why it failed. Those in the statistics only condition were told that they would be shown NPV and reliability information for two focal projects. They were told that these values were sourced from a study with a large sample. Those in the anecdote & statistics condition were shown both of these instructions and were also told that the information in the anecdote had been included in the aggregated study data. Those in the anecdote & enhanced statistics condition were shown the same instructions as those in the anecdote & statistics condition, but were also provided with the explanation of scientific thinking used by Wainberg (2018). Appendix C.1.1.2.1 shows the instructions used in Experiment 1.

6.2.1.2.2 Allocation Task In the allocation task, participants were asked to allocate a hypothetical budget to one of two projects from two different businesses. In this chapter, these projects are referred to as the *focal* projects, with one being the *target* project and the other the *comparison* project. The target project was used as the reference for the similarity manipulation. That is, the anecdote was either high or low in similarity to the target project. Further, the data analyses presented in Section 6.2.2 used allocations to the target project as the dependent variable. The comparison project was simply the other focal project to which participants were allocating. It was always a different type of project to both the target and anecdote projects.

Participants were presented with information about the name, location, integration (vertical or horizontal), and organisational structure (centralised or decentralised) of each business (see Appendix C.1.1.2.2 for an explanation of these terms). Further, they were presented with information about the features of each project that are typically available to managers prior to investment. Participants in the anecdote only condition were shown only this information (see Figure 6.1). Those in the anecdote & statistics, anecdote & enhanced statistics, and statistics only conditions were shown this information along with measures of NPV and

Target projects

Allocate your budget between the following two projects using percentage values (the two values should sum to 100):

Relevant information	**Project A**	**Project B**
Business name	Enfuel	Microxy
Investment	oil well	microchip
Location	Texas, USA	Manchester, UK
Integration	vertical	horizontal
Structure	centralised	decentralised
Predicted project features	• Oil extracted: 2200L an hour • Time the machinery lasts before requiring maintenance: 8 years • Probability of finding oil: 88% • Type of well: onshore	• Microchips produced: 4000 an hour • Usable semiconductor yield after testing: 60% • Compatible PCs in the market: 75% • Type of integrated circuit: digital

Project A allocation: %

Project B allocation: %

Figure 6.1: Focal project display for the anecdote only condition in Experiment 1. Here, Project A was the target project and Project B was the comparison project.

overall reliability rating (see Figure 6.2). Participants entered their allocation data beneath this table in two text boxes labelled *Project A allocation* and *Project B allocation*, respectively.

Target projects

Allocate your budget between the following two projects using percentage values (the two values should sum to 100):

Relevant information	**Project A**	**Project B**
Business name	Enfuel	Microxy
Investment	oil well	microchip
Location	Texas, USA	Manchester, UK
Integration	vertical	horizontal
Structure	centralised	decentralised
Predicted project features	• Oil extracted: 2200L an hour • Time the machinery lasts before requiring maintenance: 8 years • Probability of finding oil: 88% • Type of well: onshore	• Microchips produced: 4000 an hour • Usable semiconductor yield after testing: 60% • Compatible PCs in the market: 75% • Type of integrated circuit: digital
Overall reliability rating (%)	95	87
NPV ($)	900	100

Project A allocation: %

Project B allocation: %

Figure 6.2: Focal project display for the statistics only, anecdote & statistics, and anecdote & enhanced statistics conditions in Experiment 1. Here, Project A was the target project and Project B was the comparison project.

6.2.1.2.3 Anecdote Participants who were presented with an anecdote (those in either the anecdote only, anecdote & statistics, or anecdote & enhanced statistics conditions) were shown a description of another business project and an accompanying analysis. Figures 6.3 and 6.4 show the anecdotes for those in the high and low similarity conditions, respectively. The project description had a similar layout to that of the focal projects. That is, it contained information about the business name, location, integration, and organisational structure of the business. It also detailed several predicted features of the project. Beneath this description was a paragraph presenting an analysis of why the project had failed. This paragraph referenced each of the features in the description to justify the failure of the project.

Participants in the high similarity condition were shown a description of a project from a business with the same type of investment as the target project (Project A). All categorical attributes were identical to those in Project A, but all numerical attributes were lower. The analysis explained that the numerical attributes had failed because they had not reached certain cut-offs. Critically, these cut-offs were all higher than the matching values in Project A. This was done to ensure that the numerical attributes in the anecdote appeared more relevant than those in Project A. For instance, in Project A, oil extraction was set at 2,200 L/hr, and in the anecdote it was 2,000 L/hr, while the cut-off was set at 3,000 L/hr. Thus, the failure of the anecdotal project arising from insufficient oil extraction would appear more relevant because the oil extraction in both the anecdotal project and Project A was lower than the cut-off value. Note, however, that there was uncertainty about the generalisability of these cut-off values because the participants did not receive an explicit indication of whether these values were meant to generalise to other cases.

6.2.1.2.4 Follow-up Questions Participants who were shown the anecdote were subsequently presented with follow-up questions. They were asked about how similar they believed the anecdote was to the target project, how relevant it was

Case study

- Business details:
  - Business name: Refinera
  - Location: New Mexico, USA
  - Integration: vertical
- Investment: oil well
- Predicted project features:
  - Oil extracted: 2000L an hour
  - Time the machinery lasts before requiring maintenance: 7 years
  - Probability of finding oil: 80%
  - Type of well: onshore

Refinera struggled to establish itself in the regional market because of what scientists now know is a hydrocarbon shortage in the New Mexico area. A centralised organisational structure meant that key operational decisions were delayed with what needed to be a timely process. Being vertically integrated meant that these delays caused losses at the retail sites due to miscalculations of petrol supply. To make up for this, a post hoc analysis concluded that oil was needed to be extracted at a rate of 3000L an hour and sites have at least a 96% probability of finding oil before management approved the project. Further, machinery needed to have thought to last at least 10 years before requiring maintenance, because maintenance costs further offset the initial investment after the 7 years of development. Further, the well was quite susceptible to crude oil price changes due to it being an onshore well, and so added additional financial setbacks over the course of the project.

Figure 6.3: Anecdote for participants in the high similarity condition in Experiment 1.

to their allocations and how relevant it would be for their judgements about other projects of that type (see Appendix C.1.1.2.3).

6.2.1.3 Procedure

Participants were introduced to the study through the general instructions followed by the specific instructions for their condition. Participants were then presented with the allocation task and a description of the focal projects. All participants except those in the statistics only condition were also presented with the anecdote description and analysis, and the follow-up questions.

Case study

- Business details:
  - Business name: Refinera
  - Location: Zhuhai, China
  - Integration: horizontal
- Investment: oil well
- Predicted project features:
  - Oil extracted: 1400L an hour
  - Time the machinery lasts before requiring maintenance: 5 years
  - Probability of finding oil: 56%
  - Type of well: offshore

Refinera struggled to establish itself in the regional market because of what scientists now know is a hydrocarbon shortage in the Zhuhai area. A decentralised organisational structure meant that communication across relevant business units was delayed with what needed to be a timely process. Being horizontally integrated meant that these delays caused losses at the other well sites due to a drain on the collective resources. To make up for this, a post hoc analysis concluded that oil was needed to be extracted at a rate of 2100L an hour and sites have at least a 67% probability of finding oil before management approved the project. Further, machinery needed to have thought to last at least 8 years before requiring maintenance, because maintenance costs further offset the initial investment after the 5 years of development. Further, the well was quite difficult to construct due to it being an offshore well, and so added additional financial setbacks over the course of the project.

Figure 6.4: Anecdote for participants in the low similarity condition in Experiment 1.

6.2.2 Results

6.2.2.1 The Effect of Similarity on Anecdotal Bias

Anecdotal bias was tested by comparing the statistics only condition with both the high- and low-similarity anecdote and statistics conditions (see Figure 6.5). The omnibus one-way ANOVA of these three conditions was significant, $F(2, 118) = 4.19$, $p = .018$, $\hat{\eta}_p^2 = .066$. Planned comparisons show that participants in the statistics only condition allocated a higher percentage of their budget to the target project compared with participants in the high-similarity anecdote with statistics condition, $\Delta M = -12.31$, 95% CI $[-21.53, -3.09]$, $t(118) = -2.64$, $p = .009$; but not the low-similarity anecdote with statistics condition, $\Delta M = -1.48$, 95%

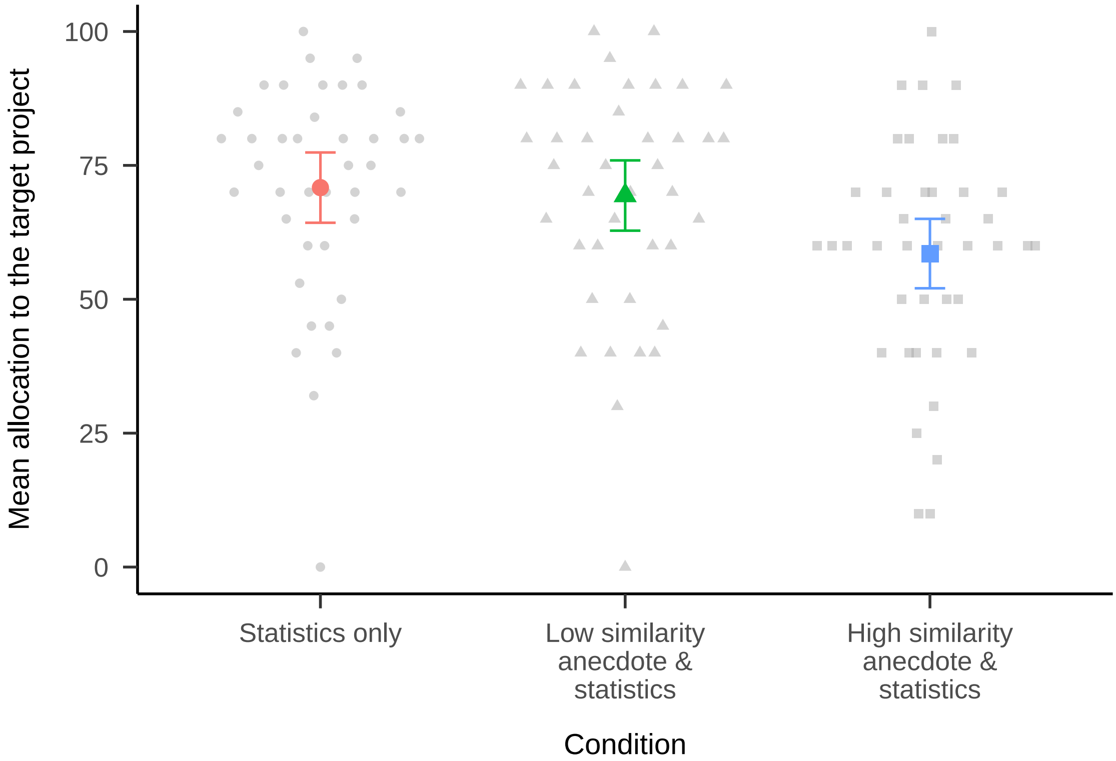


Figure 6.5: Mean allocation to the target project for the statistics only condition and the two anecdote & statistics conditions. Error bars represent 95% confidence intervals. Raw data are plotted in the background.

CI $[-10.75,\ 7.80]$, $t(118) = -0.31$, $p = .753$. These findings provide evidence of anecdotal bias in the high similarity condition only.

6.2.2.2 The Effect of Enhanced Statistics

The effect of enhanced statistics was investigated by testing the interaction of anecdote similarity and evidence type (anecdote & statistics and anecdote & enhanced statistics conditions, excluding the anecdote only and statistics only conditions). As shown in Figure 6.6, the two-way interaction was not significant, $M = 3.89$, 95% CI $[-8.86,\ 16.65]$, $t(238) = 0.60$, $p = .548$. Further, the difference between the anecdote & statistics condition and the anecdote & enhanced statistics condition (averaged over similarity conditions) was also not significant, $\Delta M = -0.12$, 95% CI $[-6.50,\ 6.26]$, $t(238) = -0.04$, $p = .971$. This suggests that providing participants with information about how to think statistically is not sufficient to facilitate a focus on statistics.

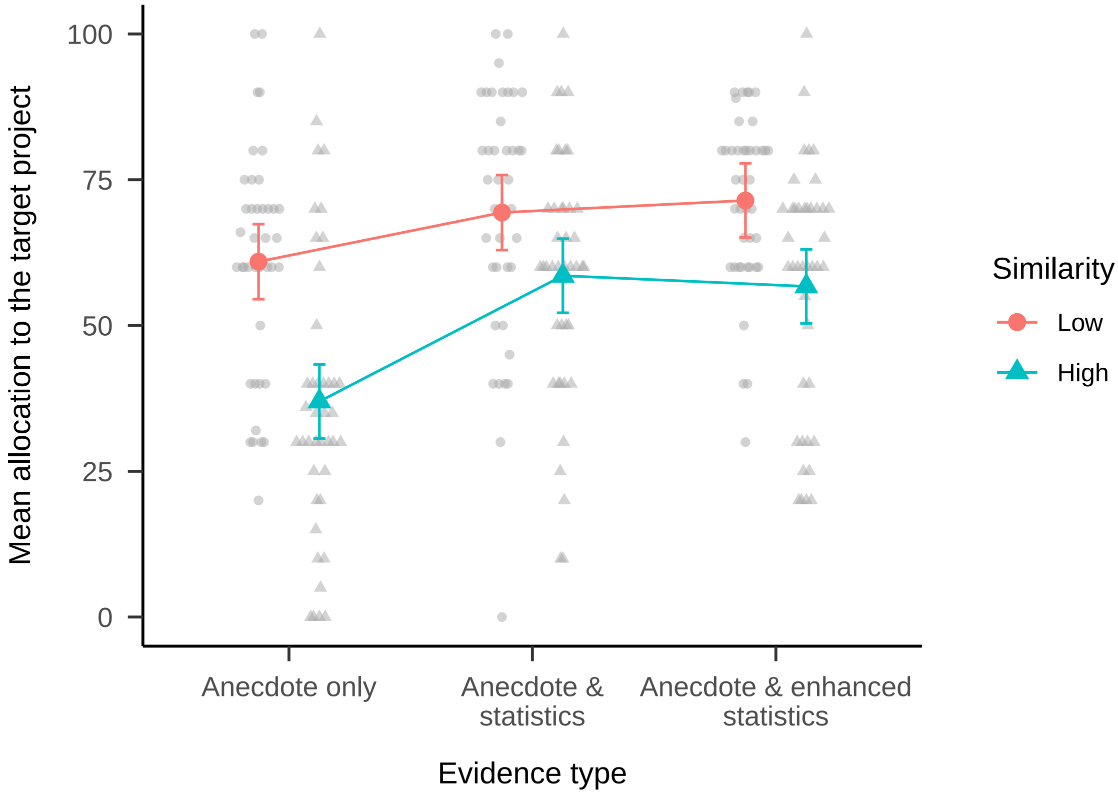


Figure 6.6: Mean allocation to the target project, by anecdote similarity and evidence type conditions (excluding the statistics only condition). Error bars represent 95% confidence intervals. Raw data are plotted in the background.

6.2.2.3 The Effect of Statistics

To identify the influence of statistics on participants' allocations, a two-way ANOVA of the interactions between anecdote similarity (low and high) and evidence type (anecdote only and anecdote & statistics conditions, excluding the anecdote & enhanced statistics and statistics only conditions) was conducted (see Figure 6.6). The interaction between anecdote condition and similarity (excluding the enhanced statistics condition) was significant, $M = -13.14$, 95% CI $[-25.93, -0.34]$, $t(238) = -2.02$, $p = .044$. Specifically, the difference in allocations between the anecdote only condition and the anecdote & statistics condition was greater when the anecdote was similar, $\Delta M = -21.56$, 95% CI $[-32.33, -10.80]$, $t(238) = -4.72$, $p < .001$; compared with when it was dissimilar, $\Delta M = -8.43$, 95% CI $[-19.32, 2.47]$, $t(238) = -1.82$, $p = .164$. These findings provide evidence of partial anecdotal bias in the high similarity condition because the anecdote &

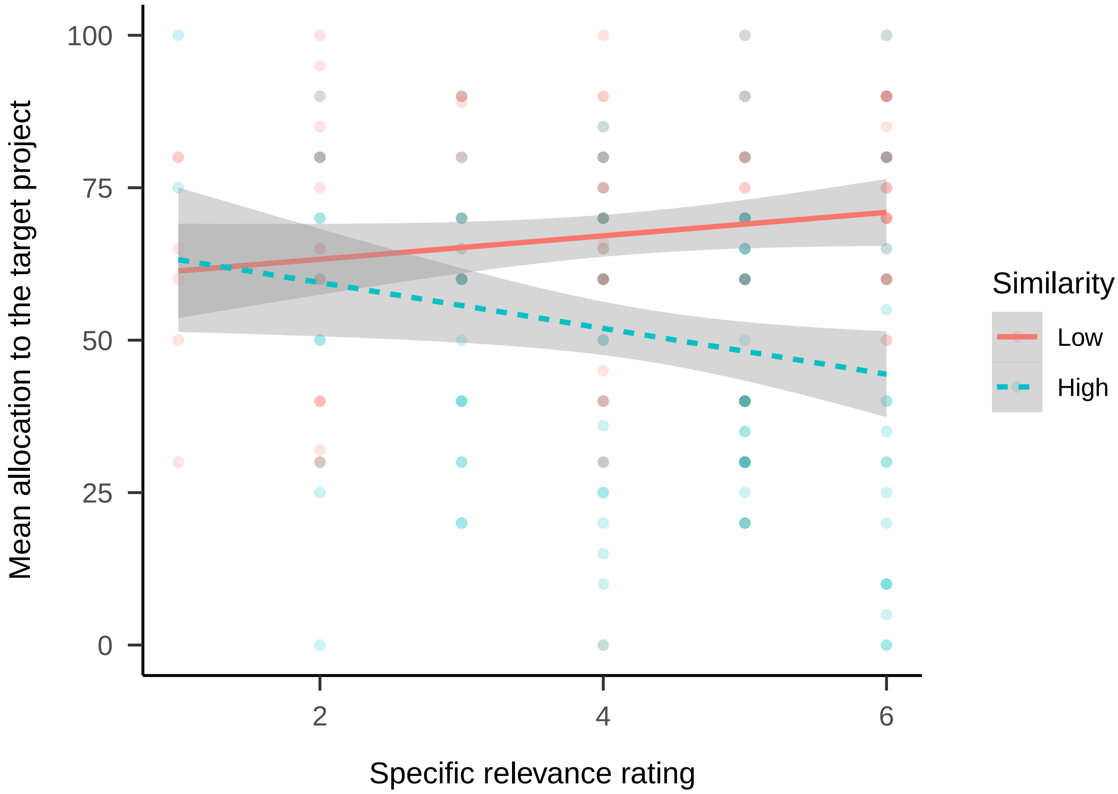


Figure 6.7: Mean allocation to the target project, by specific relevance rating and similarity condition. LOESS method was used for smoothing over trials and the shading represents 95% confidence intervals. Raw data are plotted in the background.

statistics condition was lower than the statistics only condition (shown above) but higher than the anecdote only condition.

6.2.2.4 Relevance Ratings

Regression analyses were conducted to determine the relationship between allocations and the follow-up relevance ratings. As shown in Figure 6.7, the specific relevance ratings interacted with similarity condition, $b = -2.84$, 95% CI $[-4.80, -0.87]$, $t(240) = -2.85$, $p = .005$. It appears that specific relevance ratings were related to allocations, but only in the high similarity condition. Further, there were no significant associations with the general relevance ratings. This suggests that participants applied reasoning to the connection between the anecdote and the target project as opposed to simply reacting to the failed project and associating that with that project's industry.

6.2.3 Discussion

Hypothesis 6.1 was supported. Participants in the anecdote & statistics condition allocated less capital to the target project compared with those in the statistics only condition. However, this effect depended on anecdote similarity because this only occurred in the high similarity condition, not in the low similarity condition. Thus, while anecdotal bias was evident when the anecdote was similar to the target project, participants were not influenced when the causal mechanisms did not align. Contrary to Hypothesis 6.2, despite being influenced by the anecdote, participants still made some use of the statistics. This is different from the findings of Wainberg et al. (2013), who found no difference between the anecdote only and anecdote and statistics conditions, indicating a complete anecdotal bias effect. Hypothesis 6.3 was also not supported because the added enhanced statistical instructions used to encourage participants to use the statistical information did not reduce participants' reliance on anecdotes.

Experiment 1 was limited because it only considered a *negative* anecdote; that is, a failed project. In real life, however, case studies often have a *positive* valence; that is, the story of a successful company. In fact, in business, it is possible that the anecdotes used are more likely to be positive because of survivorship bias. Jaramillo et al. (2019) found an anecdotal bias effect for negative but not positive anecdotes. This may be because the stimuli consisted of medical decisions and, in this domain, the loss of health may be more strongly noted than an equivalent gain in health. In Experiment 2 (discussed in the subsequent section) a positive anecdote was added to investigate whether anecdote valence would affect anecdotal bias.

It is unclear whether the effects found in Experiment 1 were related to participants' perceptions of the type of sampling used to select the anecdotes. The instructions in Experiment 1 did not explain how the anecdote displayed to participants was chosen. Whether sampling is believed to be intentional or random has been shown to affect people's decision-making (e.g., Hayes et al., 2019). In the present experiments, participants' sampling assumptions may have changed the extent to which they used the anecdote in their decisions. For example, it may be rational to choose the anecdote over the aggregated data if (a) the anecdote was not sampled randomly from a pool of anecdotes, and (b) the anecdote had a greater similarity to the target project compared with other anecdotes in the pool in relevant ways. That is, if the anecdote were chosen because of its high relevance to the target project, it would be irrational to ignore it. In Experiment 1, it was unclear whether participants may have held these beliefs. To control for these assumptions, in Experiment 2, the instructions further clarified that the anecdote (a) was sampled randomly from a pool of anecdotes, and (b) was not significantly more similar to the target project than any of the other anecdotes in the pool.

6.3 Experiment 2

Experiment 1 replicated the anecdotal bias effect found in the literature. That is, participants allocated less capital to a project when presented with an anecdote and conflicting statistics compared with when they were presented with the statistics only. However, this effect depended on anecdote similarity, such that anecdotal bias was stronger when the anecdote was similar to the current task compared with when it was dissimilar. A negative anecdote only was used Experiment 1 because previous research has found anecdotal bias for negative but not for positive anecdotes (Jaramillo et al., 2019). However, Jaramillo et al. (2019) investigated medical decision-making, and the effect of anecdote valence may be different in a business context. In the study by Jaramillo et al. (2019), the positive anecdote involved a treatment that led to a reduction in symptoms, while the negative anecdote involved symptoms persisting. This framing may have led participants to perceive the positive anecdote as a return to a reference point and the negative anecdote as a continuation of a reduction in wellbeing relative to the reference point. In business, however, both successful and failed business projects represent a deviation from a reference point. To test this difference further, manipulation of anecdote valence was added to Experiment 2.

To increase the experiment's power, anecdote valence and anecdote similarity were manipulated within subjects. Further, Experiment 2 did not include the anecdote & enhanced statistics condition because Experiment 1 found no evidence for its effect. All participants saw the statistics only condition, which did not contain an anecdote; therefore, this did not need to be manipulated between subjects. Therefore, each participant was shown five displays: one for the statistics only condition, and four for either the anecdote only condition or the anecdote & statistics condition. These four anecdote displays consisted of the similarity (low and high) × valence (negative and positive) conditions.

In Experiment 1, assumptions about the pool from which the anecdote was sampled were not clarified. In Experiment 2, participants were told that the anecdote was sampled randomly and that it was not uniquely similar to the target project. This was expected to lead to a reliance on statistical evidence, regardless of the anecdote's similarity. However, people often struggle to use statistical concepts presented descriptively, as seen in the enhanced statistics condition in Experiment 1, the neglect of variance shown in Chapter 4, and the lack of risk aggregation in descriptive risky decisions shown in Chapter 2. Therefore, it was expected that the results of Experiment 1 would be replicated for the negative valence condition. Further, it was expected that there would be a reverse effect in the positive valence condition. Appendix C.2 shows a simulation of the hypothesised effects. Therefore, Experiment 2 tested the following hypothesis:

Hypothesis 6.4—overall effect. The three-way similarity × valence × anecdote (excluding statistics only) interaction is significant

The main effect of interest was the effect of anecdote similarity on anecdotal bias. However, because in Experiment 2 all participants were presented with the statistics only condition, a difference score was calculated to simplify the analyses. Specifically, this was the difference between the allocation in the anecdote & statistics conditions and the relevant allocation in the statistics only condition. A score that is different from zero indicates deviation from the allocation when only statistics were shown. For positive valence, a stronger influence of anecdote is indicated by a lower difference score; whereas for negative valence, a stronger influence of anecdote is indicated by a higher difference score. Therefore, Experiment 2 tested the following hypotheses:

Hypothesis 6.5—anecdotal bias difference score for negative anecdotes. For negative anecdotes, the difference between budget allocations to the target project in the statistics only condition and the anecdote & statistics condition will be higher when the anecdote is similar to the target project compared with when it is dissimilar.

Hypothesis 6.6—anecdotal bias difference score for positive anecdotes. For positive anecdotes, the difference between budget allocations to the target project in the statistics only condition and the anecdote & statistics condition will be lower when the anecdote is similar to the target project compared with when it is dissimilar.

Contrary to both Wainberg et al. (2013) and Hypothesis 6.2, Experiment 1 found that participants do integrate statistics in their decisions to some extent. This effect was expected to be replicated in Experiment 2. Therefore, Experiment 2 tested the following hypotheses:

Hypothesis 6.7—effect of statistics for negative anecdotes. For negative anecdotes, budget allocations to the target project will be higher for the high-similarity anecdote & statistics condition than for the high-similarity anecdote only condition.

Hypothesis 6.8—effect of statistics for positive anecdotes. For positive anecdotes, budget allocations to the target project will be higher for the high-similarity anecdote only condition than for the high-similarity statistics & anecdote condition.

Table 6.2: Experiment 2 group allocation.

Evidence type	N
Anecdote & statistics	48
Anecdote only	48
Total	96

6.3.1 Method

6.3.1.1 Participants

Ninety-six participants (50 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 41.69 years (*SD* = 11.29, *min.* = 27, *max.* = 74). Participants reported an average of 7.19 years (*SD* = 8.34, *min.* = 0, *max.* = 43) working in a business setting, and an average of 3.91 years (*SD* = 7.67, *min.* = 0, *max.* = 50) of business education. The mean completion time of the task was 14.98 min (*SD* = 8.84, *min.* = 2.57, *max.* = 58.71). Table 6.2 shows the allocation of participants to the different conditions. Anecdote similarity and valence were manipulated within subjects. Therefore, each participant was assigned to one of two between-subjects evidence type conditions (anecdote only and anecdote & statistics) and saw five displays (statistics only, and one of each of the four similarity and valence conditions). Appendix C.2.1.1.1 describes the power analysis conducted to arrive at this sample size.

6.3.1.2 Materials

6.3.1.2.1 Instructions Participants were shown similar instructions to those in Experiment 1 (see Section 6.2.1.2.1). The general instructions page included a test of the basic information expressed in the instructions. This test also functioned as an attention check. As in Experiment 1, participants were also shown instructions that were specific to their condition. These were shown on the same page as the rest of the project display, above the case study and focal projects. The instructions clarified that the anecdote had been randomly sampled and that all anecdotes in the pool were equally similar to the target project. Appendix C.2.1.2.1 shows the instructions used in Experiment 2.

6.3.1.2.2 Allocation Task As in Experiment 1, the allocation task included a table describing the two focal projects and (apart from the statistics only condition) a description and analysis of an anecdote. Figures 6.8 and 6.9 show the anecdote and focal projects, respectively, for the negative valence, low similarity condition. Figures 6.10 and 6.11 show the anecdote and focal projects, respectively, for the positive valence, high similarity conditions. In the statistics only condition, participants were only shown the focal projects display. Appendix C.2.1.2.2 details the counterbalancing and randomisation used in the experiment.

6.3.1.2.3 Interstitial Page Prior to the display, participants were shown an interstitial page, which was used to (a) introduce the display and (b) check the participant's attention (given that no input was required, participants could easily skip the page without reading the text). See Appendix C.2.1.2.4.

6.3.1.2.4 Follow-up Questions Participants were shown similar follow-up questions as in Experiment 1, except that in Experiment 2, rating scales were 1–7 instead of 1–6. See Appendix C.2.1.2.3 for a sample display of the follow-up questions.

6.3.1.3 Procedure

Participants were introduced to the study via the general instructions page. They were then shown five sets (presented in a random order) containing two pages each: a page showing the allocation task and a page with follow-up questions (except for the anecdotes only condition, in which participants were not shown the follow-up questions page). Each allocation task page contained specific instructions relevant to the condition followed by the anecdote analysis and description, and the description of the two focal projects. The only exception was the statistics only display, for which there was no anecdote description or analysis.

Case study

Cweb struggled to establish itself in the regional market because of changes in privacy laws (that reduced consumer confidence in the business' apps) in the Mumbai area. A centralised organisational structure meant that poor performers took longer to be replaced, so some tasks needed considerable revision. Being vertically integrated meant that the project was reliant on in-house manufacturing and so was slow to adopt the newest technologies used by competitors. A post hoc analysis concluded that, to make up for these issues, the developers needed to write at least 800 lines a day and the the application needed to be certified with a security rating of at least 68%. Further, the number of potential first-year customers needed to be at least 2 million. Further, the problems in the application were slow to solve because of the lack of large-scale quantitative data due to it being for enterprise, and so added additional financial setbacks over the course of the project.

- Business details:
  - Business name: Cweb
  - Location: Mumbai, India
  - Integration: vertical
  - Structure: centralised
- Investment: software
- Predicted project features:
  - Code written: 600 lines a day
  - Security rating: 51%
  - Number of potential customers in first year: 2 million
  - Target users: enterprise

Figure 6.8: An example of the anecdote display in the negative valence, low similarity condition of Experiment 2.

6.3.2 Results

This section reports only the data relevant to the Experiment 2 hypotheses. See Appendix C.2.2 for manipulation check analyses and analyses of the follow-up rating data.

6.3.2.1 Overall Effect of Manipulations

As shown in Figure 6.12, the similarity × valence × evidence type interaction (excluding the statistics only condition) was not significant, $F(1, 94) = 3.42$, $p = .067$, $\hat{\eta}_p^2 = .035$. However, the similarity × valence interaction was significant,

Target projects

Allocate your budget between the following two projects using percentage values (the two values should sum to 100):

Relevant information	Project 1	Project 2
Business name	Codeck	Enfuel
Project type	software	oil well
Location	Austin, USA	Houston, USA
Integration	horizontal	vertical
Structure	decentralised	centralised
Predicted project features	• Code written: 1000 lines a day • Security rating: 85% • Number of potential customers in first year: 3 million • Target users: ordinary consumers	• Oil extracted: 2000L an hour • Time the machinery lasts before requiring maintenance: 7 years • Probability of finding oil: 80% • Type of well: onshore
Project allocation (%)	Allocation:	Allocation:
Overall reliability rating (%)	91	90
NPV ($)	901	100

Figure 6.9: An example of the focal projects in the negative valence, low similarity condition of Experiment 2. Here, Project 1 was the target project and Project 2 was the comparison project.

Case study

Microxy performed really well in the regional market because of decreased silicon taxes in the Montreal area. A decentralised organisational structure meant that the individual teams had greater autonomy to complete their tasks, increasing the efficiency of important project stages. Being horizontally integrated meant that the project can be easily marketed to the customer base of the other business units in the company. A post hoc analysis concluded that, to take advantage of these benefits, the microchips needed to be produced at a rate of at least 3200 an hour and the semiconductor yield needed to be at least 57%. Further, the percent of compatible devices needed to be at least 71%. Further, the chip has a relatively low power consumption due to it operating Reduced Instruction Set Computing, and so added additional financial resilience over the course of the project.

- Business details:
  - Business name: Microxy
  - Location: Montreal, Canada
  - Integration: horizontal
  - Structure: decentralised
- Investment: microchip
- Predicted project features:
  - Microchips produced: 4800 an hour
  - Usable semiconductor yield after testing: 63%
  - Compatible devices in the market: 79%
  - Type of chip architecture: Reduced Instruction Set Computing

Figure 6.10: An example of an anecdote display in the positive valence, high similarity condition of Experiment 2.

$F(1, 94) = 76.41$, $p < .001$, $\hat{\eta}_p^2 = .448$, as was the evidence type × valence interaction, $F(1, 94) = 10.11$, $p = .002$, $\hat{\eta}_p^2 = .097$. The analyses below elaborate on the specific hypothesised effects.

6.3.2.2 Anecdotal Bias Depends on Anecdote Similarity

To investigate whether anecdotal bias depended on anecdote similarity, the differences in budget allocations between the statistics only condition and the two anecdote & statistics conditions (high and low similarity) were calculated. The values of the statistics only condition were different for each valence condition to create equivalent comparisons. For the negative valence condition, participants

Target projects

Allocate your budget between the following two projects using percentage values (the two values should sum to 100):

Relevant information	Project 1	Project 2
Business name	Solgistics	Altchip
Project type	shipping logistics	microchip
Location	Kuala Lumpur, Malaysia	Toronto, Canada
Integration	vertical	horizontal
Structure	centralised	decentralised
Predicted project features	• Packages shipped: 800 a week • Number of orders that do not spend time in a bottleneck: 400 a day • Average accuracy of shipments: 90% • Shipping type: parcel	• Microchips produced: 4000 an hour • Usable semiconductor yield after testing: 60% • Compatible devices in the market: 75% • Type of chip architecture: Reduced Instruction Set Computing
Project allocation (%)	Allocation:	Allocation:
Overall reliability rating (%)	93	90
NPV ($)	905	105

Figure 6.11: An example of the focal projects in the positive valence, high similarity condition of Experiment 2. Here, Project 2 was the target project and Project 1 was the comparison project.

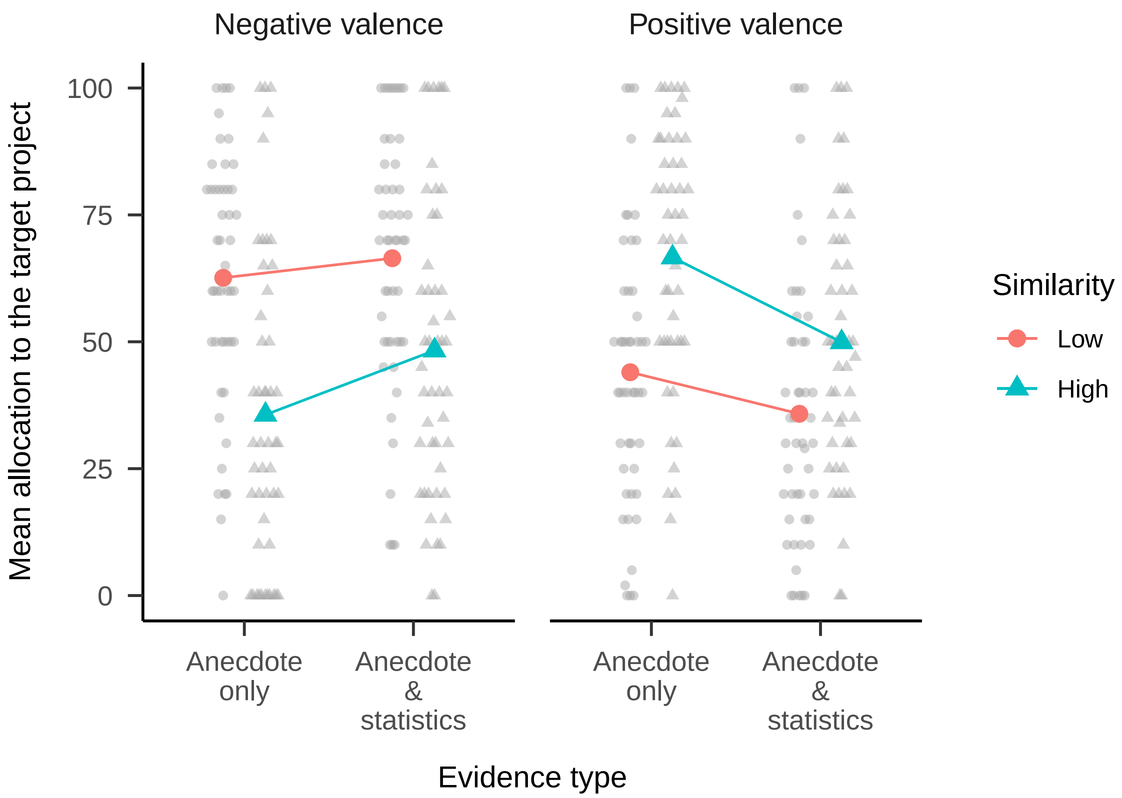


Figure 6.12: Mean allocation to the target project, by evidence type, similarity, and valence conditions. In mixed factorial designs, error bars cannot be used to make inferences by "eye" across all conditions. Therefore, error bars are not included. Raw data are plotted in the background.

allocated more money the high-NPV project; while for the positive valence condition, participants allocated more money to the low-NPV project. As shown in Figure 6.13, the similarity × valence interaction was significant, $F(1, 47) = 30.66$, $p < .001$, $\hat{\eta}_p^2 = .395$, as was the main effect of valence, $F(1, 47) = 9.85$, $p = .003$, $\hat{\eta}_p^2 = .173$. The main effect of similarity was not significant, $F(1, 47) = 0.53$, $p = .469$, $\hat{\eta}_p^2 = .011$.

The effect of the anecdote is represented differently for each valence condition. As such, the interaction was further analysed by comparing the two similarity conditions for each valence condition. For negative anecdotes, the statistical values (e.g., NPV) associated with the target project were higher than those for the comparison project. If participants were influenced by the negative anecdote they would therefore allocate less to the target. For negative anecdotes, a lower allocation to the target project is represented in Figure 6.13 as a positive value—the difference in allocation from when the participant did not see an anecdote. For positive anecdotes, the statistics were lower for the target project, so an influence of the anecdote is seen as a negative value in Figure 6.13. The hypothesised effect of negative anecdote similarity on anecdotal bias would suggest a higher difference score in high similarity than in low similarity. That is, more influence of the anecdote when it is similar than when it is dissimilar. For positive anecdotes a the hypothesised effect would suggest the reverse: a higher difference score in low similarity than in high similarity.

For negative anecdotes, the allocation difference was greater when the anecdote was similar to the target project than when it was dissimilar, $\Delta M = -18.17$, 95% CI $[-26.17, -10.17]$, $t(93.80) = -4.51$, $p < .001$. For positive anecdotes, the allocation difference was greater when the anecdote was dissimilar to the target project than when it was similar, $\Delta M = 14.10$, 95% CI $[6.10, 22.11]$, $t(93.80) = 3.50$, $p = .001$. This provides evidence that anecdotal bias depends on anecdote similarity for both negative and positive anecdotes. Participants appeared to be sensitive to the relevance of the anecdote to the target problem.

6.3.2.3 Effect of Statistics

As in Experiment 1, Experiment 2 investigated the extent to which statistical information influenced participants' allocations. As shown in Figure 6.12, for negative anecdotes, participants in the high-similarity anecdote & statistics condition allocated more to the target project than those in the high-similarity anecdote only condition, $\Delta M = -12.67$, 95% CI $[-23.53, -1.81]$, $t(336.36) = -2.29$, $p = .022$. For positive anecdotes, participants in the high-similarity anecdote only condition allocated more to the target project than those in the high-similarity anecdote & statistics condition, $\Delta M = 16.71$, 95% CI $[5.85, 27.57]$, $t(336.36) = 3.03$, $p = .003$. This provides evidence for the influence of statistics on participants' allocations for both negative and positive anecdotes.

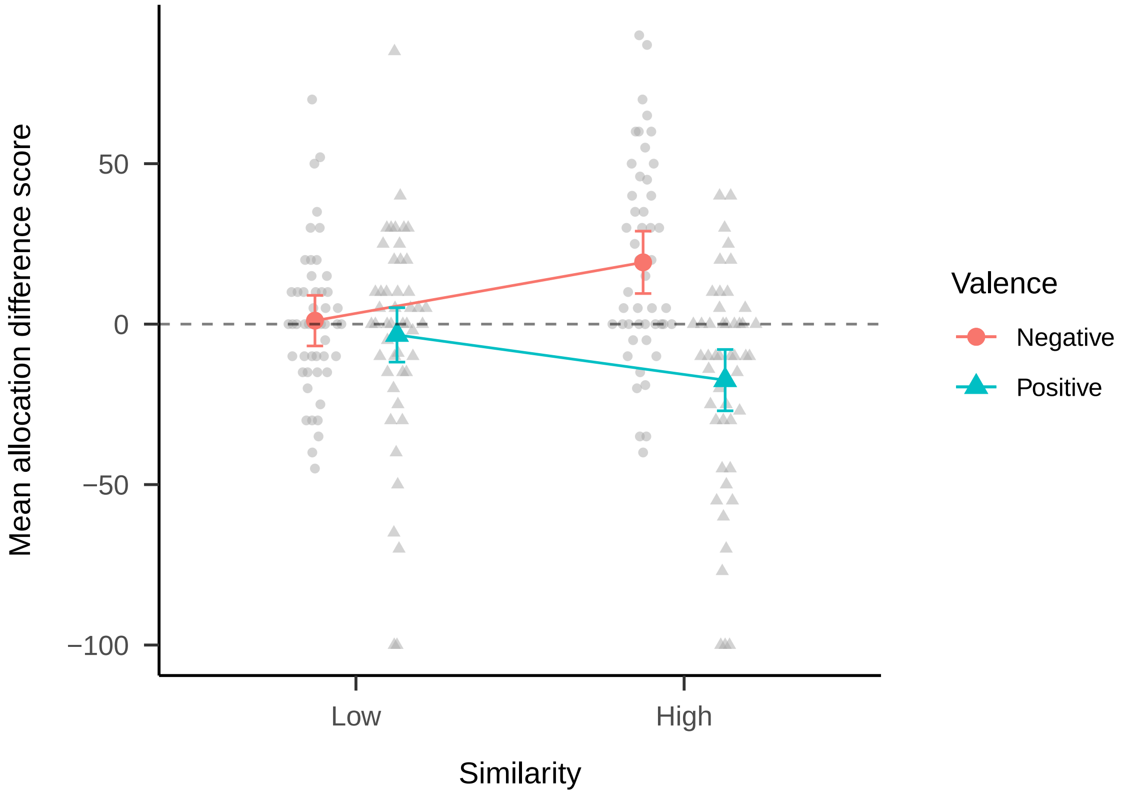


Figure 6.13: Mean allocation difference between the statistics only condition and the anecdote & statistics condition, by similarity and valence conditions. The horizontal dashed line shows the point in which the two allocations were equivalent. Values above this line show the higher allocation to the target project when participants were shown statistics only compared with when they were shown statistics with an anecdote. Error bars represent 95% confidence intervals, calculated from the within-subjects standard errors using the method from Cousineau and O'Brien (2014). Raw data are plotted in the background.

6.3.2.4 Relevance Ratings

Regression analyses were conducted to determine the relationship between allocations and the follow-up relevance ratings. Figure 6.14 shows these data. While the specific relevance ratings for negative anecdotes showed the same trends as in Experiment 1, the interaction was not significant. Similarly, the ratings trends for positive anecdotes were as hypothesised, but their interaction not significant. It appears that specific relevance ratings were related to allocations, but only in the high similarity condition. Further, there were no significant associations with the general relevance ratings. This provides limited evidence that people were explicitly reasoning about the connection between the anecdote and target.

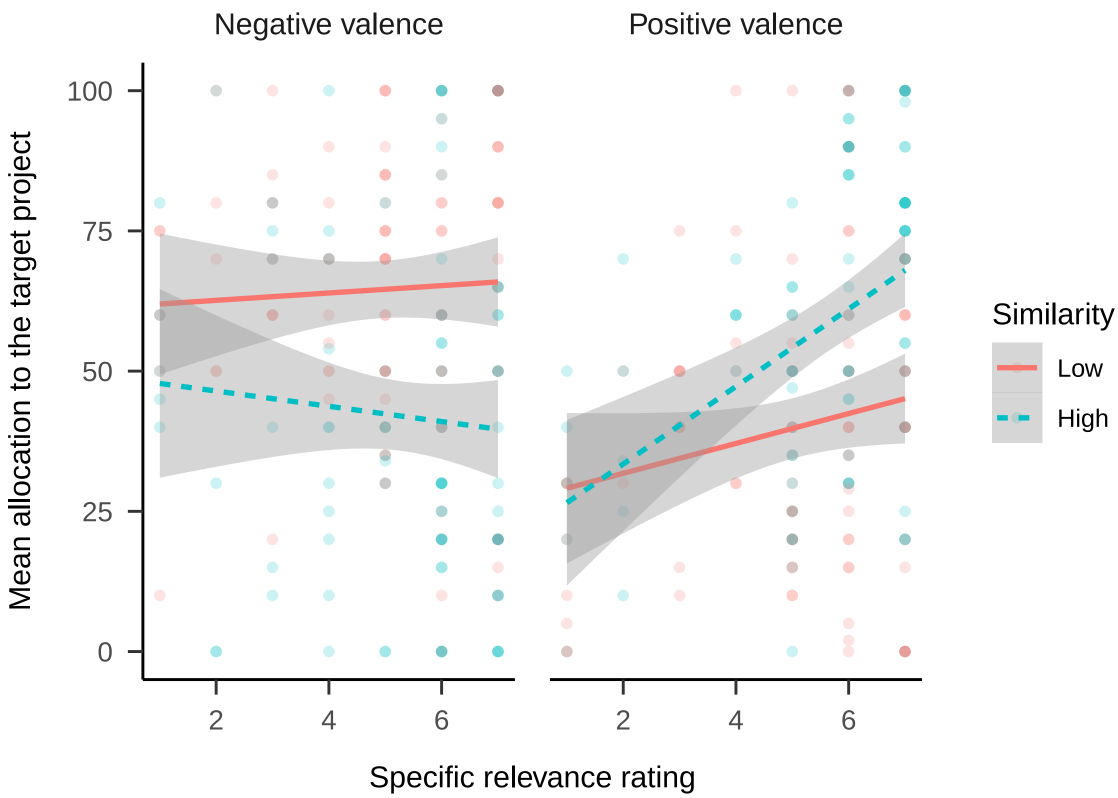


Figure 6.14: Mean allocation to the target project, by specific relevance rating, similarity condition, and valence condition. LOESS method was used for smoothing over trials and the shading represents 95% confidence intervals. Raw data are plotted in the background.

6.3.3 Discussion

Hypotheses 6.5 and 6.6 were supported because participants showed a stronger anecdotal bias effect when both positive and negative anecdotes had greater similarity to the target project. Further, as per Hypotheses 6.7 and 6.8, participants incorporated statistical information in their judgements, for both negative and positive anecdotes. Unlike in Experiment 1, the relevance rating data did not provide as clear indication that participants were using only the specific project information rather than merely its industry.

Therefore, Experiment 2 found that, unlike in the medical domain, the effect of anecdotes in financial decision-making does not depend on anecdote valence. Further, similar to the findings of Experiment 1, and unlike those of Wainberg et al. (2013), the anecdotal bias effect does not appear to be complete, with statistics still playing some role in participants' decisions, despite the effect of the anecdote.

6.4 General Discussion

Most of the hypotheses were supported. This chapter found that, in the capital allocation context, people's decisions are influenced by anecdotes, even when aggregated data are available. There were three novel findings: (a) the anecdotal bias effect was only seen when participants considered the anecdote sufficiently relevant to the target project, (b) participants integrated statistics into their decisions, and (c) these effects were found in both negative and positive anecdotes. Further, people did not consider verbal sample distribution information, which could have helped to inform their decisions. This is surprising since other work showed that generalisations are sensitive to sampling (Carvalho et al., 2021).

The first novel finding from these experiments is that participants' use of anecdotal evidence depended on the anecdote's similarity. Specifically, if the anecdote appeared relevant, participants used it in their decisions. However, when it appeared irrelevant, participants almost entirely relied on statistics. The findings for high anecdote similarity are largely congruent with findings from other work investigating anecdotal bias in business decision-making. As in Wainberg et al. (2013) and Wainberg (2018), this chapter found that people allocated less capital to a project when presented with statistical evidence and a similar but contradictory anecdote than when they were presented with statistics alone.

It appears that participants distinguished between the low- and high-similarity anecdotes based on the structure of the anecdote. The low similarity condition always included the same project type as the high similarity condition for all domains. For instance, in one variation, both the high- and low-similarity anecdotes involved oil well projects. However, the high-similarity anecdotes also matched the target project in a number of specific features. This means that participants were sensitive to the specific information in the anecdote description and analysis and did not simply use the project type for their inferences. Further, participants' answers to the follow-up questions indicated that they did not consider that the anecdote was necessarily relevant to other projects from the same industry. In other words, participants did not appear to carelessly use anecdotal evidence in their decisions; rather, they carefully considered the anecdote according to its particular causal structure.

The second novel finding from these experiments is that participants who were shown the anecdote with statistics did not completely disregard the statistical measures. Wainberg et al. (2013) found a complete anecdotal bias effect because results for the anecdote only and anecdote & statistics conditions were equivalent, meaning that the presented statistics had a negligible effect on participants' decisions. In contrast, the experiments discussed in this chapter showed a partial anecdotal bias effect, seen as a difference in allocations between the anecdote only and anecdote & statistics conditions. It appears that participants integrated both anecdotal and statistical information. This suggests that people's evaluation of evidence may be more sensitive than previously thought.

The discrepancy between these results and those in Wainberg et al. (2013) could be a result of the sampled population. Since Freling et al. (2020) found that anecdotes had a stronger effect when decisions were more personally relevant; thus, the managers recruited for the Wainberg et al. (2013) study may have simply been more personally invested in the task compared with the laypeople recruited for the experiments presented in this chapter. Similarly, Yang et al. (2015) found that anxiety increases anecdotal bias when making risky decisions. However, the discrepancy may also be attributable to the difference in the anecdote & statistics condition between the Wainberg et al. (2013) study and the present work. Specifically, the statistics presented in the anecdote & statistics condition in Wainberg et al. (2013) were not the same as those shown in the same study's statistics only condition, unlike in both the present experiments and Wainberg (2018). Instead, it was the anecdote & enhanced statistics condition that contained the same statistics as in the statistics only condition. This suggests that people only integrate statistics when they are sufficiently clear and no further interpretation is required.

The third novel finding from these experiments is that anecdotal bias was found for both negative and positive anecdotes. Most previous studies have included negative anecdotes (i.e. those with negative consequences) such as a medication that fails to reduce symptoms. However, there is little work in the literature involving positive anecdotes (those with positive consequences). Jaramillo et al. (2019) found an asymmetry in the anecdote effect—the effect of the anecdote was stronger when the medication failed to improve symptoms (negative anecdote) compared with when it did improve symptoms (positive anecdote). The present experiments found a more symmetrical effect—the effects of both anecdotal bias and statistics were found for both negative and positive anecdotes.

The difference between the findings of this chapter and those of Jaramillo et al. (2019) may be attributable to the latter's negative anecdote representing a persistence in a negative shift from the status quo (i.e. good health). In the business domain, both positive and negative anecdotes represent shifts from the status quo (a company's financial position). Nevertheless, it was surprising to find no asymmetry given the predictions of prospect theory. Loss aversion suggests that participants will avoid projects that are similar to negative anecdotes more than they will choose those similar to positive anecdotes. However, each choice was associated with conflicting statistical information, so this may have cancelled out the change from the reference point. Future research should use more realistic incentives to investigate this effect further. Doing so will also increase the ecological validity of the findings.

6.4.1 Theoretical Implications

The findings presented in this chapter add to the current understanding of the way in which people use different types of evidence in their decision-making. Previous research mostly investigated the relative influence of statistics and anecdotes by comparing anecdotal with statistical conditions. The current work shows that comparing a joint anecdote & statistics condition with both an anecdote only and statistics only condition enables a more specific investigation of participants' anecdotal bias. The influence of anecdotes can be seen in the comparison of the statistics only and the anecdote & statistics conditions, while the effect of statistics can be seen in the comparison of the anecdote & statistics condition and the anecdote only condition. These two effects enable the determination of the independent influences of anecdote & statistics. Use of such a design in future research may help to further the understanding of conditions under which these types of evidence are used.

Some of the anecdotal bias literature is based on the assumption that using anecdotal evidence over statistical evidence is necessarily irrational. This is likely to have arisen from examples in the medical domain in which such decisions are indeed irrational (e.g., believing that vaccines cause certain disorders, despite the available evidence). In such cases, people over-rely on anecdotes and should be relying more on aggregated data. However, a case could be made for the rational use of an anecdote based on its similarity to the target problem. For instance, there are times when an anecdote is so similar to the target situation (e.g., the identical twin example discussed in Section 6.1.3) that it would be unwise not to consider it. That is, the use of anecdote should depend on both (a) the extent of underlying structural similarity to the target problem and (b) the distribution of this similarity across the pool from which the anecdote was sampled. People should use anecdotes if their casual structures are significantly more relevant compared with other cases in the available data.

However, similarity can also be misleading. For instance, if a case appears highly similar but differs in terms of a key hidden dimension that is the real causal mechanism, then using the anecdote may be the wrong thing to do. What appears to be important is being sensitive to relational rather than surface similarity. Future research should investigate how varying participants' assumptions about sampling from a data set of anecdotes influences their anecdotal bias. Such assumptions can include the size of the sample, the shape of the distribution, and where in the distribution the anecdote came from. Prior work found that people are sensitive to distributional properties when generalizing (Carvalho et al., 2021), but it is not clear if this will replicate with descriptive cues such as in the experiments in this chapter.

6.4.2 Practical Implications

The current work contributes to managerial decision-making by providing insights into how managers make better decisions when using case studies and statistical information. Managers of large companies are often in a difficult position; they have incomplete information and are in an uncertain environment. Despite this, different biases and responses to those biases may be anticipated for different levels of uncertainty. For instance, a manager may be presented with both a convincing case study that suggests a certain course of action as well as aggregated data. The manager needs to be able to weigh the evidence accordingly.

The work in this chapter suggests that there are three elements to consider: (a) the quality of aggregated data (determined by factors such as sample size), (b) the relative similarity of the cases in the data pool to the target situation, and (c) the similarity of the anecdote to the target problem. For instance, an anecdote that is similar to the target situation in terms of relevance and is significantly more similar than other cases in the data set should carry more weight than an anecdote that comes from a pool of cases that are all equally similar to the target problem. Lovallo et al. (2012) found that similarity judgements increase prediction accuracy beyond a simple regression model. Taking into account a project's relative similarity to other cases is likely to further increase predictive validity.

When aggregated data are not available, however, managers should rely more on anecdotes that have greater similarities in terms of causal structure. That is, they should be wary of merely using surface similarities to make inferences and instead consider the underlying relational structures. The present data suggest that laypeople can do this to some extent, with participants not being completely swayed by the mere similarity of type of business project. However, future research should investigate this further to better understand the boundaries of people's analogical reasoning in capital allocation decisions.

References

Al Khatib, K., Wachsmuth, H., Hagen, M., & Stein, B. (2017). Patterns of Argumentation Strategies across Topics. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 1351–1357. https://doi.org/10/gjscsq

Allen, M., & Preiss, R. W. (1997). Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports*, *14*(2), 125–131. https://doi.org/10/djqrp7

Carvalho, P. F., Chen, C.-h., & Yu, C. (2021). The distributional properties of exemplars affect category learning and generalization. *Scientific Reports*, *11*(1), 11263. https://doi.org/10.1038/s41598-021-90743-0

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, *46*(4), 1149–1151. https://doi.org/10/f6vdsw

Freling, T. H., Yang, Z., Saini, R., Itani, O. S., & Rashad Abualsamh, R. (2020). When poignant stories outweigh cold hard facts: A meta-analysis of the anecdotal bias. *Organizational Behavior and Human Decision Processes*, *160*, 51–67. https://doi.org/10/gg4t2f

Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*(8), 691–712. https://doi.org/10/b64gsr

Gavetti, G., & Rivkin, J. W. (2005). How Strategists Really Think. *Harvard Business Review*, *83*(4), 54–63.

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10/dw52z8

Gilboa, I., & Schmeidler, D. (1995). Case-Based Decision Theory. *The Quarterly Journal of Economics*, *110*(3), 605–639. https://doi.org/10/c7tz7x

Hayes, B. K., Navarro, D. J., Stephens, R. G., Ransom, K., & Dilevski, N. (2019). The diversity effect in inductive reasoning depends on sampling assumptions. *Psychonomic Bulletin & Review*, *26*(3), 1043–1050. https://doi.org/10/gjscss

Hoeken, H. (2001). Convincing citizens: The role of argument quality. In D. Janssen & R. Neutelings (Eds.), *Reading and writing public documents: Problems, solutions, and characteristics* (pp. 147–169). John Benjamins Publishing Company. https://doi.org/10.1075/ddcs.1.08hoe

Hoeken, H., & Hustinx, L. (2009). When is Statistical Evidence Superior to Anecdotal Evidence in Supporting Probability Claims? The Role of Argument Type. *Human Communication Research*, *35*(4), 491–510. https://doi.org/10/fgtwjd

Hornikx, J. (2005). A review of experimental research on the relative persuasiveness of anecdotal, statistical, causal, and expert evidence. *Studies in Communication Sciences*, *5*(1), 205–216. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.725.6516&rep=rep1&type=pdf

Hornikx, J. (2018). Combining Anecdotal and Statistical Evidence in Real-Life Discourse: Comprehension and Persuasiveness. *Discourse Processes*, *55*(3), 324–336. https://doi.org/10/gjscrf

Jaramillo, S., Horne, Z., & Goldwater, M. (2019, May 15). *The impact of anecdotal information on medical decision-making* (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/r5pmj

Lassaline, M. E. (1996). Structural alignment in induction and similarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 754–770. https://doi.org/10/fq9fww

Lovallo, D., Clarke, C., & Camerer, C. (2012). Robust analogizing and the outside view: Two empirical tests of case-based decision making. *Strategic Management Journal*, *33*(5), 496–512. https://doi.org/10/dnkh8m

Markman, A. B., & Medin, D. L. (1995). Similarity and Alignment in Choice. *Organizational Behavior and Human Decision Processes*, *63*(2), 117–130. https://doi.org/10/c8z7r9

Ratcliff, C. L., & Sun, Y. (2020). Overcoming Resistance Through Narratives: Findings from a Meta-Analytic Review. *Human Communication Research*, *46*(4), 412–443. https://doi.org/10/gjscrm

Reinard, J. C. (1988). The Empirical Study of the Persuasive Effects of Evidence The Status After Fifty Years of Research. *Human Communication Research*, *15*(1), 3–59. https://doi.org/10/ccb67v

Reinhart, A. M. (2006). *Comparing the persuasive effects of narrative versus statistical messages: A meta -analytic review* (Ph.D.). State University of New York at Buffalo. New York, US. https://www.researchgate.net/profile/Amber-Reinhart/publication/34707525_Comparing_the_persuasive_effects_of_narrative_versus_statistical_messages_electronic_resource_A_meta-analytic_review/links/57335b6108ae9f741b26120c/Comparing-the-persuasive-effects-of-narrative-versus-statistical-messages-electronic-resource-A-meta-analytic-review.pdf

Shen, F., Sheer, V. C., & Li, R. (2015). Impact of Narratives on Persuasion in Health Communication: A Meta-Analysis. *Journal of Advertising*, *44*(2), 105–113. https://doi.org/10/gfkwj7

Wainberg, J. S. (2018, January 1). Stories vs Statistics: The Impact of Anecdotal Data on Managerial Decision Making. *Advances in Accounting Behavioral Research* (pp. 127–141). Emerald Publishing Limited. https://doi.org/10.1108/S1475-148820180000021006

Wainberg, J. S., Kida, T., David Piercey, M., & Smith, J. F. (2013). The impact of anecdotal data in regulatory audit firm inspection reports. *Accounting, Organizations and Society*, *38*(8), 621–636. https://doi.org/10/gjscqz

Winterbottom, A., Bekker, H. L., Conner, M., & Mooney, A. (2008). Does narrative information bias individual's decision making? A systematic review. *Social Science & Medicine*, *67*(12), 2079–2088. https://doi.org/10/cfpr4z

Yang, Z., Saini, R., & Freling, T. (2015). How Anxiety Leads to Suboptimal Decisions Under Risky Choice Situations. *Risk Analysis*, *35*(10), 1789–1800. https://doi.org/10/f7wtcx

Zebregs, S., van den Putte, B., Neijens, P., & de Graaf, A. (2015). The Differential Impact of Statistical and Narrative Evidence on Beliefs, Attitude, and Intention: A Meta-Analysis. *Health Communication*, *30*(3), 282–289. https://doi.org/10/ghk97p

Work ... primarily concerned with the psychological processes that govern judgment and inference ... portrayed people as fallible, not irrational.

—Amos Tversky

7
Discussion

Contents

This thesis investigated the psychology of capital allocation decisions. The influence of psychological factors on such decisions has not been sufficiently considered in the literature despite their importance to the performance of hierarchical organisations. This discrepancy is likely due to a greater focus of the role of organisational influences on firm performance in the management literature. The thesis did not investigate expertise effects, but instead focused largely on participants without management experience. This allowed a study of the specific cognitive processes without the potential confound of experience. Though, it is also worth noting that, in the one case where the work examined people with management experience, the pattern of results was largely the same as with naive participants. Each of the empirical chapters investigated distinct but related processes that are relevant to the capital allocation process. These chapters investigated whether people were able to account for the benefits of aggregation when considering multiple projects (Chapter 2), the influence of project feature alignability and metric variance when comparing projects directly (Chapter 4), and the influence of project anecdote similarity when the anecdote conflicts with statistical evidence (Chapter 6). Section 7.1 will first summarise the results of the empirical chapters, and Sections 7.2 and 7.3 will then discuss their theoretical and practical implications, respectively. Section 7.4 will conclude the thesis.

7.1 Summary of Results

Chapter 2 investigated participants' choice of risky business projects, when these are displayed sequentially and without feedback in between decisions. This design modelled the real-life situation that managers face in hierarchical organisations: an evaluation of a set of separate business project proposals over time with no immediate indication of the performance of those projects. Aggregating a portfolio of such projects is likely to show a lower chance of potential loss overall than might be originally assumed. The results from this chapter showed that people not only did not do this spontaneously, but also were not facilitated by manipulations that encouraged grouping choices together as a portfolio. People only seemed to recognise the benefits of aggregation when they were presented with an outcome probability distribution of the aggregated set of projects. There was no strong evidence that more subtle manipulations aimed at encouraging aggregation worked. Specifically, presenting projects together, specifying the total number of projects, and presenting projects that were all from the same industry did not reliably encourage aggregation.

Chapter 4 investigated capital allocation when projects were evaluated jointly and capital was allocated as a proportion of the budget, rather than a binary choice. The main manipulation was whether all the project attributes were alignable, or only the abstract financial metric (NPV) was alignable. NPV was also manipulated to be considered as either a reliable metric or not. This information was expressed either as explicit verbal instruction or as numerical ranges. The results showed that when reliability information was presented verbally, participants used NPV appropriately when all project attributes were completely alignable. That is, they used it when it was reliable and used the intrinsic project features when it was unreliable. When only NPV was alignable, participants relied on it almost exclusively. However, when reliability information was presented numerically, participants' allocation did not depend on the ranges—participants used NPV even when they had an opportunity to use the intrinsic features of the project. Overall, however, participants relied on NPV more when projects were low in alignment than when they were high in alignment.

Chapter 6 investigated the effect of anecdote similarity on allocations when the anecdote conflicted with the statistical data. Participants were asked to allocate a hypothetical budget between two projects. One of the projects (the target project) was clearly superior in terms of the provided statistical measures, but some of the participants also saw a description of a project with a conflicting outcome (the anecdotal project). This anecdotal project was always in the same industry as the target project. The anecdote description, however, either contained substantive connections to the target or not. Further, the anecdote conflicted with the statistical measures because it was either successful (positive anecdote) or unsuccessful (negative anecdote). The results showed that participants' decisions were influenced by anecdotes only when they believed that they were actually relevant to the target project. Further, they still incorporated the statistical measures into their decision. This was found for both positive and negative anecdotes. Further, participants were given information about the way that the anecdotes were sampled that suggested that the statistical information should have been used in all cases. Participants did not use this information in their decisions and still showed an anecdotal bias effect. Therefore, people seem to appropriately use anecdotes based on their relevance, but do not understand the implications of certain statistical concepts.

Together, these results show the bounds of people's decision-making capability in capital allocation. The participants in these experiments in general behaved rationally but struggled to incorporate certain statistical concepts into their decisions. Further, when confronted with multi-attribute choice, participants tended to allocate capital using a trade-off strategy. This was seen in the conflict between intrinsic project attributes and NPV in Chapter 4 and the conflict between the anecdotal and statistical evidence in Chapter 6. Participants' allocations were informed by relevant factors when these were sufficiently clear (as in the verbal reliability condition in Chapter 4). However, participants struggled to do this when the factor involved using a relatively basic statistical concept. Each empirical chapter included such a concept: risk aggregation in Chapter 2, metric variance in Chapter 4, and sample distribution in Chapter 6. The aggregated distribution in Chapter 2 and the verbal reliability manipulation in Chapter 4 showed that a formal understanding of such concepts is not always necessary if they are expressed explicitly.

The statistical concepts used in these studies are all likely accessible for people without much formal mathematical knowledge. A basic concept of risk aggregation is clearly available to laypeople as seen in the responses to multi-play gambles (e.g., one vs. 100 gambles). Further, people certainly have a basic understanding of numerical ranges and that a wider range means more spread. Despite likely having this understanding, participants in the above experiments were unable to use it in the decisions. Similarly, other work has shown that generalisations are sensitive to sampling (Carvalho et al., 2021). Therefore, it is unlikely that the people in the thesis experiments simply lacked any understanding of these statistical concepts or (at least sensitivity to this kind of information). Instead there appear to be important contextual factors that critically support or prevent people from showing their intuitive understanding. Unfortunately, the methods used in this thesis more closely resemble real decisions managers make than the prior research that showed people do reason with these kinds of statistical concepts. Further, it is not clear that these effects will simply disappear with just more maths knowledge and business experience. Previous work showed that expertise does not always remove biases and in some cases it seems to augment such effects (e.g., Haigh & List, 2005).

7.2 Theoretical Implications

The main theoretical contribution of this thesis is the addition of evidence that further specifies the conditions under which people make rational decisions in capital allocation scenarios. People made good decisions most of the time, but sometimes do not use relevant information in their decisions. Amos Tversky explained in his response to Cohen (1981, p. 355) that the work on heuristics and biases "portrayed people as fallible, not irrational." That is, people are not constantly making mistakes, but often behave rationally, largely due to adaptive heuristics. However, sometimes shortcuts that are usually helpful can fail. Studying such biases is similar to the way that optical illusions help understand the visual system. In both cases, these are systems that most of the time function properly, but sometimes reveal deficits.

Similarly, Simon (1955) identified human rationality as *bounded*, meaning that people's cognitive processes are limited. The main aim of the thesis was to contribute evidence for the ways that capital allocation decisions are bounded. To this end, in each experiment, participants were given capital allocation scenarios alongside both cues that describe their options and cues that frame the options in different ways. Identifying which cues were used by participants in their decisions, which cues were ignored, and which cues were integrated allowed to specify the bounds of people's cognitive capacity in these decisions. The experiments showed that people struggle to use certain statistical concepts in their decisions, but that they are also capable of making nuanced trade-offs and can be assisted by decision aides. Understanding how decision-making in capital allocation is constrained and biased is important in order to improve decision-making. Even if decisions are largely consistent with normative concepts, falling prey to the biases identified in this thesis can have severe consequences for organisations.

7.2.1 Statistical Concepts

Chapter 2 presented participants with a capital allocation situation in which an understanding of risk aggregation would have led to beneficial outcomes. Investing in all the hypothetical projects would have led to a much higher chance of gaining money than losing any. Each choice bracketing manipulation provided a hint of the possibility of combining the choices in this way. However, participants did not need to compute the aggregated value of the prospects themselves. An intuitive understanding of aggregation involved considering that some of the gambles will pay-off and make up for those that lost. However, this was not seen, with only weak evidence that people were influenced by the more subtle choice bracketing manipulations. Instead, people only seemed to respond to the concept of aggregation when it was explicitly showcased. Showing people a distribution of the outcome probabilities explicitly visualised the extent to which an aggregation of the risks can lead to an incredibly low chance of loss.

In Chapter 4, the NPVs that participants saw were critical to the allocation task. In the low alignment condition, NPV was the only alignable attribute in the comparison. In the high alignment condition, however, NPV was in competition with the intrinsic project feature values. An understanding of how to use numerical variance would have allowed participants to allocate capital according to the implied reliability of the comparison metric. In the low alignment condition, NPV was the only easy way to compare across projects, so it was a more useful cue than the rest of the non-alignable values. However, in the high alignment condition, the extent of numerical variance associated with each NPV could have been used to determine NPV reliability. There were two ways to do this: (a) noticing that in the low numerical reliability condition the ranges were all overlapping, and (b) noticing the difference in the width of the ranges between the two within-subjects reliability level conditions. By doing this, participants would have then been able to know to (in the high alignment condition) use NPV when ranges were narrow and use the intrinsic values more or exclusively when ranges were wider and overlapping.

Participants were able to do this sort of conditional allocation when reliability was expressed explicitly as words, but not when it was expressed numerically.

In Chapter 6, participants did not make use of descriptive information about the anecdote sample distribution. As in Chapter 4, participants were confronted with a conflict of cues: statistical information vs. a potentially relevant anecdote. Regardless of the similarity manipulation, a consideration of the sample from which the anecdote was sampled should have informed how the anecdote was used. Imagine a distribution that represents the similarity of all the individual projects in the sample. That is, the x-axis represents the similarity to the target project and the y-axis is the frequency of projects that represent each level of similarity. Even if the sampled anecdote appears very relevant to the target project, if the underlying distribution of the sample is highly negatively skewed, such that most projects in the sample are equivalently similar to the target, then the sampled anecdote is not necessarily more informative than the aggregated measure. On the other hand, if the underlying distribution is positively skewed, normally distributed, or even uniform, then the fact that the sampled anecdote appears highly relevant to the target project may actually mean that it is more informative than the aggregated measure. Prior work shows that people can reason about distributions effectively when experiencing the sampling directly (e.g., Carvalho et al., 2021; Hertwig et al., 2004). Chapter 6 shows that people struggle to use this information when it is described verbally.

While people struggled to understand and use certain statistical concepts they still seemed to be able to integrate multiple cues and create trade-offs. As discussed in Chapter 5, both Chapters 4 and 6 provided participants with more than one cue to use for project evaluation. In Chapter 4, people seemed to strike a trade-off between NPV and the intrinsic project features as opposed to choosing one or the other with a consistent strategy. In Chapter 6, the anecdotal and statistical evidence provided conflicting cues for each target project. However, participants allocated as if both the anecdotes and statistics had some relevance. Similar to the above, participants appeared to integrate the influence of these two cues, as opposed to picking a consistent evidence reliance strategy for their allocation decisions. These findings might be explained through satisficing (Simon, 1955) or a constraint satisfaction model (e.g., Glöckner et al., 2014). Future research can test these explanations, as well as further clarify to what extent constructs such as need for cognition or mathematical skill may explain individual differences.

7.2.2 Decision Aides

While trade-offs allow people to integrate multiple cues, decision aides allow people to use statistical concepts for more nuanced decision-making. Chapter 2 found that people's understanding of risk aggregation was facilitated when the mathematical work was done for them and the aggregated values were displayed visually as a distribution. However, a follow-up experiment to Chapter 4 (detailed in Appendix B.7) found that even explicit instructions sometimes do not work. That is, even explaining the way that ranges can be used as reliability information and telling participants how to implement this in the capital allocation task did not facilitate proper use of ranges.

Future work should investigate the impact of visualisation on people's use of variance information in these situations. Much work has investigated visualising uncertainty information (Bostrom et al., 2008; Brodlie et al., 2012; T. J. Davis & Keller, 1997; Johnson & Sanderson, 2003; Kinkeldey et al., 2017; Kox, 2018; Lapinski, 2009; Lipkus & Hollands, 1999; Lipkus, 2007; MacEachren, 1992; Padilla et al., 2018; Pang et al., 1997; Potter et al., 2012; Ristovski et al., 2014; Spiegelhalter et al., 2011; Torsney-Weir et al., 2015). For instance, a Hypothetical Outcome Plot (Hullman et al., 2015; Kale et al., 2019) expresses variance information as dynamic plots and is one method that is likely to be beneficial to people's understanding of ranges as used in this thesis. Visualisation could also apply to the work in Chapter 6. Using a visual array as in Jaramillo et al. (2019) is likely to facilitate people's understanding of the importance of statistical evidence over anecdotes. However, an additional visualisation of the distribution of the underlying similarity to the target may also be necessary to facilitate understanding of the relevance of the sample distribution. Ultimately, visualisations of the effects of certain statistical concepts may be necessary for people to use them appropriately.

7.2.3 How Bounded is Bounded Rationality?

The boundary between the cues that participants were able to use and the statistical concepts that they did not use is unclear. That is, the cues that they were able to use were not trivial, and the concepts that they were not able to use are relatively basic. For instance, the finding in Chapter 6 that people were able to use relevance information to guide their allocations shows an ability to quite specific information to inform choice. On the other hand, the statistical concepts that participants ignored in each empirical chapter are all relatively intuitive. While concepts of aggregation, variance, and sample distribution are typically studied at a tertiary level, they can be understood when acted out or experienced.

Clark and Karmiloff-Smith (1993) proposed a distinction between two levels of representing knowledge. At the *implicit* level an individual can only make use of a certain system of knowledge, while it is only at the *explicit* level that they develop insight into that system. For instance, young children can use closed class words such as "the" or "to", but only identify them as words later in development. Further, children's play often implicitly contains many mathematical concepts, despite the children's struggle to explicitly reason with the exact same concepts in more formal problem-solving (Sarama & Clements, 2009). Adults may have a similar distinction in knowledge representation. Concepts that can be used when experienced directly, such as in risky choice from experience, are not represented in a way that they can be used when presented descriptively, such as in risky choice from description. This kind of distinction may explain why participants in the thesis experiments failed to use concepts that have been shown to be accessible to laypeople.

7.2.4 Expertise Effects

Future research should investigate the potential expertise effects that may influence the findings of the thesis. This is important because of the potential downstream effects of biased managerial decision-making. For instance, it is unclear to what extent psychological factors such as the ones discussed in this thesis may account for the finding that undiversified firms often perform better than diversified firms. On the one hand, business professionals tend to work with numbers, so the effects found in this thesis may be less pronounced for them. For instance, Smith and Kida (1991) reviewed the heuristics and biases literature and concluded that certain cognitive biases are not as strong for accounting professionals as they are for naive participants.

On the other hand, these effects may actually be stronger in managers. For instance, Haigh and List (2005) found that professional traders show more myopic loss aversion than students. Chapter 2 showed that people tend to consider risky choices one at a time and therefore tend to be more risk averse to a set of projects than they would be if the risks were aggregated. Managers might be even more risk averse in these situations because of the increased stakes for their jobs. Lovallo et al. (2020) discussed the ways in which managers tend to have a blind spot for such project evaluations: they aggregate their personal stock market portfolio, but not their intra-firm project portfolio.

Chapter 4 found evidence of variance neglect for both laypeople and Master of Management students. Further, in the case of the work in Chapter 6, it is possible that business managers prefer anecdotal cases to inform their decisions because of their higher salience, compared to statistical data. Managers are also more likely to feel as if the situation is relevant to them, which acording to Freling et al. (2020) would predict more anecdotal bias.

7.3 Practical Implications

The findings of this thesis have a number of potential implications for managerial decision-making. Despite the uncertainty about potential expertise effects, this section assumes that the findings of the thesis generalise to experienced managers, if not in degree, at least qualitatively. Management researchers have suggested ways of overcoming psychological biases in managerial decision-making ever since such biases were identified. Many practitioner-oriented papers have used the findings of the judgement and uncertainty literature both to explain managerial decision-making processes and to suggest ways of reducing bias (Courtney et al., 1997; Courtney et al., 2013; Hall et al., 2012; Koller et al., 2012; Lovallo & Sibony, 2014; Sibony et al., 2017), with only some specifically focused on capital allocation decisions (Birshan et al., 2013). This section will review some of the implications the findings of this thesis may have on both organisational policies and manager decision-making.

The findings of Chapter 2 show that the framing of business project proposals is important for the way that people perceive their risk. Specifically, in order to better account for the risks of business projects it is important to (a) make it easier for managers to group projects together, and (b) aggregate a portfolio of projects for them. This suggests implementing organisational changes that will facilitate the capital allocation process. For instance, Lovallo et al. (2020) suggested that companies change the frequency that they evaluate projects to better allow for an aggregation of the projects. Doing this will enable an explicit computation of the aggregated values and therefore a visualisation of the outcome probability distribution. Such a process could facilitate aggregation without a need to rely on managers' intuition during sequential project evaluation decisions.

One implication of Chapter 4 is that it is important to expose the variance that underlies abstract financial measures. Further, translating such numerical variance estimates into clear verbal information would help facilitate managers' understanding and implementation of such estimates. Organisational changes could include reducing diversification so that there is less reliance on abstract metrics. This would allow for more of a comparison between alignable project attributes, potentially reducing forecast error. Koller et al. (2017) found that companies with more similar business units report faster growth and greater profitability than competitors, compared with companies with dissimilar business units. Further, companies can also work to develop better metrics and establish norms about how much to discount a metric given its underlying variance.

The main implication of Chapter 6 is that managers should pay attention to the way that they compare target projects to other cases. It is important to collect prior cases that are relevant, and to have as many such cases as possible. Ideally, each such prior case should be weighed by similarity (Lovallo et al., 2012). If this is done, the prior distribution of the similarity of the sample would be taken into account when computing subsequent aggregation. When identifying such similarity ratings, it is important to focus on relevant underlying structure. This would reduce any erroneous connections to cases that only have a mere surface similarity. This distinction is also relevant in a situation in which only one prior case can be found. Research on analogy shows that analogical comparison helps expose the underlying relational structure between objects (e.g., Kurtz et al., 2013; Markman & Gentner, 1993). Therefore, managers should take care to first identify such relational structures first before making subsequent inferences.

Addressing these psychological effects will help eliminate some of the biases in the capital allocation process, but will not address other related biases. For instance, the above effects all involve decisions that require an evaluation of financial forecast estimates such as future cash flows and the related uncertainty. Therefore, a further source of error could arise from the initial estimation of these probability and cash flow values. For instance, such estimates could be influenced by optimism or confidence biases. These biases, however, can in turn also be addressed (Flyvbjerg et al., 2018).

7.4 Conclusion

Capital allocation decisions are consequential for large organisations. This thesis tested the conditions under which people behave rationally or are fallible when allocating capital. The experiments found that participants struggle to incorporate concepts such as risk aggregation, estimate variance, and sample distribution into their decisions. Participants only seemed to be able to do this when the concept was expressed visually very explicitly. However, when there were multiple cues for choice evaluation, the results also showed that participants were capable of integrating conflicting information in their decisions. Identifying such cognitive bounds helps to better understand how people evaluate multiple choices and helps future research develop methods to facilitate better decisions.

References

Birshan, M., Engel, M., & Sibony, O. (2013). Avoiding the quicksand: Ten techniques for more agile corporate resource allocation. *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/avoiding-the-quicksand

Bostrom, A., Anselin, L., & Farris, J. (2008). Visualizing Seismic Risk and Uncertainty: A Review of Related Research. *Annals of the New York Academy of Sciences*, *1128*(1), 29–40. https://doi.org/10/c32d9k

Brodlie, K., Allendes Osorio, R., & Lopes, A. (2012). A Review of Uncertainty in Data Visualization. In J. Dill, R. Earnshaw, D. Kasik, J. Vince, & P. C. Wong (Eds.), *Expanding the Frontiers of Visual Analytics and Visualization* (pp. 81–109). Springer London. https://doi.org/10.1007/978-1-4471-2804-5_6

Carvalho, P. F., Chen, C.-h., & Yu, C. (2021). The distributional properties of exemplars affect category learning and generalization. *Scientific Reports*, *11*(1), 11263. https://doi.org/10.1038/s41598-021-90743-0

Clark, A., & Karmilff-Smith, A. (1993). The cognizer's innards. *Mind & Language*, *8*(4), 487–519. https://doi.org/10/csfpck

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, *4*(3), 317–331. https://doi.org/10/fn9rpc

Courtney, H., Kirkland, J., & Viguerie, P. (1997). Strategy under uncertainty. *Harvard Business Review*, *75*(6), 67–79. https://hbr.org/1997/11/strategy-under-uncertainty

Courtney, H., Lovallo, D., & Clarke, C. (2013). Deciding How to Decide. *Harvard Business Review*, *91*(11), 62–70.

Davis, T. J., & Keller, C. P. (1997). Modelling and visualizing multiple spatial uncertainties. *Computers & Geosciences*, *23*(4), 397–408. https://doi.org/10.1016/S0098-3004(97)00012-5

Flyvbjerg, B., Ansar, A., Budzier, A., Buhl, S., Cantarelli, C., Garbuio, M., Glenting, C., Holm, M. S., Lovallo, D., Lunn, D., Molin, E., Rønnest, A., Stewart, A., & van Wee, B. (2018). Five things you should know about cost overrun. *Transportation Research Part A: Policy and Practice*, *118*, 174–190. https://doi.org/10/ghdgv4

Freling, T. H., Yang, Z., Saini, R., Itani, O. S., & Rashad Abualsamh, R. (2020). When poignant stories outweigh cold hard facts: A meta-analysis of the anecdotal bias. *Organizational Behavior and Human Decision Processes*, *160*, 51–67. https://doi.org/10/gg4t2f

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account. *Cognition*, *133*(3), 641–666. https://doi.org/10/f6q9fj

Haigh, M. S., & List, J. A. (2005). Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis. *The Journal of Finance*, *60*(1), 523–534. https://doi.org/10/c7jn9k

Hall, S., Lovallo, D., & Musters, R. (2012). How to put your money where your strategy is. *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-to-put-your-money-where-your-strategy-is

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), 534–539. https://doi.org/10/b274n8

Hullman, J., Resnick, P., & Adar, E. (2015). Hypothetical Outcome Plots Outperform Error Bars and Violin Plots for Inferences about Reliability of Variable Ordering. *PLOS ONE*, *10*(11), e0142444. https://doi.org/10/f3tvsd

Jaramillo, S., Horne, Z., & Goldwater, M. (2019, May 15). *The impact of anecdotal information on medical decision-making* (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/r5pmj

Johnson, C., & Sanderson, A. (2003). A Next Step: Visualizing Errors and Uncertainty. *IEEE Computer Graphics and Applications*, *23*(5), 6–10. https://doi.org/10/df8kvd

Kale, A., Nguyen, F., Kay, M., & Hullman, J. (2019). Hypothetical Outcome Plots Help Untrained Observers Judge Trends in Ambiguous Data. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 892–902. https://doi.org/10/gghfzn

Kinkeldey, C., MacEachren, A. M., Riveiro, M., & Schiewe, J. (2017). Evaluating the effect of visually represented geodata uncertainty on decision-making: Systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science*, *44*(1), 1–21. https://doi.org/10/f3m63m

Koller, T., Lovallo, D., & Williams, Z. (2012). Overcoming a bias against risk. *McKinsey Quarterly*, (4). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/overcoming-a-bias-against-risk

Koller, T., Lovallo, D., & Williams, Z. (2017). Should assessing financial similarity be part of your corporate portfolio strategy? *McKinsey on Finance*, (64). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/should-assessing-financial-similarity-be-part-of-your-corporate-portfolio-strategy

Kox, E. (2018). *Evaluating the effectiveness of uncertainty visualizations: A user-centered approach* (Masters thesis). University of Utrecht. http://dspace.library.uu.nl/handle/1874/367380

Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1303. https://doi.org/10/gjvc63

Lapinski, A.-L. S. (2009, October 1). *A Strategy for Uncertainty Visualization Design.* Defence R&D Canada – Atlantic. https://apps.dtic.mil/sti/citations/ADA523694

Lipkus, I. M., & Hollands, J. G. (1999). The Visual Communication of Risk. *JNCI Monographs*, *1999*(25), 149–163. https://doi.org/10/gd589v

Lipkus, I. M. (2007). Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations. *Medical Decision Making*, *27*(5), 696–713. https://doi.org/10/b8p3gf

Lovallo, D., Clarke, C., & Camerer, C. (2012). Robust analogizing and the outside view: Two empirical tests of case-based decision making. *Strategic Management Journal*, *33*(5), 496–512. https://doi.org/10/dnkh8m

Lovallo, D., Koller, T., Uhlaner, R., & Kahneman, D. (2020). Your Company Is Too Risk-Averse. *Harvard Business Review*, *98*(2), 104–111.

Lovallo, D., & Sibony, O. (2014). Is your budget process stuck on last year's numbers? *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/is-your-budget-process-stuck-on-last-years-numbers
MacEachren, A. M. (1992). Visualizing Uncertain Information. *Cartographic Perspectives*, (13), 10–19. https://doi.org/10/gjscq9
Markman, A. B., & Gentner, D. (1993). Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, *25*(4), 431–467. https://doi.org/10/cqtx7q
Padilla, L. M., Creem-Regehr, S. H., Hegarty, M., & Stefanucci, J. K. (2018). Decision making with visualizations: A cognitive framework across disciplines. *Cognitive Research: Principles and Implications*, *3*(1), 29. https://doi.org/10/ggrtng
Pang, A. T., Wittenbrink, C. M., & Lodha, S. K. (1997). Approaches to uncertainty visualization. *The Visual Computer*, *13*(8), 370–390. https://doi.org/10/fdnbmw
Potter, K., Kirby, R. M., Xiu, D., & Johnson, C. R. (2012). Interactive visualization of probability and cumulative density functions. *International Journal for Uncertainty Quantification*, *2*(4), 397–412. https://doi.org/10/ghhdw2
Ristovski, G., Preusser, T., Hahn, H. K., & Linsen, L. (2014). Uncertainty in medical visualization: Towards a taxonomy. *Computers & Graphics*, *39*, 60–73. https://doi.org/10/f5v59d
Sarama, J., & Clements, D. H. (2009). Building Blocks and Cognitive Building Blocks: Playing to Know the World Mathematically. *American Journal of Play*, *1*(3), 313–337. Retrieved June 23, 2021, from https://eric.ed.gov/?id=EJ1069014
Sibony, O., Lovallo, D., & Powell, T. C. (2017). Behavioral Strategy and the Strategic Decision Architecture of the Firm. *California Management Review*, *59*(3), 5–21. https://doi.org/10/gcp2w3
Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99. https://doi.org/10/dw3pfg
Smith, J. F., & Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. *Psychological bulletin*, *109*(3), 472–489. https://doi.org/10/fwv6z6
Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, *333*(6048), 1393–1400. https://doi.org/10.1126/science.1191181
Torsney-Weir, T., Sedlmair, M., & Möller, T. (2015). Decision making in uncertainty visualization. http://eprints.cs.univie.ac.at/4598/

A
Chapter 2 Appendix

Contents

This appendix contains supplementary materials and analyses for the two experiments reported in Chapter 2. In addition, it also report two experiments that were conducted to follow-up the findings in Experiments 1 and 2. Both follow-up experiments tested project choice as in the first two experiments, but Experiment 3 further investigated the effect of similarity, and Experiment 4 further investigated the effect of awareness.

All four experiments featured probability outcome distributions. These were

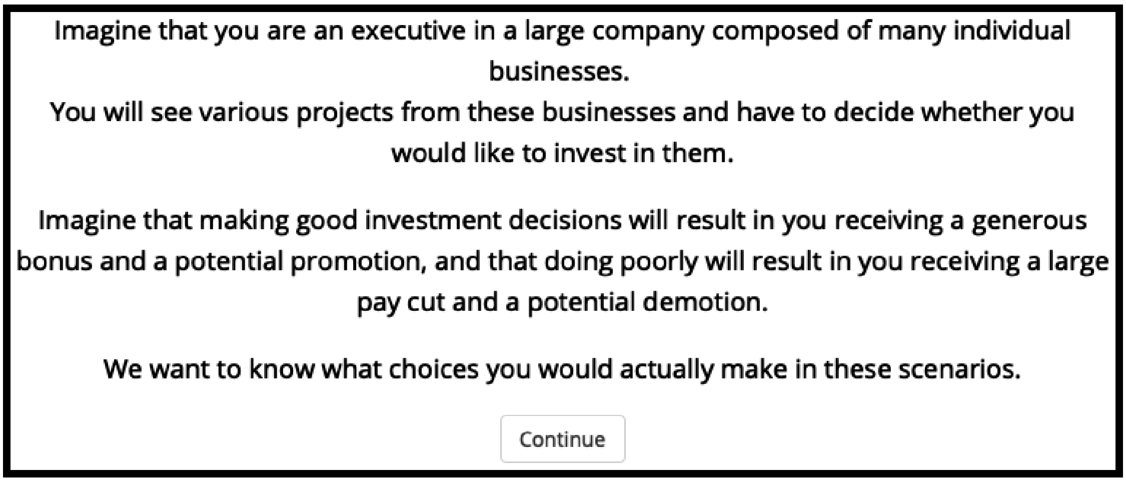


Figure A.1: Experiment 1 instructions.

Poisson binomial distributions that were calculated using the R package `poibin`, which uses calculations described in Hong (2013).

A.1 Experiment 1

A.1.1 Method

A.1.1.1 Materials

A.1.1.1.1 Instructions Participants were shown the instructions in Figure A.1.

A.1.1.1.2 Outcome Distribution Decision Figure A.2 shows the outcome distribution display that participants saw in Experiment 1.

A.1.1.1.3 Follow-up Gambles

Negative EV Gambles It was important to make sure that participants were generally making decisions that were in line with EV theory and that the sample was not abnormally risk tolerant. As such, participants saw two project decisions that had a negative EV. Out of the 396 negative EV gambles included (two per participant), all but four were rejected.

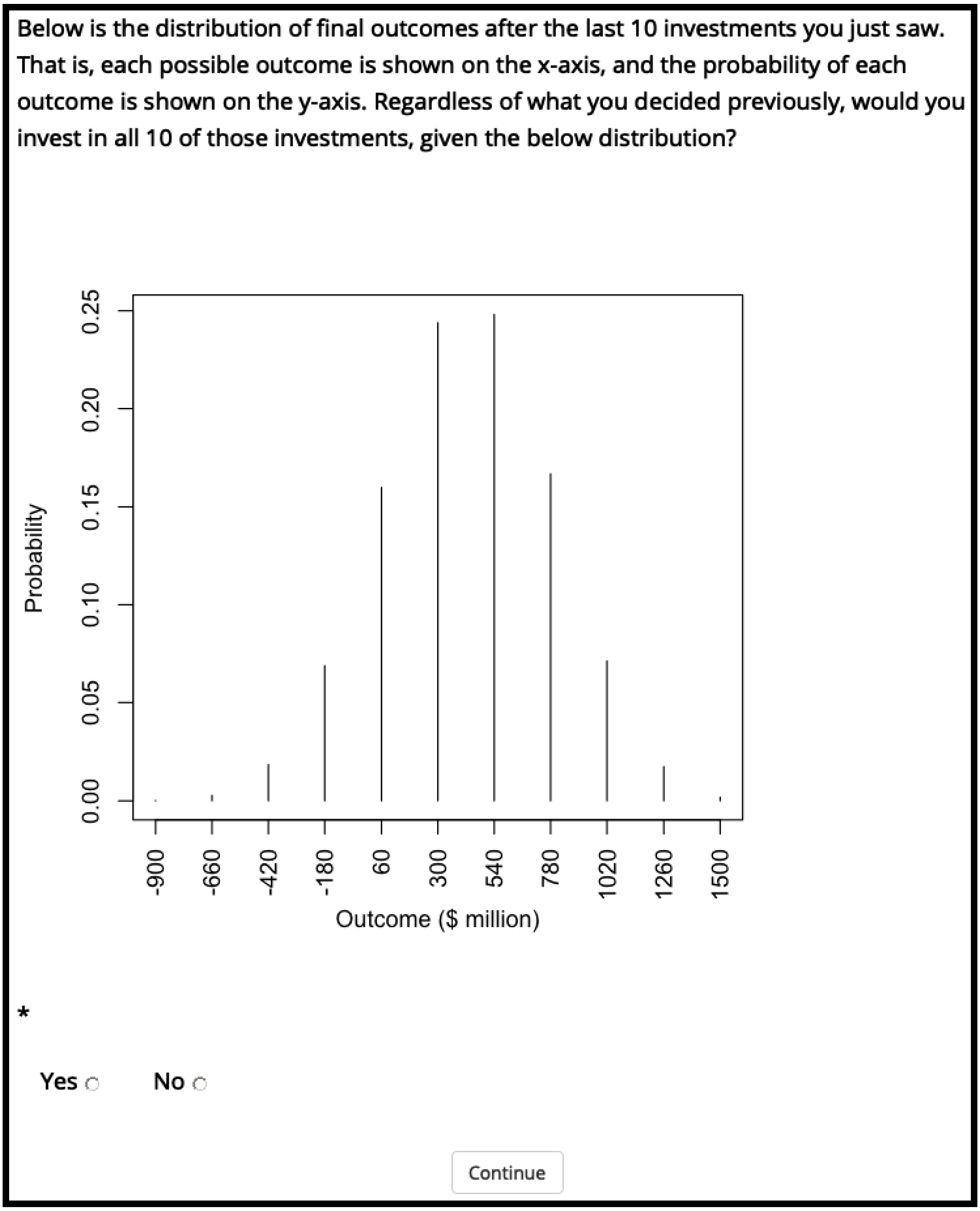


Figure A.2: The outcome distribution of the 10 gambles used in Experiment 1.

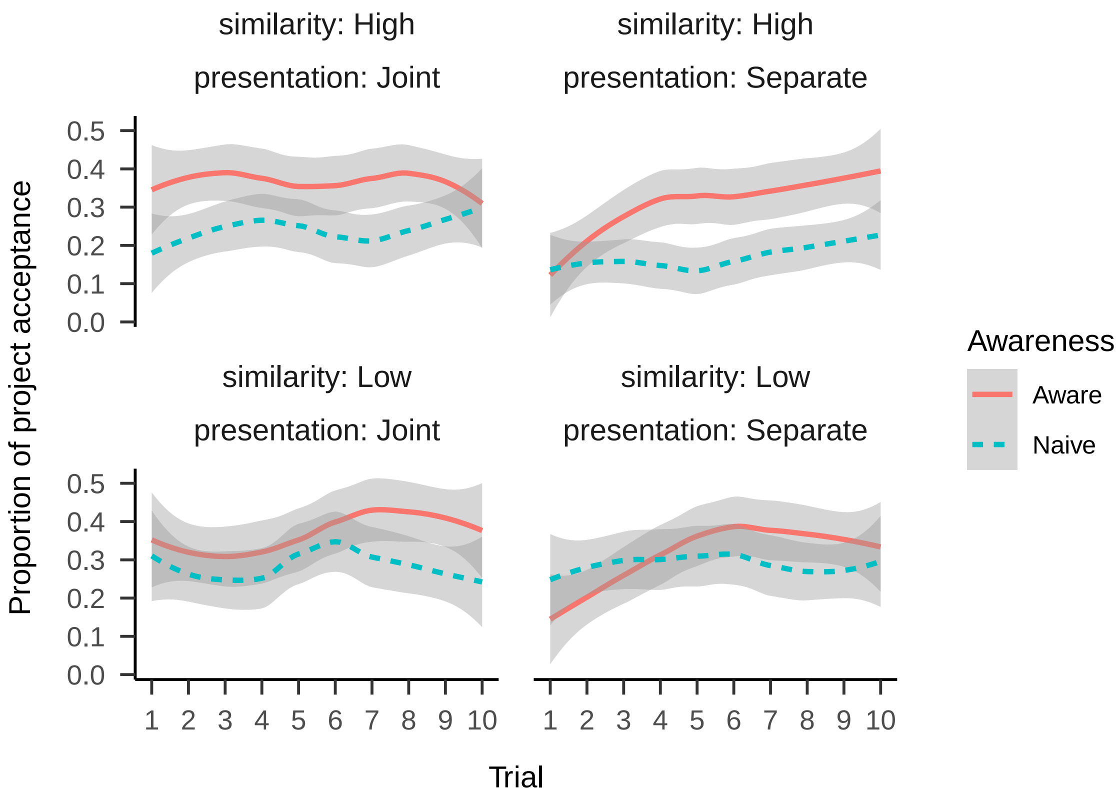


Figure A.3: Proportion of project acceptance by trial, similarity, awareness, and presentation conditions. LOESS is used for smoothing over trials, and the shading represents 95% confidence intervals.

Samuelson (1963) Gambles Participants saw the original Samuelson (1963) gamble, were asked whether they would accept 10 of that gamble, and whether they would accept those 10 given the associated outcome distribution. They then saw the same three questions, but using outcome magnitudes that were similar to the ones in the risky investment task. That is, $100 million instead of $100.

Redelmeier and Tversky (1992) Gambles Participants saw the same three types of gambles (single, 10, and aggregated), but with the values from the gambles that were used by Redelmeier and Tversky (1992).

A.1.2 Results

A.1.2.1 Trial-by-Trial Analysis

Figure A.3 shows proportions of project acceptance across all conditions and trials.

A.1.2.2 Outcome Distribution

A paired-samples t-test was conducted to compare participants' decision to invest in the 10 projects while seeing an aggregated distribution, and their decisions to invest in the projects individually, without the distribution. Participants invested in the 10 projects more when seeing the distribution both in the separate presentation phase, $t(197) = 5.48$, $p < .001$, $d_z = 0.50$, 95% CI $[0.31, 0.68]$; and in the joint presentation phase, $t(197) = 4.17$, $p < .001$, $d_z = 0.37$, 95% CI $[0.19, 0.56]$.

However, it was subsequently discovered that the code that generated this distribution mistakenly flipped the outcome values. This means that although it appeared from the distribution that the probability of loss was 0.09, the actual probability of loss of the underlying values given the correct distribution was 0.26. As such, even though Experiment 1 found an effect of distribution, it was unclear if the effect was driven by participants actually accurately assessing the riskiness of the individual gambles, and therefore showing a difference between the isolated and aggregated gambles in a normative way.

A.2 Experiment 2

A.2.1 Method

A.2.1.1 Participants

A.2.1.1.1 Power Analysis The power analysis was conducted using the `pwr` package (Champely, 2020a), based on the presentation effect size from Experiment 1, since it was the smallest effect. The analysis suggested that a minimum sample size of 164 $(41 \cdot 4)$ was required for the presentation effect with an expected power of at least 80%.

A.2.1.2 Materials

A.2.1.2.1 Follow-up Figure A.4 shows the project number question. The maximum value that they could enter was set to 20. Figures A.5 and A.6 ask participants whether they are willing to take all or none of the projects; and how

Figure A.4: Experiment 2 project number question.

many projects would they choose if they could pick randomly (maximum value was set to 20). Those in the distribution absent condition were asked the same questions, but without the distribution and its explanation.

A.2.2 Results

A.2.2.1 Follow-up

A.2.2.1.1 Project Number Participants were asked how many projects they thought they saw. Figure A.7 shows that overall people correctly estimated the number of projects, with more accuracy for those in the aware condition.

A.2.2.1.2 Portfolio Choice - Binary Participants were then asked if they would rather invest in all or none of the projects. As Figure A.8 shows, the difference between presentation conditions was not significant, $\hat{\beta} = 0.15$, 95% CI $[-0.29, 0.60]$, $z = 0.67$, $p = .500$. The awareness effect was also not significant, $\hat{\beta} = 0.28$, 95% CI $[-0.17, 0.72]$, $z = 1.21$, $p = .225$. However, those that that saw a distribution chose to invest in all 10 projects significantly more (71.43%) than those that did not see a distribution (36.59%), .

A.2.2.1.3 Portfolio Choice - Number Subsequently, participants were asked how many projects they would invest in out of the 10 that they saw. As Figure A.9 shows, the difference between presentation conditions was not significant, $d_s = 0.08$, 95% CI [-0.35, 0.52], $t(80) = 0.38$, $p = .706$. The awareness effect was also not significant, $d_s = 0.09$, 95% CI [-0.34, 0.53], $t(79) = 0.42$, $p = .678$. However, those that that saw a distribution chose to invest in significantly more projects than those that did not see a distribution, $d_s = 0.60$, 95% CI [0.15, 1.03], $t(81) = 2.70$, $p = .009$.

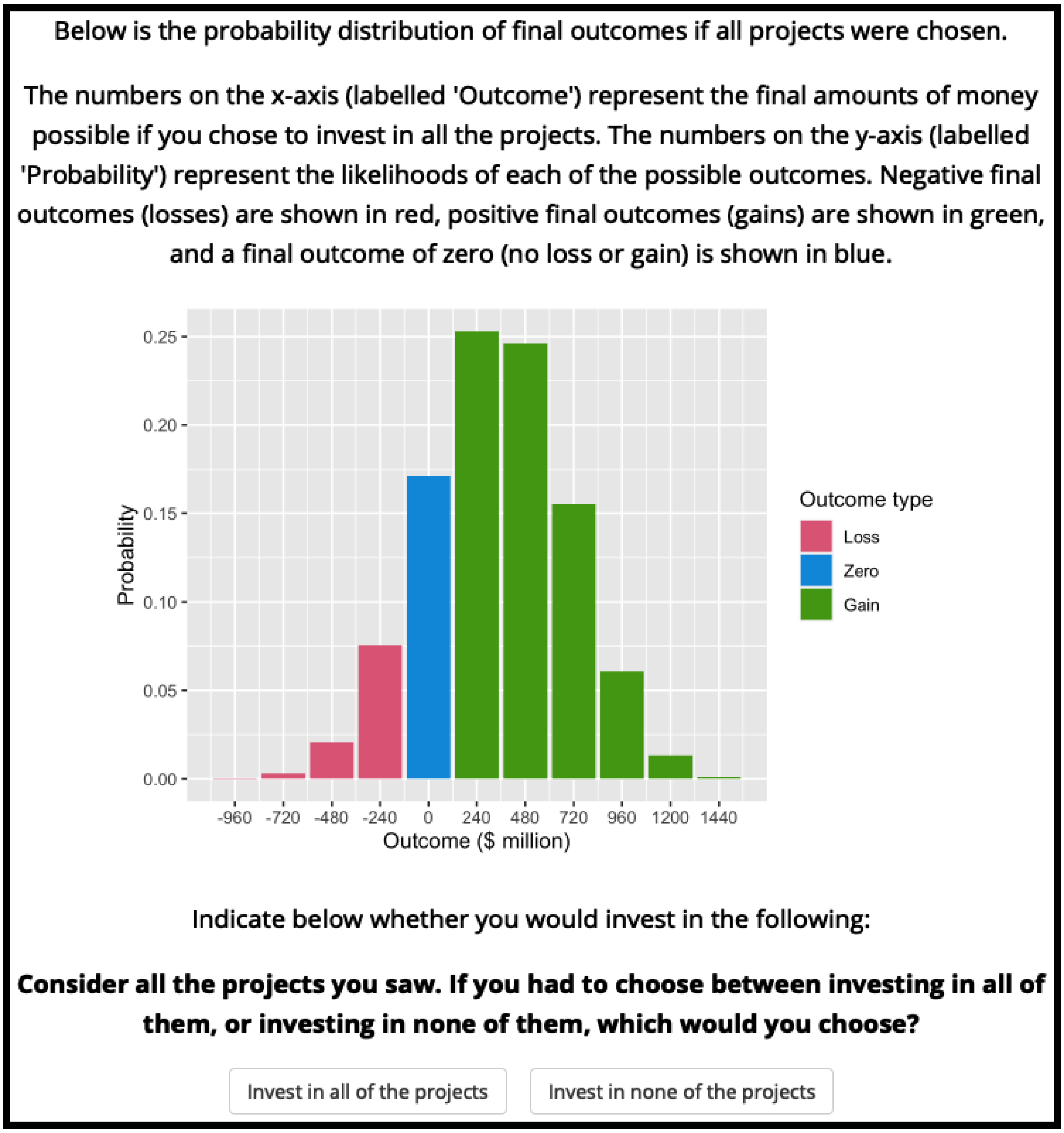


Figure A.5: Experiment 2 binary portfolio question.

A.2.2.2 Gambles

Figures A.10 and A.11 show that the overall people seemed to prefer gambles with higher probabilities of gain, sometimes regardless of expected value or value of the gain.

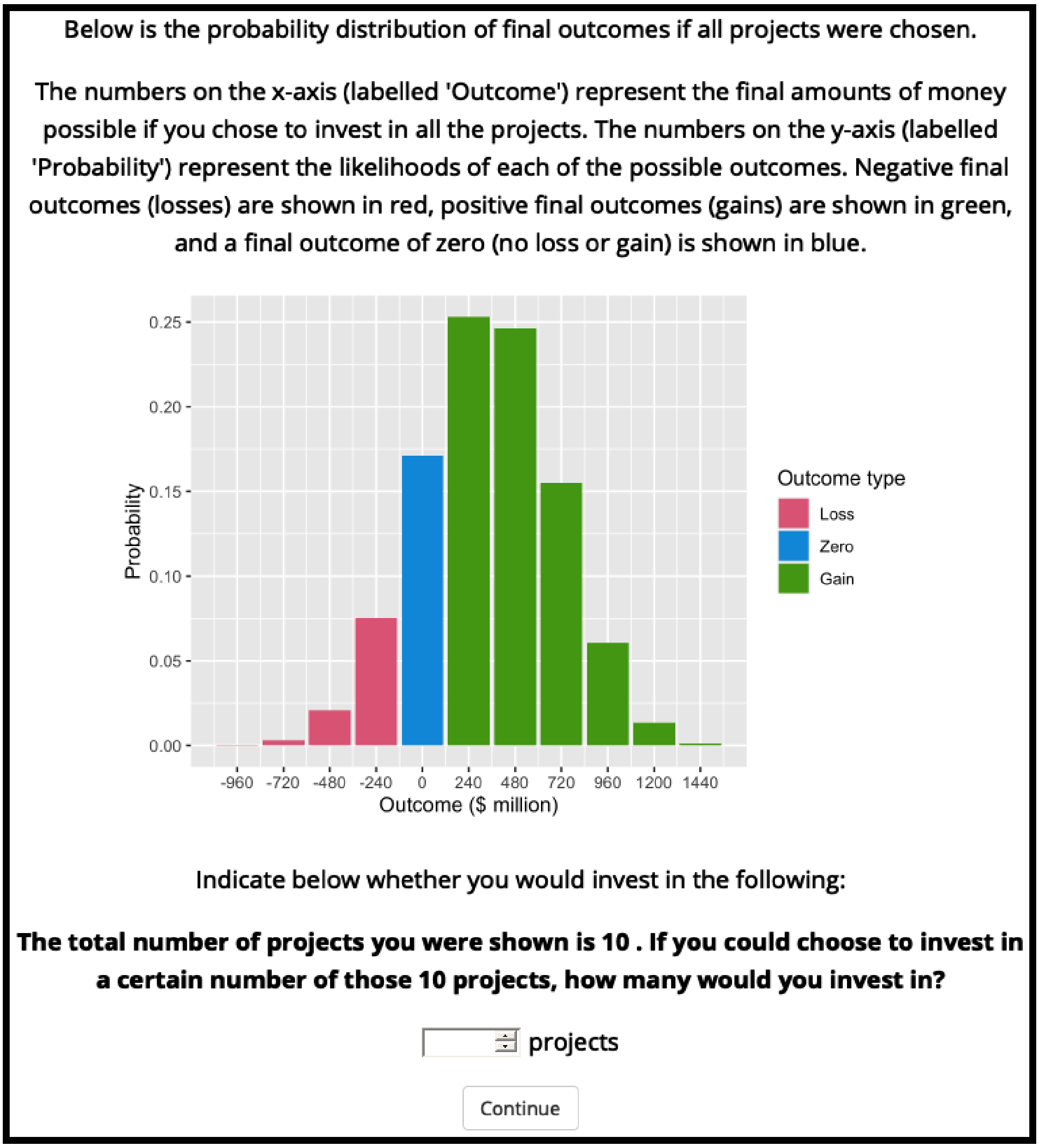


Figure A.6: Experiment 2 numerical portfolio question.

A.3 Experiment 3

Experiment 3 investigated the effect of similarity on project choice. The previous experiments did not counterbalance the project domain when displaying the 10 projects to participants. Experiment 3 used 10 different potential business domains when constructing the project descriptions in order to reduce any potential effect that the specific domain may have on people's choice. Therefore, Experiment 3

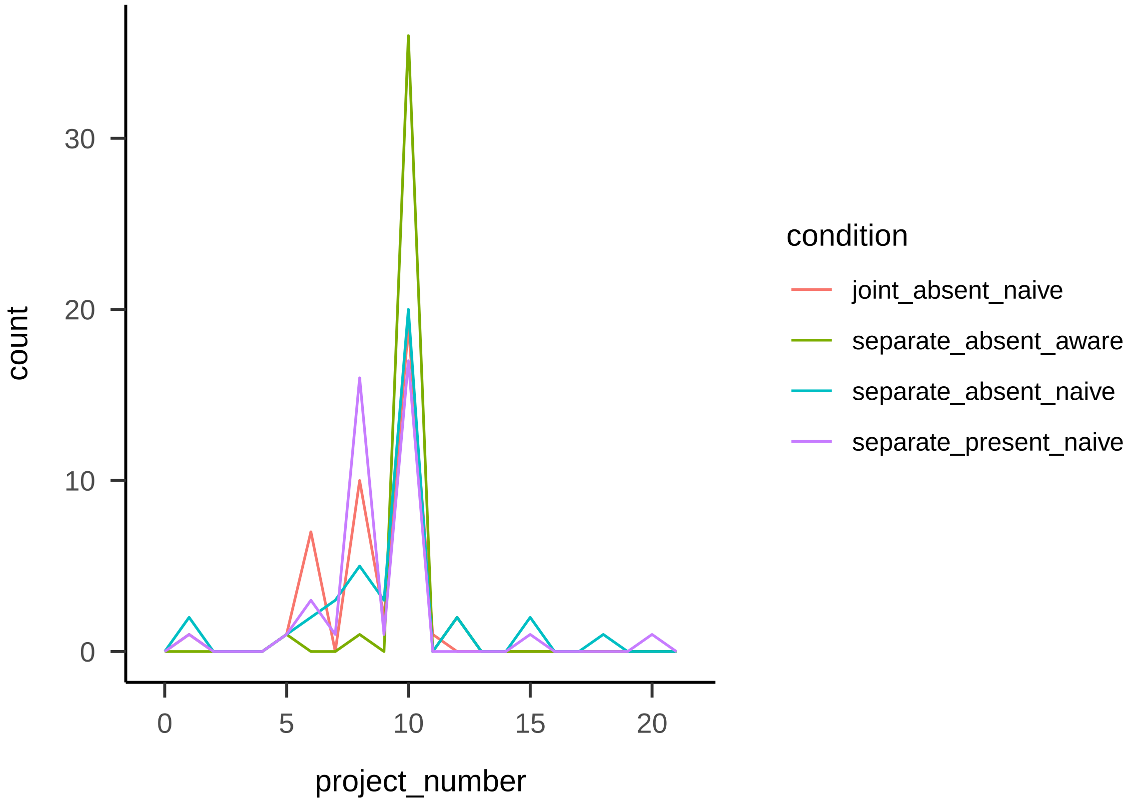


Figure A.7: Number of projects participants reported seeing, by condition.

again tested Hypothesis 2.3.

A.3.1 Method

A.3.1.1 Participants

Two hundred and sixty-six participants (127 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 39.56 years ($SD = 8.77$, $min. = 25$, $max. = 71$). Participants reported an average of 5.64 years ($SD = 6.45$, $min. = 0$, $max. = 40$) working in a business setting, and an average of 3.28 years ($SD = 4.92$, $min. = 0$, $max. = 30$) of business education. The mean completion time of the task was 9.23 min ($SD = 7.2$, $min. = 1.41$, $max. = 65.46$). Table A.1 shows the allocation of participants to the different conditions.

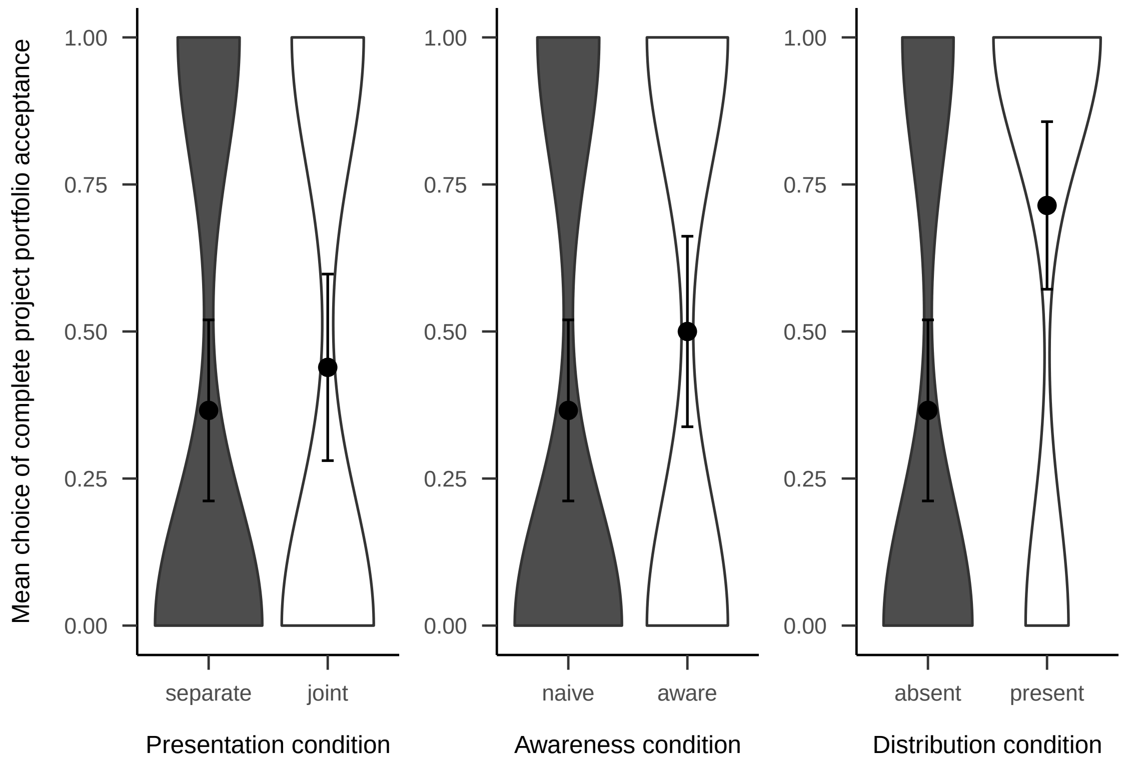


Figure A.8: Mean choice of investing in all 10 projects for the presentation, awareness, and distribution effects. Note, the condition on the left of each effect is the reference condition (separate presentation, naive awareness, distribution absent). As such, it is identical for the three effects.

Table A.1: Experiment 3 group allocation.

Similarity	N
High	133
Low	133
Total	266

A.3.1.2 Materials

A.3.1.2.1 Instructions Participants were shown the same instructions as in Experiment 1 (see Section 2.2.1.2.1).

A.3.1.2.2 Risky Investment Task Participants saw displays with the same gamble values as those in Experiment 2 (see Section 2.3.1.2.2), but with some changes in wording and sentence structure. The gamble information was the same,

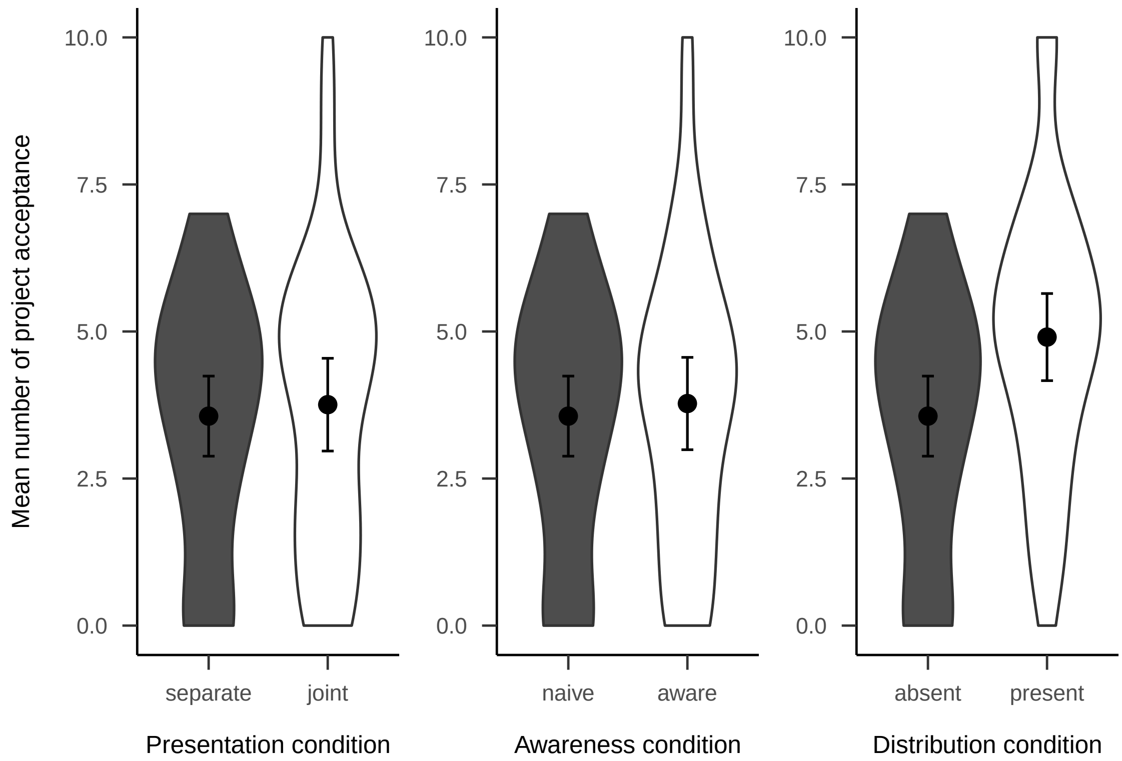


Figure A.9: Mean number of projects chosen in the follow-up for the presentation, awareness, and distribution effects. Note, the condition on the left of each effect is the reference condition (separate presentation, naive awareness, distribution absent). As such, it is identical for the three effects.

but extra prose was added to describe the projects. Further, the order of the sentences was randomised, so that the descriptions would not appear so similar. See Figure A.12 for an example.

The similarity manipulation was as in Experiment 1. However, project domain was varied so that in the high similarity condition participants saw one of ten project domains.

A.3.1.2.3 Follow-up The follow-up questions were similar to those in Experiment 2 (see Section 2.3.1.2.3), except in the portfolio number question participants were also shown the total number of projects that they saw (10). Further, another question was added, asking how many projects participants were expecting to see at the beginning of the experiment (see Figure A.13).

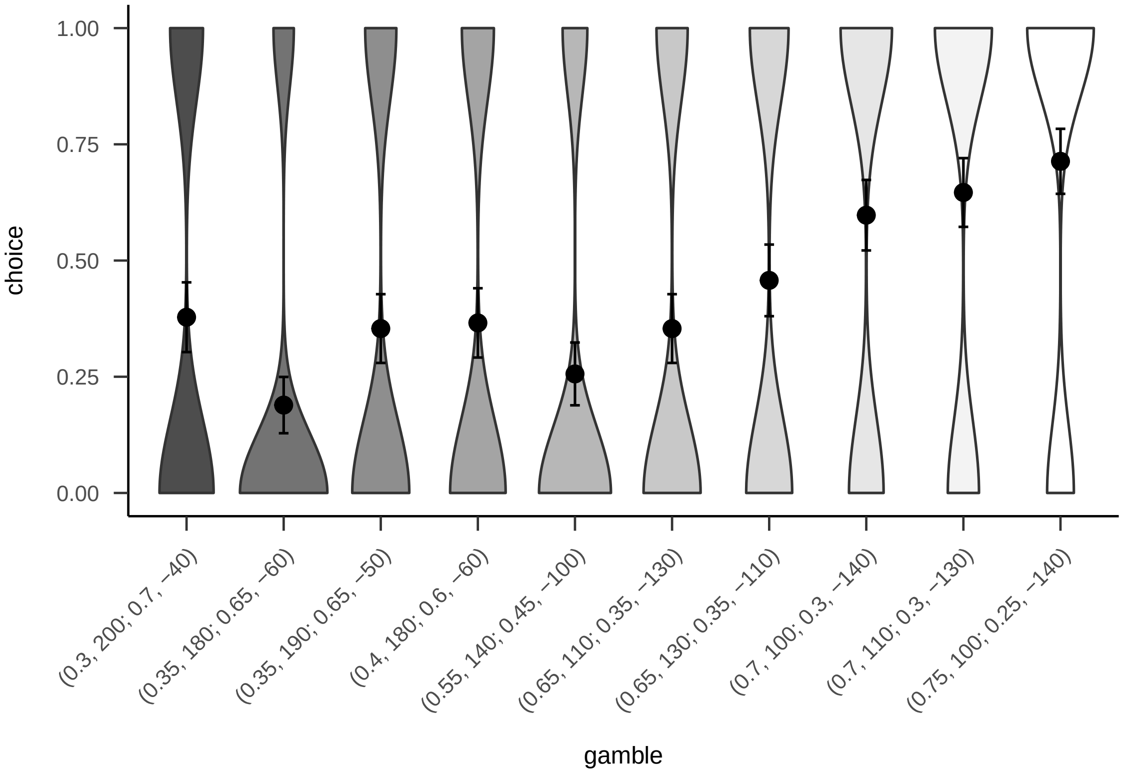


Figure A.10: Mean project acceptance for the 10 gambles. The format of the labels indicates: (gain probability, gain value; loss probability, loss value).

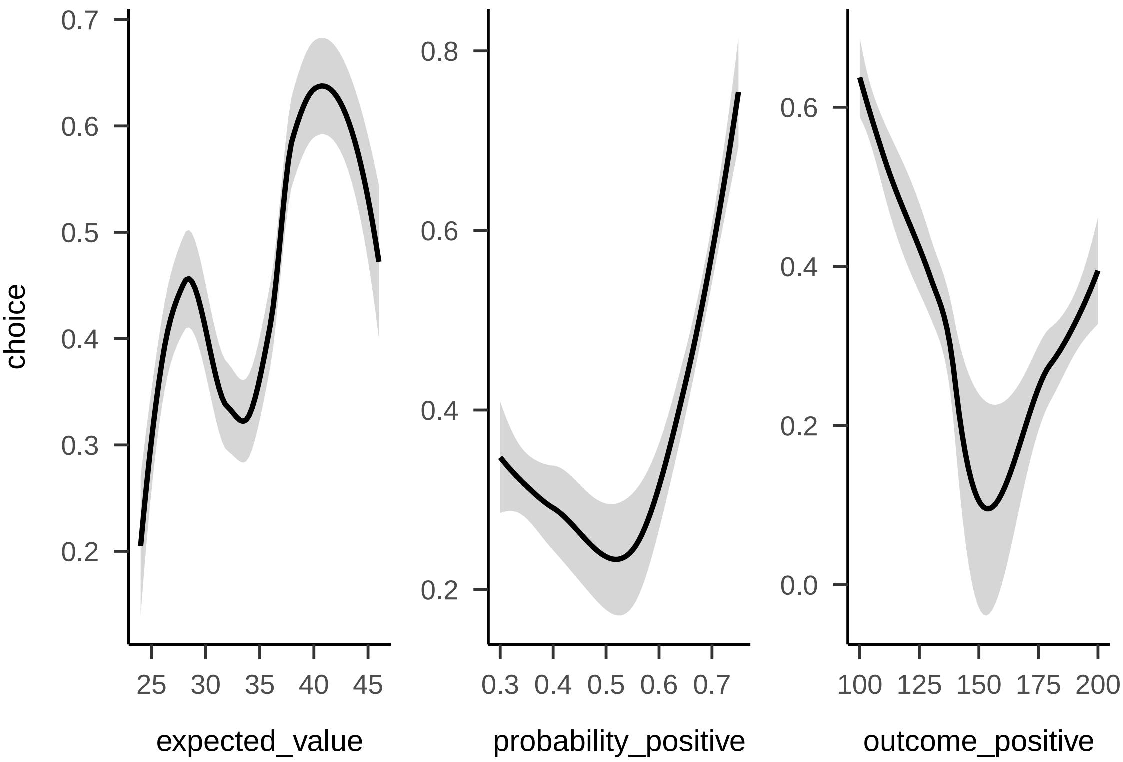


Figure A.11: Mean project acceptance for the gambles' expected value, positive probability, and positive outcome.

Indicate below whether you would invest in the following:

To summarise this investment, there is a 30% chance of gaining $200 million (the forecasted revenue minus the cost amount) and a 70% chance of losing $40 million. The company would make $240 million if the forecasted concentration and quality of recoverable hydrocarbons at the site eventuates. The estimate for the anticipated chance of gain is based on a geological and seismic study of the site, and an analysis of previous similar sites. Refinera is a business in your company that proposes to construct an oil well project. Specifically, they want to establish an exploration site at an onshore location in Houston, US in order to see if the hydrocarbon supply is sufficient to establish a more permanent well. Refinera's research team has been investigating a possible site in an as yet unexplored area. Due to the location and size of the site, and consultant fees (e.g., geologists), they forecast the entire project to cost $40 million (the loss amount).*

Yes

No

Continue

Figure A.12: An example of a project display in Experiment 3.

At the begining of the experiment, before you saw any projects, how many projects did you expect to see?

project(s)

Continue

Figure A.13: Experiment 3 project expectation question.

A.3.1.3 Procedure

Participants read the instructions and completed the risky investment task in their respective conditions. After seeing the individual projects, participants were then asked the four follow-up questions.

A.3.2 Results

A.3.2.1 Project Investment

The project investment data were analysed as in Experiment 2 (see Section 2.3.2). Figures A.14 and A.15 show the choice and proportion data, respectively. The

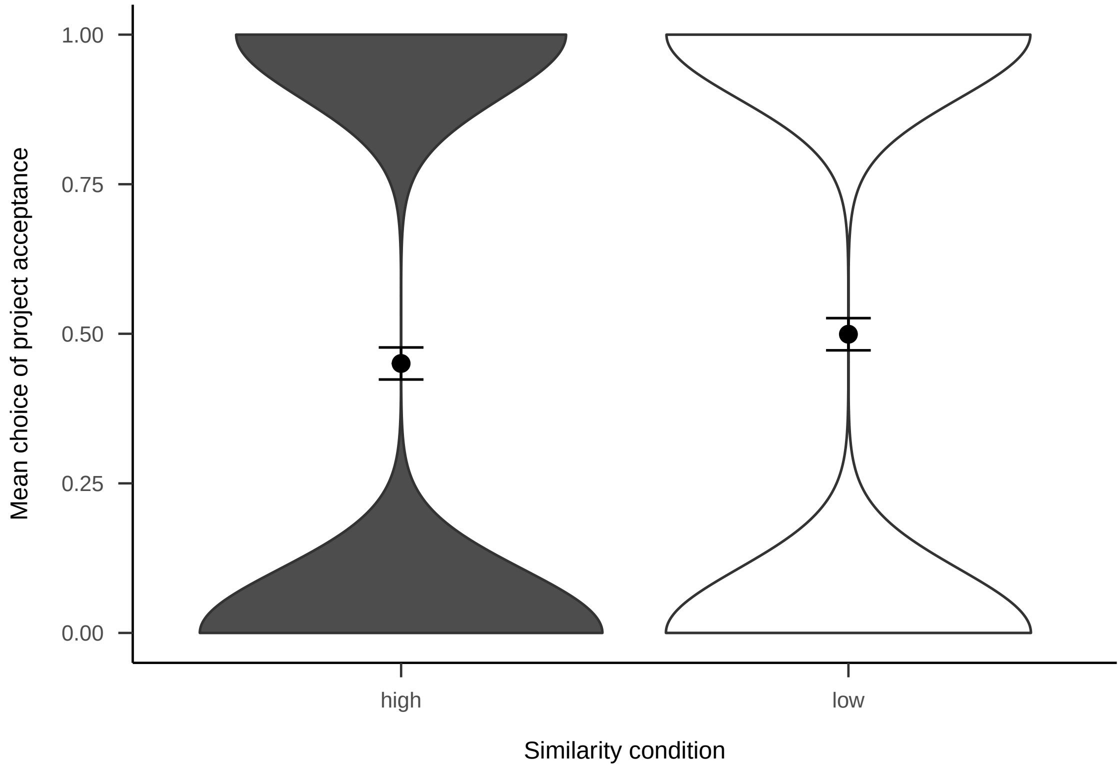


Figure A.14: Mean project acceptance for the similarity effect.

Table A.2: Logistic regression table of project acceptance by similarity and trial.

Term	$\hat{\beta}$	95% CI	z	p
Intercept	0.01	[-0.20, 0.22]	0.07	.944
Similarity1	-0.02	[-0.23, 0.18]	-0.22	.826
Project order	-0.02	[-0.05, 0.01]	-1.52	.127
Similarity1 × Project order	-0.02	[-0.05, 0.01]	-1.07	.284

difference between similarity conditions was not significant, both in the logistic regression $b = 0.00$, 95% CI $[-0.18, 0.17]$, $z = -0.04$, $p = .966$, and in the t-test, d_s = -0.21, 95% CI [-0.45, 0.03], $t(264)$ = -1.69, p = .093.

Further, Figure A.16 shows the choice data as a function of the order of the project in the sequence. As Table A.2 shows, there were no main effects or interactions.

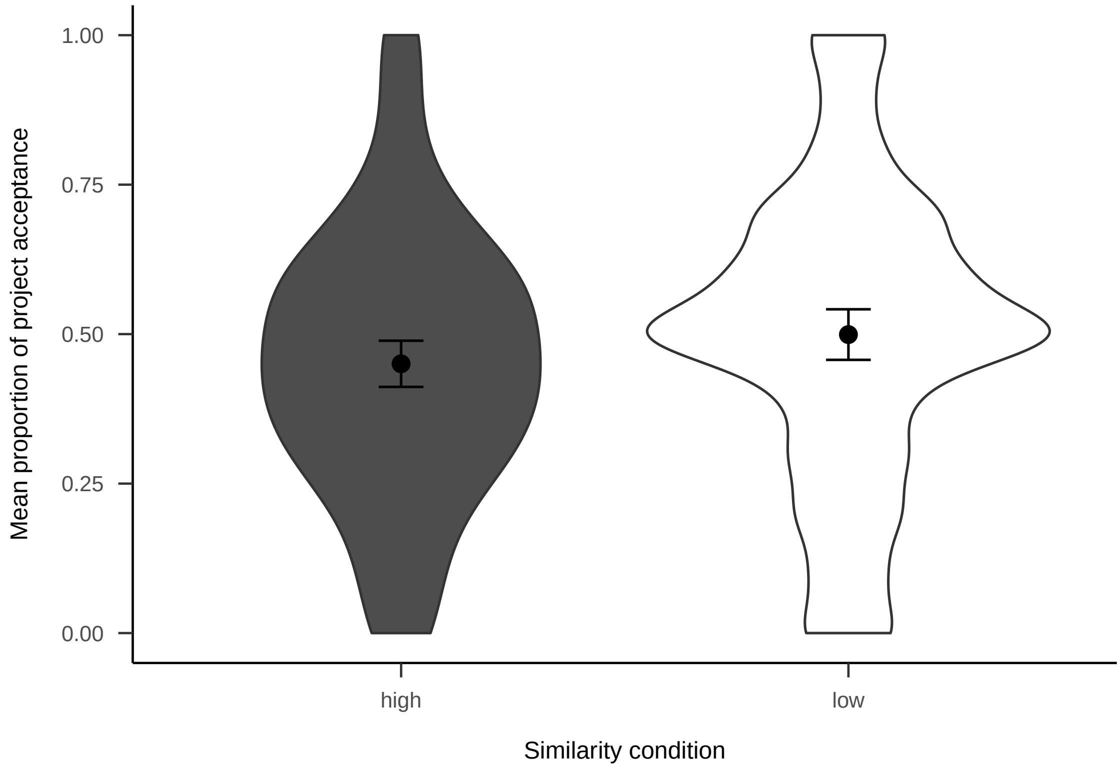


Figure A.15: Mean proportion of project acceptance for the similarity effect.

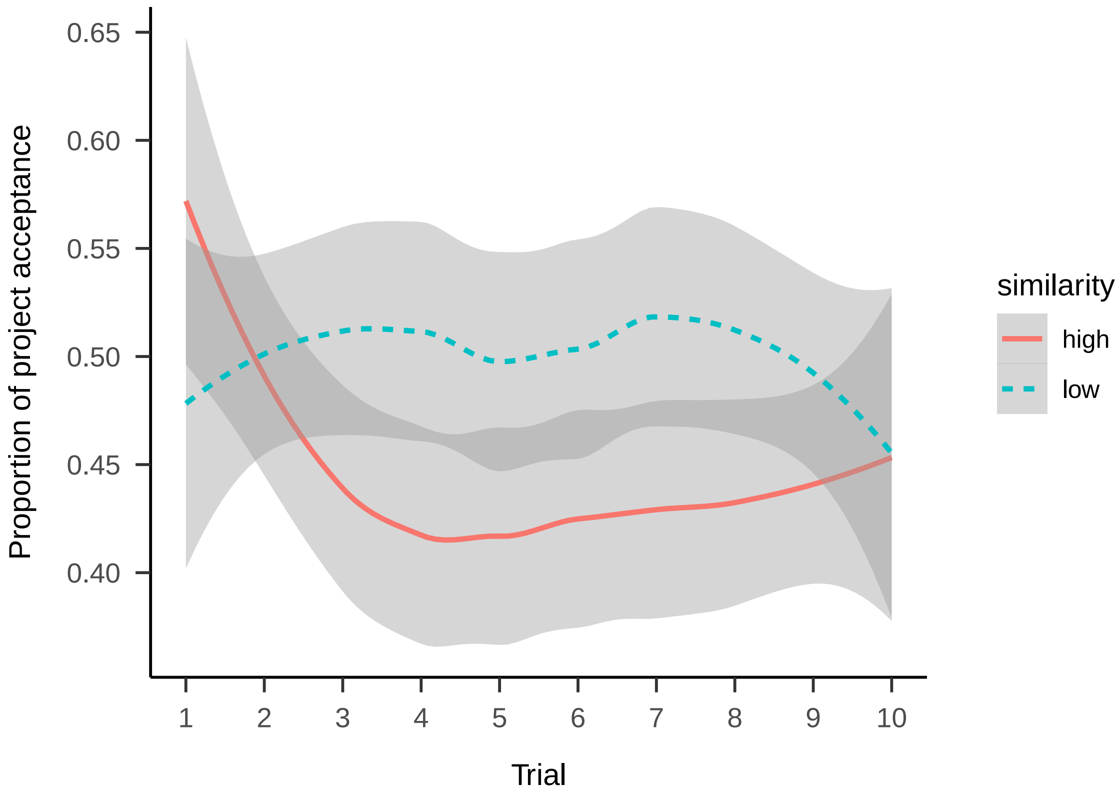


Figure A.16: Mean project acceptance by similarity and trial.

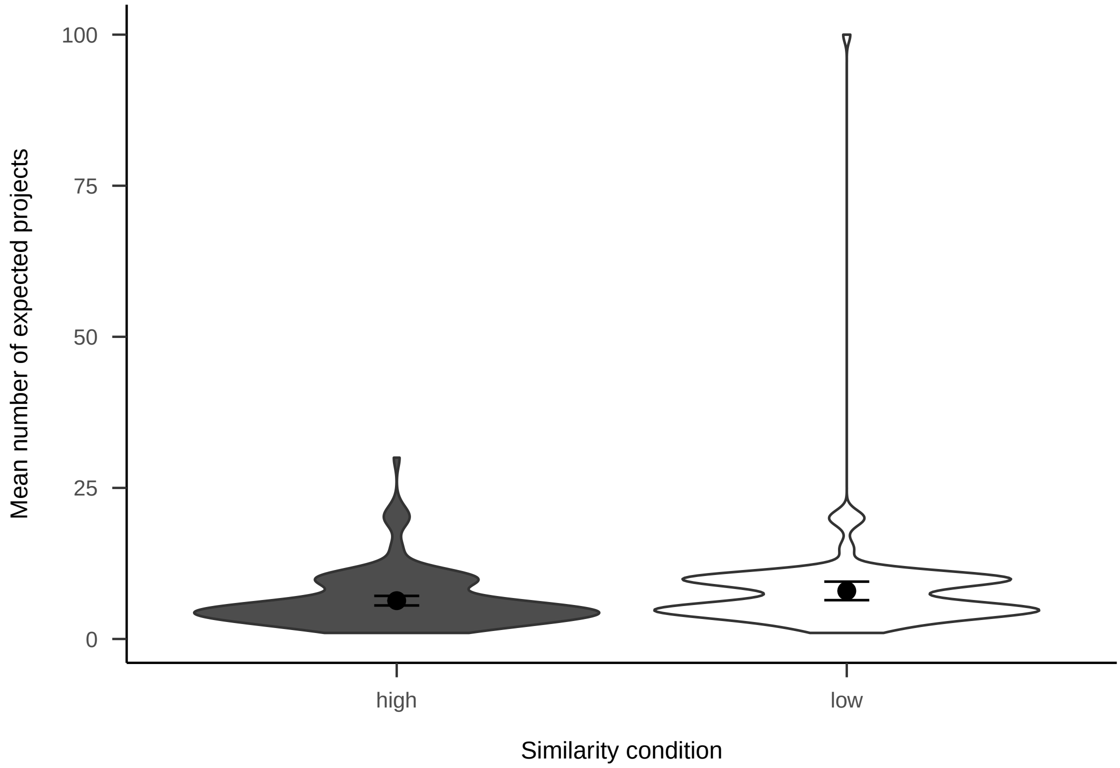


Figure A.17: Number of projects participants expected to see, by similarity.

A.3.2.2 Follow-up

A.3.2.2.1 Project Expectation Participants were asked how many projects they expected to see. As Figure A.17 shows, the difference between similarity conditions was not significant, d_s = -0.23, 95% CI [-0.47, 0.01], $t(264)$ = -1.85, p = .065.

A.3.2.2.2 Project Number Participants were asked how many projects they thought they saw. Figure A.18 shows that overall people correctly estimate the number of projects.

A.3.2.2.3 Portfolio Choice - Binary Participants were then asked if they would rather invest in all or none of the projects. As Figure A.19 shows, those in the low similarity condition were significantly more likely to want to invest in all of the projects, $b = -0.26$, 95% CI $[-0.51, -0.02]$, $z = -2.10$, $p = .036$.

A.3.2.2.4 Portfolio Choice - Number Subsequently, participants were asked how many projects they would invest in out of the 10 that they saw. As Figure A.20

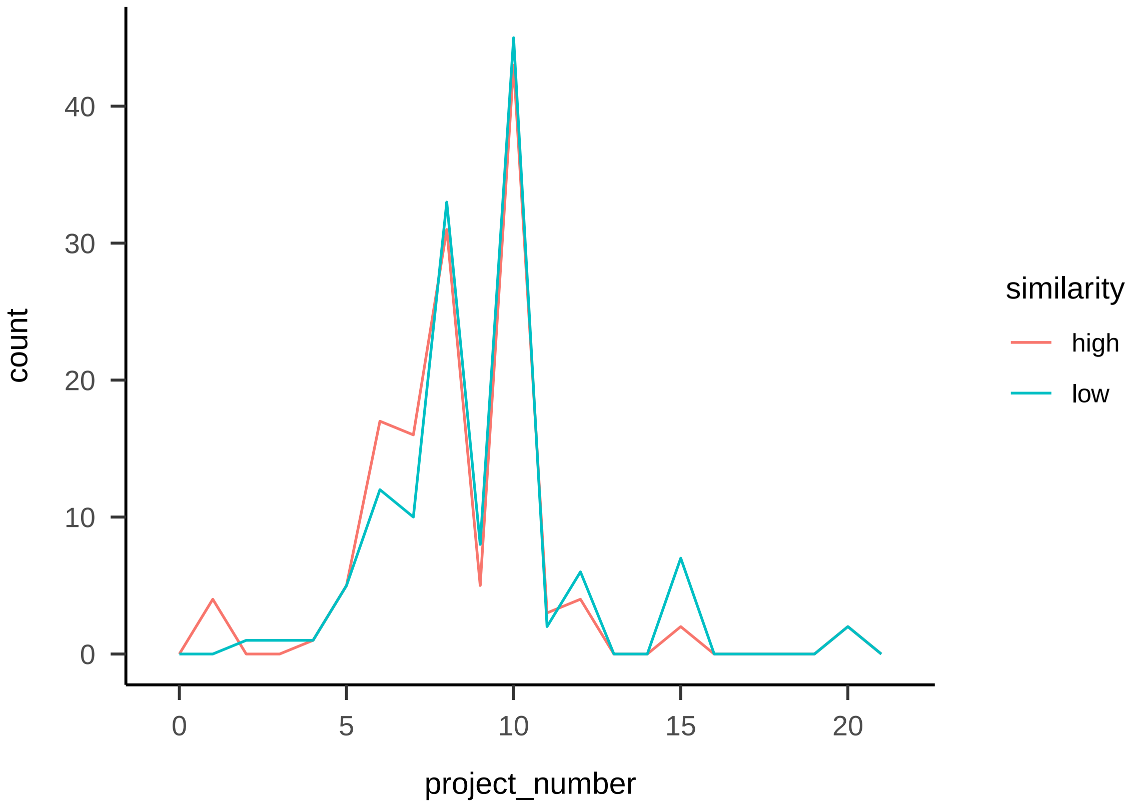


Figure A.18: Number of projects participants reported seeing, by similarity.

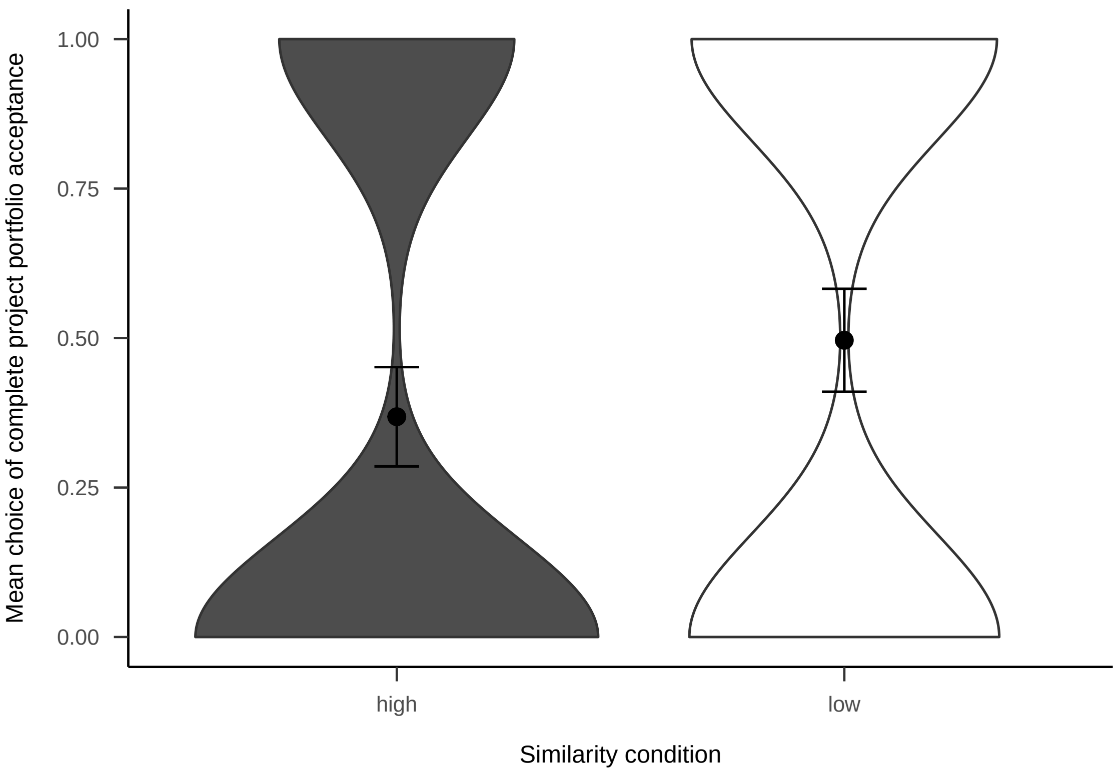


Figure A.19: Mean choice of investing in all 10 projects for the similarity effect.

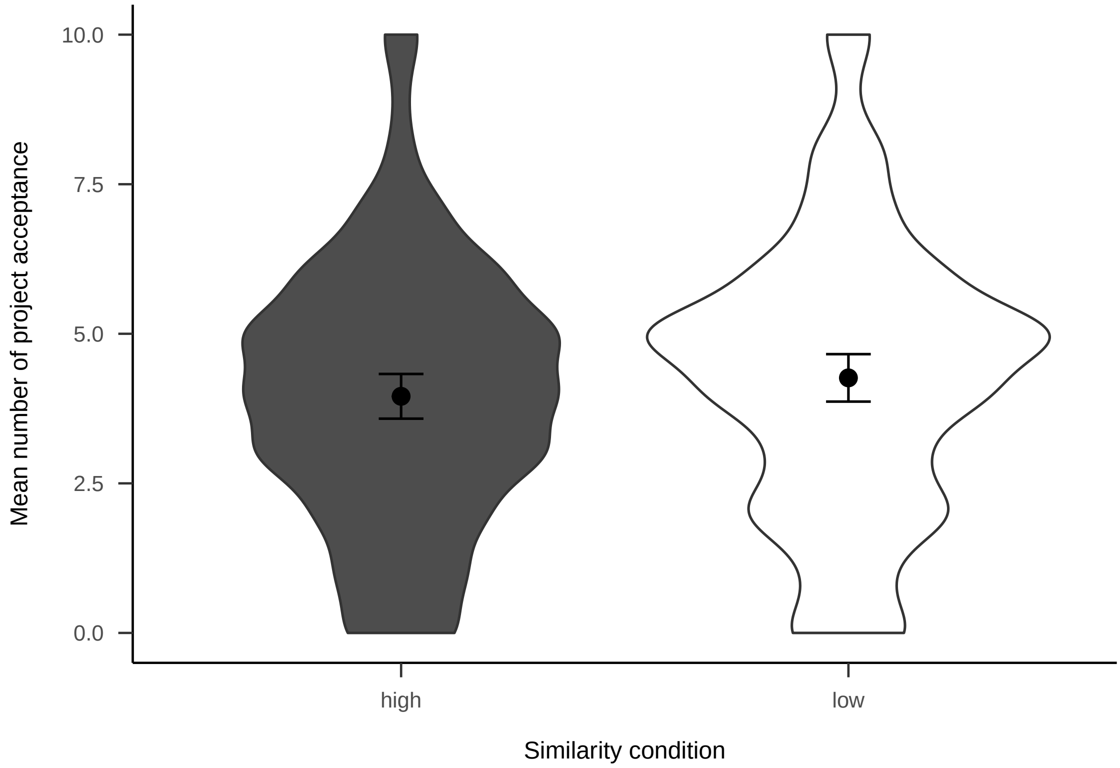


Figure A.20: Mean number of projects chosen in the follow-up for the similarity effect.

shows, the difference between similarity conditions was not significant, d_s = -0.14, 95% CI [-0.38, 0.10], $t(264)$ = -1.12, p = .264.

A.3.2.3 Gambles

Figures A.21 and A.22 show the overall people seemed to prefer gambles with higher probabilities of gain, sometimes regardless of expected value or value of the gain.

A.3.3 Discussion

Experiment 3 found some evidence for the effect of similarity on project choice, but it was in the opposite direction to the one hypothesised. Specifically, the results showed that when considering projects individually, participants' risk aversion did not differ between similarity conditions, but when offered a portfolio of the projects, those that saw the dissimilar projects were more likely to invest.

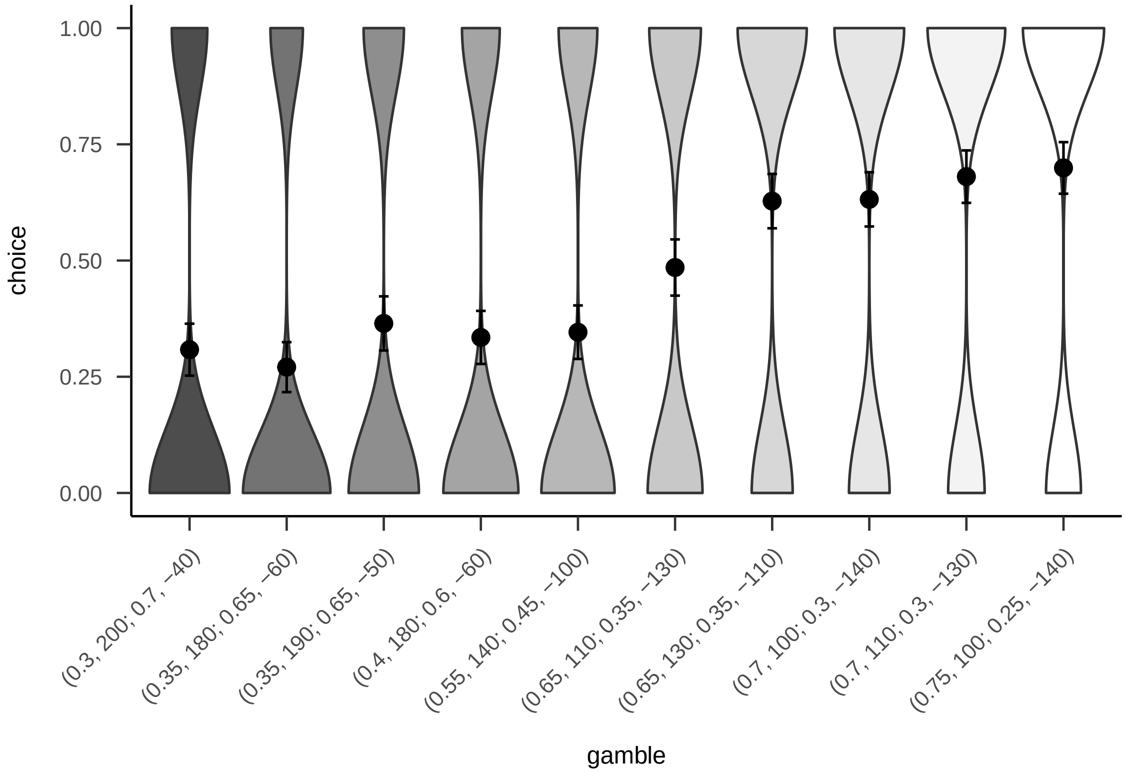


Figure A.21: Mean project acceptance for the 10 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value).

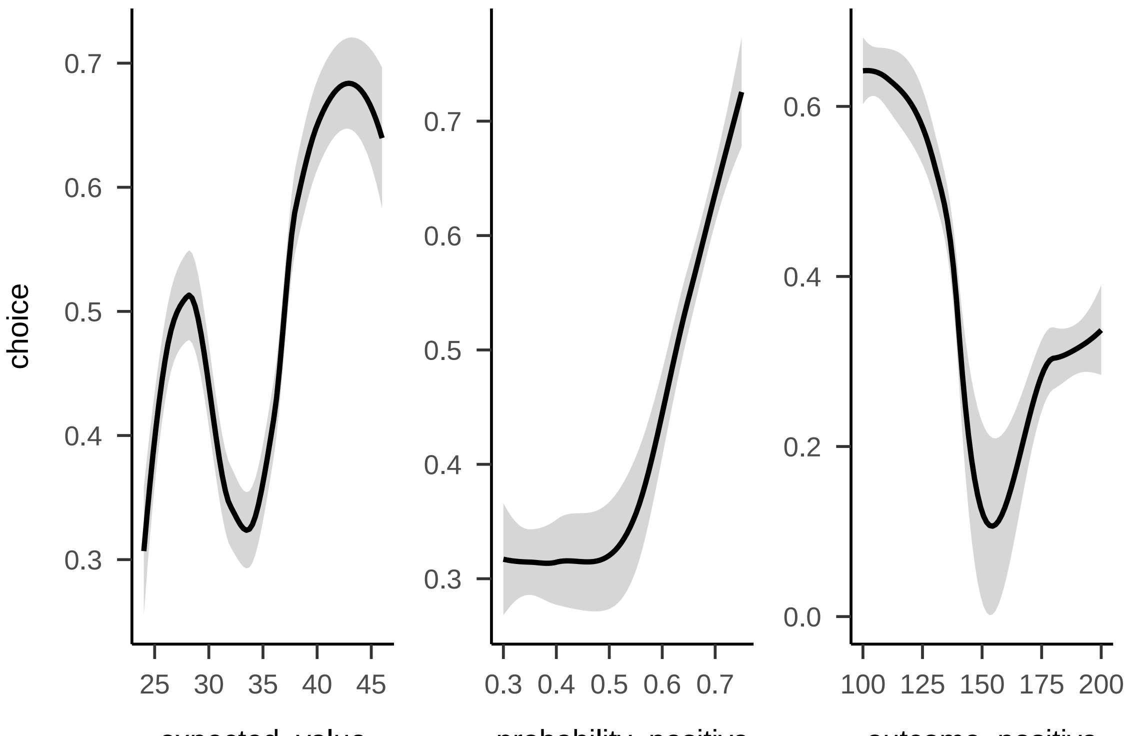


Figure A.22: Mean project acceptance for the gambles' expected value, positive probability, and positive outcome.

These results provide evidence for the naive diversification account expressed above (see Section 2.2.3.3). Specifically, participants may really be naively diversifying, but only when they are explicitly given an opportunity to do so. This is similar to the multi-play effects because the question itself provides a sort of choice bracketing. That is, the gambles are grouped together as a portfolio by the question. Together, this suggests that people are not naively aggregating when viewing gambles in isolation, but when the choices are bracketed explicitly, then the choice seems to be driven by a naive diversification.

A.4 Experiment 4

Experiment 4 investigated the effect of awareness on project choice. Experiment 1 found an effect of awareness in the trial-by-trial data that was not replicated in Experiment 2. Above, this effect was explained through the law of small numbers: people may have been anticipating less risky gambles towards the end of the set. As such, the effect could be seen with more trials. Experiment 4 attempted to replicate the effect from Experiment 1 with 20 projects. The *naive* condition attempted to encourage participants to focus on projects one at a time and did not reveal the total number of projects. The *aware* condition attempted to encourage participants to think of all 20 projects. This was done by revealing the total number of projects in the beginning of the task and by identifying at each project display its order in the sequence. Experiment 4 again tested Hypothesis 2.4.

A.4.1 Method

A.4.1.1 Participants

Two hundred and sixty-six participants (110 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 40.62 years (*SD* = 9.59, *min.* = 25, *max.* = 74). Participants reported an average of 7.45 years (*SD* = 7.8, *min.* = 0, *max.* = 47) working in a business setting, and an average of 5.52 years (*SD* = 7.27, *min.* = 0, *max.* = 48) of business education. The mean

Table A.3: Experiment 4 group allocation.

Awareness	N
Aware	133
Naive	133
Total	266

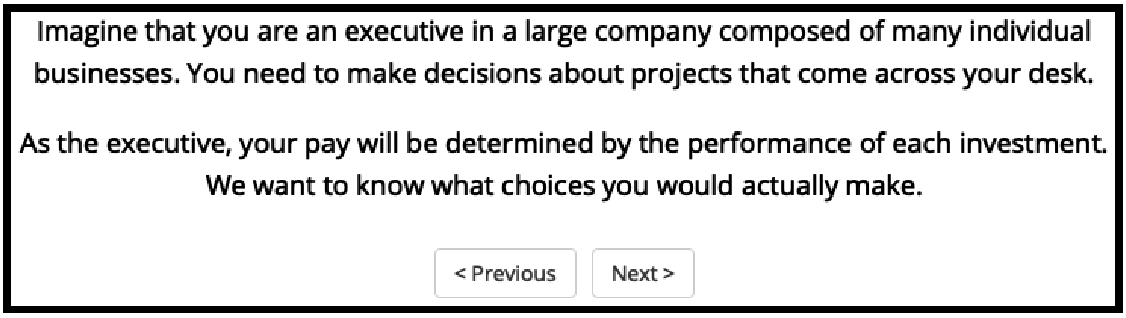


Figure A.23: Instructions for those in the naive condition of Experiment 4.

completion time of the task was 12.66 min (SD = 8.26, $min.$ = 1.48, $max.$ = 53.47). Table A.3 shows the allocation of participants to the different conditions.

A.4.1.2 Materials

A.4.1.2.1 Instructions Participants were shown similar instructions to Experiment 1 (see Section 2.2.1.2.1), except that the awareness manipulation was incorporated into the text. Participants in the naive condition saw the instructions in Figure A.23, and those in the aware condition saw the instructions in Figure A.24.

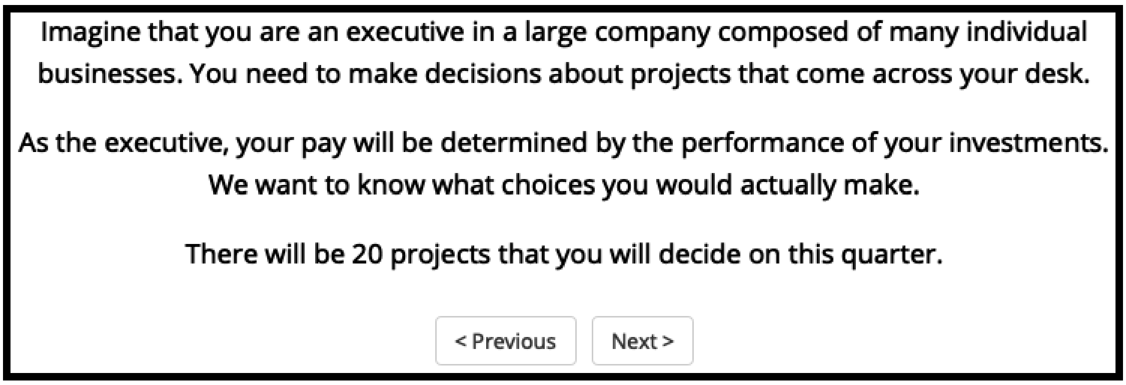
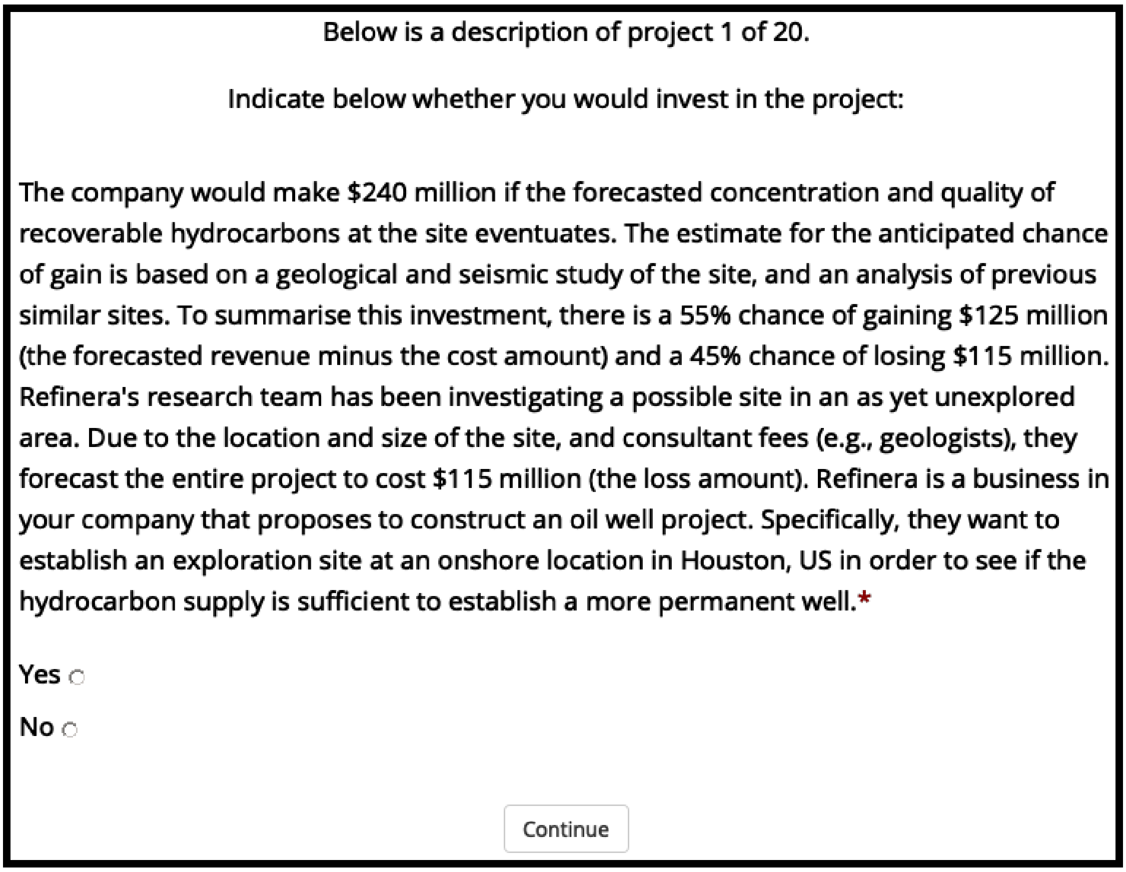


Figure A.24: Instructions for those in the aware condition of Experiment 4.



Below is a description of project 1 of 20.

Indicate below whether you would invest in the project:

The company would make $240 million if the forecasted concentration and quality of recoverable hydrocarbons at the site eventuates. The estimate for the anticipated chance of gain is based on a geological and seismic study of the site, and an analysis of previous similar sites. To summarise this investment, there is a 55% chance of gaining $125 million (the forecasted revenue minus the cost amount) and a 45% chance of losing $115 million. Refinera's research team has been investigating a possible site in an as yet unexplored area. Due to the location and size of the site, and consultant fees (e.g., geologists), they forecast the entire project to cost $115 million (the loss amount). Refinera is a business in your company that proposes to construct an oil well project. Specifically, they want to establish an exploration site at an onshore location in Houston, US in order to see if the hydrocarbon supply is sufficient to establish a more permanent well.*

Yes

No

Continue

Figure A.25: An example of a project display in Experiment 4.

A.4.1.2.2 Risky Investment Task Participants saw similar displays to those in Experiment 3 (see Section A.3.1.2.2). However, here participants viewed 20 projects, so while the gamble constrains explained above were still applied, the actual gamble values were different. Further, those in the aware condition saw an added sentence that identified the number of the project they were currently considering in the context of the total 20. See Figure A.25 for an example. Those in the naive condition saw the same display without this sentence.

A.4.1.2.3 Follow-up The follow-up questions were identical to those in Experiment 3 (see Section A.3.1.2.3), except that the portfolio number question identified the number of projects they saw as 20.

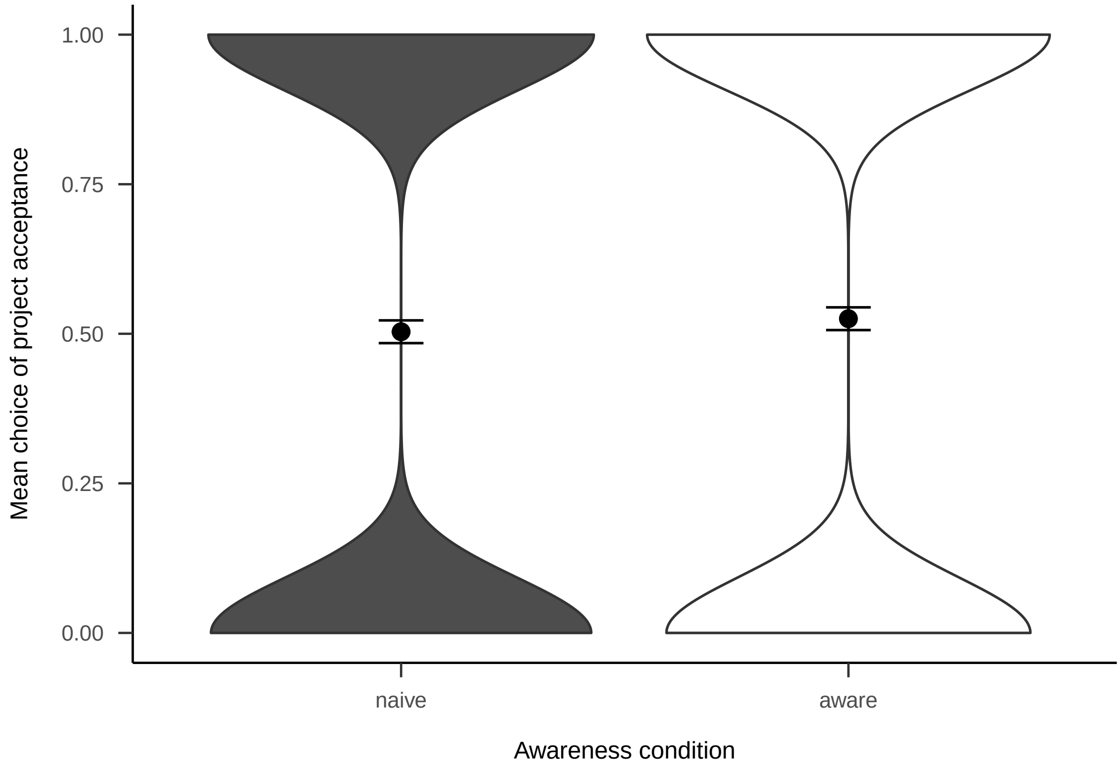


Figure A.26: Mean project acceptance for the awareness effect.

A.4.1.3 Procedure

Participants read the instructions and completed the risky investment task in their respective conditions. After seeing the individual projects, participants were then asked the four follow-up questions.

A.4.2 Results

A.4.2.1 Project Investment

The project investment data were analysed as in Experiment 2 (see Section 2.3.2). Figures A.26 and A.27 show the choice and proportion data, respectively. The difference between awareness conditions was not significant, both in the logistic regression $b = -0.05$, 95% CI $[-0.22, 0.13]$, $z = -0.53$, $p = .595$, and in the t-test, d_s = -0.09, 95% CI [-0.33, 0.15], $t(264)$ = -0.73, p = .464.

Further, Figure A.28 shows the choice data as a function of the order of the project in the sequence. As Table A.4 shows, there were no main effects or interactions.

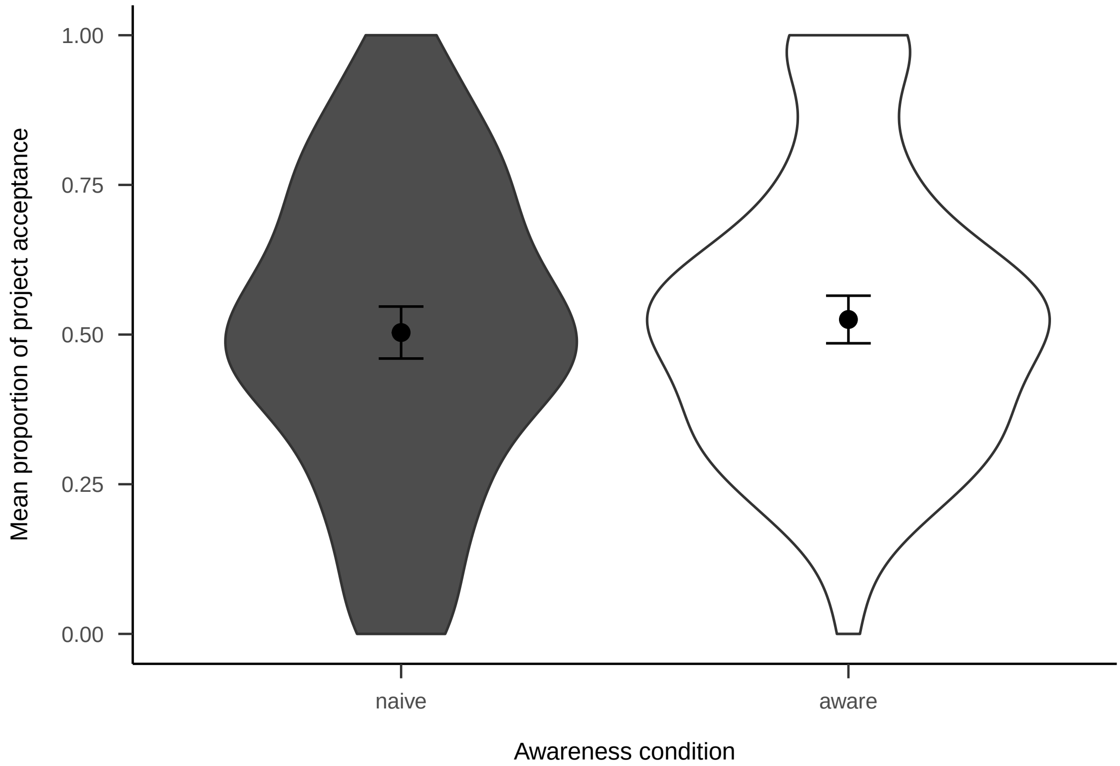


Figure A.27: Mean proportion of project acceptance for the awareness effect.

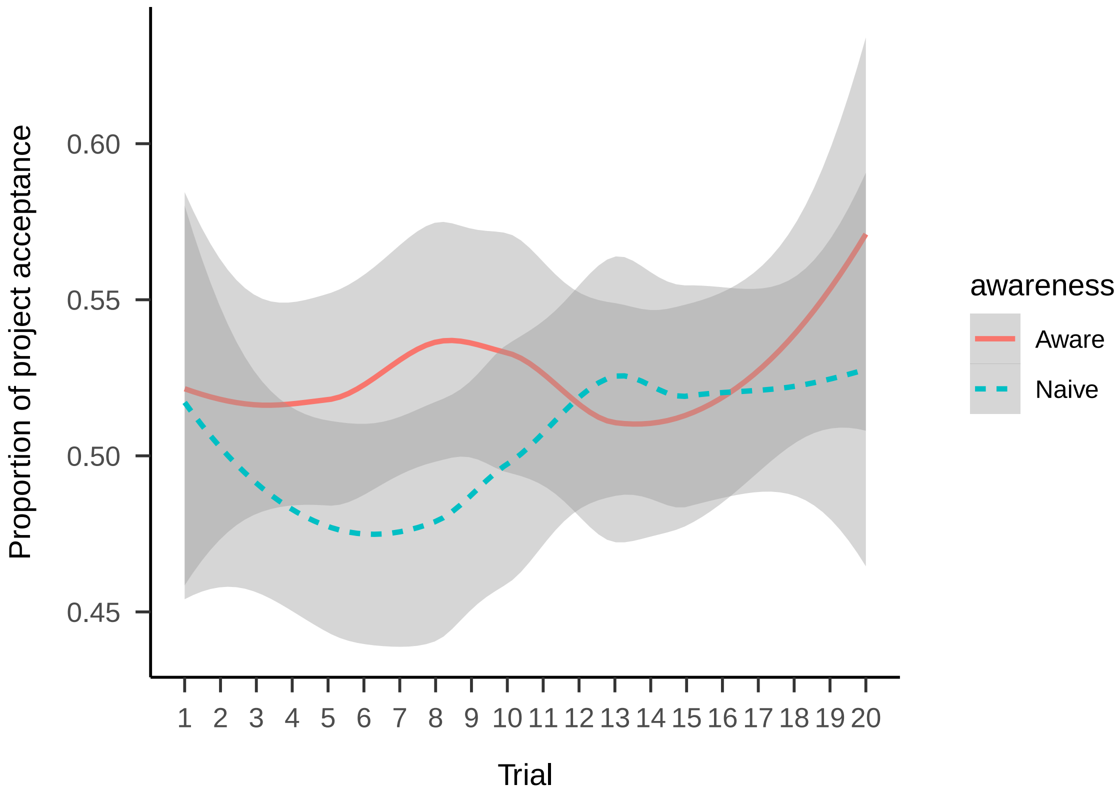


Figure A.28: Mean project acceptance by awareness and trial.

Table A.4: Logistic regression table of project acceptance by awareness and trial.

Term	$\hat{\beta}$	95% CI	z	p
Intercept	-0.01	[-0.20, 0.17]	-0.12	.907
Awareness1	-0.10	[-0.28, 0.09]	-1.05	.293
Project order	0.01	[0.00, 0.02]	1.66	.096
Awareness1 × Project order	0.00	[-0.01, 0.01]	0.29	.775

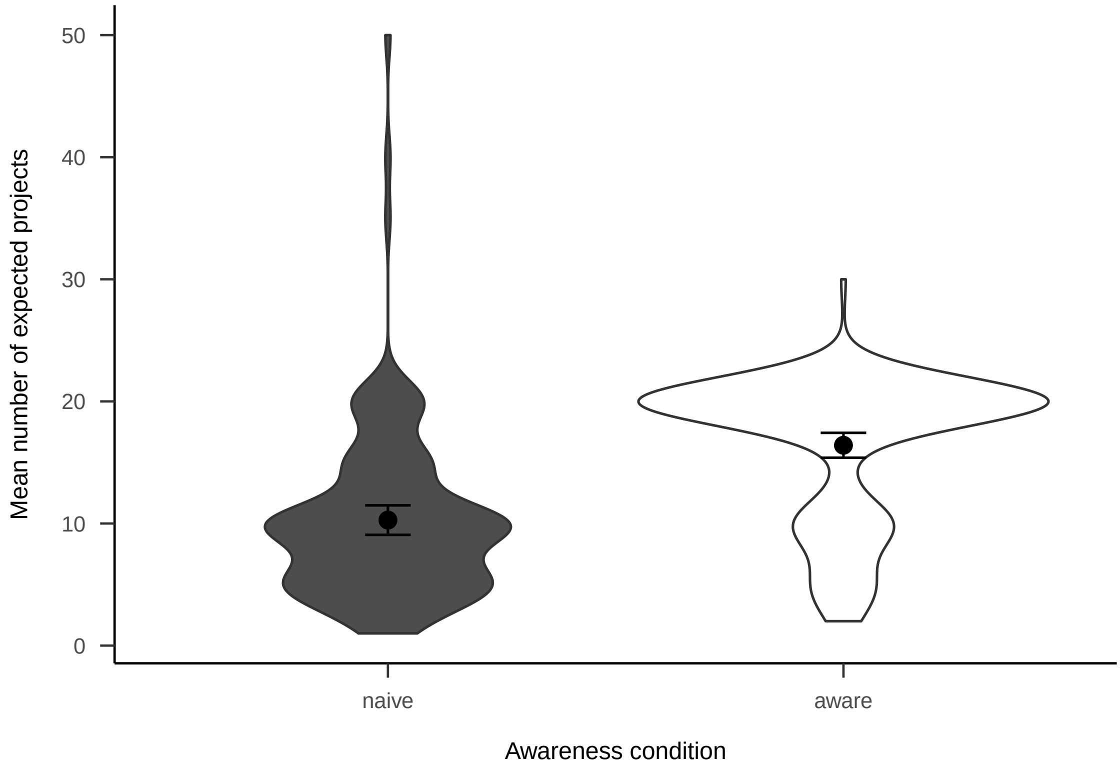


Figure A.29: Number of projects participants expected to see, by awareness.

A.4.2.2 Follow-up

A.4.2.2.1 Project Expectation Participants were asked how many projects they expected to see. Figure A.29 shows that those in the aware condition reportedly expect to see more, d_s = -0.94, 95% CI [-1.19, -0.69], $t(264)$ = -7.67, p < .001. However, this is likely to be due to the fact that they were told how many projects there were.

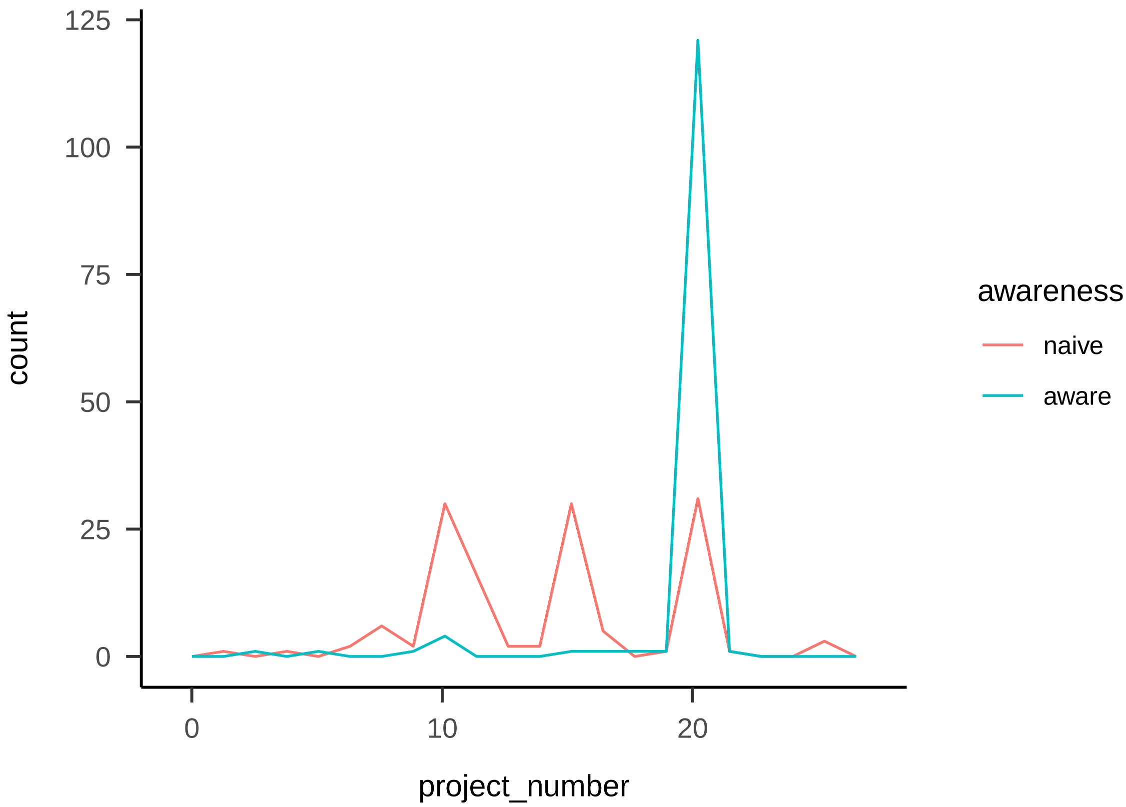


Figure A.30: Number of projects participants reported seeing, by awareness.

A.4.2.2.2 Project Number Participants were asked how many projects they thought they saw. Figure A.30 shows that overall people correctly estimated the number of projects, with higher accuracy for those in the aware condition.

A.4.2.2.3 Portfolio Choice - Binary Participants were then asked if they would rather invest in all or none of the projects. As Figure A.31, there was no significant difference between awareness conditions in wanting to invest in all of the projects, $b = -0.09$, 95% CI $[-0.33, 0.15]$, $z = -0.74$, $p = .460$.

A.4.2.2.4 Portfolio Choice - Number Subsequently, we asked participants how many projects they would invest in out of the 20 that they saw. As Figure A.32 shows, the difference between awareness conditions was not significant, d_s = -0.12, 95% CI [-0.36, 0.12], $t(264)$ = -0.97, p = .334.

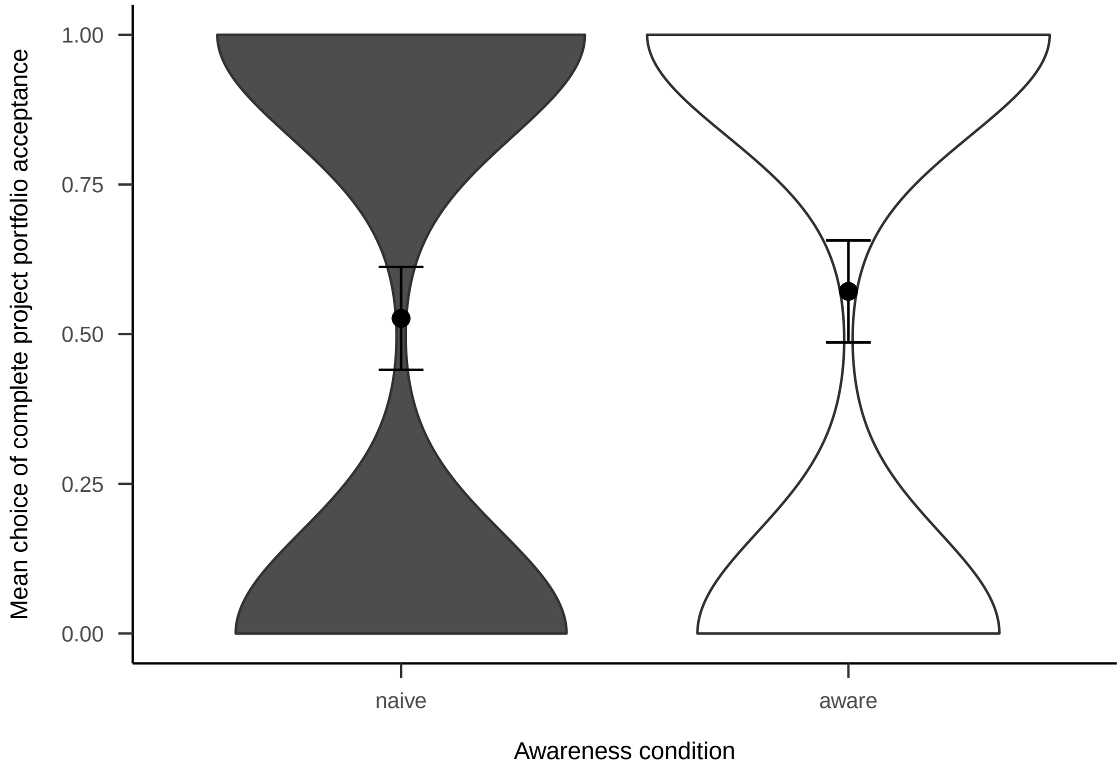


Figure A.31: Mean choice of investing in all 20 projects for the awareness effect.

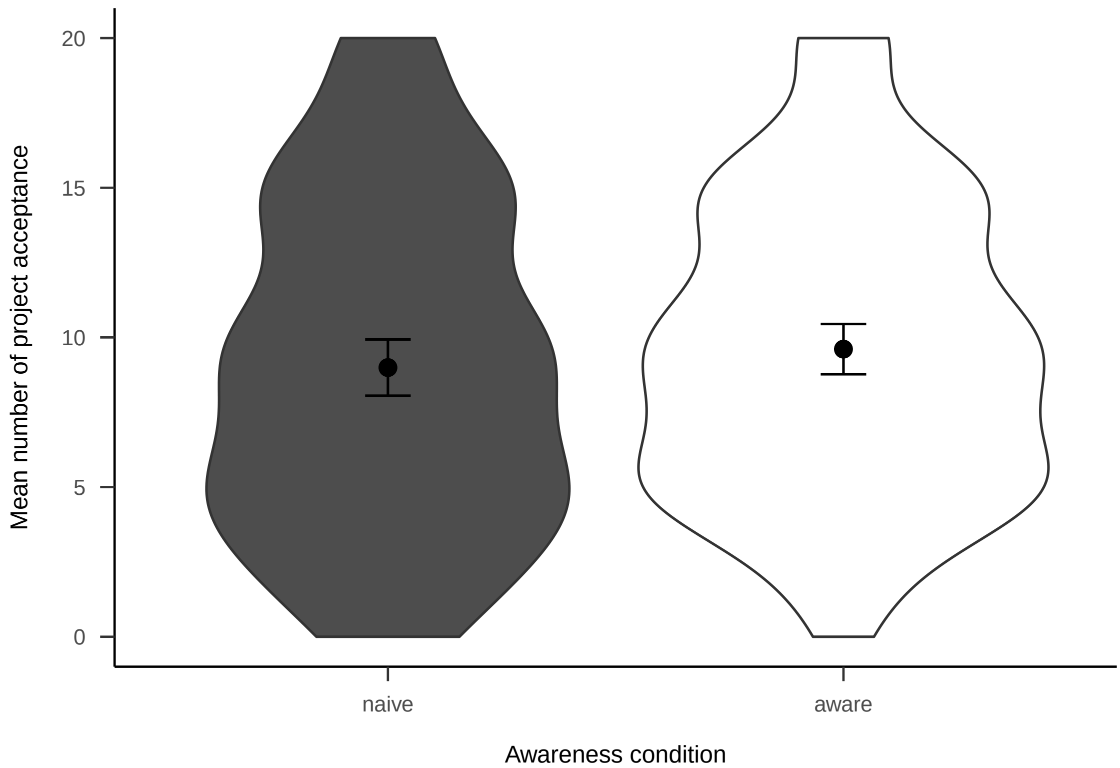


Figure A.32: Mean number of projects chosen in the follow-up for the awareness effect.

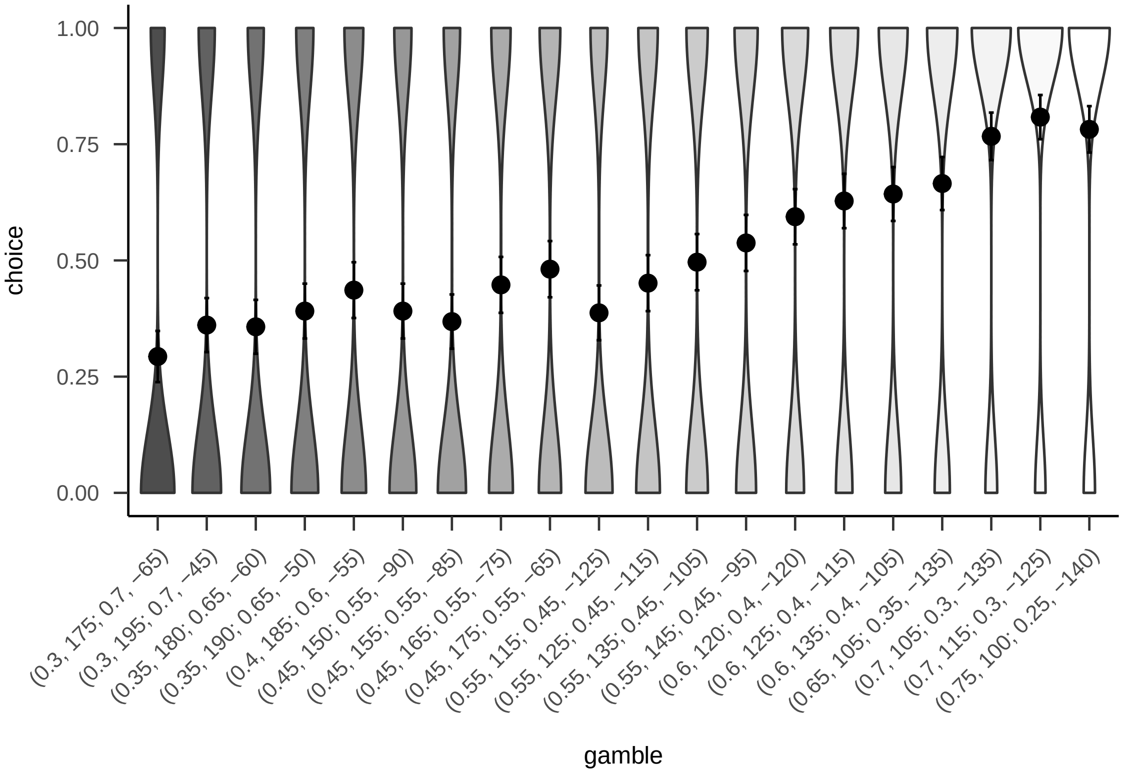


Figure A.33: Mean project acceptance for the 20 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value).

A.4.2.3 Gambles

Figures A.33 and A.34 show the overall people seemed to prefer gambles with higher probabilities of gain, sometimes regardless of expected value or value of the gain.

A.4.3 Discussion

Experiment 4 did not find evidence for Hypothesis 2.4. There was no significant effect of awareness on project choice by trial. Participants in the aware condition were expected to become less risk averse as they continued with the experiment if they were using a strategy similar to the law of small numbers. The fact that this effect was not replicated in Experiment 4 might mean that the finding in Experiment 1 was due to the specific gambles used in that experiment, or statistical chance.

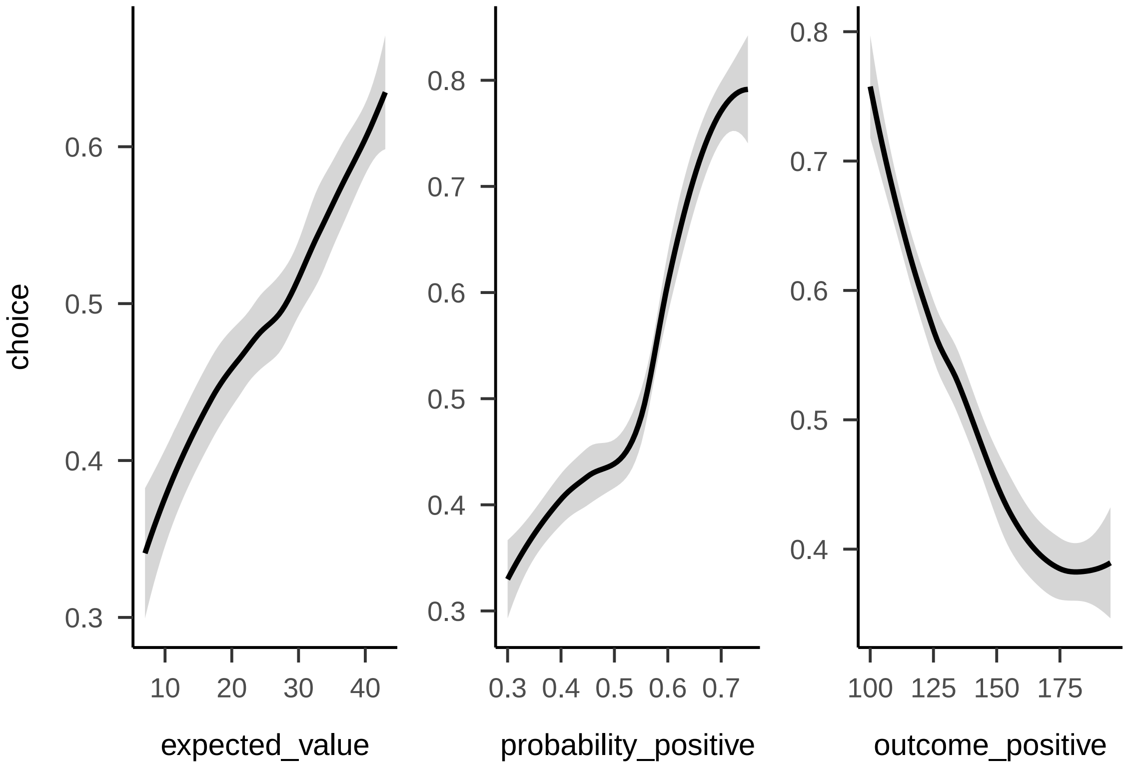


Figure A.34: Mean project acceptance for the gambles' expected value, positive probability, and positive outcome.

B
Chapter 4 Appendix

Contents

This appendix contains supplementary materials and analyses for the three experiments reported in Chapter 4. In addition, five related experiments are reported. Experiment 4 was identical to Experiment 1, except that alignment was manipulated within-subjects, it did not include a no NPV condition, and there was no forecasting measure. Experiment 5 replicated Experiment 1, but only tested the forecasting effect and did so with a sample that had investing experience. Experiment 6 replicated Experiment 5 but with a larger sample size and a lay sample. Experiment 7 attempted to facilitate a use of numerical reliability through explicit hints. Experiment 8 tested both verbal and numerical reliability effects in an all within-subjects design. However, unlike Experiment 3, the design of Experiment 8 did not allow for a direct comparison of alignment conditions.

B.1 Experiment 1

In addition to the allocation measure, participants were also asked to rank the projects and forecast their future returns. The ranking task was included before the allocation task in order to encourage alignment and to have another measure of participants' decision-making. The forecasting task was added (described further below in Section B.1.1.1.2) in order to test whether the variance in people's forecasts is affected by alignment and NPV reliability.

Hypothesis B.1. All allocation effects will replicate in the ranking measure.

Hypothesis B.2. All allocation effects will replicate in the forecasting mean measure.

In the forecasting measures, more alignable differences were expected to bring about more certainty about forecasting decisions, since participants will have more easily comparable information. As such, people's forecasting should be less variable when comparing projects with alignable differences, than when comparing projects with non-alignable differences.

Hypothesis B.3. The standard deviation of participants' forecasts will be higher, on average, in the low alignment condition than in the high alignment condition.

B.1.1 Method

B.1.1.1 Materials

B.1.1.1.1 Instructions Figures B.1, B.2, and B.3 show the instructions given to those in the low NPV reliability, high NPV reliability, and no NPV condition, respectively.

B.1.1.1.2 Forecasting Participants were asked to respond to a forecasting task (adapted from Long et al., 2018), seen in Figure B.4. Participants were asked to predict each project's rate of return after one month. This allowed to calculate each participant's forecasting mean and standard deviation (the latter as inversely proportional to forecasting precision).

B.1.1.1.3 Ranking As shown in Figure B.5, participants were asked to rank the projects in order of investment priority.

B.1.1.1.4 Confidence As Figure B.6 shows, participants were asked to indicate how confident they were about each of their allocation decisions on a scale from 0 ("Not confident at all") to 100 ("Extremely confident").

B.1.1.1.5 Justification As Figure B.7 shows, participants were asked to justify their allocation decision in a free-response text-box.

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided. In addition to those numbers, you will find each project's net present value (NPV), which is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. **The higher the NPV, the better the expectations for each project.** However, it is important to note that NPV is a very noisy measure relative to the other more specific measures because it relies on future forecasting. As such, **NPV is very unreliable and should be relied upon only as a last result; the specific project's measures should be used instead.**

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest objective value**. The features of the products that are listed matter because they reflect the direct value of the product, whereas financial measures such as NPV may reflect other factors, thus making it noisier, as mentioned above.

You will see a set of five different projects for which you must predict the investment returns after one month. For example, how likely is it that the project will return more than 9% after one month, how likely is it that the project will return 7% to 9%, etc.

You will also decide how to rank the projects in order of investment priority, and decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects.

Figure B.1: Experiment 1 low reliability instructions.

B.1.2 Results

B.1.2.1 Ranking

A mixed factorial ANOVA was conducted to investigate the effects of alignment and verbally-instructed NPV reliability on participants' rankings of the target project. As shown in Figure B.8, the alignment × reliability level × NPV interaction was significant, $F(6.62, 370.54) = 2.70$, $p = .011$, $\hat{\eta}_p^2 = .046$. This effect seems to be driven by the differences between the no NPV condition and the conditions with NPV across the two alignment conditions. Specifically, in

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided. In addition to those numbers, you will find each project's net present value (NPV), which is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. **The higher the NPV, the better the expectations for each project.** However, it is important to note that NPV is a very noisy measure relative to the other more specific measures because it relies on future forecasting. As such, NPV is very unreliable and should be relied upon only as a last result; the specific project's measures should be used instead.NPV is a very useful measure relative to the other more specific measures because it can be calculated regardless of the type of product. As such, **NPV is very reliable in most cases.**

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest objective value.**

You will see a set of five different projects for which you must predict the investment returns after one month. For example, how likely is it that the project will return more than 9% after one month, how likely is it that the project will return 7% to 9%, etc.

You will also decide how to rank the projects in order of investment priority, and decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects.

Figure B.2: Experiment 1 high reliability instructions.

the low alignment condition, the linear NPV trend was significantly lower in the no NPV condition than both the low reliability condition, $M = -6.56$, 95% CI $[-10.26, -2.85]$, $t(112) = -3.50$, $p = .001$, and the high reliability condition, $M = -7.38$, 95% CI $[-10.83, -3.93]$, $t(112) = -4.24$, $p < .001$. However, in the high alignment condition, the linear NPV trend was only significantly lower in the no NPV condition than the high reliability condition, $M = -8.37$, 95% CI $[-11.85, -4.88]$, $t(112) = -4.76$, $p < .001$, and not the low reliability condition, $M = -1.71$, 95% CI $[-5.54, 2.13]$, $t(112) = -0.88$, $p = .380$.

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided.

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest objective value**. The features of the products that are listed matter because they reflect the direct value of the product, whereas financial measures may reflect other factors.

You will see a set of five different projects for which you must decide how to rank in order of investment priority, and decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects.

Figure B.3: The instructions for the no NPV condition in Experiment 1.

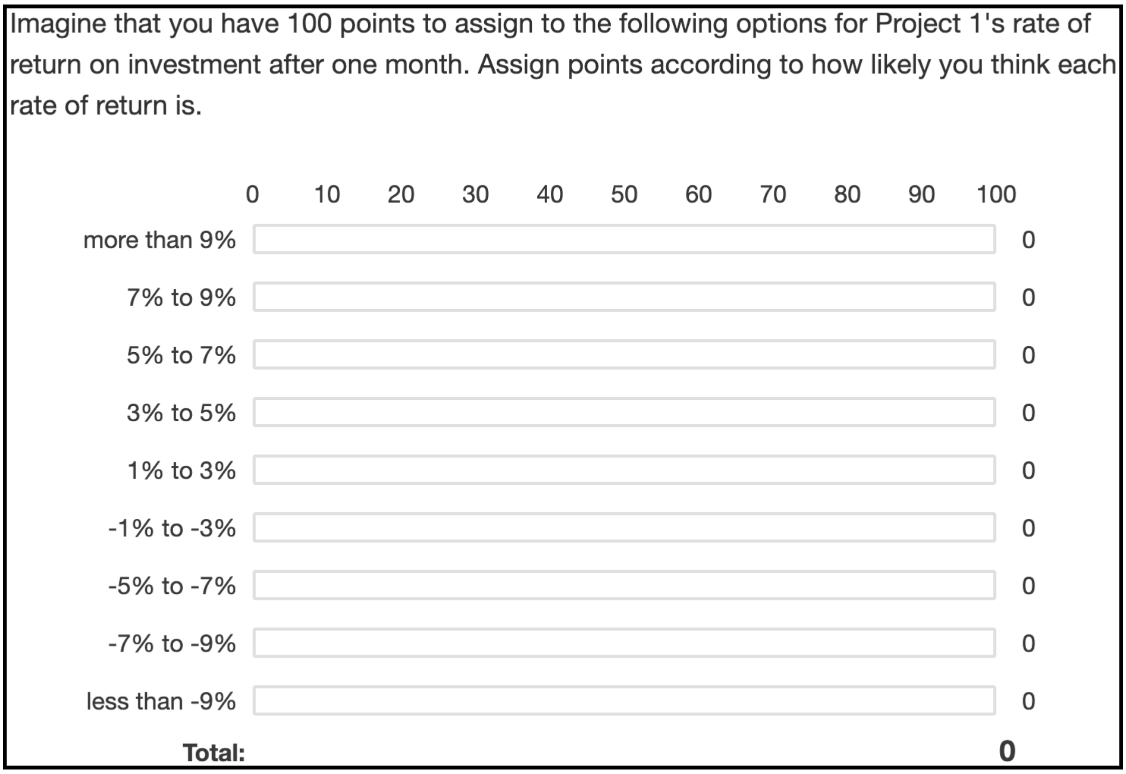


Figure B.4: The forecasting task.

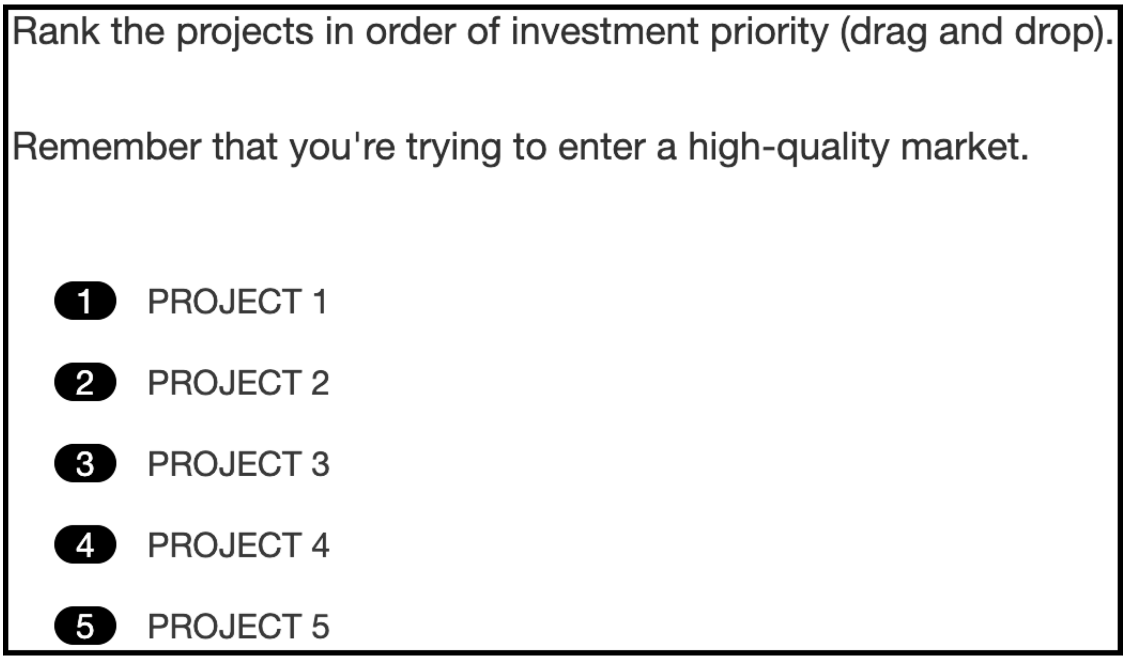


Figure B.5: The ranking task.

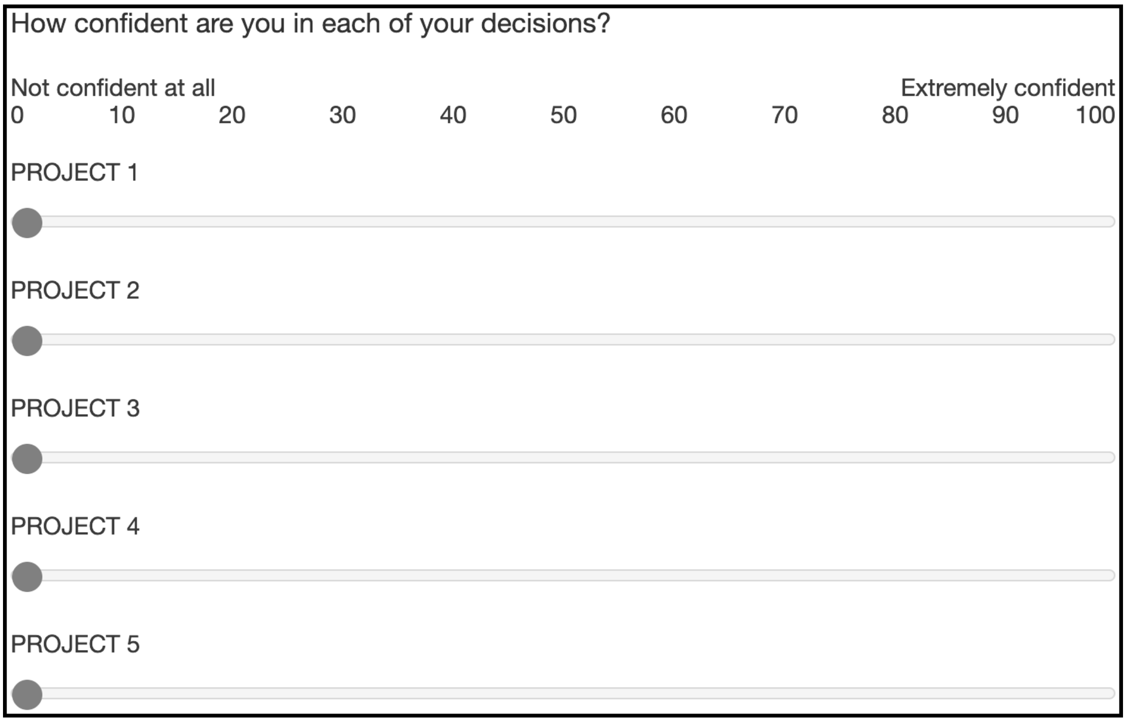


Figure B.6: The confidence task.

Figure B.7: The justification task.

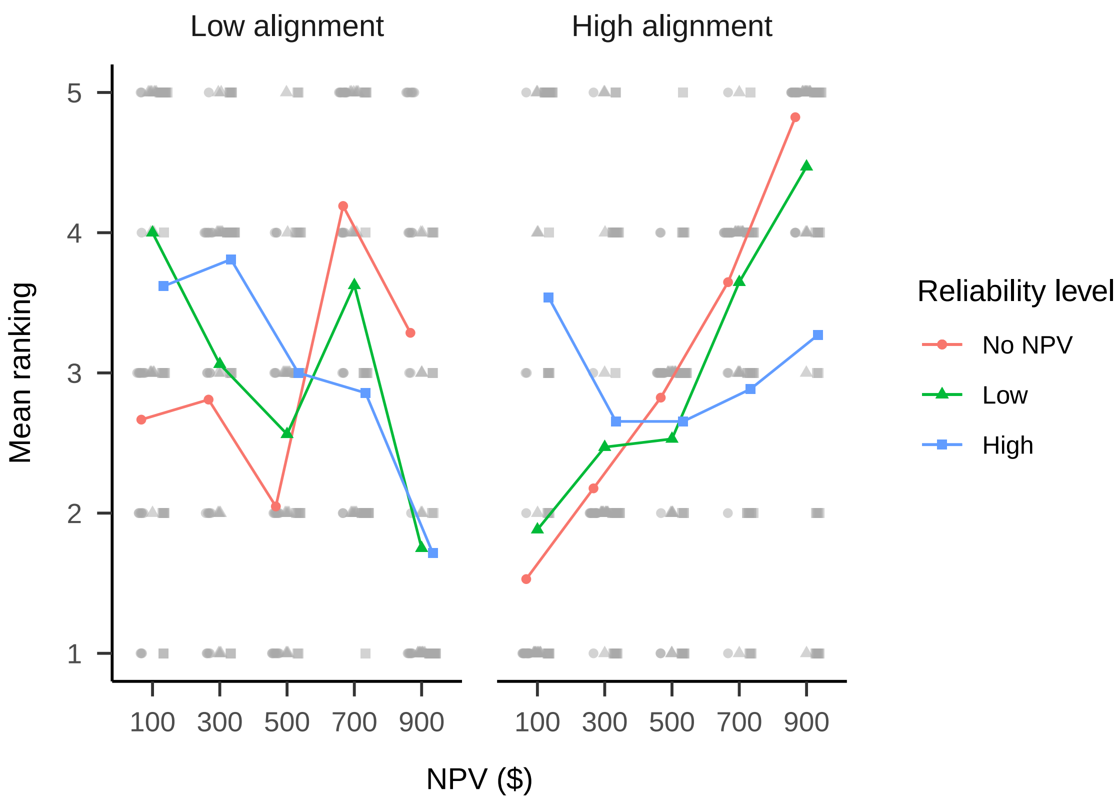


Figure B.8: Mean ranking.

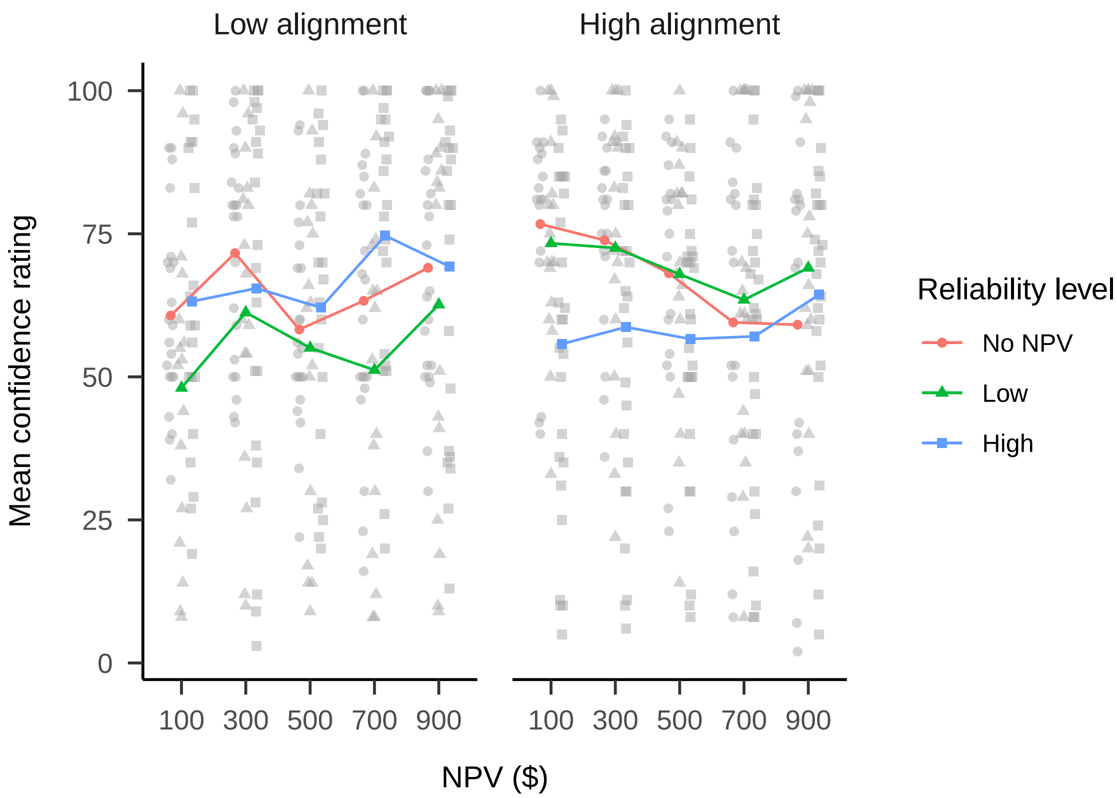


Figure B.9: Mean confidence.

B.1.2.2 Confidence

A mixed factorial ANOVA was conducted to investigate the effects of alignment and verbally-instructed NPV reliability on participants' confidence rating of their decisions. As shown in Figure B.9, the alignment × reliability level × NPV interaction was not significant, $F(7.47, 418.08) = 1.26$, $p = .267$, $\hat{\eta}_p^2 = .022$. Contrary to the allocation and ranking data, in the low alignment condition, there were no significant differences in the linear NPV trend between the no NPV condition and low reliability condition, $M = 10.73$, 95% CI $[-30.15,\ 51.61]$, $t(112) = 0.52$, $p = .604$, nor the high reliability condition, $M = 13.05$, 95% CI $[-24.97,\ 51.07]$, $t(112) = 0.68$, $p = .498$. However, as above, in the high alignment condition, the linear NPV trend was significantly lower in the no NPV condition than the high reliability condition, $M = 65.14$, 95% CI $[26.72,\ 103.57]$, $t(112) = 3.36$, $p = .001$, and not the low reliability condition, $M = 31.88$, 95% CI $[-10.38,\ 74.14]$, $t(112) = 1.49$, $p = .138$.

B.1.2.3 Forecast Mean

A mixed factorial ANOVA was conducted to investigate the effects of alignment and verbally-instructed NPV reliability on participants' forecast means. As seen in Figure B.10, the alignment × reliability level × NPV interaction was not significant, $F(5.26, 142.10) = 1.89$, $p = .095$, $\hat{\eta}_p^2 = .066$. However, the alignment × NPV interaction was significant, $F(2.63, 142.10) = 2.89$, $p = .044$, $\hat{\eta}_p^2 = .051$; as well as the reliability level × NPV interaction, $F(5.26, 142.10) = 7.91$, $p < .001$, $\hat{\eta}_p^2 = .227$. The simple effects appear to be as above. Specifically, in the low alignment condition, the linear NPV trend was significantly lower in the no NPV condition than both the low reliability condition, $M = 0.19$, 95% CI [0.09, 0.30], $t(54) = 3.63$, $p = .001$, and the high reliability condition, $M = 0.16$, 95% CI [0.04, 0.28], $t(54) = 2.75$, $p = .008$. However, in the high alignment condition, the linear NPV trend was only significantly lower in the no NPV condition than the high reliability condition, $M = 0.22$, 95% CI [0.11, 0.32], $t(54) = 4.04$, $p < .001$, and not the low reliability condition, $M = 0.08$, 95% CI [−0.04, 0.21], $t(54) = 1.30$, $p = .198$.

B.1.2.4 Forecast SD

A mixed factorial ANOVA was conducted to investigate the effects of alignment and verbally-instructed NPV reliability on participants' forecast SDs. As seen in Figure B.11, the alignment × reliability level × NPV interaction was significant, $F(6.87, 185.42) = 2.91$, $p = .007$, $\hat{\eta}_p^2 = .097$. However, none of the linear NPV trends were significantly different from each other as above. Of relevance, the low alignment condition on average had higher SDs than those in the high alignment condition, $F(1, 54) = 5.77$, $p = .020$, $\hat{\eta}_p^2 = .097$.

B.1.3 Discussion

Hypothesis B.4 was not supported, as there was no evidence of a main effect of alignment on participants' confidence in their allocation decisions. Instead, exploratory analyses showed that the difference in confidence between reliability conditions is greater in the low alignment condition. This may reflect participants'

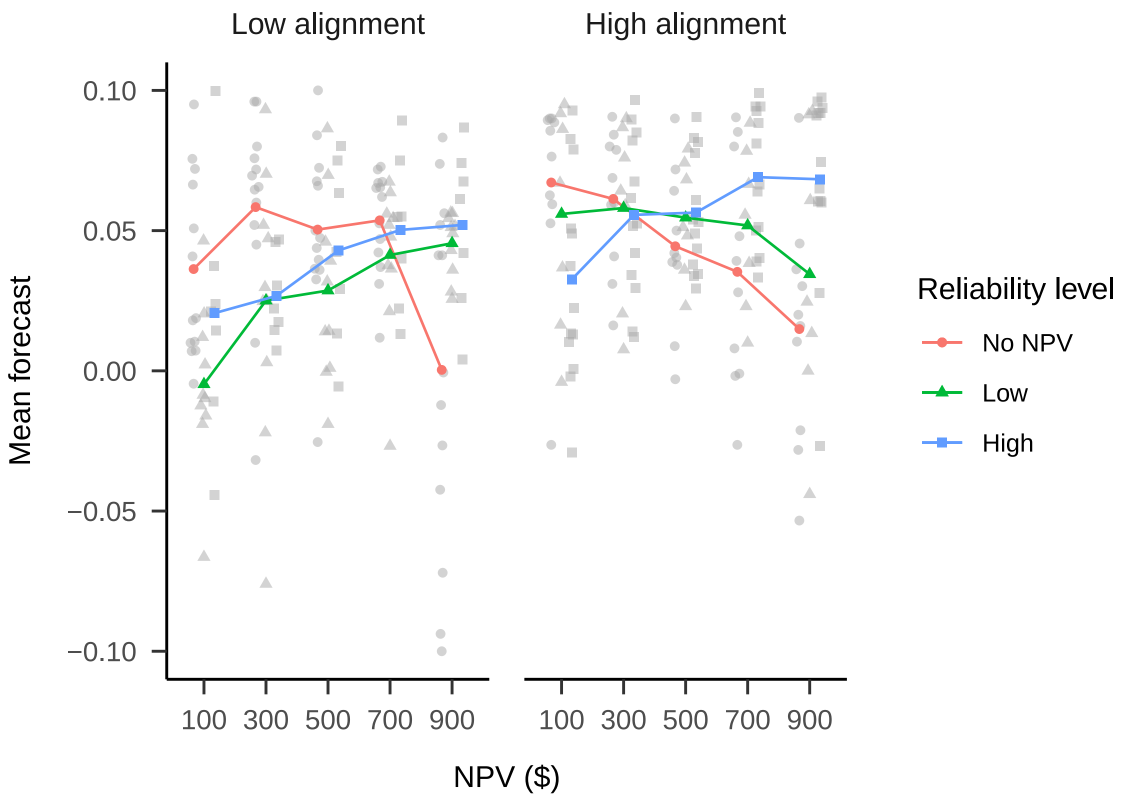


Figure B.10: Mean forecasts.

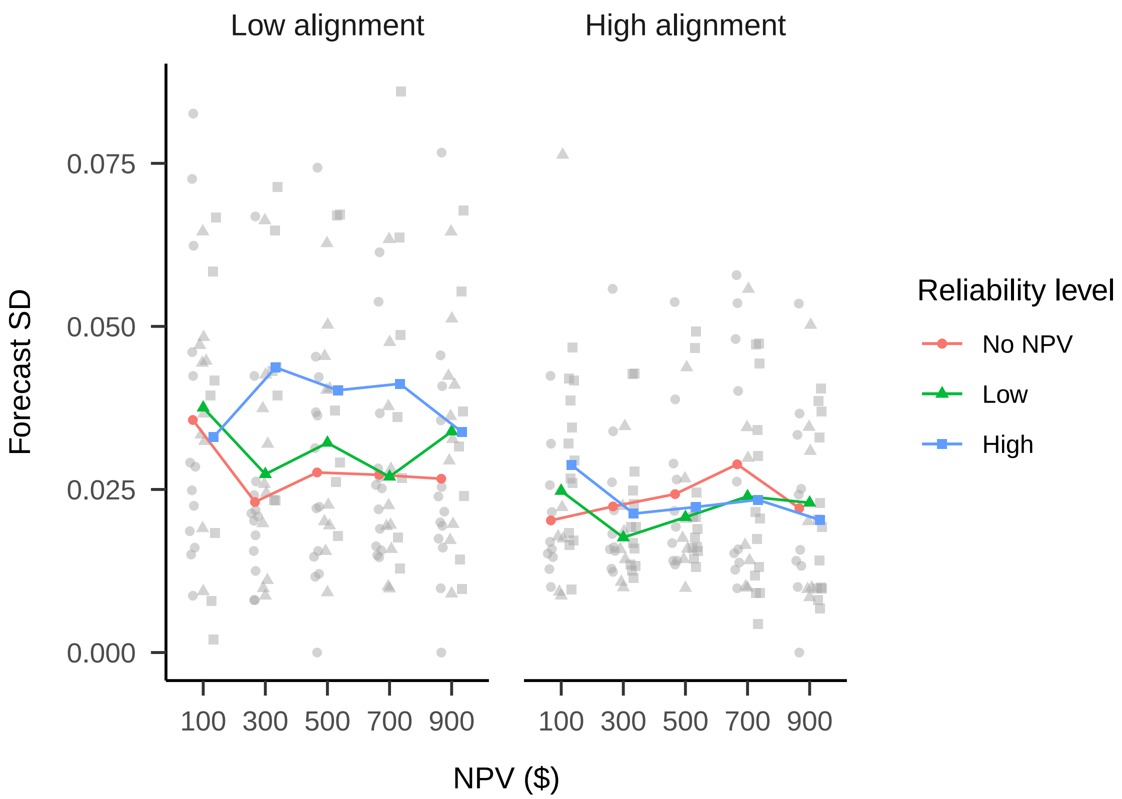


Figure B.11: Mean forecast SD.

difficulty in making sense of their choices when alignment was low, given more confidence when assured of the reliability of NPV. In the high alignment condition, on the other hand, regardless of reliability condition, they had a way of using the reliability information. Further, confidence also seemed to increase more with NPV, on average, more when projects were dissimilar, which provides evidence for their reliance on NPV in this situation. There was limited evidence for the effect of alignment on forecast variability. Experiments 5 and 6 attempted to replicate this result with more participants.

B.2 Experiment 2

B.2.1 Method

B.2.1.1 Materials

B.2.1.1.1 Instructions Figure B.12 shows the instructions.

B.2.1.1.2 NPV Test Participants were given more extensive information about NPV than in the previous experiment and were tested on their ability to calculate simple averages from given numerical ranges, as shown in Figures B.13 and B.14.

B.2.1.1.3 NPV Knowledge Ratings A similar design to Long et al. (2018, Study 1) was used to test whether this sample may be overconfident in their understanding on NPV. Therefore, participants were asked to rate their knowledge of NPV in various points in the study (see the procedure in Section 4.3.1.3). Figure B.15 shows an example of one such display.

Investment task

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided. In addition to those numbers, you will find each project's projected cash inflow for each year, and the net present value (NPV) that was calculated using those figures. The discount rate will always be 10% and the initial investment will always be $5000. These are taken into account in the NPV calculations.

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest intrinsic quality.**

You will decide how to rank the projects in order of investment priority, and decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects.

Importantly, each page's set of five projects should be treated independently of the other pages' project sets.

Figure B.12: Experiment 2 instructions.

B.2.1.1.4 Variance Lecture See below the slides for the variance lecture.

B.2.2 Results

B.2.2.1 Ranking

A mixed factorial ANOVA was conducted to investigate the effects of NPV, alignment, and numerical NPV reliability on participants' project rankings. Figure B.30 shows these data. The alignment × reliability level × NPV interaction was not significant, $F(3.00, 159.10) = 2.44$, $p = .066$, $\hat{\eta}_p^2 = .044$. However, the alignment × NPV interaction was significant, $F(3.31, 370.54) = 21.00$, $p < .001$, $\hat{\eta}_p^2 = .158$; as well as the reliability amount × NPV interaction, $F(6.62, 370.54) = 9.73$, $p < .001$, $\hat{\eta}_p^2 = .148$. As in the allocation data, the linear NPV trend did

Understanding NPV

Net Present Value (NPV) is used as a measure of a project's potential profitability. A positive value indicates that the project is profitable, while a negative value indicates that a project is not profitable.

When calculating NPV, the potential future cash inflows are converted to their "present values". This is important, because we know that an amount of money is more valuable in the present than it is in the future. The time value of money is accounted for by dividing each year's cash inflow by the discount rate. Finally, the sum of all the present values is deducted from the value of the initial investment.

To calculate the NPV you need the following components:

1. The cash inflow for each year of the project
2. The discount rate
4. The initial investment

Below is the generic formula for calculating NPV:

$$NPV = \frac{Cash\ inflow\ for\ year\ 1}{(1 + discount\ rate)^1} + \frac{Cash\ inflow\ for\ year\ 2}{(1 + discount\ rate)^2} + \frac{Cash\ inflow\ for\ year\ 3}{(1 + discount\ rate)^3} \ldots - Initial\ investment$$

Some of the time, it might be unclear exactly what the future cash inflow is, so it might be given as a range of possible values.

Below is an example of an NPV calculation with the discount rate calculations and initial investment already filled in. Notice that instead of a single cash inflow value, a range is provided (assume that the distribution is uniform). The value that should be used as the cash inflow is the mid point of these values. This is done by calculating the average of the two values.

For this session, you will get some practise in calculating NPV. However, we will give you the value that is in the denominator (the discount rate calculation) and the initial investment. All you need to do is calculate each year's cash inflow and enter them into the formula.

Example 1

$$NPV = \frac{[range: 1500 - 2500]}{1.1} + \frac{[range: 750 - 1250]}{1.21} + \frac{[range: 1875 - 3125]}{1.331} - 3000$$

Please calculate the mid-points for these ranges and type them in below:

Year 1 cash inflow

Year 2 cash inflow

Year 3 cash inflow

Figure B.13: Experiment 2 NPV test.

The range for Year 1 was $1500-$2500.
You calculated the Year 1 cash inflow to be $2000.

The range for Year 2 was $750-$1250.
You calculated the Year 2 cash inflow to be $1000.

The range for Year 3 was $1875-$3125.
You calculated the Year 3 cash inflow to be $2500.

Therefore, NPV = $1522.92

Figure B.14: Experiment 2 NPV test answers.

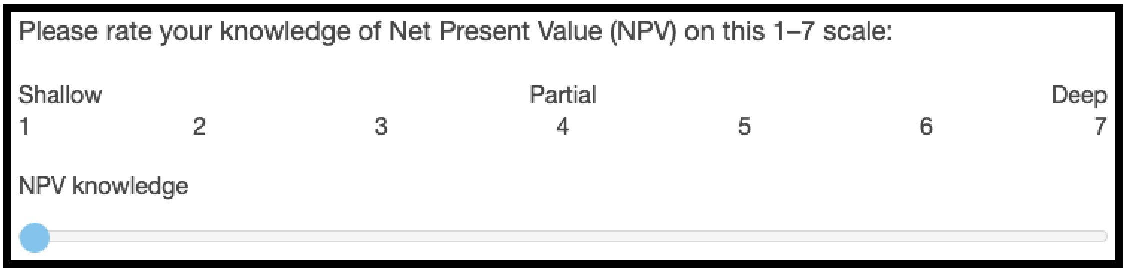


Figure B.15: Experiment 2 NPV knowledge rating task.

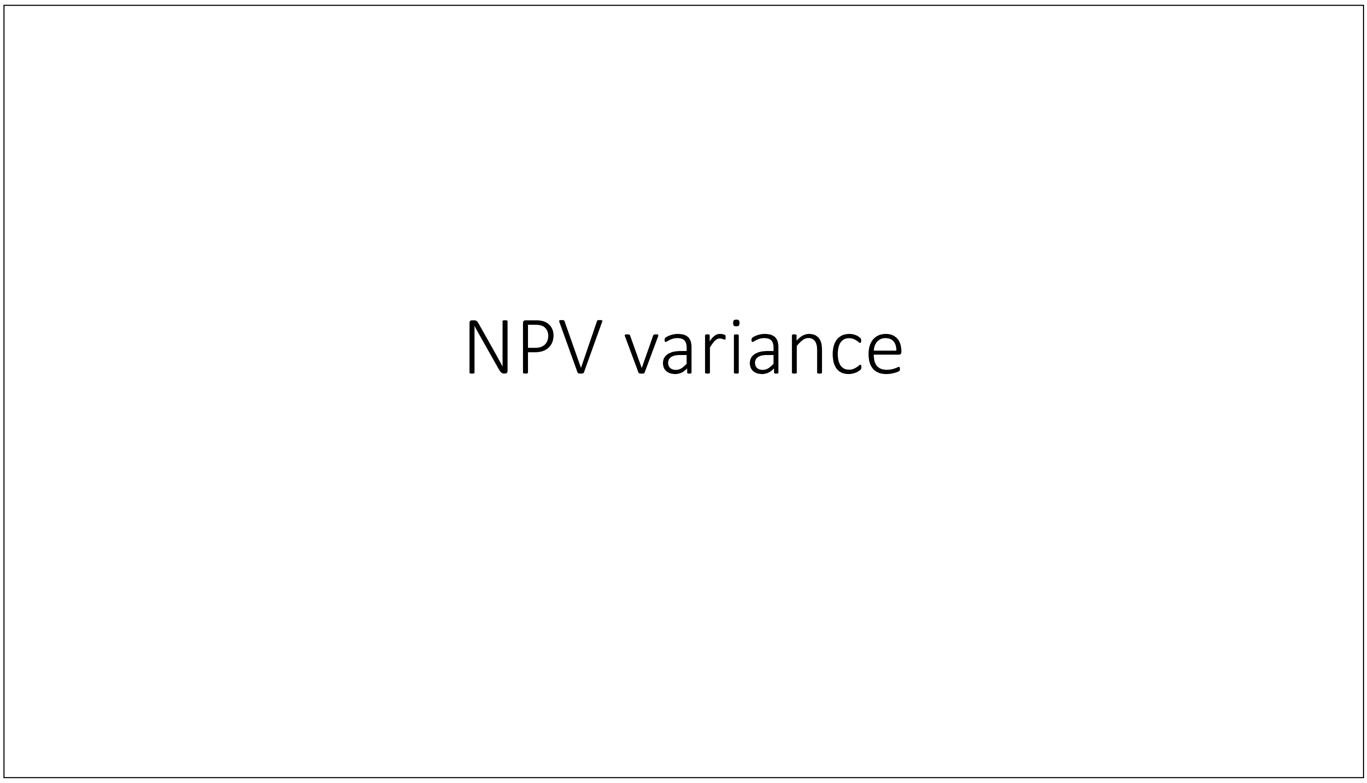


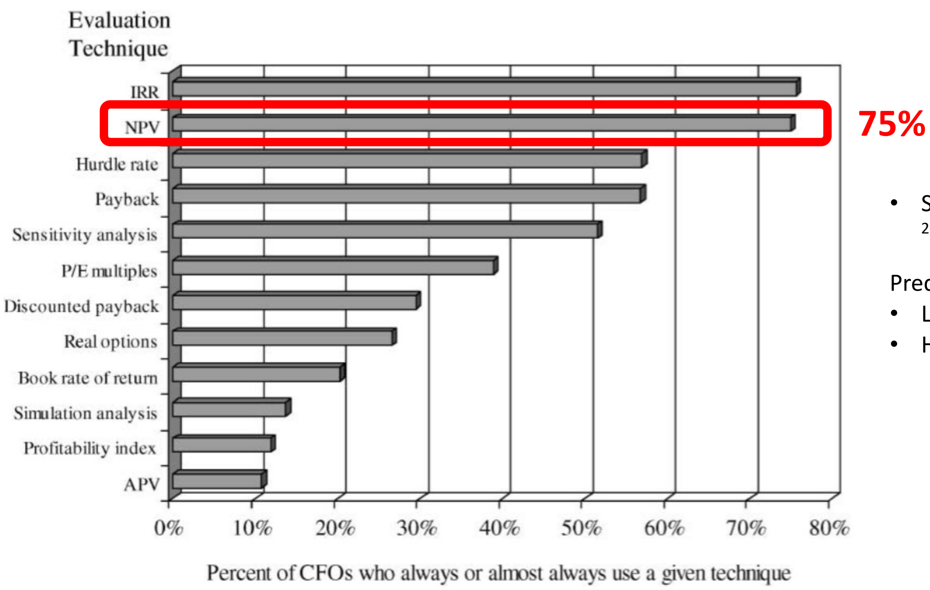
Figure B.16: Variance lecture slide 1.

NPV

$$NPV = \frac{Year\ 1\ inflow}{(1{+}discount\ \%)^1} + \frac{Year\ 2\ inflow}{(1{+}discount\ \%)^2} + \frac{Year\ 3\ inflow}{(1{+}discount\ \%)^3} \ldots - Initial\ investment$$

Figure B.17: Variance lecture slide 2.

NPV is used very frequently



- Sample: 392 CFOs (Graham & Harvey, 2001)

Predictors:

- Large firms > small firms
- High debt ratio > small debt ratio

Figure B.18: Variance lecture slide 3.

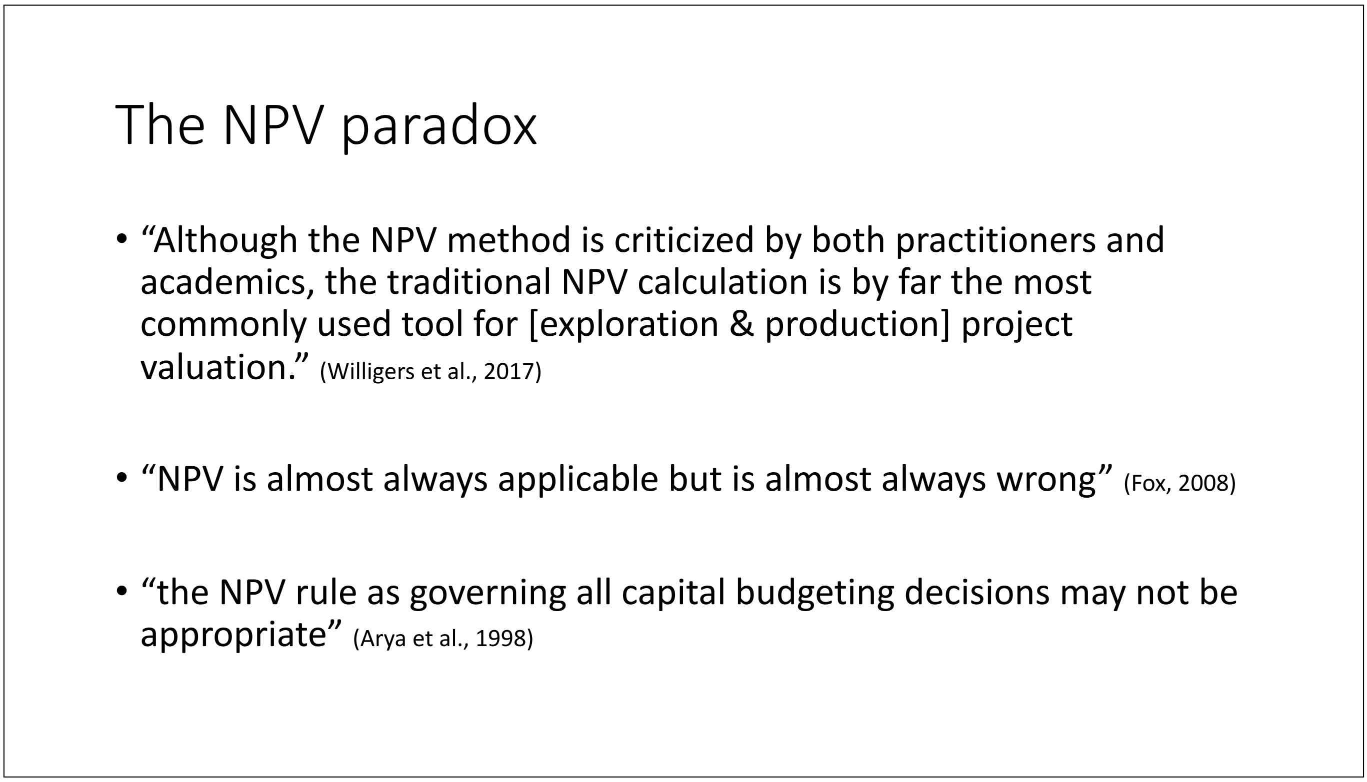


Figure B.19: Variance lecture slide 4.

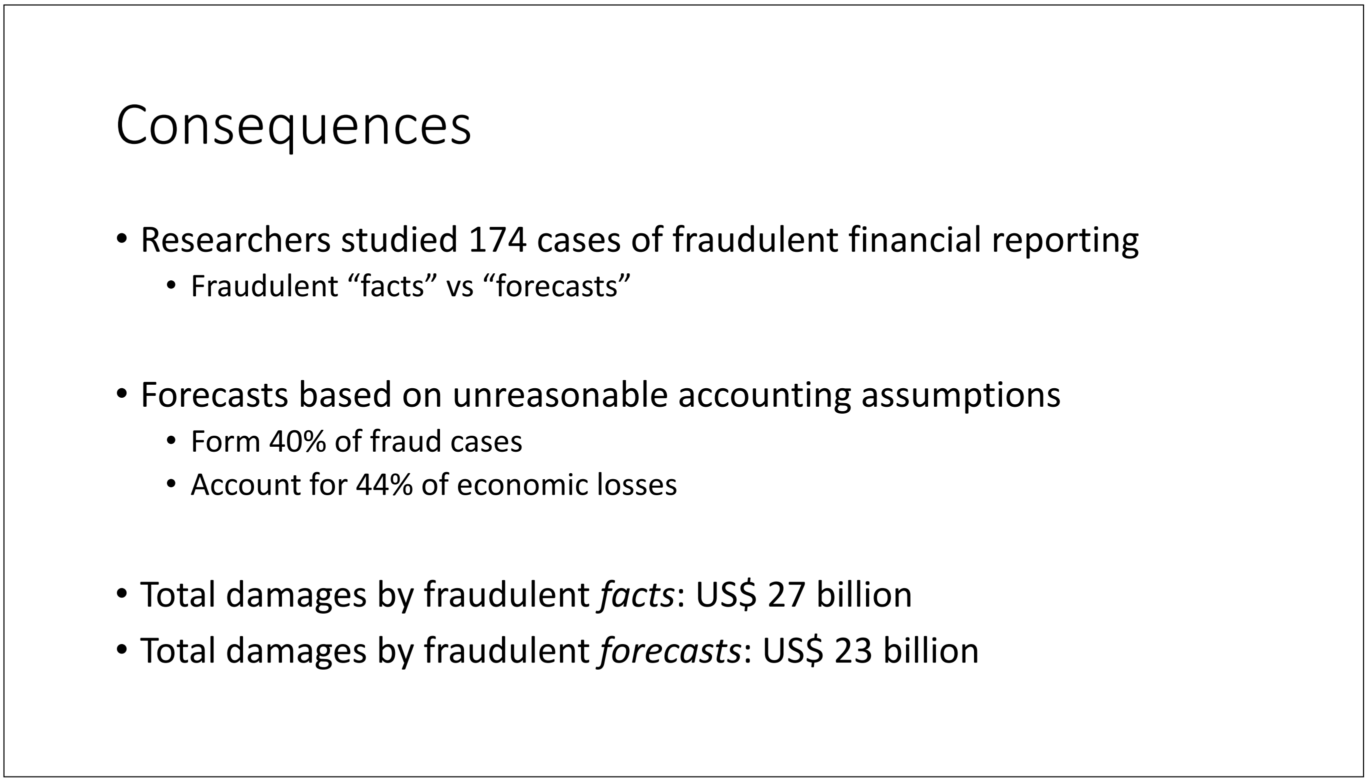


Figure B.20: Variance lecture slide 5.

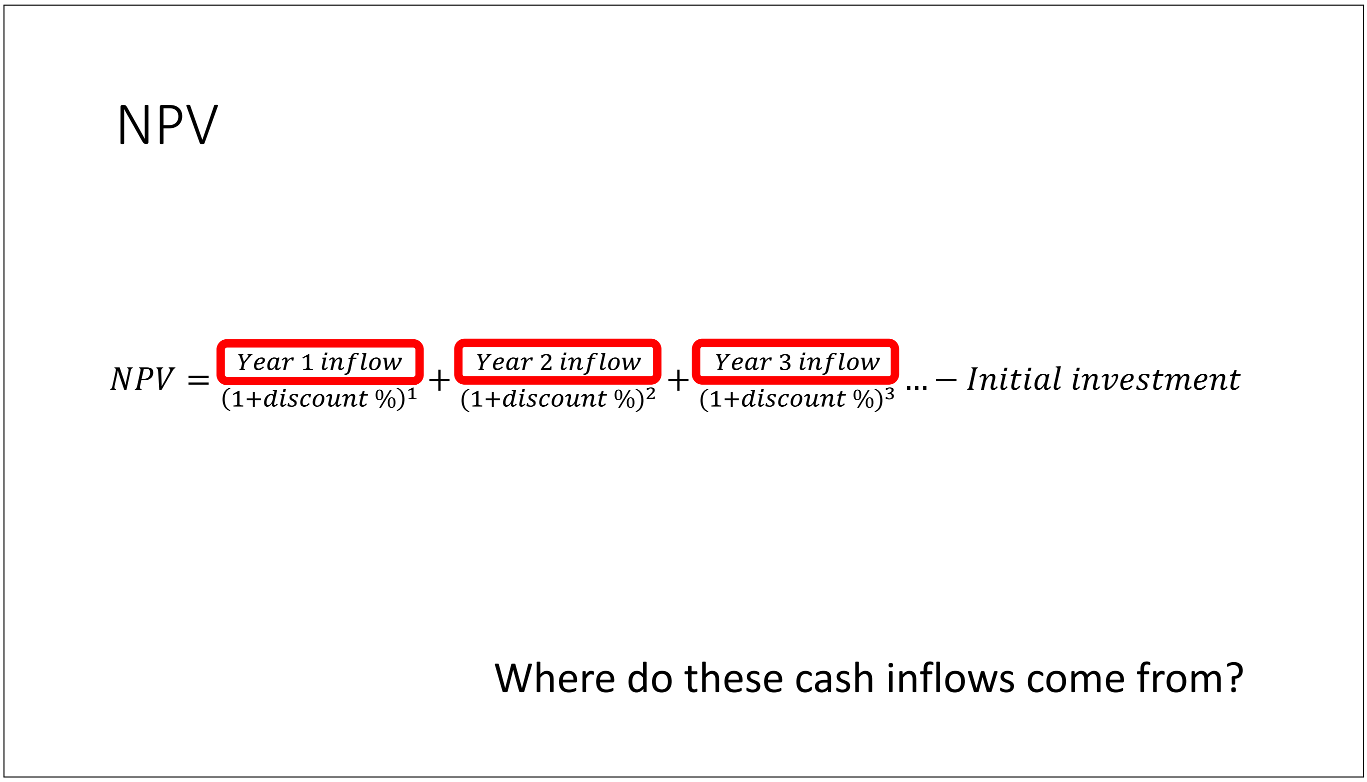


Figure B.21: Variance lecture slide 6.

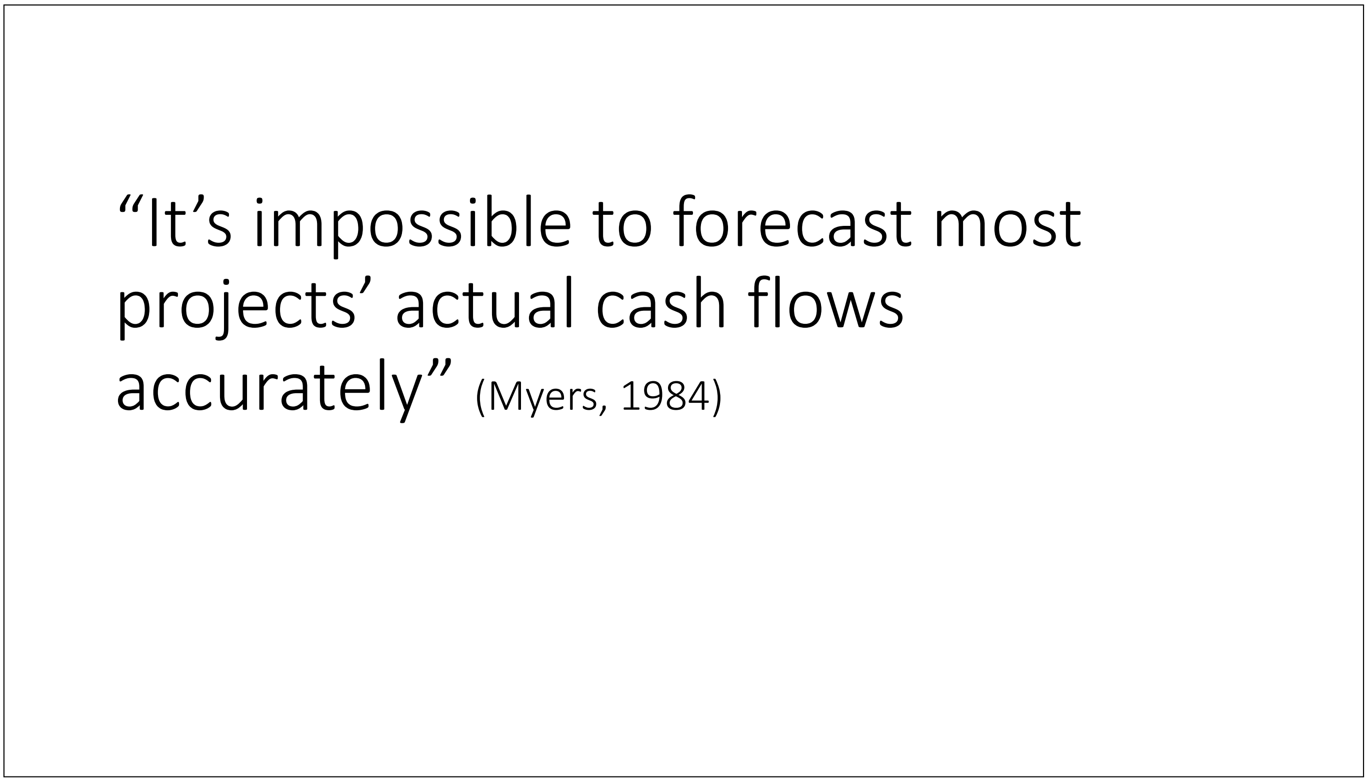


Figure B.22: Variance lecture slide 7.

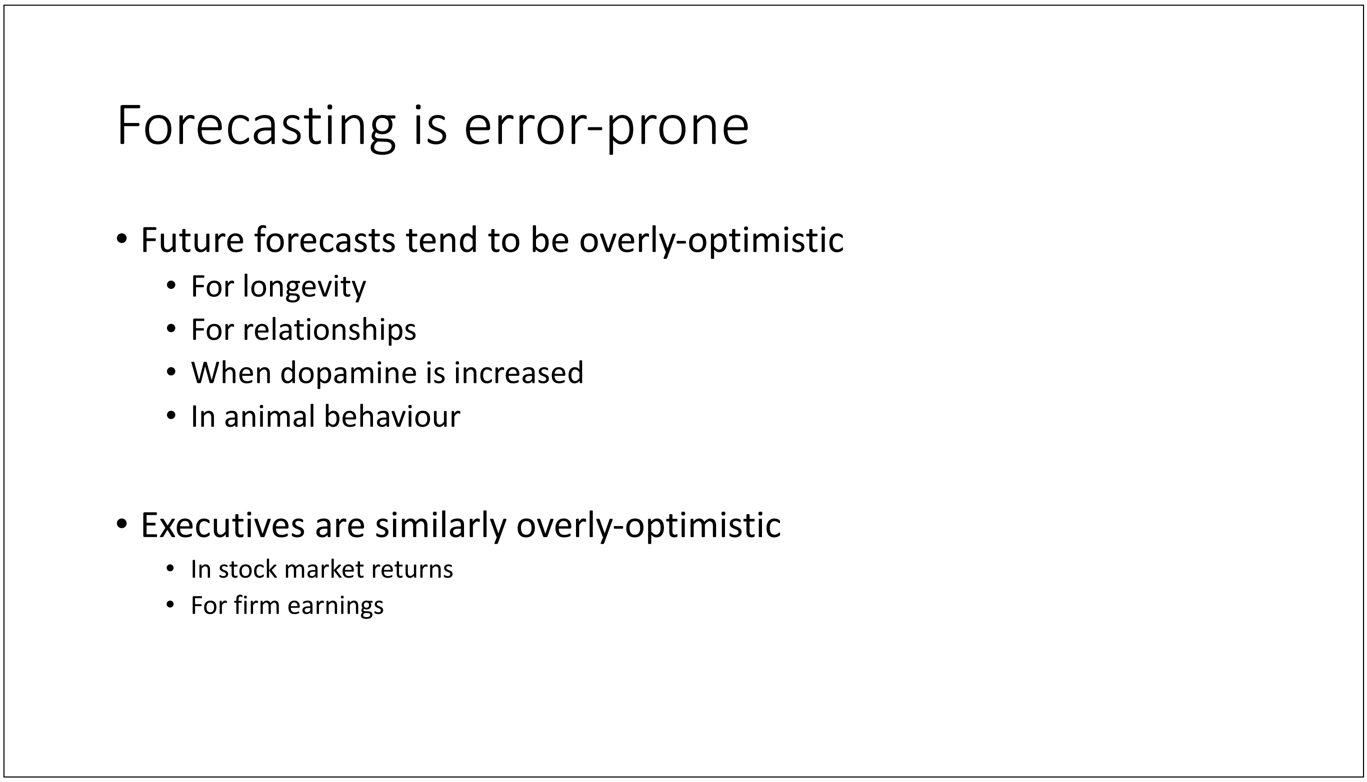


Figure B.23: Variance lecture slide 8.

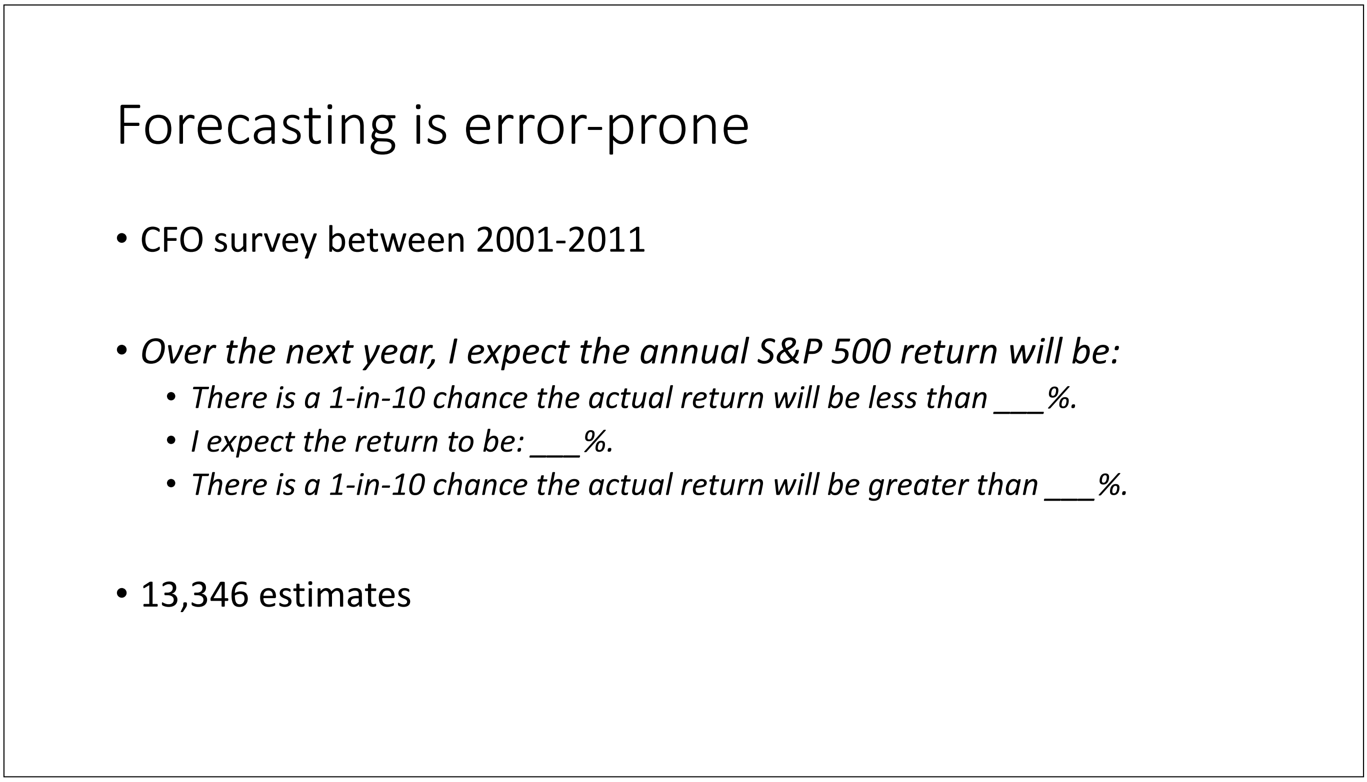


Figure B.24: Variance lecture slide 9.

Forecasting is error-prone

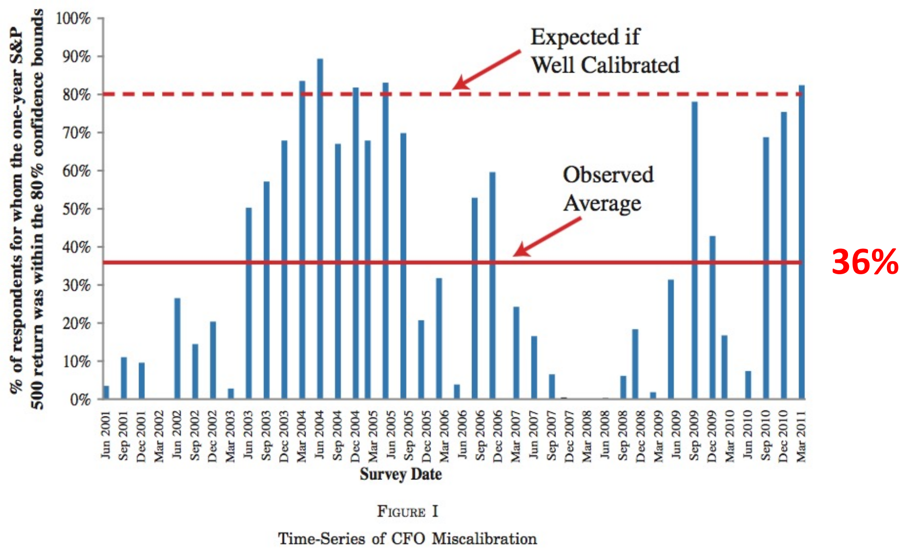


Figure B.25: Variance lecture slide 10.

Paying attention to variance

- Ranges are frequently used for forecast estimates
  - 80% of the time between 2002-2010
- Taking account of variance increases forecasting accuracy

Figure B.26: Variance lecture slide 11.

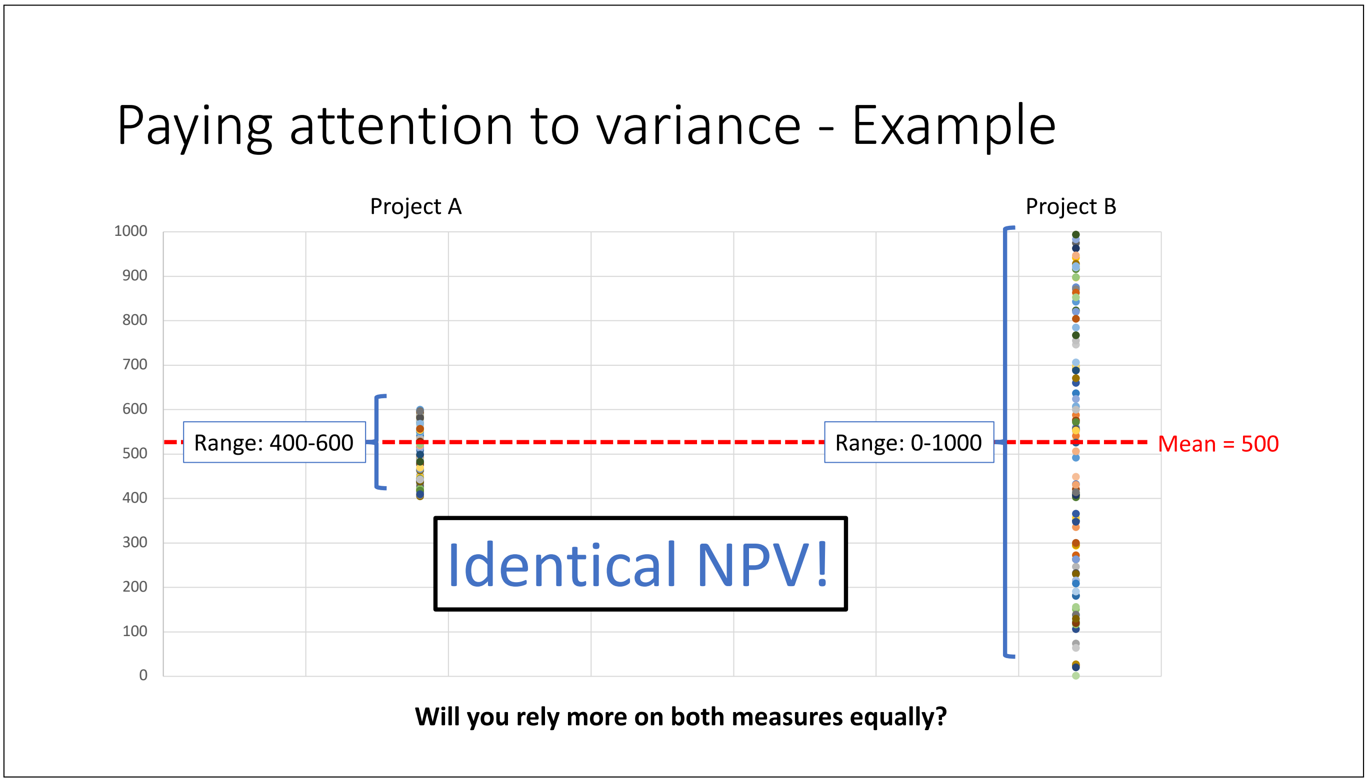


Figure B.27: Variance lecture slide 12.

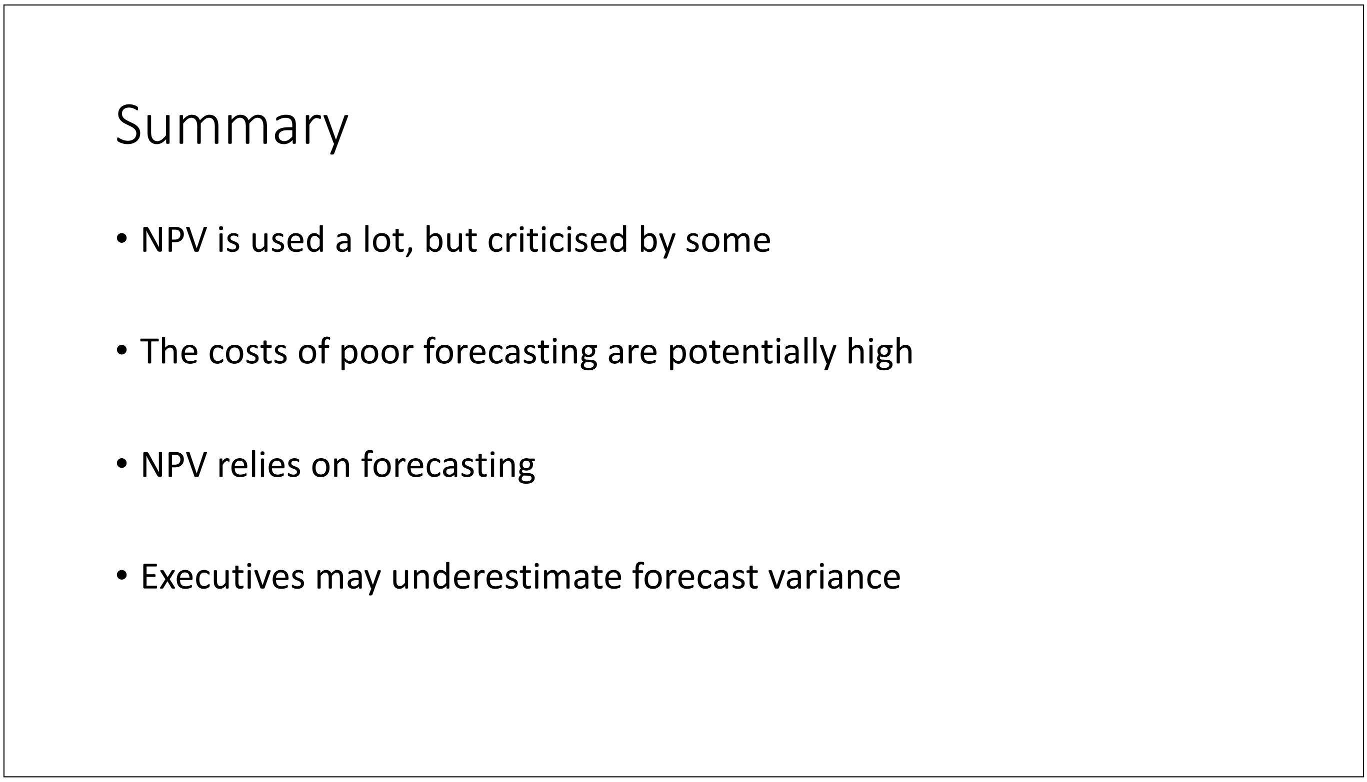


Figure B.28: Variance lecture slide 13.

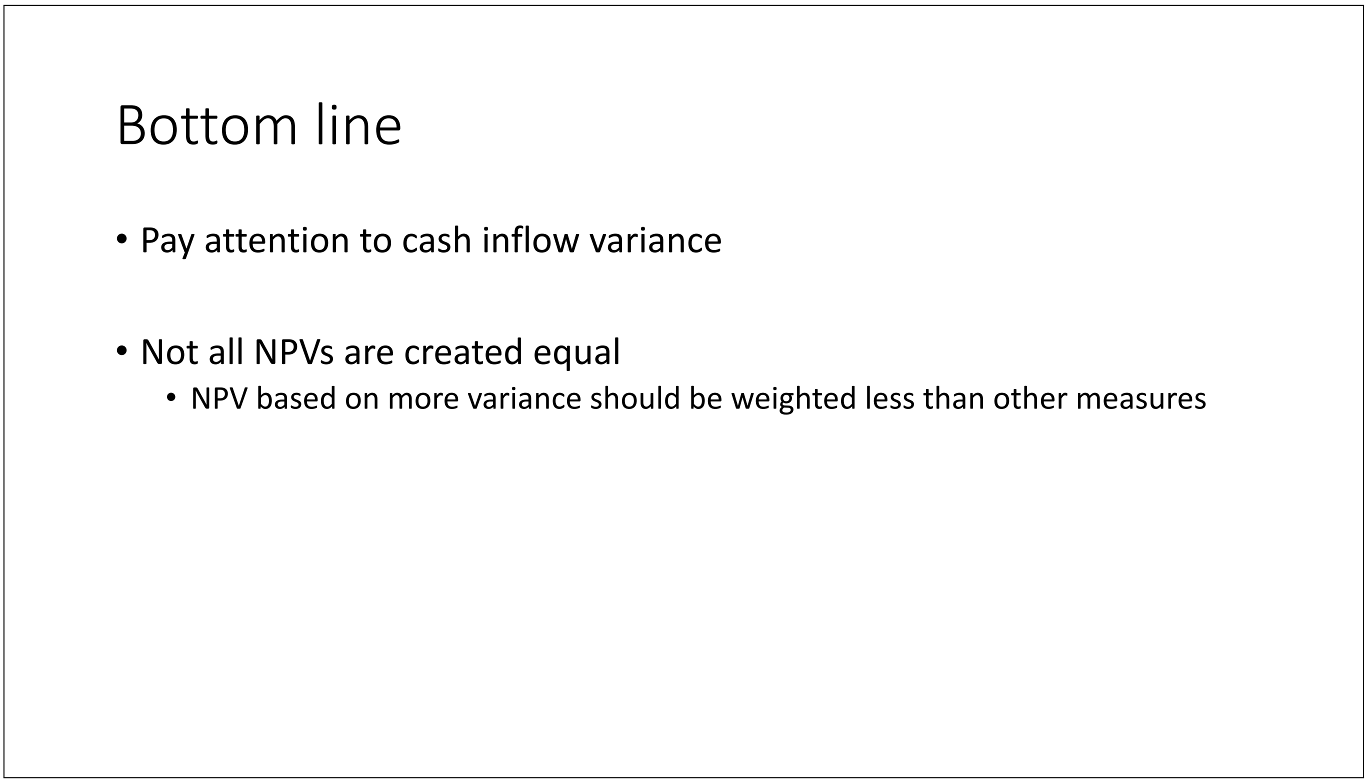


Figure B.29: Variance lecture slide 14.

not differ between reliability level condition in neither the low alignment condition, $\Delta M = 0.43$, 95% CI $[-0.77,\ 1.63]$, $t(53) = 0.71$, $p = .480$, nor the high alignment condition, $\Delta M = 0.46$, 95% CI $[-0.92,\ 1.84]$, $t(53) = 0.67$, $p = .504$. However, averaging over reliability level, the linear NPV trend was higher in the low alignment condition than in the high alignment condition, $\Delta M = -4.54$, 95% CI $[-6.39,\ -2.68]$, $t(53) = -4.91$, $p < .001$.

B.2.2.2 Confidence

A mixed factorial ANOVA was conducted to investigate the effects of NPV, alignment, and numerical NPV reliability on participants' confidence ratings. Figure B.31 shows these data. Only the main effect of NPV was significant, $F(2.62, 139.08) = 2.97$, $p = .041$, $\hat{\eta}^2_p = .053$.

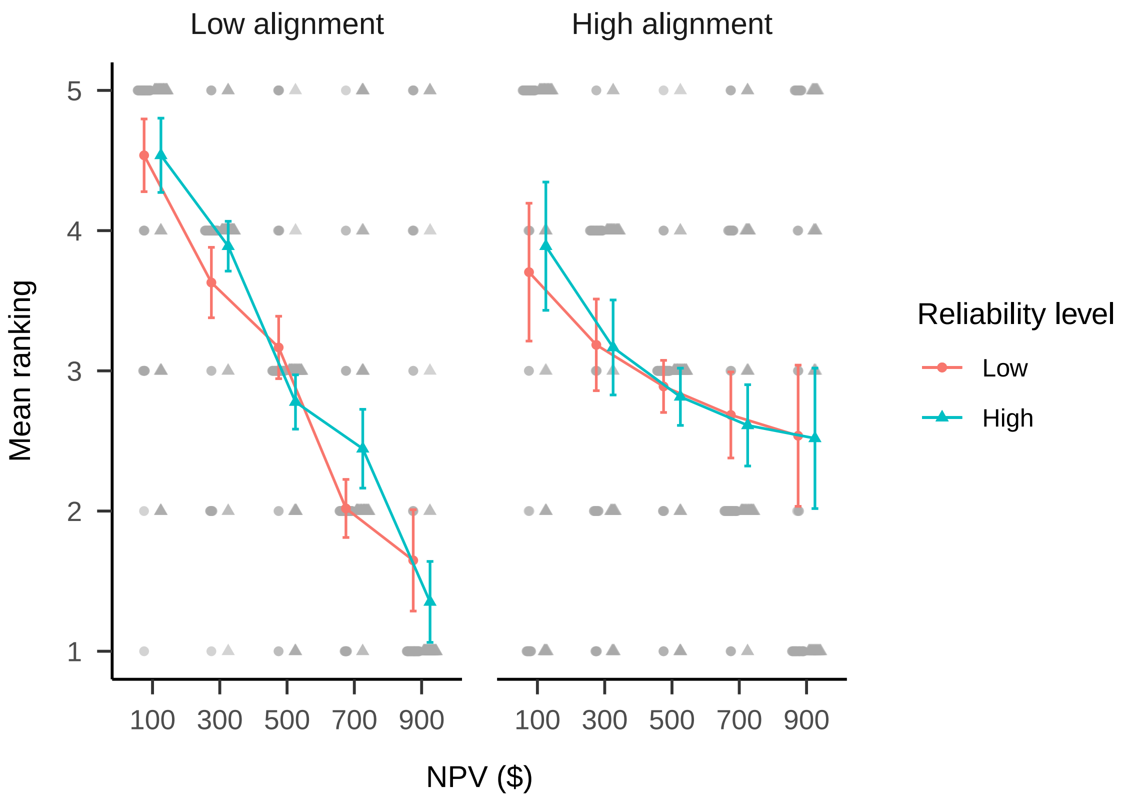


Figure B.30: Mean ranking.

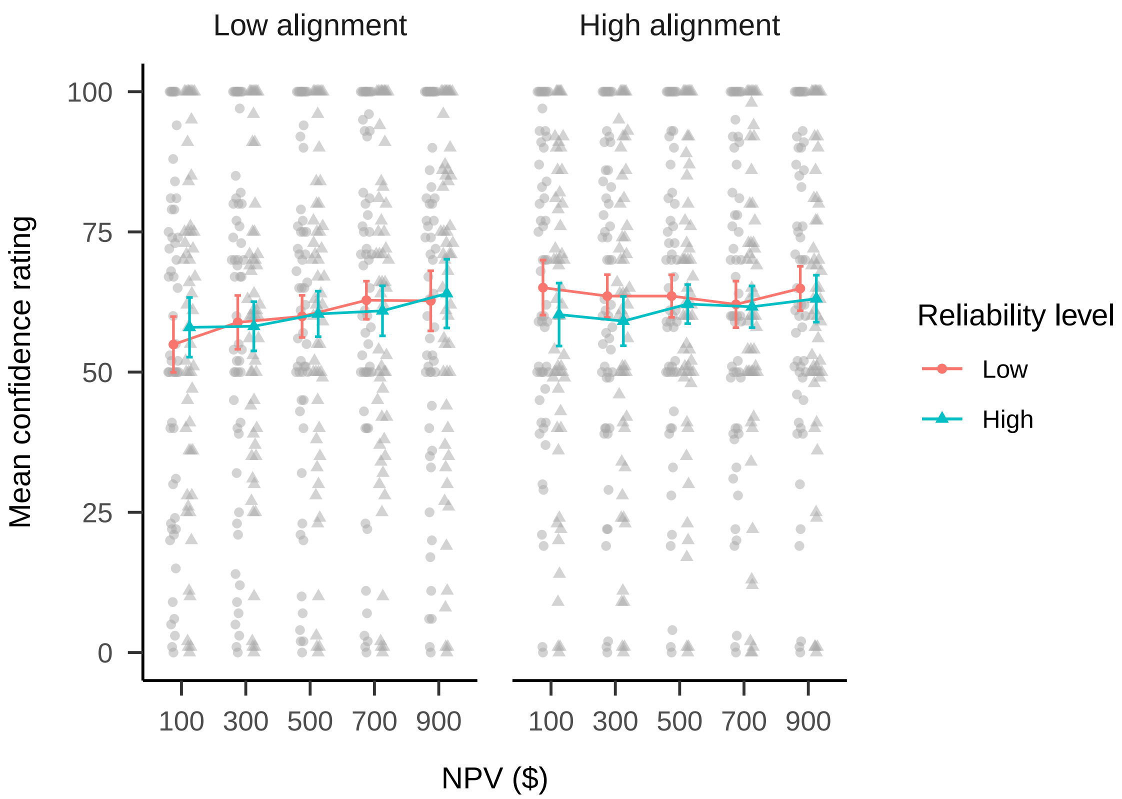


Figure B.31: Mean confidence.

B.2.2.3 Variance Lecture

The allocation and ranking data show that participants were affected by the similarity of options, but were not affected by variance information. After the main task of this experiment, participants were shown a short lecture about the importance of variance information when making allocation decisions. They were then presented with half of their previous allocations and gave them an opportunity to amend their allocations. It was hypothesised that participants will be more sensitive to variance after the educational intervention.

A mixed factorial ANOVA was conducted to investigate the effects of phase on participants' project allocations. As shown in Figure B.32, the four-way interaction was not significant, $F(2.56, 133.09) = 1.74$, $p = .169$, $\hat{\eta}^2_p = .032$. Further, the NPV $\times$ phase $\times$ reliability level interactions were not significant for either the low alignment condition, $\Delta M = 4.43$, 95% CI $[-23.71,\ 32.58]$, $t(52) = 0.32$, $p = .753$; or the high alignment conditions, $\Delta M = -11.92$, 95% CI $[-43.39,\ 19.55]$, $t(52) = -0.76$, $p = .451$. In the low alignment condition, the linear NPV trend (averaged over reliability level) was significantly weaker after the lecture, compared with the linear NPV trend before the lecture, $\Delta M = -12.85$, 95% CI $[-24.08,\ -1.62]$, $t(52) = -2.30$, $p = .026$. However, this comparison was not significant in the high alignment condition, $\Delta M = -6.37$, 95% CI $[-18.93,\ 6.18]$, $t(52) = -1.02$, $p = .313$. These results suggest that participants did not become better informed by NPV numerical reliability after the variance lecture. There was, however, some reduction in reliance on NPV overall when projects were dissimilar.

B.2.2.4 NPV Knowledge

A repeated-measures ANOVA was conducted to investigate the effects of experiment phase condition on participants' NPV knowledge rating. Figure B.33 shows these data. The main effect of phase was significant, $F(2.43, 128.59) = 7.80$, $p < .001$, $\hat{\eta}^2_p = .128$. The post-explanation rating was significantly higher than the pre-explanation rating, $\Delta M = -0.59$, 95% CI $[-0.92,\ -0.26]$, $t(53) = -5.07$,

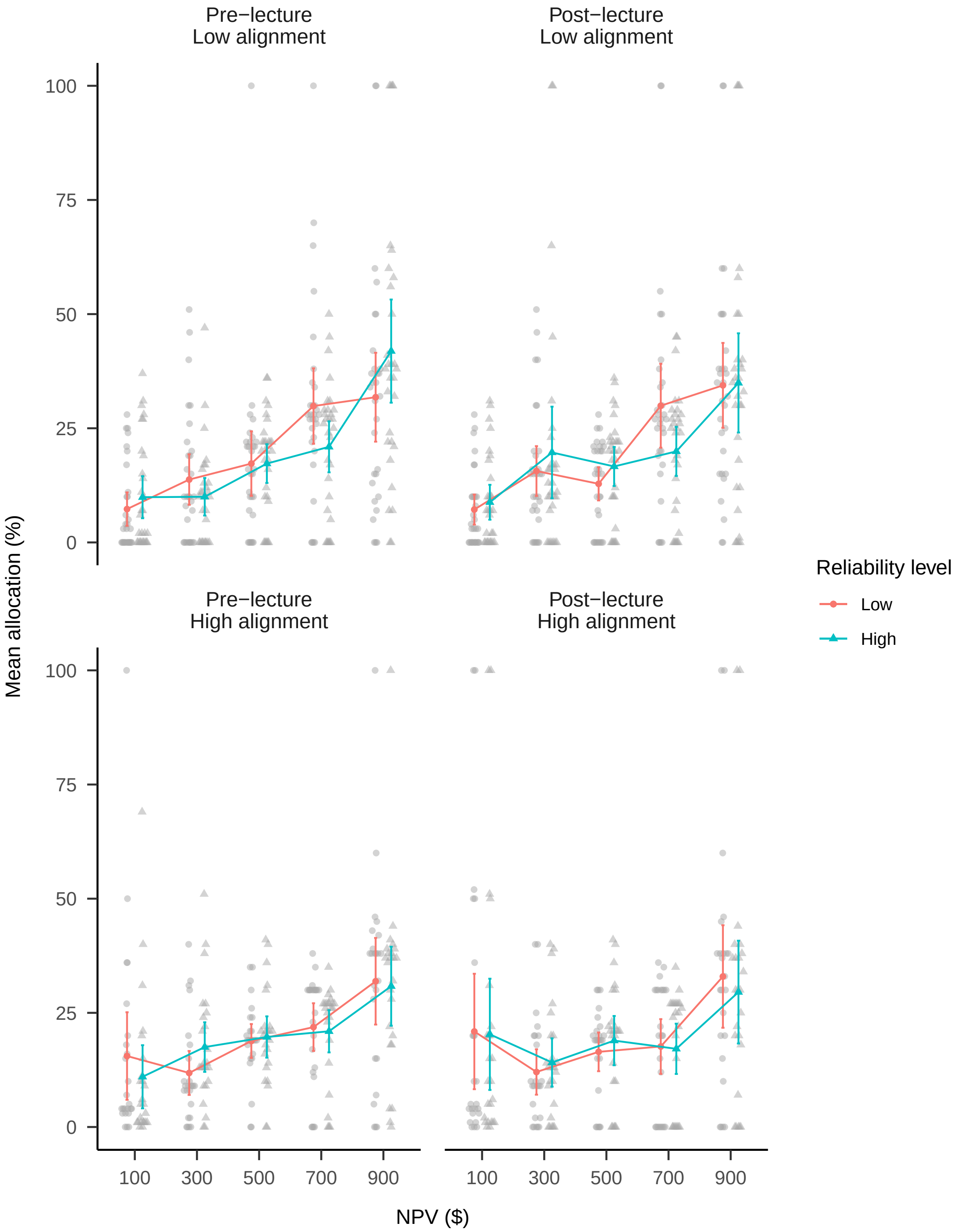


Figure B.32: Mean allocation by NPV, reliability level, alignment, and phase.

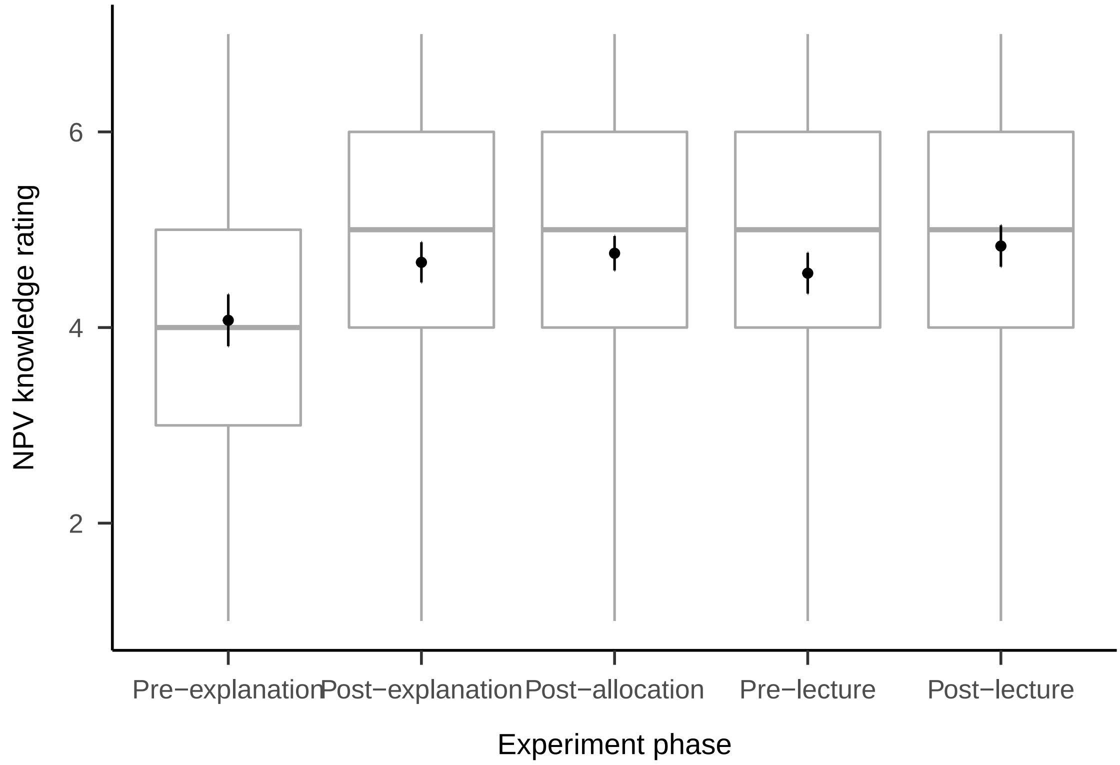


Figure B.33: Mean NPV knowledge rating.

$p < .001$. However, there were no significant differences in rating between any of the later phases.

B.3 Experiment 3

Figure B.34 shows the simulated hypothesised effects for Experiment 3. These effects were constructed as a composite of Experiment 1 data (without the no NPV condition) for the verbal reliability type condition, and data from a pilot study (see Appendix B.8) for the numerical reliability type condition. Variance was removed to see the effects clearer.

B.3.1 Method

B.3.1.1 Participants

B.3.1.1.1 Power Analysis A power analysis was conducted through simulation of the effects hypothesised in Experiment 3 (and the simple effects implied by them). The simulated data used the same regression coefficients as Experiment 2 for

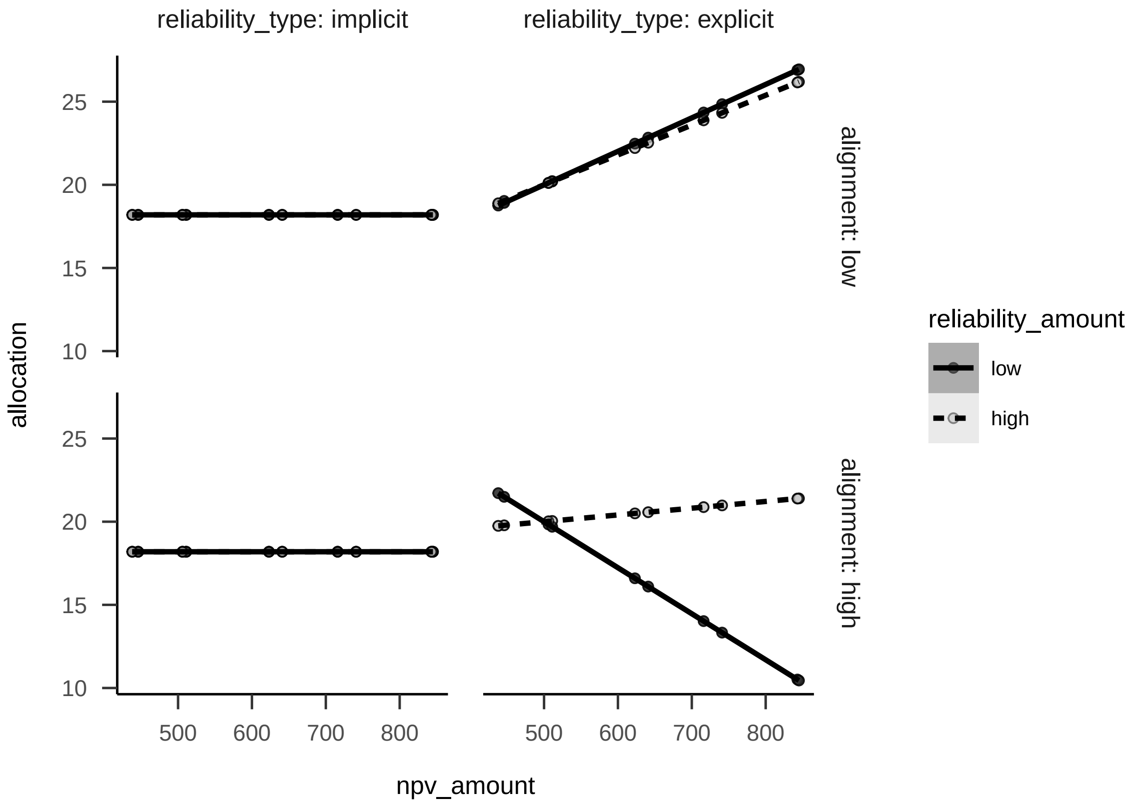


Figure B.34: Experiment 3 predicted data.

the explicit condition, no effects for the implicit condition (as shown in Figure B.34), and the intercept and residual variance of Experiment 2. The null effects were analysed using the two one-sided tests (TOST) procedure, or *equivalence* testing (Lakens et al., 2018), and setting the smallest effect size of interest to the smallest difference that leads to a significant equivalence between low and high implicit reliability for low alignment in Experiment 8 (see Appendix B.8). Figure B.35 shows the resulting power curve. The analysis suggests a total sample size of 448 $(112 \cdot 4)$.

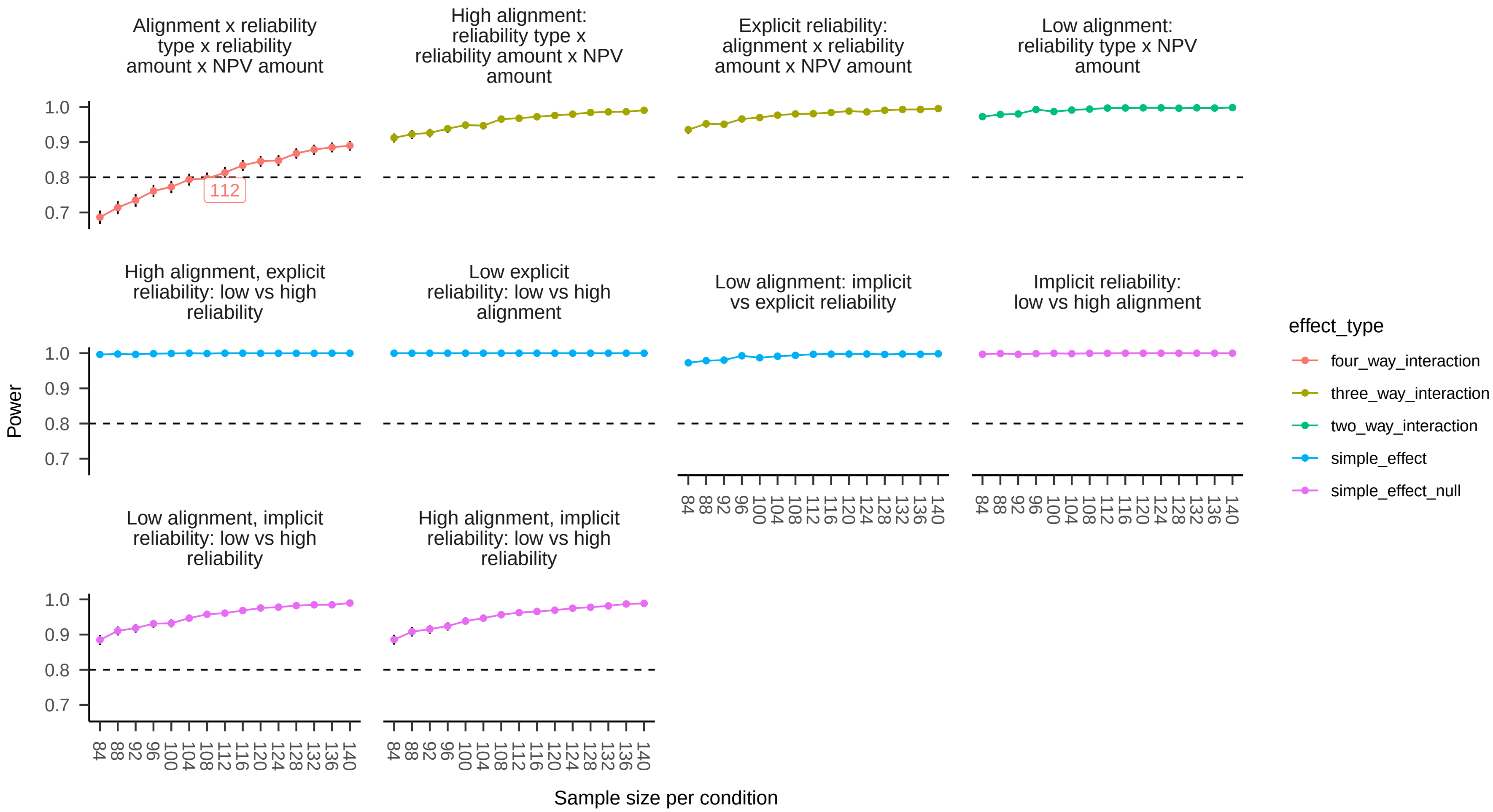


Figure B.35: Alignment Experiment 3 power curve. Labels indicate lowest sample size above 80% power.

B.3.1.2 Materials

B.3.1.2.1 Instructions Figures B.36 and B.37 show the instructions for the verbal and numerical reliability conditions, respectively.

B.3.1.2.2 Interstitial Display Figure B.38 shows an example of an interstitial display.

B.3.2 Results

B.3.2.1 Allocation

The three-way interaction (reliability level × NPV × reliability type) in the high alignment condition was significant, $\Delta M = 35.43$, 95% CI $[20.74,\ 50.12]$, $t(444) = 4.74$, $p < .001$. The NPV × reliability type (averaging over reliability level) in the low alignment condition was significant, $\Delta M = 11.48$, 95% CI $[0.19,\ 22.77]$, $t(444) = 2.00$, $p = .046$. The association between allocation and NPV for those in the explicit low reliability condition was significantly stronger for those in the low alignment condition, than for those in the high alignment condition, $\Delta M = 35.68$, 95% CI $[22.27,\ 49.09]$, $t(444) = 5.23$, $p < .001$. The linear NPV trend for those in the low alignment condition was significantly stronger for those in the explicit reliability condition, than for those in the implicit reliability condition (averaging over reliability level), $\Delta M = 11.48$, 95% CI $[0.19,\ 22.77]$, $t(444) = 2.00$, $p = .046$. The linear NPV trend for those in the implicit reliability condition was not significantly "equivalent" between those in the low and high reliability conditions for both those in the low alignment $\Delta M = 1.64$, 95% CI $[-8.74,\ 12.03]$, $t(444) = 0.31$, $p = .620$ and high alignment conditions $\Delta M = -1.21$, 95% CI $[-11.59,\ 9.18]$, $t(444) = 0.22$, $p = .589$. However, this is likely to be because the "lowest effect size of interest" estimate originated from an analysis used before data collection that was different to the one that one used after data collection. Specifically, a univariate linear model was originally used (treating NPV as a continuous predictor), whereas the data were ultimately analysed using a multivariate linear model (treating NPV as a repeated measures factor). In the numerical reliability condition, a pilot

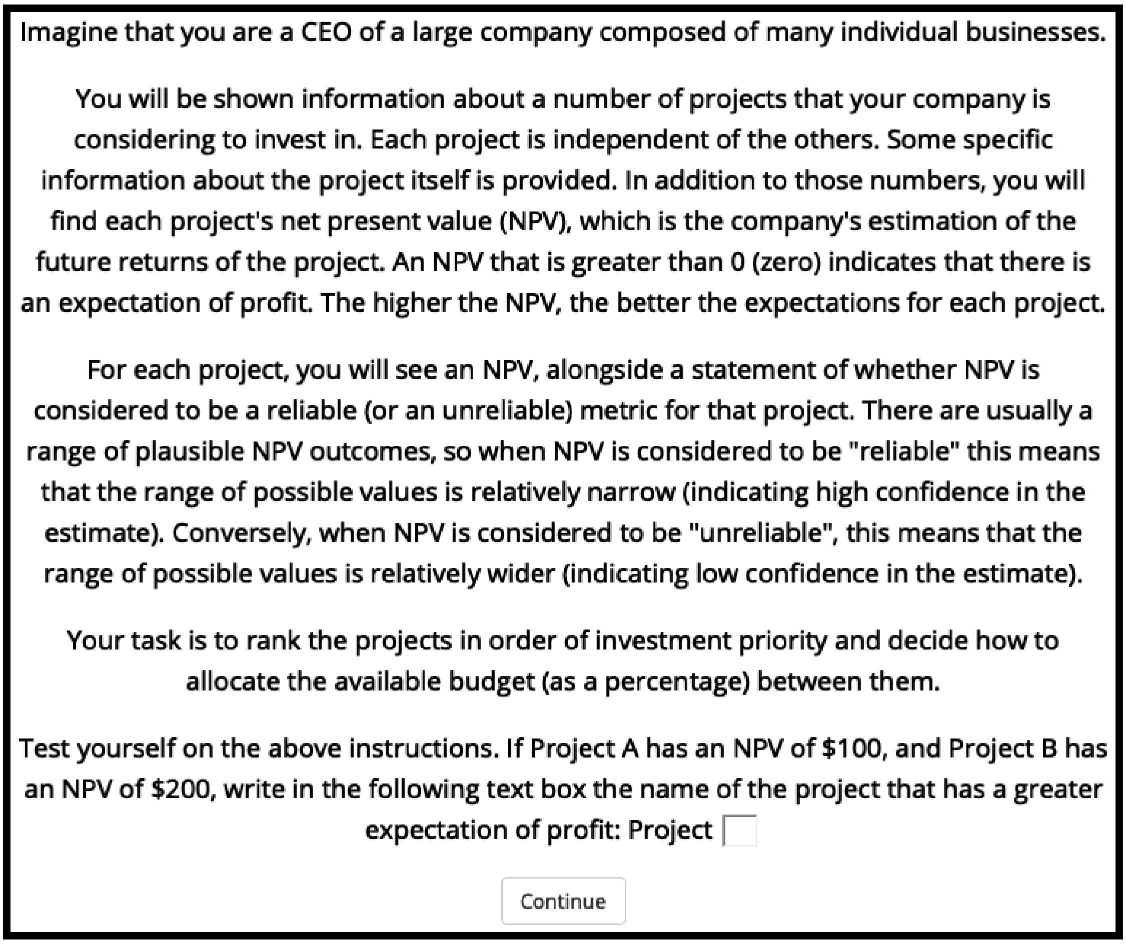
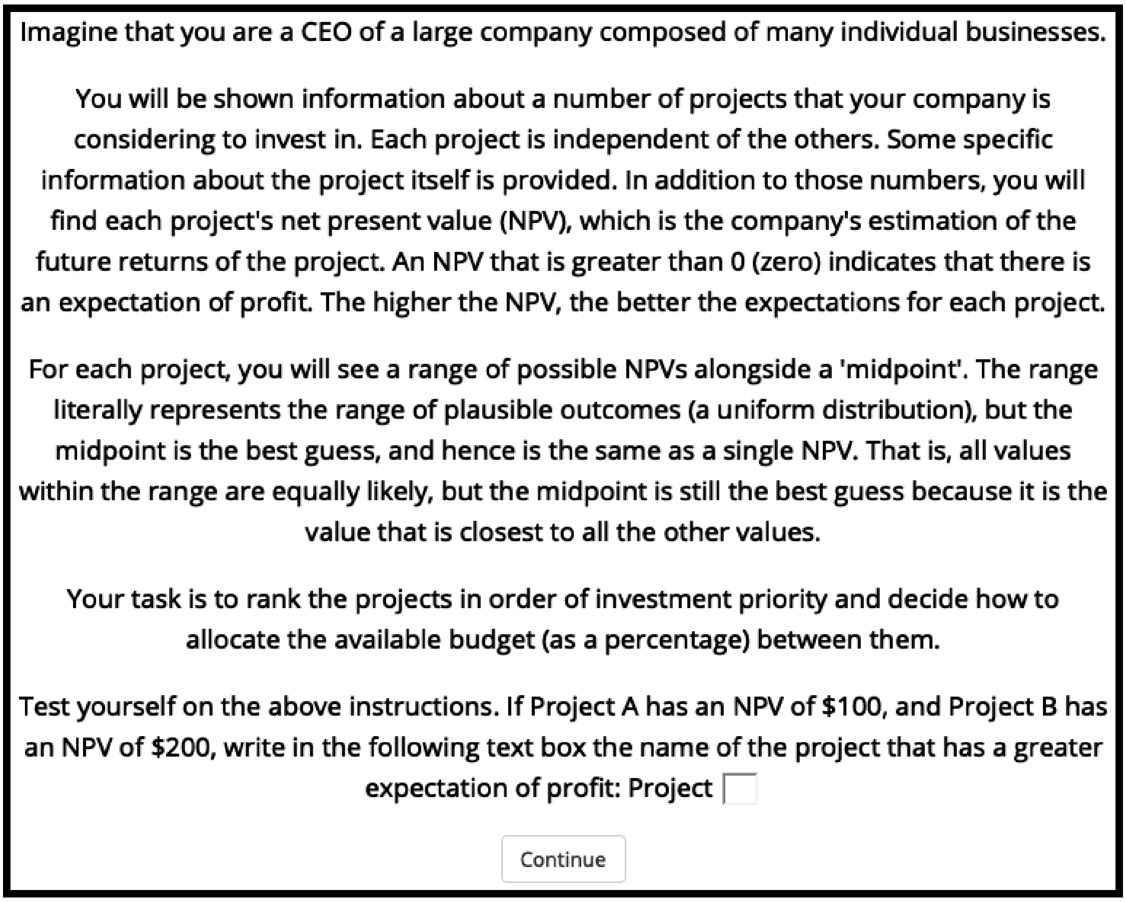


Figure B.36: Experiment 3 verbal reliability instructions.

experiment (see Appendix B.8) suggested that the linear NPV trend would be equivalent between those in the low and high alignment conditions, averaged over reliability level. However, the test of equivalence was not significant, $\Delta M = 15.19$, 95% CI [3.90, 26.48], $t(444) = 2.64$, $p = .996$.

B.4 Experiment 4

Experiment 4 further investigated the effects of alignment and verbal NPV reliability information on capital allocation decisions. Experiment 4 used the same methodology as in Experiment 1 (see Section 4.3.1), except for two main changes. First, the alignment conditions were manipulated within subjects. Second, the no NPV condition in the NPV reliability variable was removed.



Imagine that you are a CEO of a large company composed of many individual businesses.

You will be shown information about a number of projects that your company is considering to invest in. Each project is independent of the others. Some specific information about the project itself is provided. In addition to those numbers, you will find each project's net present value (NPV), which is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. The higher the NPV, the better the expectations for each project.

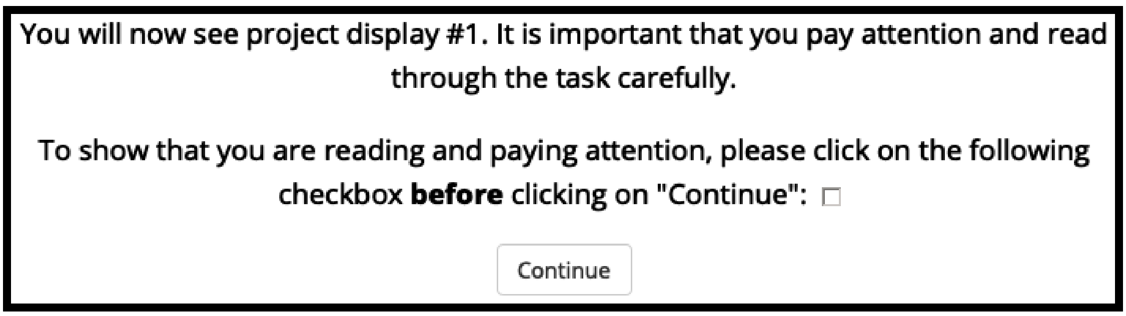
For each project, you will see a range of possible NPVs alongside a 'midpoint'. The range literally represents the range of plausible outcomes (a uniform distribution), but the midpoint is the best guess, and hence is the same as a single NPV. That is, all values within the range are equally likely, but the midpoint is still the best guess because it is the value that is closest to all the other values.

Your task is to rank the projects in order of investment priority and decide how to allocate the available budget (as a percentage) between them.

Test yourself on the above instructions. If Project A has an NPV of $100, and Project B has an NPV of $200, write in the following text box the name of the project that has a greater expectation of profit: Project

Continue

Figure B.37: Experiment 3 numerical reliability instructions.



You will now see project display #1. It is important that you pay attention and read through the task carefully.

To show that you are reading and paying attention, please click on the following checkbox **before** clicking on "Continue":

Continue

Figure B.38: An example of an interstitial display in Experiment 3.

Table B.1: Experiment 4 group allocation.

Reliability level of net present value (NPV)	N
High	34
Low	37
Total	71

The results of Experiment 1 were expected to replicate (see Section 4.3.2). Specifically, it was expected that in the high alignment condition, participants will be able to respond to each reliability condition, whereas, in the low alignment condition, they will rely more on NPV regardless of reliability condition.

In addition to the all-project allocation data analysed above, analyses for just the "target project" are also reported. This refers to allocation of capital to the project that had the highest NPV, but the lowest value on concrete measures intrinsic to the actual product (e.g., the capacity of a laptop in gigabytes). Therefore, a higher allocation value indicated a higher reliance on NPV. Further, the method and analyses for the confidence measure are also reported.

Hypothesis B.4. Participants will be more confident about their decisions in the high alignment condition than in the low alignment condition.

B.4.1 Method

B.4.1.1 Participants

Seventy-one participants (44 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 33.27 years (*SD* = 10.21, *min.* = 18, *max.* = 65). Table B.1 shows the allocation of participants to the different conditions. The two alignment conditions (low and high) were presented within subjects and the order of their presentation was randomised. Further, NPV was varied within subjects.

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided. In addition to those numbers, you will find each project's net present value (NPV), which is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. **The higher the NPV, the better the expectations for each project.** However, it is important to note that NPV is a very noisy measure relative to the other more specific measures because it relies on future forecasting. As such, **NPV is very unreliable and should be relied upon only as a last result; the specific project's measures should be used instead.**

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest objective value**. The features of the products that are listed matter because they reflect the direct value of the product, whereas financial measures such as NPV may reflect other factors, thus making it noisier, as mentioned above.

You will see a set of five different projects in each page, and for each set you must decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects. Critically, treat each set of projects as independent of one another; one page's project set allocation does not impact another page's allocation.

Figure B.39: Experiment 4 low reliability instructions.

B.4.1.2 Materials

The project display, allocation task, and confidence task were the same as in Experiment 1 (see Section 4.2.1.2).

B.4.1.2.1 Instructions Participants were shown similar instructions to Experiment 1 (see Section 4.2.1.2.1), except for the addition of references to the multiple displays and the removal of an explanation about the forecasting task. Figures B.39 and B.40 show the instructions for each NPV reliability condition.

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided. In addition to those numbers, you will find each project's net present value (NPV), which is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. **The higher the NPV, the better the expectations for each project.** NPV is a very useful measure relative to the other more specific measures because it can be calculated regardless of the type of product. As such, **NPV is very reliable in most cases**.

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest objective value.**

You will see a set of five different projects in each page, and for each set you must decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects. Critically, treat each set of projects as independent of one another; one page's project set allocation does not impact another page's allocation.

Figure B.40: Experiment 4 high reliability instructions.

B.4.1.3 Procedure

The procedure was the same as in Experiment 1, except that there were no forecasting or ranking tasks.

B.4.2 Results

A mixed factorial ANOVA was conducted to investigate the effects of alignment, verbal NPV reliability, and NPV on participants' project allocations. As seen in Figure B.41, the alignment × reliability level × NPV interaction was not significant, $F(3.64, 250.93) = 1.71$, $p = .153$, $\hat{\eta}_p^2 = .024$. This is most likely due to the fact that the reliability level × NPV interaction was significant in the high alignment condition, $\Delta M = -64.82$, 95% CI $[-102.70,\ -26.93]$, $t(69) = -3.41$, $p = .001$, the low alignment condition, $\Delta M = -37.74$, 95% CI $[-70.92,\ -4.56]$, $t(69) =$

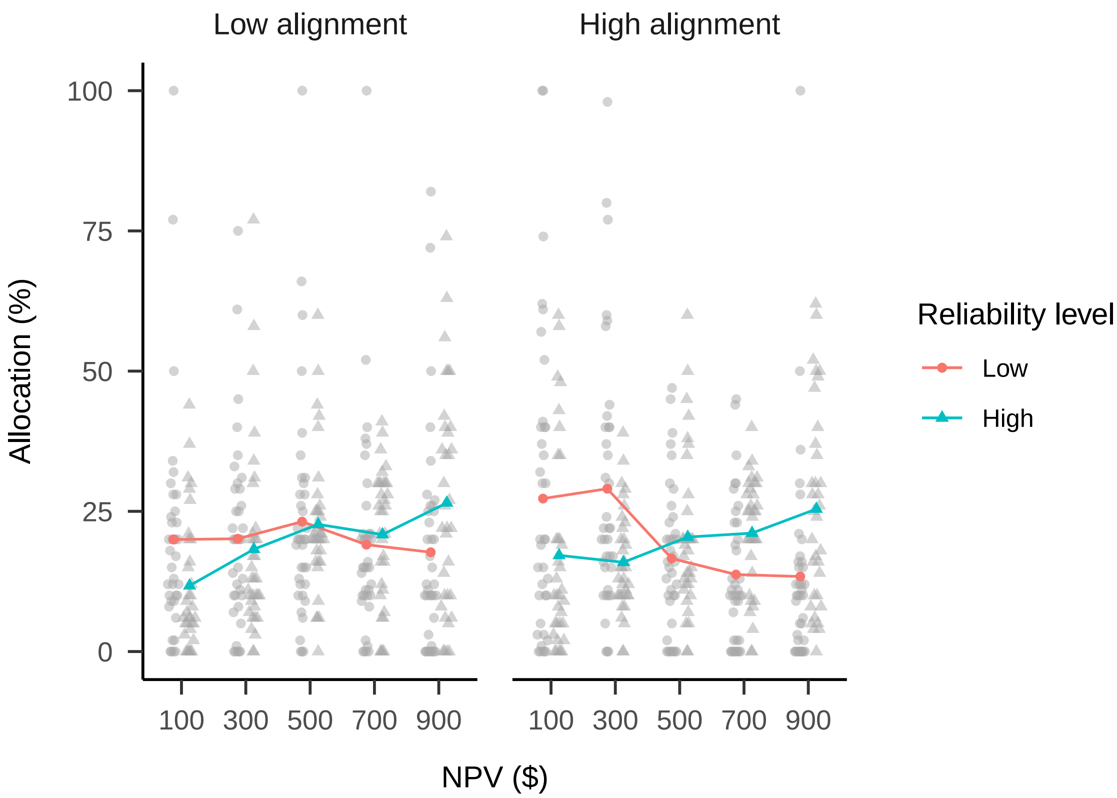


Figure B.41: Mean project allocation in Experiment 4. Error bars represent 95% confidence intervals based on the multivariate model. Note that this mixed factorial design does not allow for using confidence intervals to make inferences by "eye" across conditions.

-2.27, $p = .026$, as well as averaging over alignment conditions, $F(2.98, 205.65) = 4.90$, $p = .003$, $\hat{\eta}_p^2 = .066$. Despite this, the alignment $\times$ NPV interaction was significant, $F(3.64, 250.93) = 3.19$, $p = .017$, $\hat{\eta}_p^2 = .044$, such that the linear trend of NPV was stronger in the low alignment, $\Delta M = 13.28$, 95% CI $[-3.31,\ 29.87]$, $t(69) = 1.60$, $p = .115$ than in the high alignment condition, $\Delta M = -10.67$, 95% CI $[-29.62,\ 8.27]$, $t(69) = -1.12$, $p = .265$. However, neither of these trends were individually significant.

B.4.2.1 Confidence

A mixed factorial ANOVA was conducted to investigate the effects of alignment, verbal NPV reliability, and NPV on participants' confidence in their allocations. As shown in Figure B.42, the difference between alignment conditions was not significant, $F(1, 69) = 2.76$, $p = .101$, $\hat{\eta}_p^2 = .038$. However, the reliability $\times$

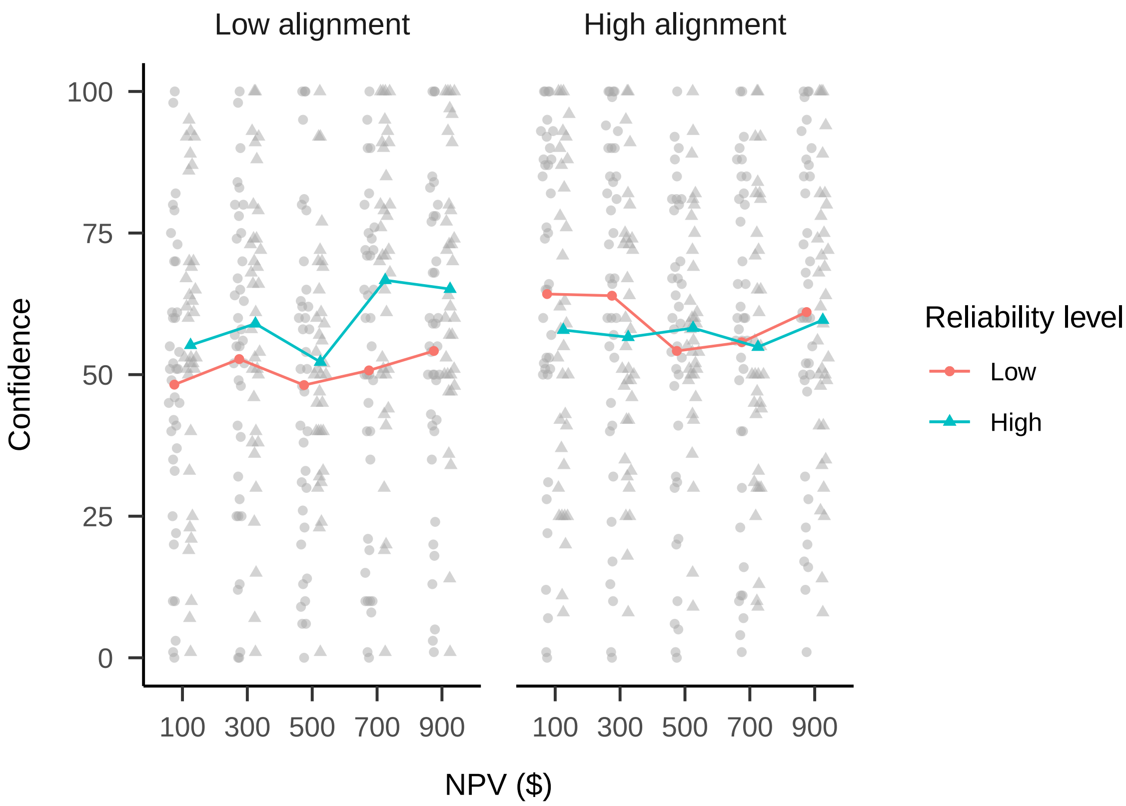


Figure B.42: Mean confidence. Error bars represent 95% confidence intervals based on the multivariate model. Note that this mixed factorial design does not allow for using confidence intervals to make inferences by "eye" across conditions.

alignment interaction was significant, as well as the NPV $\times$ alignment interaction. An exploratory analysis was conducted of the relevant simple effects for each interaction, applying a Šidák correction to the p values for each effect. None of the simple effects were significant after the correction.

The raw mean differences indicated that there was a greater difference between reliability conditions in the low alignment condition, $\Delta M = -8.83$, 95% CI $[-17.84,\ 0.18]$, $t(69) = -1.95$, $p = .055$ compared to the high alignment condition, $\Delta M = 2.37$, 95% CI $[-8.65,\ 13.40]$, $t(69) = 0.43$, $p = .669$. Further, there was a stronger linear trend of NPV in the low alignment condition, $\Delta M = 18.70$, 95% CI $[-0.87,\ 38.26]$, $t(69) = 2.44$, $p = .067$ compared to the high alignment condition, $\Delta M = -6.40$, 95% CI $[-26.84,\ 14.04]$, $t(69) = -0.80$, $p = .891$.

B.4.3 Discussion

Experiment 4 found evidence for most of the hypotheses. As per Hypothesis 4.4, laypeople responded appropriately to verbal reliability instructions in the high alignment condition. Contrary to Hypothesis 4.5, however, participants also did this in the low reliability condition. That is, regardless of the type of project display, participants tended to use NPV more when they were told that it was reliable and tended to use it less when they were told that it was unreliable. Further, there was no evidence that this effect was depended on alignment condition, contrary to Hypothesis 4.3. However, the linear NPV trend was higher in the high than low alignment condition, when averaging over reliability level, as predicted in Hypothesis 4.2. This suggests that overall participants still make more use of NPV information when it is hard to compare between projects.

Hypothesis B.4 was not supported, as there was no evidence of a main effect of alignment on participants' confidence in their allocation decisions. Instead, exploratory analyses showed that the difference in confidence between reliability conditions was greater in the low alignment condition. This may reflect participants' difficulty in making sense of their choices when alignment was low, given more confidence when assured of the reliability of NPV. In the high alignment condition, on the other hand, regardless of reliability condition, they had a way of using the reliability information. Further, confidence also seemed to increase more with NPV, on average, more when projects were dissimilar, which provides evidence for their reliance on NPV in this situation.

B.5 Experiment 5

Experiment 5 further investigated the effects of alignment and explicit NPV Presence information on forecasting. The goal of this experiment was to replicate the forecasting results of Experiment 1, but with a sample that has investing experience. As before, the hypothesis was that people's forecasting would be less

Table B.2: Experiment 5 group allocation.

Project alignment	Reliability level of net present value (NPV)	N
High	Absent	19
High	Present	17
Low	Absent	14
Low	Present	10
Total		60

variable when comparing projects with alignable differences, than when comparing projects with non-alignable differences.

B.5.1 Method

B.5.1.1 Participants

Sixty participants (2 female) were recruited from Reddit. Participants were compensated with a virtual Gold Award, which gives the recipient a week of a premium version of Reddit and 100 virtual coins. The average age was 28.17 years (SD = 8.73, *min.* = 16, *max.* = 61). Table B.2 shows the allocation of participants to the different conditions.

B.5.1.2 Materials

B.5.1.2.1 Risky Investment Task The only task that was used was the forecasting task used in Experiment 1, except that it was fixed by adding the relevant percentage intervals that were left out in Experiment 1, seen in Figure B.43.

B.5.1.3 Procedure

The procedure was the same as in Experiment 1, except participants only completed the forecasting task.

B.5.2 Results

B.5.2.1 Forecast Mean

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecasts. As shown in Figure B.44, the align-

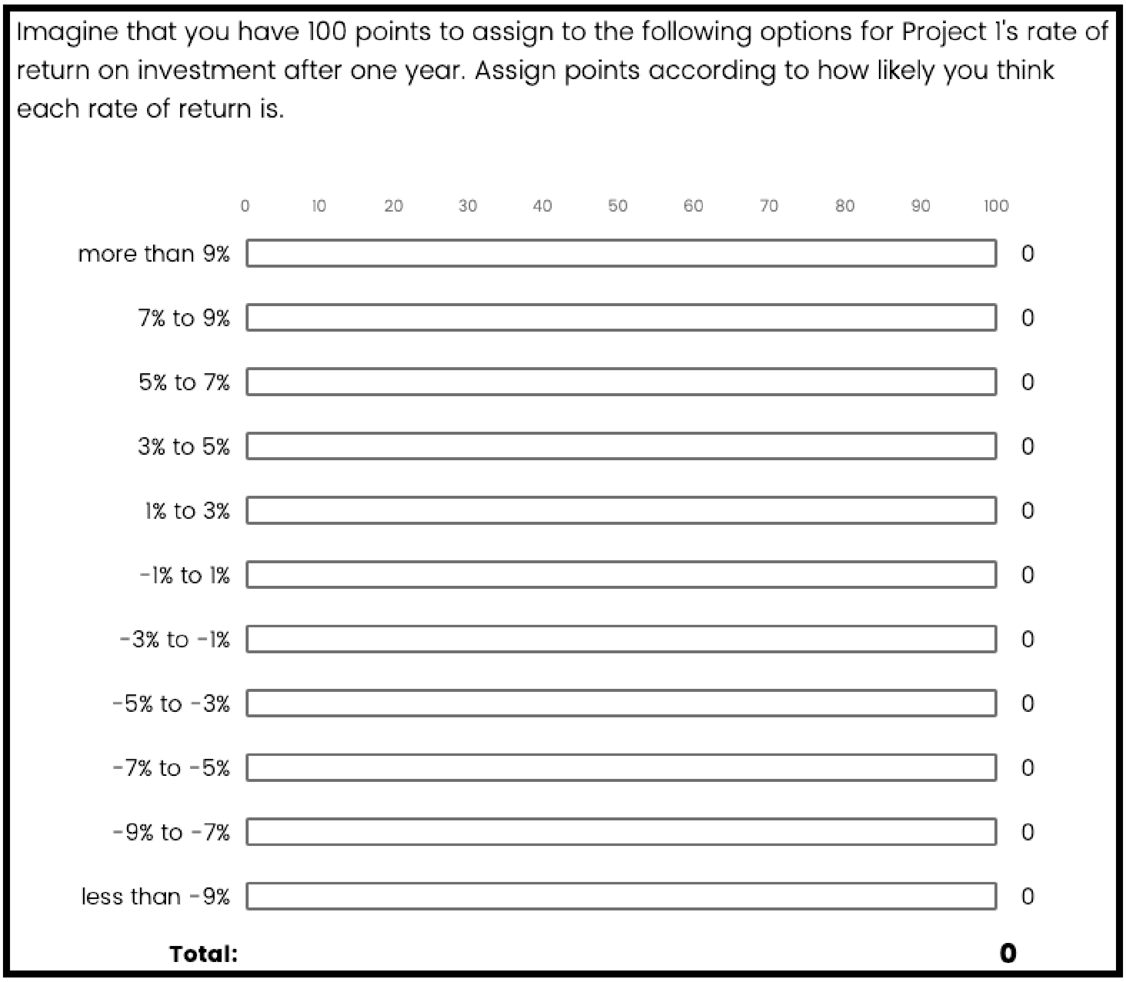


Figure B.43: An example of the forecasting task in Experiment 5.

ment × reliability level × NPV interaction was not significant, $F(2.75, 154.16) = 0.72$, $p = .531$, $\hat{\eta}_p^2 = .013$. Despite this, as in the previous experiments, the interaction between the linear NPV trend and NPV presence was significant in the high alignment condition, $M = -0.12$, 95% CI $[-0.21, -0.02]$, $t(56) = -2.50$, $p = .015$, but not in the low alignment condition, $M = -0.05$, 95% CI $[-0.16, 0.07]$, $t(56) = -0.81$, $p = .424$.

B.5.2.2 Forecast SD

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecast SDs. As shown in Figure B.45, there were no significant differences between alignment conditions, $F(1, 56) = 0.41$, $p = .522$, $\hat{\eta}_p^2 = .007$. The alignment × reliability level × NPV interaction was

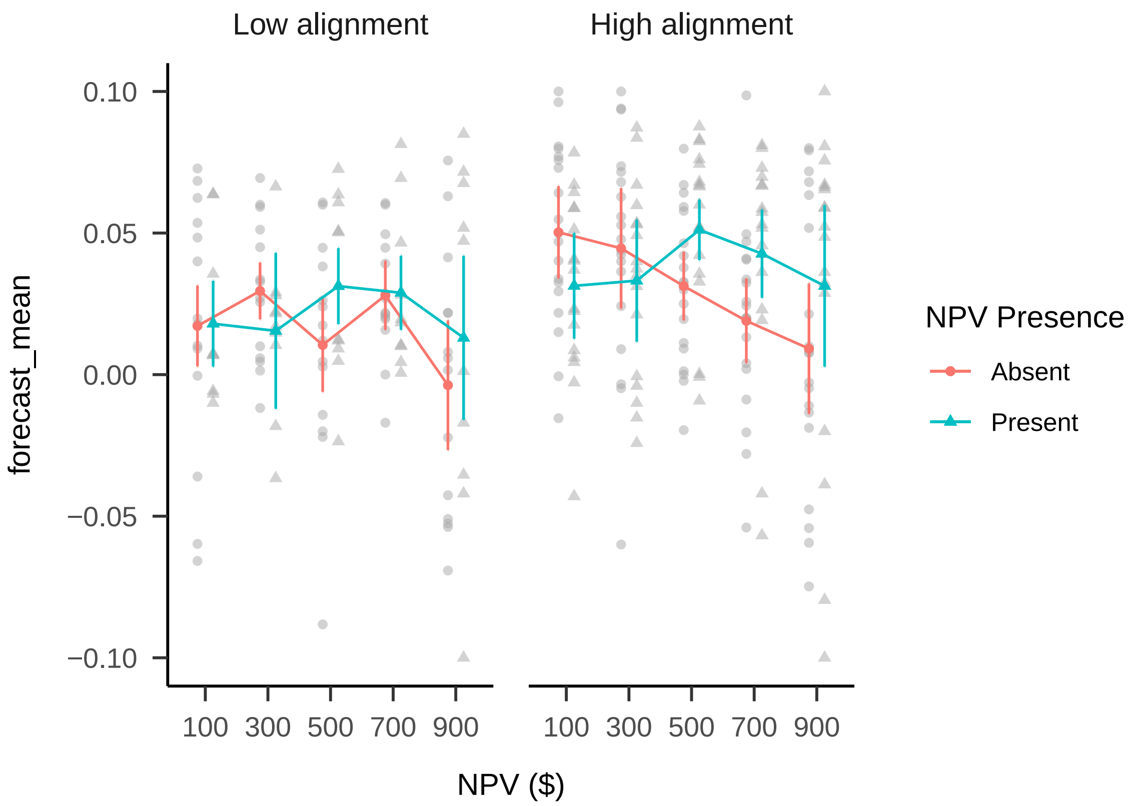


Figure B.44: Mean forecasts.

not significant, $F(2.99, 167.18) = 1.27$, $p = .287$, $\hat{\eta}_p^2 = .022$. However, as above, the interaction between the linear NPV trend and NPV presence was significant in the high alignment condition, $M = 0.02$, 95% CI $[0.00,\ 0.04]$, $t(56) = 2.06$, $p = .045$, but not in the low alignment condition, $M = 0.01$, 95% CI $[-0.02,\ 0.03]$, $t(56) = 0.38$, $p = .709$.

B.5.3 Discussion

Experiment 5 found that people with some investing experience responded to alignable information in the form of NPV when it is given, but did not show the same effect of alignment on forecast SD that was seen in Experiment 1.

B.6 Experiment 6

Experiment 6 further investigated the effects of alignment and NPV Presence information on forecasting. Experiment 5 did not clearly replicate the forecasting

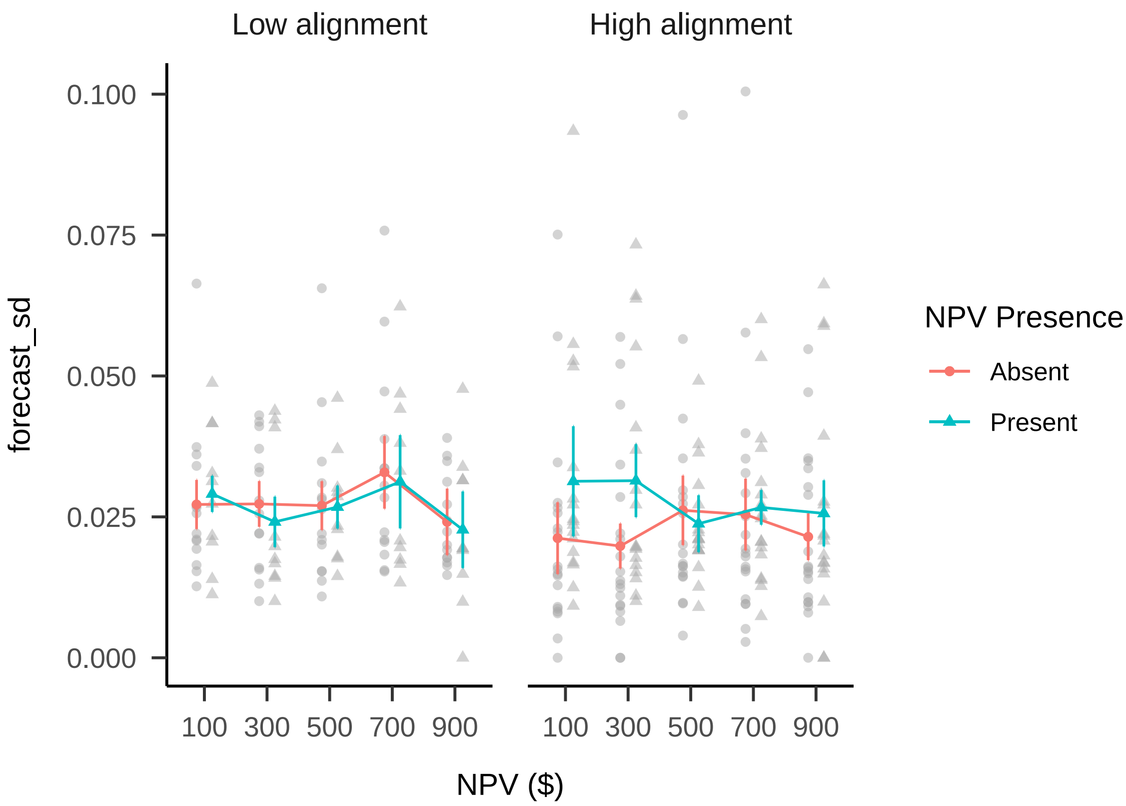


Figure B.45: Mean forecast SD.

results of Experiment 1, potentially due to low power, so this experiment collected a much larger sample size. As before, it was hypothesised that people's forecasting would be less variable when comparing projects with alignable differences, than when comparing projects with non-alignable differences.

B.6.1 Method

B.6.1.1 Participants

Three hundred and eighty-nine participants (170 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 32.39 years (*SD* = 11.89, *min.* = 18, *max.* = 75). Table B.3 shows the condition allocation.

B.6.1.2 Materials

The materials were the same as in Experiment 5.

Table B.3: Experiment 6 group allocation.

Project alignment	Reliability level of net present value (NPV)	N
High	Absent	97
High	Present	87
Low	Absent	101
Low	Present	104
Total		389

B.6.1.3 Procedure

The procedure was the same as in Experiment 5.

B.6.2 Results

B.6.2.1 Forecast Mean

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecasts. As shown in Figure B.46, the alignment $\times$ reliability level $\times$ NPV interaction was significant, $F(3.08, 1, 186.45) = 3.13$, $p = .024$, $\hat{\eta}_p^2 = .008$. As in the previous experiments, the interaction between the linear NPV trend and NPV presence was significant in both the high alignment condition, $M = -0.13$, 95% CI $[-0.16, -0.09]$, $t(385) = -6.57$, $p < .001$, and in the low alignment condition, $M = -0.06$, 95% CI $[-0.09, -0.02]$, $t(385) = -3.28$, $p = .001$.

B.6.2.2 Forecast SD

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecast SDs. As shown in Figure B.47, the alignment $\times$ reliability level $\times$ NPV interaction was not significant, $F(3.45, 1, 328.06) = 0.82$, $p = .496$, $\hat{\eta}_p^2 = .002$. The main effect of alignment was not significant, $F(1, 385) = 0.64$, $p = .424$, $\hat{\eta}_p^2 = .002$.

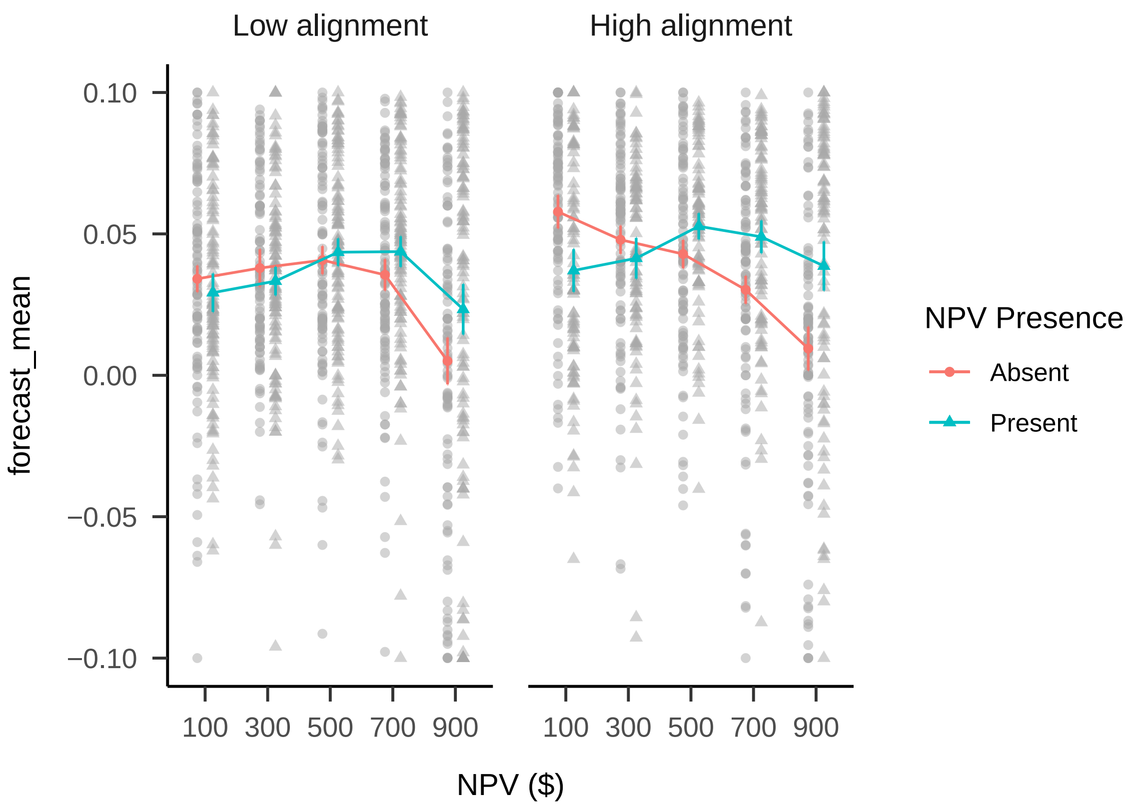


Figure B.46: Mean forecasts.

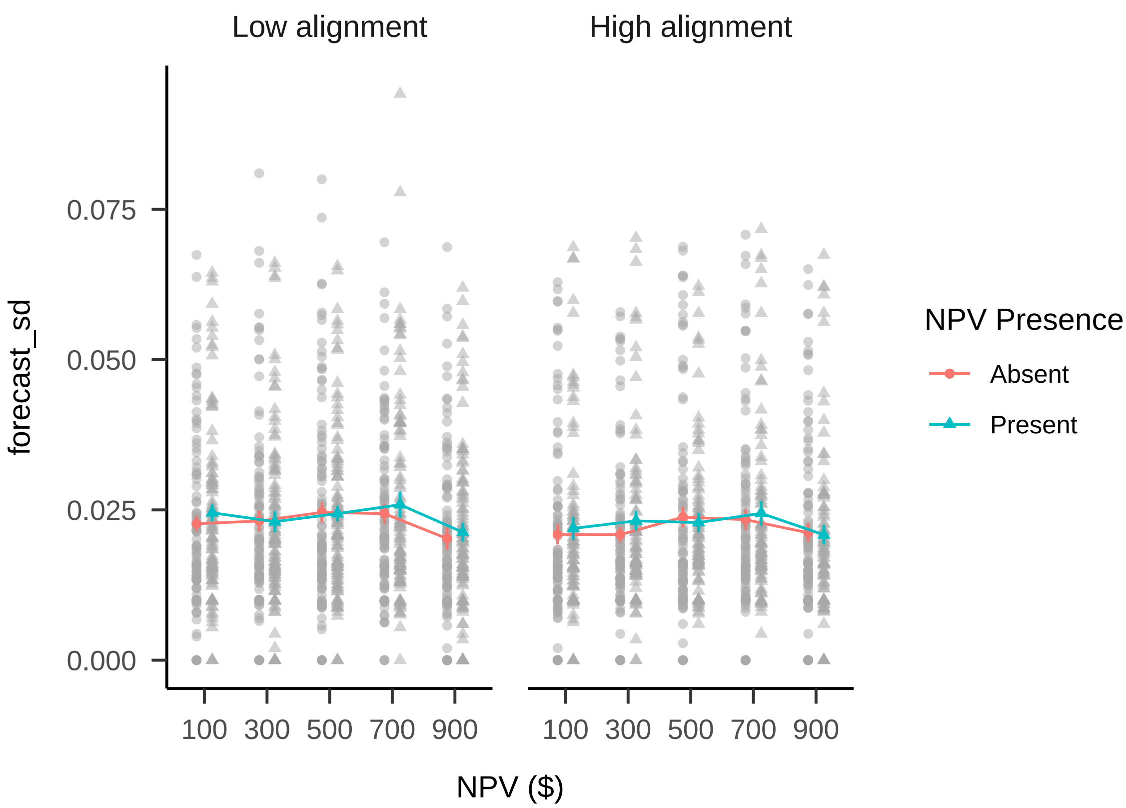


Figure B.47: Mean forecast SD.

B.6.3 Discussion

Experiment 6 did not replicate the effect of alignment on forecast SD seen in Experiment 1. However, participants still seemed to pay attention to the task, as seen in their higher forecasts for the high NPV project when NPV was present.

B.7 Experiment 7

Experiment 7 investigated potential ways to facilitate people's use of variance in capital allocation. Arguably, people's decisions should depend on variance, especially with a small set of projects. That is, when considering between two potential measures to use for capital allocation, measures with narrow ranges should be relied upon more than those with wider ranges. As such, this experiment presented participants with the same capital allocation scenario as in Experiment 2, but only in low numerical reliability displays. Experiment 7 varied both the variance associated with NPV, and the extent to which participants were explicitly hinted to use the variance information. It was predicted that participants' allocations would be more likely to be informed by variance when told explicitly to do so with increased salience for variance, than when only salience is increase, or when no hint is given.

B.7.1 Method

B.7.1.1 Participants

Seventy-nine participants (35 female) were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour (Prolific is based in the UK). The average age was 31.15 years ($SD = 11.11$, $min. = 16$, $max. = 71$). Table B.4 shows the allocation of participants to the different conditions.

B.7.1.2 Instructions

As shown in Figure B.48, participants in the no hint condition saw the same instructions as in Experiment 1. As shown in Figure B.49, those in the salience

Table B.4: Experiment 7 group allocation.

Hint	Variance	N
Hint salience	High	11
Hint salience	Low	11
No hint	High	9
No hint	Low	13
Salience only	High	19
Salience only	Low	16
Total		79

only condition saw the instructions along with a sentence that drew attention to the *Cash inflow range* row. As shown in Figure B.50, those in the salience + hint condition saw the instructions along with a specific description of how to use the variance information in their allocation decisions.

B.7.1.3 Project Display

The project displays were the same as Experiment 2 (see Figure B.51).

B.7.1.4 Procedure

Participants read the instruction page as per their hint condition, and then proceeded to complete one set of ranking and allocations.

B.7.2 Results

B.7.2.1 Allocation

A mixed factorial ANOVA was conducted to investigate the effects of hint and NPV variance on participants' allocations. As shown in Figure B.52, none of the interactions or main effects were significant.

B.7.2.2 Ranking

A mixed factorial ANOVA was conducted to investigate the effects of hint and NPV variance on participants' project rankings. As shown in Figure B.53, only the main effect of NPV was significant, $F(2.03, 148.33) = 7.59$, $p = .001$, $\hat{\eta}_p^2 = .094$.

Allocation task

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided.

In addition to those numbers, you will find each project's projected cash inflow for the first year (the money that it is expected to generate), and the net present value (NPV) that was calculated using those figures. It is usually unclear exactly what the future cash inflow is, so instead of a single cash inflow value, it will be given as a range (assume that all the values in that range are equally likely). Also assume that all the other elements that are required to calculate NPV (i.e., the discount rate and initial investment) are identical for all projects.

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest intrinsic quality.**

You will decide how to rank the projects in order of investment priority, and decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects.

Figure B.48: Instructions for the no hint condition.

B.7.3 Discussion

Experiment 7 found that explicitly telling participants how to use variance information to inform their allocations did not help them do so. However, there was an increased reliance on NPV with more hints in the ranking data. This suggests that the hint manipulations potentially simply increase participants' attention to NPV. It is possible that the study was under-powered, as there was substantial variance in both the allocation and ranking data. Future work should attempt to replicate this experiment with a larger sample.

Allocation task

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided.

In addition to those numbers, you will find each project's projected cash inflow for the first year (the money that it is expected to generate), and the net present value (NPV) that was calculated using those figures. It is usually unclear exactly what the future cash inflow is, so instead of a single cash inflow value, it will be given as a range (assume that all the values in that range are equally likely). Also assume that all the other elements that are required to calculate NPV (i.e., the discount rate and initial investment) are identical for all projects.

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest intrinsic quality.**

You will decide how to rank the projects in order of investment priority, and decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects.

Pay special attention to the cash inflow ranges as they are important to the decision making process.

Figure B.49: Instructions for the salience only condition.

B.8 Experiment 8

Experiment 8 tested the alignment and reliability effects found in the previous experiments, while addressing their limitations. Experiments 1 and 4 found a verbal reliability effect. That is, laypeople allocated more capital to a high NPV project, depending on how reliable they were told NPV was as a measure. Experiment 2 found a lack of a numerical reliability effect. That is, business students allocated an equivalent amount of capital to projects associated with a high variance NPV, as projects with a low NPV. Testing these two effects in two different populations

Allocation task

You will be shown information about a number of projects that a consumer products firm is considering to invest in. Some specific information about the product itself is provided.

In addition to those numbers, you will find each project's projected cash inflow for the first year (the money that it is expected to generate), and the net present value (NPV) that was calculated using those figures. It is usually unclear exactly what the future cash inflow is, so instead of a single cash inflow value, it will be given as a range (assume that all the values in that range are equally likely). Also assume that all the other elements that are required to calculate NPV (i.e., the discount rate and initial investment) are identical for all projects.

We would like you to take the role of the manager in charge of capital allocation for the firm. This firm is specifically interested in investing in the development of high-end goods, so your valuations should reflect this. That is, even though there might be a market for the lower-end products in the descriptions that you will see, **you should be aiming to invest in the products with the highest intrinsic quality.**

You will decide how to rank the projects in order of investment priority, and decide how to allocate the capital available for investment this year among the different projects. Note that this is not the operational budget (advertising, etc.), but rather the funds to be used for investment in developing the new products. You will do this by selecting a percentage value for each project, such that the budget is allocated completely among each set of projects.

Pay special attention to the cash inflow ranges, because they imply the extent to which you should be relying on that particular NPV. NPVs with higher variance (greater cash inflow ranges) should be relied upon less. For instance, imagine two NPVs, one with a future cash flow range of $100-$1900 (range of ±90% around the average), and one with a range of $900-$1100 (range ±10% around the average). The average of each range is the same ($1000), and yet the first estimate is more uncertain than the second. As such, with the first estimate, other factors should be used more in the decision making than the NPV, while with the second estimate, the NPV can be relied on more confidently. In general, ranges of less than 10% either way of the average are considered very low variance, and those more than 80% are considered very high variance.

Figure B.50: Instructions for the salience + hint condition.

	Project 1	Project 2	Project 3	Project 4	Project 5
Product	Laptop	Laptop	Laptop	Laptop	Laptop
RAM (GB)	4	8	32	2	16
Hard drive (GB)	500	750	2000	250	1000
Resolution (px)	900	1080	1440	768	1200
Processor (Ghz)	2.4	3.2	3.8	1.6	3.6
Cash inflow range for Year 1 ($)	$5,890-$6,510	$5,738-$6,342	$5,244-$5,796	$6,137-$6,783	$5,538.50-$6,121.50
NPV ($)	$636.36	$490.91	$18.18	$872.73	$300.00

Figure B.51: The projects display.

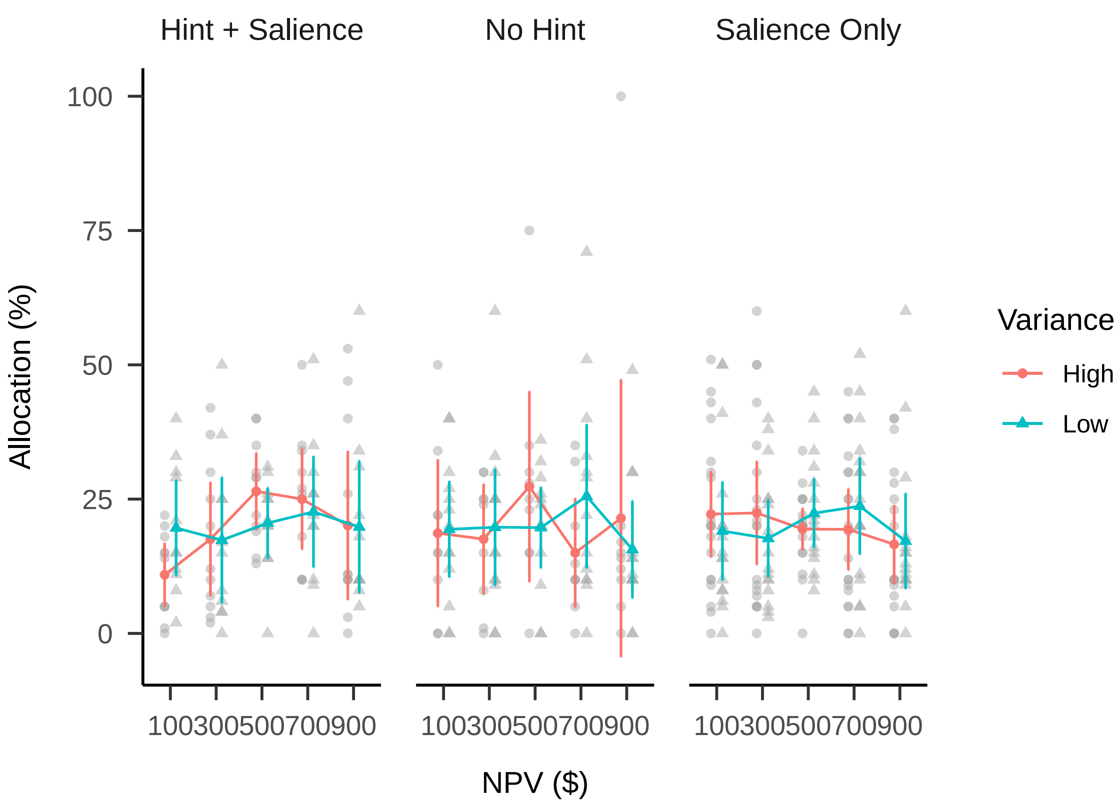


Figure B.52: Mean allocation.

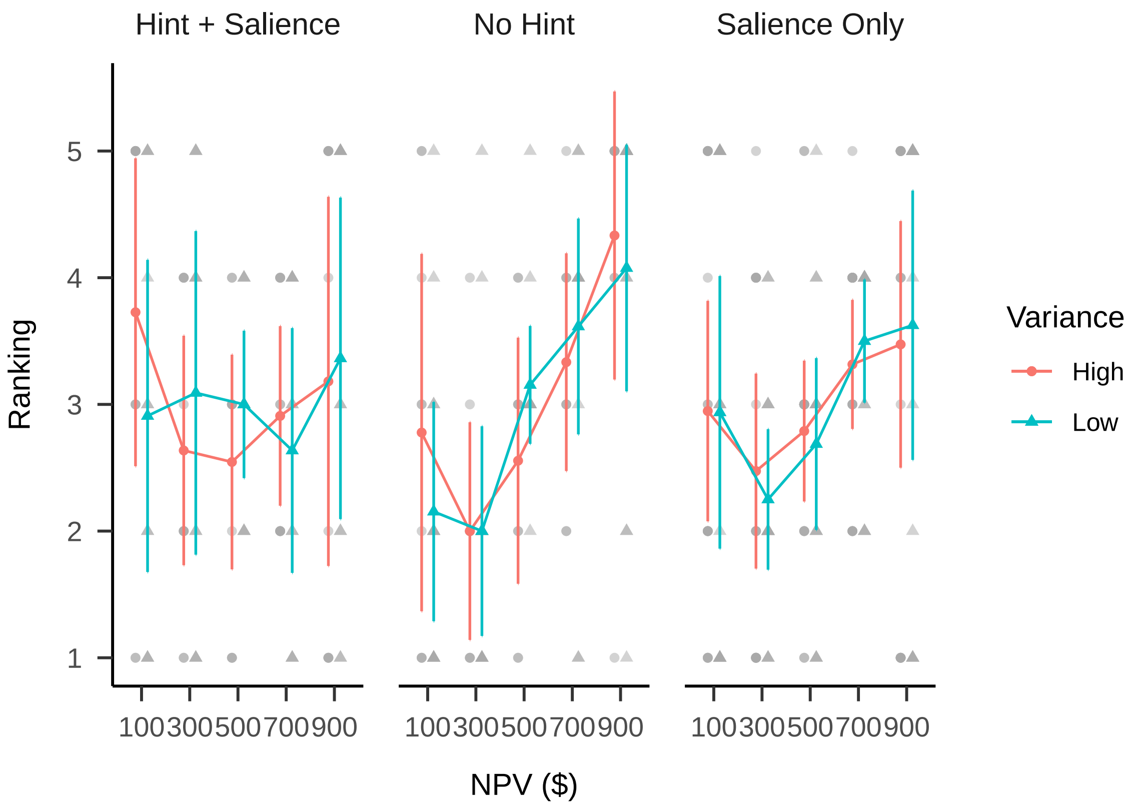


Figure B.53: Mean ranking.

did not account for potential expertise effects. As such, Experiment 8 tested both effects with a naive sample. Further, Experiment 8 used projects whose features more clearly indicate their profitability, and included more project domains.

B.8.1 Method

B.8.1.1 Participants

Fifty-two participants (33 female) were recruited from both the online recruitment platform Prolific and a cohort of psychology undergraduates at The University of Sydney. Participants from Prolific were compensated at a rate of £5 an hour (Prolific is based in the UK), and participants from the undergraduate sample were compensated with course credit. The average age was 24.46 years (*SD* = 7.77, *min.* = 18, *max.* = 68). Participants reported an average of 2.63 years (*SD* = 4.16, *min.* = 0, *max.* = 25) working in a business setting, and an average of 0.81 years (*SD* = 1.39, *min.* = 0, *max.* = 5) of business education. The mean completion time of the task was 35.57 min (*SD* = 71.96, *min.* = 7.36, *max.* = 511.74). All conditions were presented within-subjects: alignment (low and high), NPV reliability type (numerical and verbal), NPV (low and high), and NPV reliability level (low and high).

B.8.1.2 Materials

B.8.1.2.1 Instructions Participants saw instructions similar to the previous experiments.

B.8.1.2.2 Project Display Participants saw and responded to four webpage displays. At the top of each display was a text preamble, and underneath this a table that contained project descriptions. The two columns to the right of each description contained text boxes for participants to enter a value for the project ranking and budget allocation. Alignment was manipulated by asking participants to either compare between each of the project pairs (high alignment), or across all eight projects in the display (low alignment). For instance, in the high alignment display, participants had to compare between two railway projects, and then separately between two logistics projects, etc. However, in the low alignment display, participants had to compare railway projects to logistics projects directly. This was manipulated within-subjects, such that project descriptions were identical across alignment conditions and only the type of comparison (and the associated preamble text) varied.

Figures B.54, B.55, B.56, B.57 show the four conditions that participants saw (counterbalanced). Each description provided the name of the business involved in the project, the type of project, three specific features of the project, an NPV, and an indication of reliability (either numerical through ranges or verbal through explicit labels).

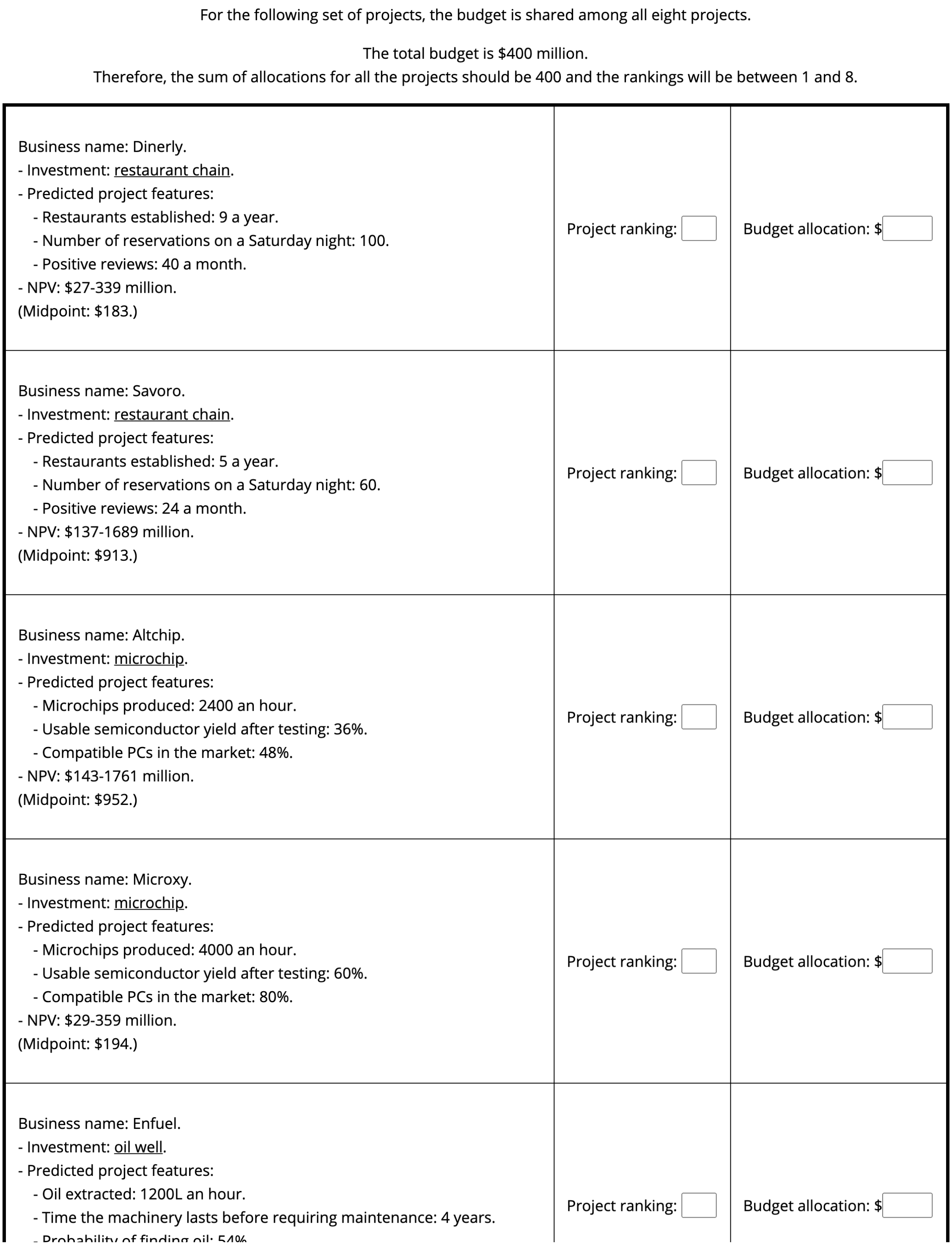
The value of each type of reliability was also manipulated. Explicit reliability was manipulated by varying whether participants were told that a project pair was in an industry in which NPV is considered a reliable or unreliable measure. Implicit reliability was manipulated by presenting NPVs alongside numerical ranges instead

For the following set of projects, the budget is shared among all eight projects.

The total budget is $400 million.
Therefore, the sum of allocations for all the projects should be 400 and the rankings will be between 1 and 8.

Business name: FreightCog. - Investment: railway. - Predicted project features: - Railway lines built: 5 a decade. - Number of seats filled by paying customers at peak hour: 2000. - Time before the train carriages will need to be serviced: 12 years. - NPV: $128 million. (In this particular industry, NPV is a **reliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: Railmont. - Investment: railway. - Predicted project features: - Railway lines built: 3 a decade. - Number of seats filled by paying customers at peak hour: 1200. - Time before the train carriages will need to be serviced: 7 years. - NPV: $974 million. (In this particular industry, NPV is a **reliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: Pharmacore. - Investment: pharmaceutical. - Predicted project features: - Pills pressed: 180000 an hour. - Shelf life: 12 months. - Probability of symptom reduction after a week: 54%. - NPV: $952 million. (In this particular industry, NPV is an **unreliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: Biotechly. - Investment: pharmaceutical. - Predicted project features: - Pills pressed: 300000 an hour. - Shelf life: 20 months. - Probability of symptom reduction after a week: 90%. - NPV: $194 million. (In this particular industry, NPV is an **unreliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: Pressbloom. - Investment: national newspaper. - Predicted project features: - Newspapers printed: 30000 a day. - Number of weekly advertisers: 48. - Ink that is not discarded due to impurities: 3000L a day	Project ranking:	Budget allocation: $

Figure B.54: Experiment 8 low alignment, verbal reliability display. Cropped for space (full display had eight projects).



For the following set of projects, the budget is shared among all eight projects.

The total budget is $400 million.
Therefore, the sum of allocations for all the projects should be 400 and the rankings will be between 1 and 8.

Business name: Dinerly. - Investment: restaurant chain. - Predicted project features: - Restaurants established: 9 a year. - Number of reservations on a Saturday night: 100. - Positive reviews: 40 a month. - NPV: $27-339 million. (Midpoint: $183.)	Project ranking:	Budget allocation: $
Business name: Savoro. - Investment: restaurant chain. - Predicted project features: - Restaurants established: 5 a year. - Number of reservations on a Saturday night: 60. - Positive reviews: 24 a month. - NPV: $137-1689 million. (Midpoint: $913.)	Project ranking:	Budget allocation: $
Business name: Altchip. - Investment: microchip. - Predicted project features: - Microchips produced: 2400 an hour. - Usable semiconductor yield after testing: 36%. - Compatible PCs in the market: 48%. - NPV: $143-1761 million. (Midpoint: $952.)	Project ranking:	Budget allocation: $
Business name: Microxy. - Investment: microchip. - Predicted project features: - Microchips produced: 4000 an hour. - Usable semiconductor yield after testing: 60%. - Compatible PCs in the market: 80%. - NPV: $29-359 million. (Midpoint: $194.)	Project ranking:	Budget allocation: $
Business name: Enfuel. - Investment: oil well. - Predicted project features: - Oil extracted: 1200L an hour. - Time the machinery lasts before requiring maintenance: 4 years. - Probability of finding oil: 54%	Project ranking:	Budget allocation: $

Figure B.55: Experiment 8 low alignment, numerical reliability display. Cropped for space (full display had eight projects).

For the following set of projects, the budget is split up evenly between each industry pair, i.e., projects with the same type of "Investment".

The total budget is $400 million. Therefore, the sum of allocations in each pair should be 100 and the rankings will be between 1 and 2.

Business name: Erectic. - Investment: high-rise construction. - Predicted project features: - High-rises built: 5 a year. - Probability that the builders complete construction within a month of the due date: 42%. - Number of tenant expressions of interest: 60. - NPV: $913 million. (In this particular industry, NPV is an **unreliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: Refit. - Investment: high-rise construction. - Predicted project features: - High-rises built: 8 a year. - Probability that the builders complete construction within a month of the due date: 70%. - Number of tenant expressions of interest: 100. - NPV: $183 million. (In this particular industry, NPV is an **unreliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: Pressbloom. - Investment: national newspaper. - Predicted project features: - Newspapers printed: 30000 a day. - Number of weekly advertisers: 48. - Ink that is not discarded due to impurities: 3000L a day. - NPV: $964 million. (In this particular industry, NPV is a **reliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: Grown Media. - Investment: national newspaper. - Predicted project features: - Newspapers printed: 50000 a day. - Number of weekly advertisers: 80. - Ink that is not discarded due to impurities: 5000L a day. - NPV: $192 million. (In this particular industry, NPV is a **reliable** predictor of project success.)	Project ranking:	Budget allocation: $
Business name: FreightCog. - Investment: railway. - Predicted project features: - Railway lines built: 5 a decade.	Project ranking:	Budget allocation: $

Figure B.56: Experiment 8 high alignment, verbal reliability display. Cropped for space (full display had eight projects).

For the following set of projects, the budget is split up evenly between each industry pair, i.e., projects with the same type of "Investment".

The total budget is $400 million. Therefore, the sum of allocations in each pair should be 100 and the rankings will be between 1 and 2.

Business name: Enfuel. - Investment: oil well. - Predicted project features: - Oil extracted: 1200L an hour. - Time the machinery lasts before requiring maintenance: 4 years. - Probability of finding oil: 54%. - NPV: $916-1012 million. (Midpoint: $964.)	Project ranking:	Budget allocation: $
Business name: Refinera. - Investment: oil well. - Predicted project features: - Oil extracted: 2000L an hour. - Time the machinery lasts before requiring maintenance: 7 years. - Probability of finding oil: 90%. - NPV: $182-202 million. (Midpoint: $192.)	Project ranking:	Budget allocation: $
Business name: Altchip. - Investment: microchip. - Predicted project features: - Microchips produced: 2400 an hour. - Usable semiconductor yield after testing: 36%. - Compatible PCs in the market: 48%. - NPV: $143-1761 million. (Midpoint: $952.)	Project ranking:	Budget allocation: $
Business name: Microxy. - Investment: microchip. - Predicted project features: - Microchips produced: 4000 an hour. - Usable semiconductor yield after testing: 60%. - Compatible PCs in the market: 80%. - NPV: $29-359 million. (Midpoint: $194.)	Project ranking:	Budget allocation: $
Business name: Solgistics. - Investment: shipping logistics. - Predicted project features: - Packages shipped: 480 a week. - Number of packages that do not spend time in a bottleneck: 240 a day. - Average accuracy of shipments: 57%.	Project ranking:	Budget allocation: $

Figure B.57: Experiment 8 high alignment, numerical reliability display. Cropped for space (full display had eight projects).

of verbal reliability information about them, and varying whether the range was high or low. Both of these were manipulated within-display, such that NPV was reliable for four projects in each display, and NPV was unreliable for the other four.

Each project had an associated NPV, which was crossed with each project pair's intrinsic features. That is, each pair had one project with a high NPV and low intrinsic feature values, and one project with a low NPV and high intrinsic feature values. As such, a reliance on NPV was inferred if participants allocated the high NPV project more capital, or a reliance on the intrinsic features if participants allocated the low NPV project more capital.

B.8.1.3 Procedure

Participants viewed the instructions and then completed the ranking and allocation tasks in the four sets of project descriptions. The order of the display was counterbalanced, and the order of the project pairs on each page was randomised.

B.8.2 Results

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV reliability type on participants project allocations. A direct comparison of the two alignment conditions was not possible due to the different allocation input scales, so the NPV reliability level × NPV interaction was tested separately in each alignment condition (see Figures B.58 and B.59). This interaction was significant for both the high alignment condition, $F(1, 51) = 27.81$, $p < .001$, $\hat{\eta}_p^2 = .353$; and the low alignment condition, $F(1, 51) = 7.63$, $p = .008$, $\hat{\eta}_p^2 = .130$. However, there was a significant effect of NPV in the low verbal reliability condition in high alignment, $\Delta M = 18.69$, 95% CI $[2.87,\ 34.52]$, $t(113.10) = 3.17$, $p = .012$; but not in low alignment, $\Delta M = 6.04$, 95% CI $[-9.24,\ 21.32]$, $t(121.35) = 1.06$, $p > .999$.

B.8.3 Discussion

Experiment 8 found that when variance was presented verbally, participants allocated according to the reliability information, for both low and high alignment

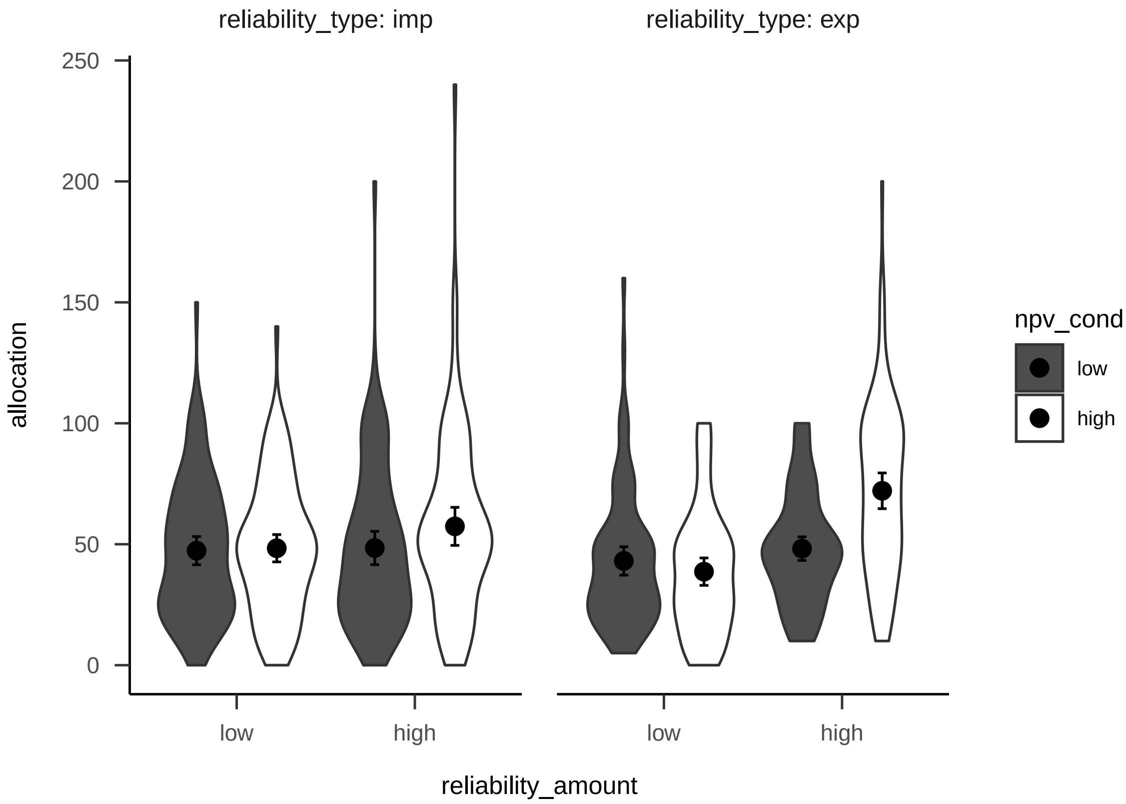


Figure B.58: Mean project allocation, for the low alignment condition. Error bars represent 95% confidence intervals.

conditions. When variance was presented numerically, there were no differences in allocations, for both low and high alignment conditions. Further, there was an effect of NPV in low reliability for the high alignment condition, but not the low alignment condition. This effect shows that people still relied on NPV more than they should when comparing across dissimilar projects.

This experiment shows that similar to the previous experiments, when controlling for presentation and domain, people still find it easier to allocate capital based on explicit reliability information when projects are comparable. However, due to the difference in scale across alignment conditions, a direct alignment effect was more difficult to test than with the previous experiments. Further, similar to Experiment 2, Experiment 8 showed that people without much business experience also struggle to use range information in capital allocation to such an extreme extent that they do not seem to be using any coherent allocation strategy.

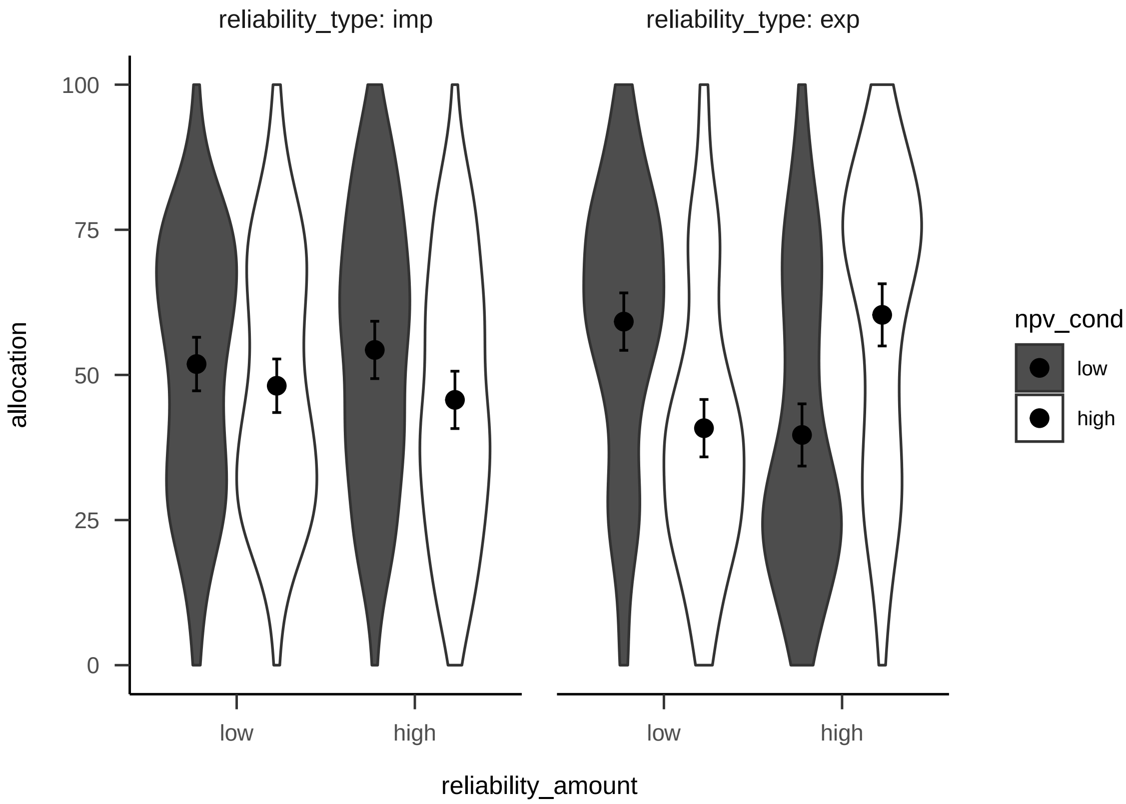


Figure B.59: Mean project allocation, for the high alignment condition. Error bars represent 95% confidence intervals.

C
Chapter 6 Appendix

Contents

This appendix contains supplementary materials and analyses for the two experiments reported in Chapter 6.

C.1 Experiment 1

Below are hypotheses that were tested, but were not sufficiently relevant for Chapter 6 to be reported in the main text.

Hypothesis C.1—Allocation similarity manipulation check for negative anecdote. For negative anecdotes, allocations for the anecdote only low similarity condition will be higher than those in the anecdote only high similarity condition.

Hypothesis C.2—Relationship between allocation and perceived similarity for negative anecdote. In the negative valence condition, the correlation between allocation and similarity rating will be negative

Hypothesis C.3—Relationship between allocation and specific-relevance for negative anecdote. In the negative valence condition, there will be no correlation between allocation and specific-relevance rating in the low similarity condition, but a negative correlation in the high similarity condition.

After the allocation task, participants were asked to rate the relevance of the anecdote to the target project. It was predicted that those that saw only an anecdote would be more influenced by the similarity of the anecdote than those that saw an anecdote as well as statistics. Therefore, the following hypotheses were tested:

Hypothesis C.4. The similarity effect on specific relevance will be greater in the anecdote only condition than in the anecdote + statistics condition.

Hypothesis C.5. The similarity effect on specific relevance will be greater in the statistics + anecdote condition than in the anecdote + enhanced statistics condition.

Further, participants were asked to rate the relevance of the anecdote to other projects in the same industry. It was predicted that those that saw only an anecdote would be more influenced by the similarity of the anecdote than those that saw an anecdote as well as statistics. Therefore, the following hypotheses are tested:

Hypothesis C.6. The similarity effect on general relevance will be greater in the anecdote only condition than in the anecdote + statistics condition.

Hypothesis C.7. The similarity effect on general relevance will be greater in the statistics + anecdote condition than in the anecdote + enhanced statistics condition.

C.1.1 Method

C.1.1.1 Participants

C.1.1.1.1 Power Analysis The sample size for Experiment 1 was determined by conducting power analyses using the `Superpower` package (Lakens & Caldwell, 2019). The package uses experimental design, and predicted means and standard deviation, to conduct a priori power calculations. Data from Wainberg (2018), Jaramillo et al. (2019), and Hoeken and Hustinx (2009, Study 3) was used to determine realistic means and standard deviations for the evidence and similarity factors. According to the power functions, the resulting sample size is assumed to allow for an expected power of at least 80%.

Data from Wainberg (2018) were used to determine the predicted means for the anecdote conditions. Specifically, the values for the high similarity condition were taken from the anecdote & statistics, anecdote & enhanced statistics, and statistics only conditions for the corresponding anecdote conditions. This was done because in Wainberg (2018) the anecdote was always of a similar case. Wainberg (2018) did not use an anecdote only condition, but Wainberg et al. (2013) did and found no significant differences between the anecdote only condition and the anecdote & statistics condition. As such, the same mean value was used for both conditions.

It was hypothesised that there will only be an effect of similarity for the anecdote only and anecdote & statistics conditions. As such, the data from Hoeken and Hustinx (2009, Study 3) were used to determine the corresponding mean values for the low similarity condition. Specifically, each predicted mean was multiplied by the Cohen's d_z of the similarity effect in Hoeken and Hustinx (2009, Study 3).

To determine the predicted standard deviation, the data from Jaramillo et al. (2019) Experiment 2 and Hoeken and Hustinx (2009, Study 3) were re-analysed to determine the coefficient of variation (CV) of each condition. Each CV was then converted to a standard deviation value in the relevant scale by multiplying the mean of the CV values by the predicted means from above.

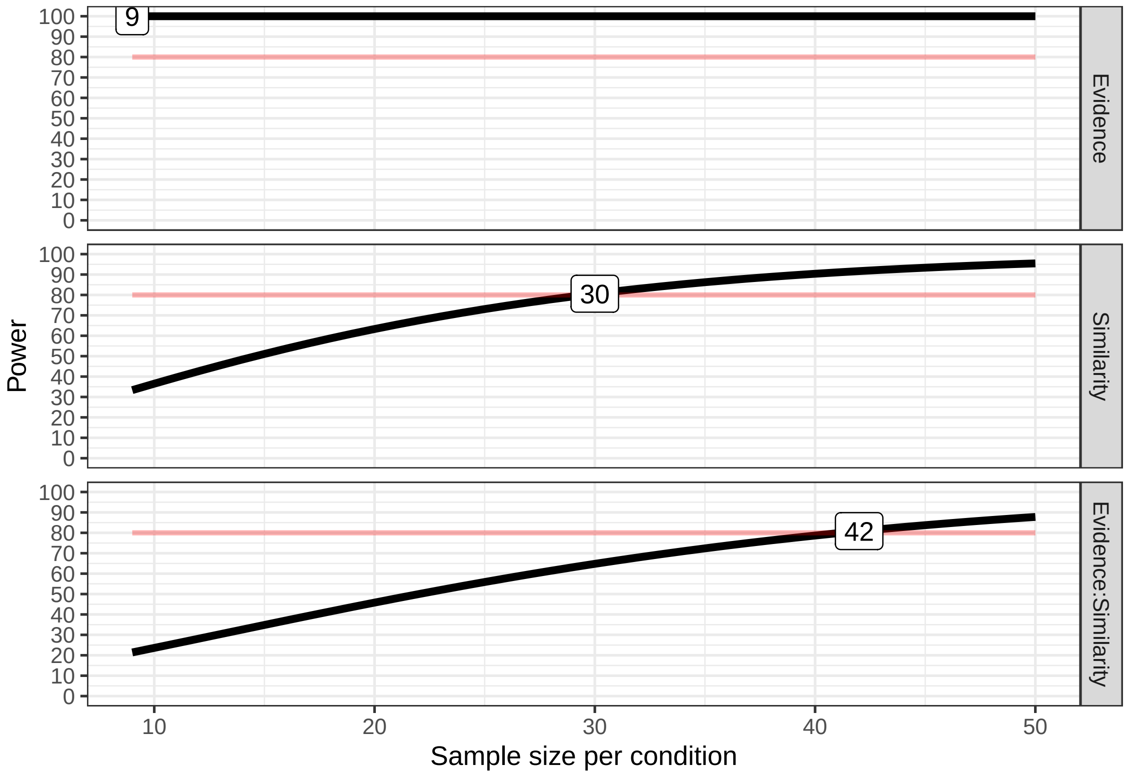


Figure C.1: Power curves for the similarity and anecdote effects.

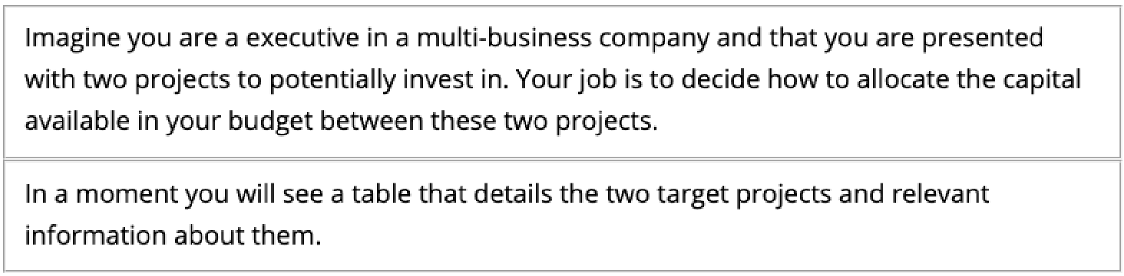


Figure C.2: Experiment 1 general instructions. The two boxes were split between two separate web-pages.

As shown in Figure C.1, the power analysis suggested that a minimum sample size of 294 ($42 \cdot 7$) is required for the interaction effect with an expected power of at least 80%.

C.1.1.2 Method

C.1.1.2.1 Instructions Figure C.2 shows the general instructions all participants received, and Figures C.3, C.4, C.5, and C.6 show the condition-specific instructions.

Managers often find it useful to consult with previous case studies before making important decisions. As well as seeing the two target projects, you will also be provided with an example of a failed project with some information that was available just before the company decided to invest in it. Further, you are also provided with an analysis of this investment decision after it became clear that the project will not meet its expected return on investment.

Figure C.3: Experiment 1 specific instructions for those in the anecdotes only condition.

Managers often find it useful to consult with previous case studies before making important decisions. As well as seeing the two target projects, you will also be provided with an example of a failed project with some information that was available just before the company decided to invest in it. Further, you are also provided with an analysis of this investment decision after it became clear that the project will not meet its expected return on investment.

As a part of the relevant information that will be provided for each target project, you will be provided with measures of overall reliability and Net Present Value (NPV). The NPV is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. The higher the NPV, the better the expectations for each project. Both these measures were collected as part of a research study conducted by an international consulting company that aggregated data from thousands of other projects in relevant industries.

Note that the project in the case study was included in the research study, so its features are subsumed in the aggregated data.

Figure C.4: Experiment 1 specific instructions for those in the anecdote & statistics condition.

C.1.1.2.2 Allocation Task A horizontally integrated company is one which is made up of multiple businesses that operate in similar markets, and may have previously been competitors (Gaughan, 2012a). A vertically integrated company, on the other hand, is one which is made up of multiple business than operate in the same market, but in different levels of the supply chain (Gaughan, 2012b). A centralised organisational structure is one in which a company decisions tend to come from a specific business unit or leader, whereas a decentralised structure is

Managers often find it useful to consult with previous case studies before making important decisions. As well as seeing the two target projects, you will also be provided with an example of a failed project with some information that was available just before the company decided to invest in it. Further, you are also provided with an analysis of this investment decision after it became clear that the project will not meet its expected return on investment.

As a part of the relevant information that will be provided for each target project, you will be provided with measures of overall reliability and Net Present Value (NPV). The NPV is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. The higher the NPV, the better the expectations for each project. Both these measures were collected as part of a research study conducted by an international consulting company that aggregated data from thousands of other projects in relevant industries.

Note that the project in the case study was included in the research study, so its features are subsumed in the aggregated data.

Alongside its results, the research study also encouraged managers to use 'scientific thinking'.

Scientific thinking can be characterized as a process of objectively analyzing information about a given topic. A scientific thinker is one who very carefully considers the quality of each piece of information so as not to be unduly swayed by insignificant and/or less significant facts.

Progress in science is generally achieved via the deliberate process of obtaining quantifiable evidence through observation and/or experimentation. The scientific method requires that experimental and observational findings be reproducible and cautions against drawing strong conclusions from any single study or observation. You may recall from statistics that this scientific principle is consistent with the fact that small samples of observations tend to have a higher probability of error while larger samples tend to be more accurate. Scientific knowledge is therefore based on an accumulation of carefully designed studies or observations which lend support to a given assertion.

Figure C.5: Experiment 1 specific instructions for those in the anecdote & enhanced statistics condition.

> As a part of the relevant information that will be provided for each target project, you will be provided with measures of overall reliability and Net Present Value (NPV). The NPV is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. The higher the NPV, the better the expectations for each project. Both these measures were collected as part of a research study conducted by an international consulting company that aggregated data from thousands of other projects in relevant industries.

Figure C.6: Experiment 1 specific instructions for those in the statistics only condition.

one in which decisions can be made by separate units or people independently (Kenton, 2021).

C.1.1.2.3 Follow-up Figure C.7 shows the follow-up questions.

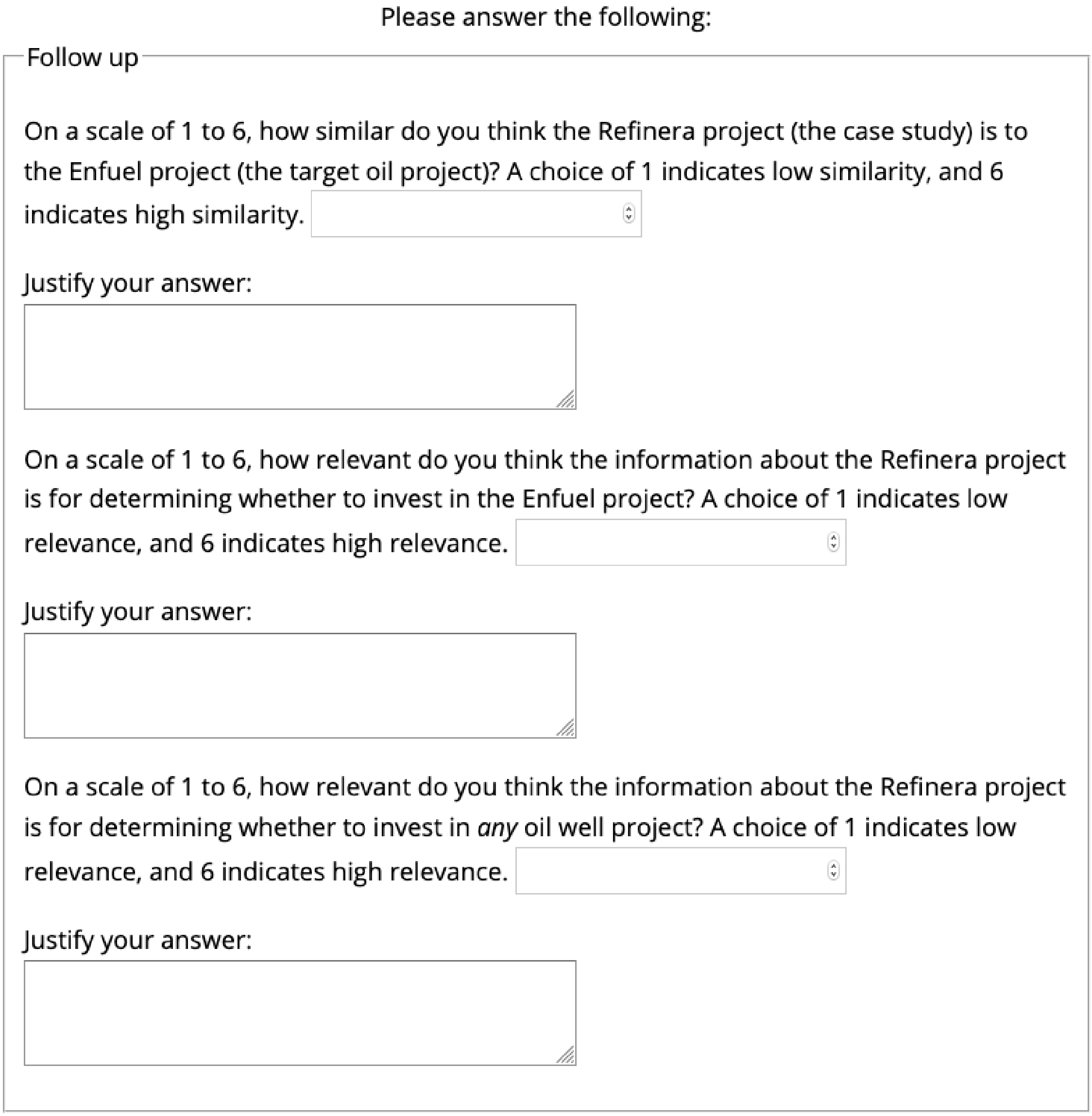
C.1.2 Results

C.1.2.1 Allocation

A two-way ANOVA was conducted to investigate the interaction of similarity (low and high) and anecdote conditions (anecdote only, statistics & anecdote, anecdote & enhanced statistics). The main text reports the more relevant interaction that excludes the enhanced statistics condition. There was a main effect of anecdote type, $F(2, 238) = 14.47$, $p < .001$, $\hat{\eta}^2_p = .108$; and a main effect of similarity, $F(1, 238) = 38.91$, $p < .001$, $\hat{\eta}^2_p = .141$. However, the interaction was not significant, $F(2, 238) = 2.16$, $p = .118$, $\hat{\eta}^2_p = .018$. The difference between the anecdote only condition and the anecdote & enhanced statistics condition was not significant, $M = -9.24$, 95% CI $[-22.00,\ 3.51]$, $t(238) = -1.43$, $p = .155$.

C.1.2.2 Manipulation Check

Figure C.8 shows participants' ratings of the similarity of the anecdote to the target project. As intended, participants in the high similarity condition rated the anecdote as more similar to the target project than those in the low similarity condition, $F(1, 238) = 27.01$, $p < .001$, $\hat{\eta}^2_p = .102$.



Please answer the following:

Follow up

On a scale of 1 to 6, how similar do you think the Refinera project (the case study) is to the Enfuel project (the target oil project)? A choice of 1 indicates low similarity, and 6 indicates high similarity.

Justify your answer:

On a scale of 1 to 6, how relevant do you think the information about the Refinera project is for determining whether to invest in the Enfuel project? A choice of 1 indicates low relevance, and 6 indicates high relevance.

Justify your answer:

On a scale of 1 to 6, how relevant do you think the information about the Refinera project is for determining whether to invest in *any* oil well project? A choice of 1 indicates low relevance, and 6 indicates high relevance.

Justify your answer:

Figure C.7: Follow-up questions in Experiment 1.

C.1.2.3 Follow-up

Figure C.9 shows participants' ratings of the specific relevance question. There was no significant effect of evidence type $F(2, 238) = 0.96$, $p = .383$, $\hat{\eta}_p^2 = .008$; or similarity, $F(1, 238) = 1.54$, $p = .216$, $\hat{\eta}_p^2 = .006$. The interaction was also not significant, `r results_anecdotes_1$relevance_specific$anecdote_alignment`.

Figure C.10 shows participants' ratings of the general relevance question. There was no main effect of similarity, $F(1, 238) = 3.32$, $p = .070$, $\hat{\eta}_p^2 = .014$, or interaction of similarity and evidence type, `r results_anecdotes_1$relevance_general$anecdote_alig` However, there was an unexpected main effect of evidence type, $F(2, 238) = 3.80$,

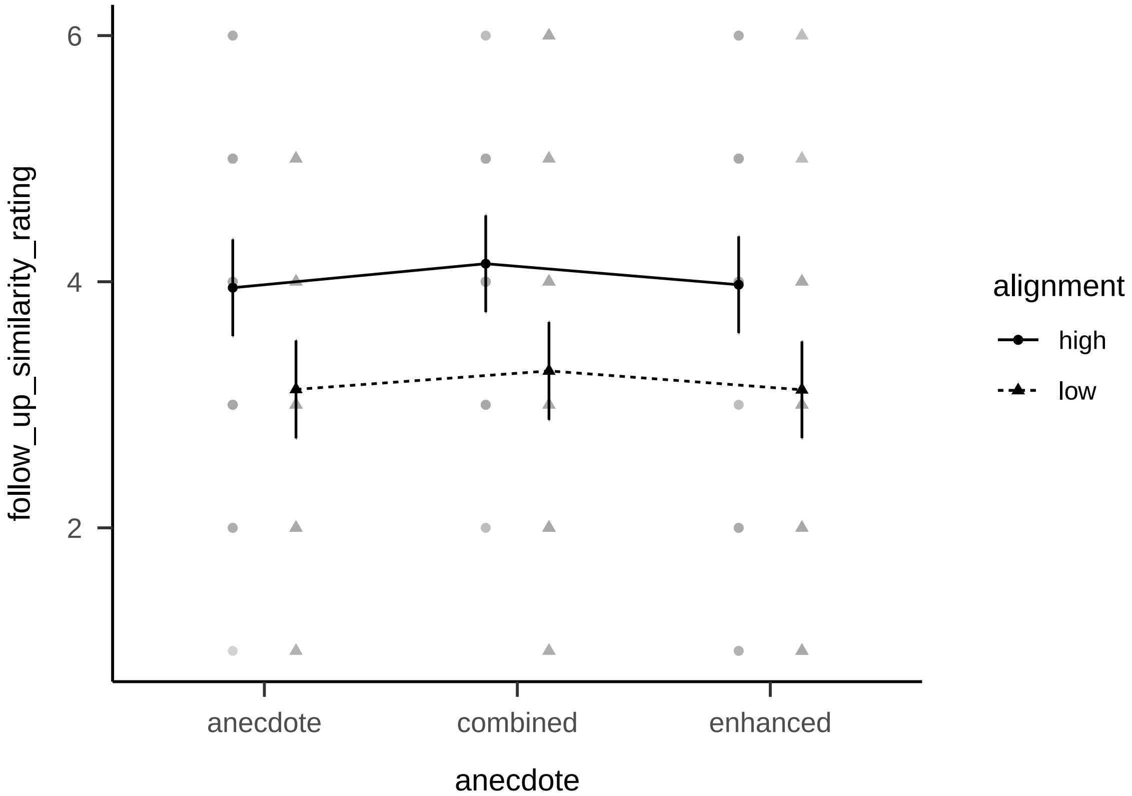


Figure C.8: Mean similarity rating of Project A (the target project) to the anecdote. Error bars represent 95% confidence intervals.

$p = .024$, $\hat{\eta}_p^2 = .031$. A contrast analysis with Bonferroni correction revealed that the anecdote only condition was rated significantly higher than the anecdote & statistics condition, $\Delta M = 0.58$, 95% CI $[0.06,\ 1.10]$, $t(238) = 2.71$, $p = .022$. However, the difference between the two anecdote & statistics conditions was not significant, $\Delta M = -0.39$, 95% CI $[-0.90,\ 0.13]$, $t(238) = -1.81$, $p = .212$.

Regression analyses were conducted to determine the relationship between allocations and the follow-up ratings of similarity and relevance. As shown in Figure C.11, similarity ratings were negatively correlated to allocations, $b = -3.53$, 95% CI $[-5.70, -1.37]$, $t(242) = -3.21$, $p = .002$. Finally, as shown in Figure C.12 similarity ratings were positively correlated to specific relevance ratings, $b = 0.30$, 95% CI $[0.17, 0.43]$, $t(242) = 4.59$, $p < .001$.

Participants' justifications for the ratings were not analysed, so are not reported.

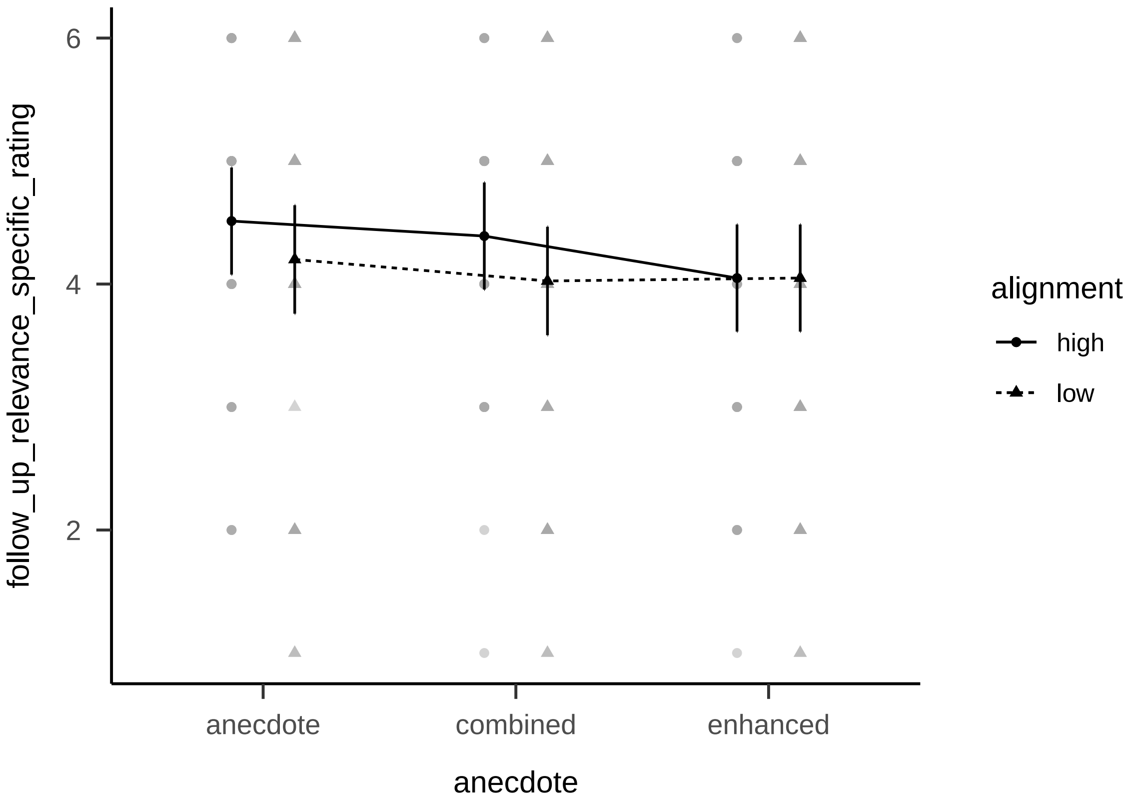


Figure C.9: Mean rating of how relevant participants thought the anecdote was to Project A (the target project). Error bars represent 95% confidence intervals.

C.2 Experiment 2

Figures C.13 and C.14 show the simulated data for the negative and positive valence conditions, respectively. These figures are different from the equivalent figures in the main text. Here, the same statistics only value was used for both valence conditions, whereas in the main text the relevant values for each condition were used. Further, the main text reports the difference score from the relevant statistics only values, whereas here the raw means are shown.

Hypothesis C.8—Allocation similarity manipulation check for positive anecdote. For positive anecdotes, allocations for the anecdote only high similarity condition will be higher than those in the anecdote only low similarity condition.

The rating effects found in Experiment 1 were expected to replicate in the Experiment 2 negative valence condition. The reverse effects were expected to be found in the positive valence condition.

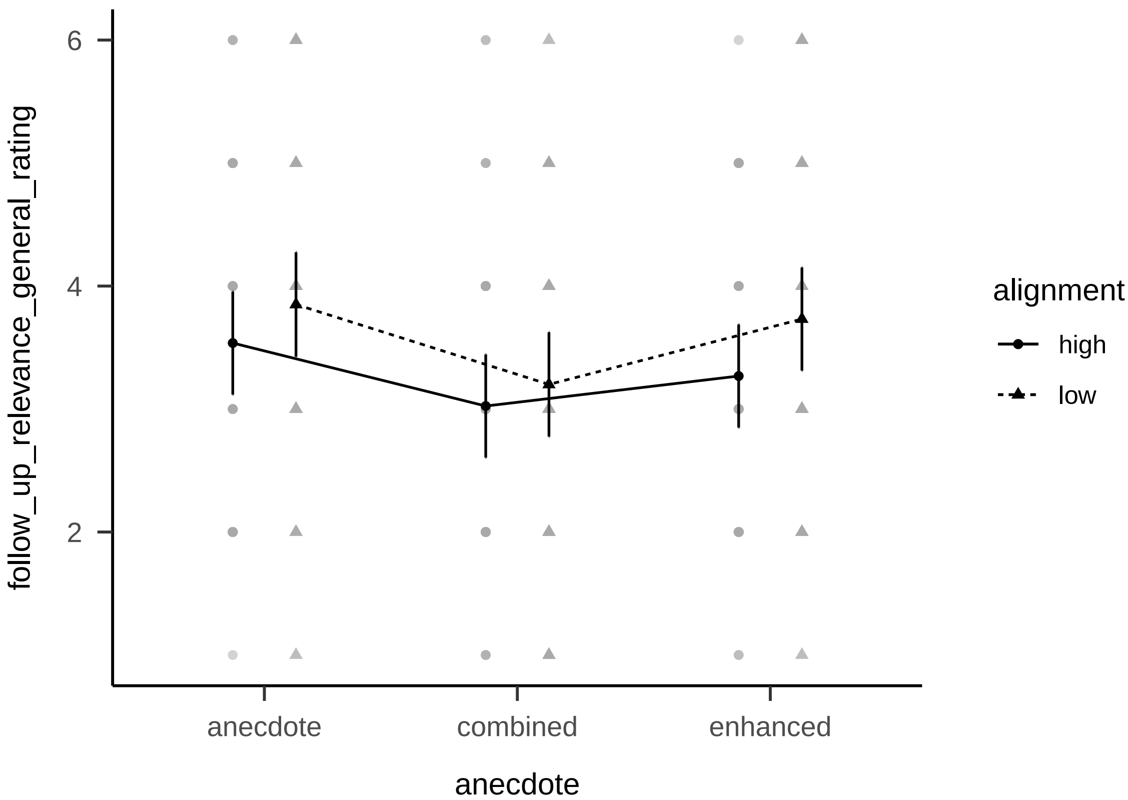


Figure C.10: Mean rating of how relevant participants thought the anecdote was to other oil projects. Error bars represent 95% confidence intervals.

Hypothesis C.9—Relationship between allocation and perceived similarity for positive anecdote. In the positive valence condition, the correlation between allocation and similarity rating will be positive

Hypothesis C.10—Relationship between allocation and specific-relevance for positive anecdote. In the positive valence condition, there will be no correlation between allocation and specific-relevance rating in the low similarity condition, but a positive correlation in the high similarity condition.

Hypothesis C.11—Relationship between allocation and general-relevance for positive anecdote. There will be no significant correlations between allocation and general-relevance rating

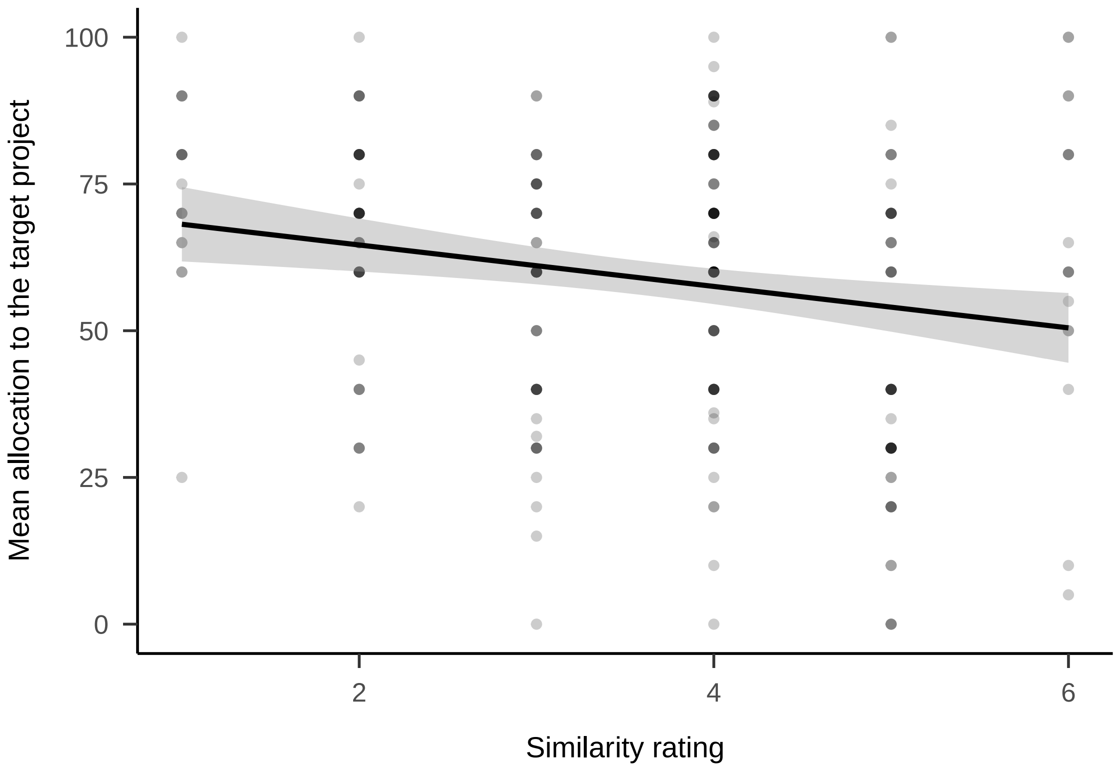


Figure C.11: Mean allocation to the target project by similarity rating. The shading represents 95% confidence intervals.

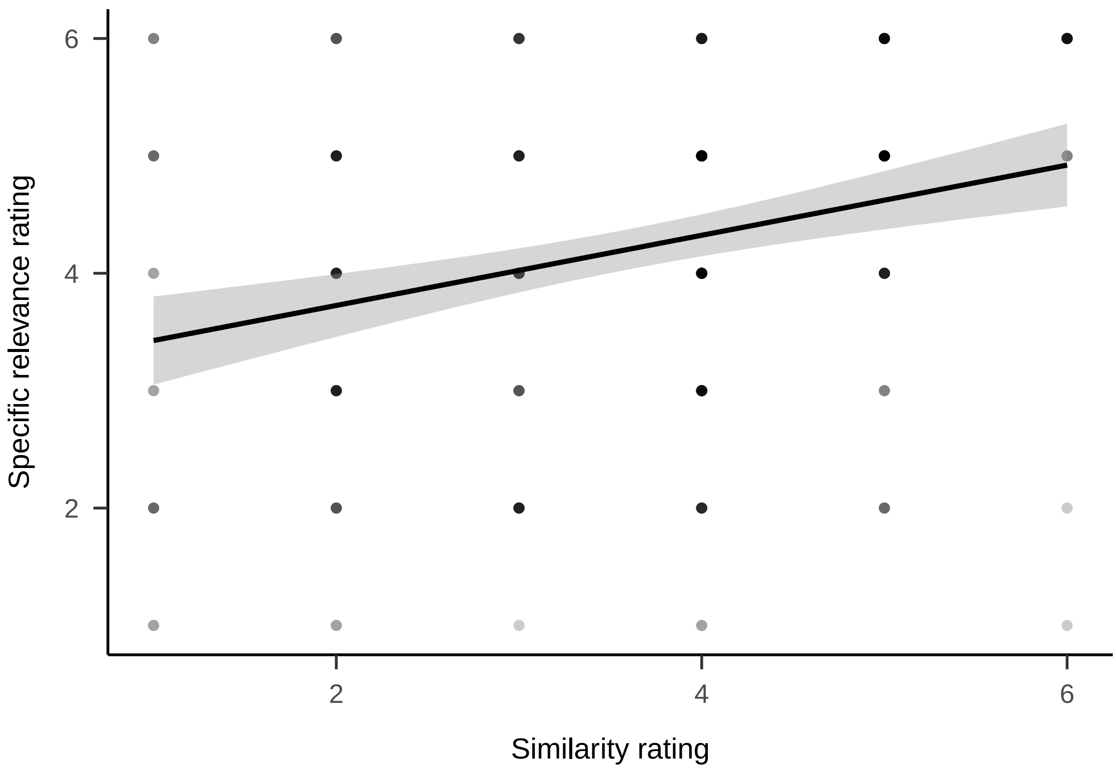


Figure C.12: Rating of how relevant participants considered the anecdote to the target project, by similarity rating. The shading represents 95% confidence intervals.

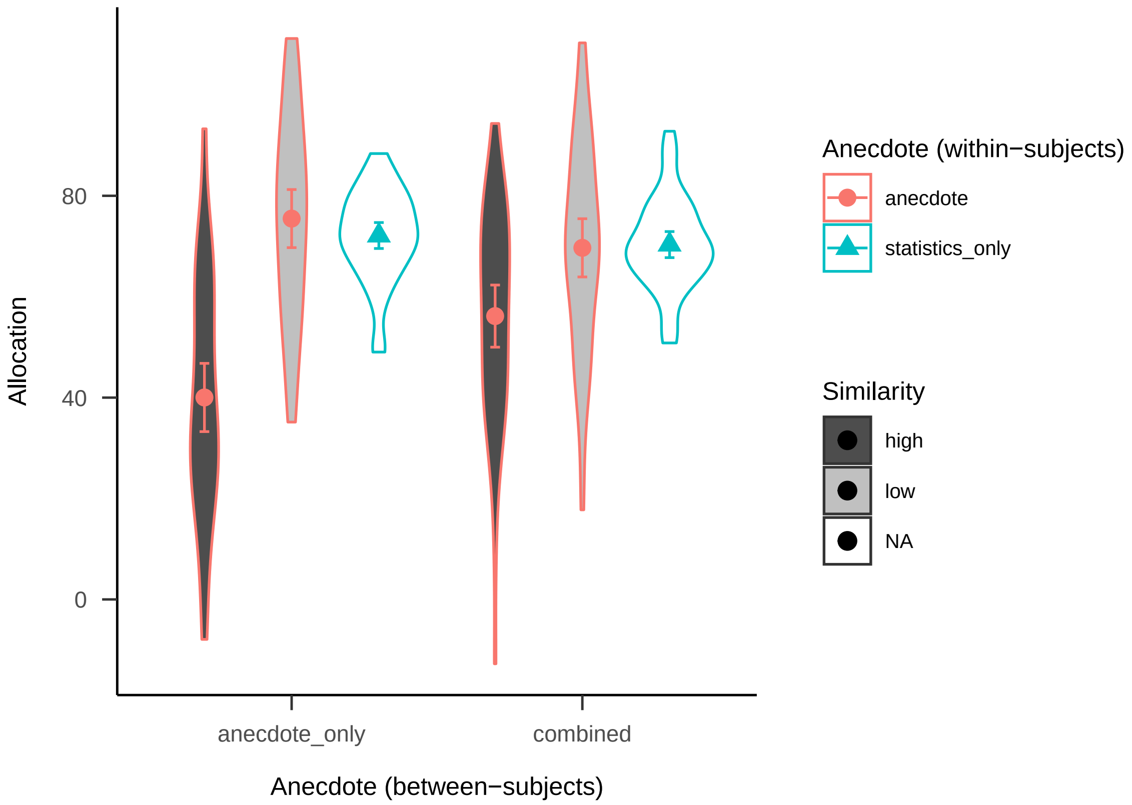


Figure C.13: Anecdotes Experiment 2 predicted data for the negative valence condition

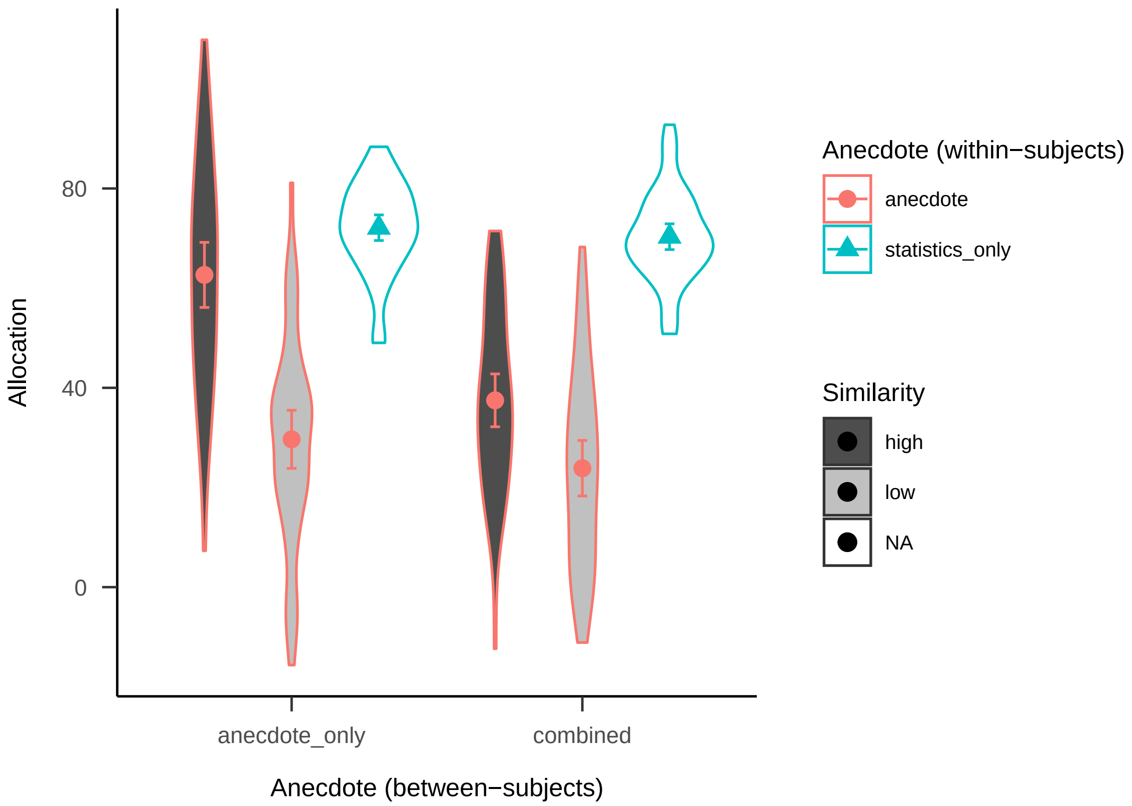


Figure C.14: Anecdotes Experiment 2 predicted data for the positive valence condition

C.2.1 Method

C.2.1.1 Participants

C.2.1.1.1 Power Analysis A power analysis was conducted through simulation of the effects implied by the hypotheses in Experiment 2. Data were simulated with the same mean values as Experiment 1 for the effects that were previously significant (i.e., similarity, statistics, and interaction effects), and no effect for the differences that were non-significant (as shown in Figures C.13 and C.14). The null effect was analysed using the two one-sided tests (TOST) procedure, or *equivalence* testing (Lakens et al., 2018), and setting the smallest effect size of interest to the smallest difference that leads to a significant equivalence between the combined low similarity and statistics only conditions in Experiment 1. Figure C.15 shows the results of this analysis, which suggested a total sample size of 92 (46×2).

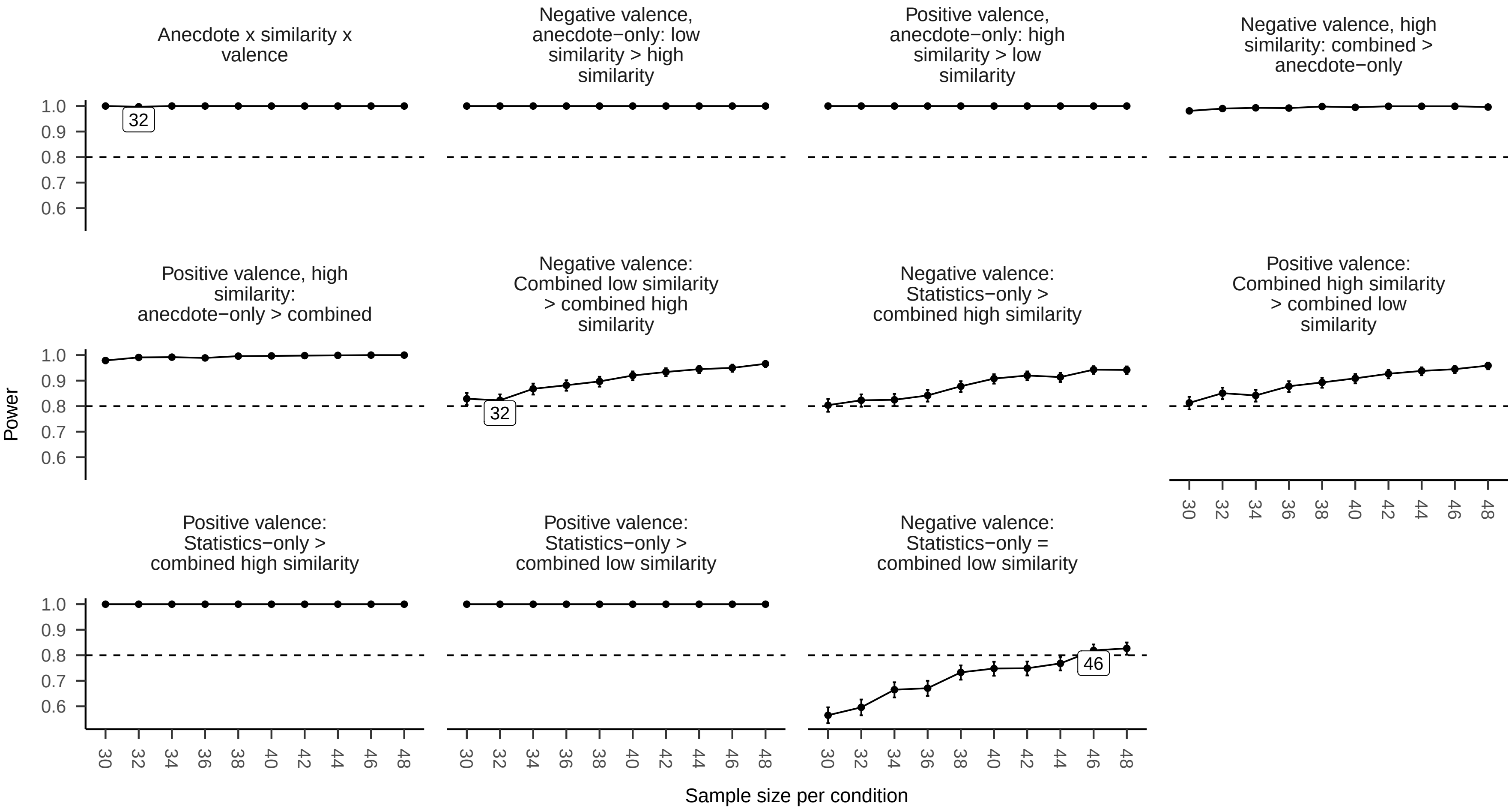


Figure C.15: Anecdotes Experiment 2 power curve. Labels indicate lowest sample size above 80% power.

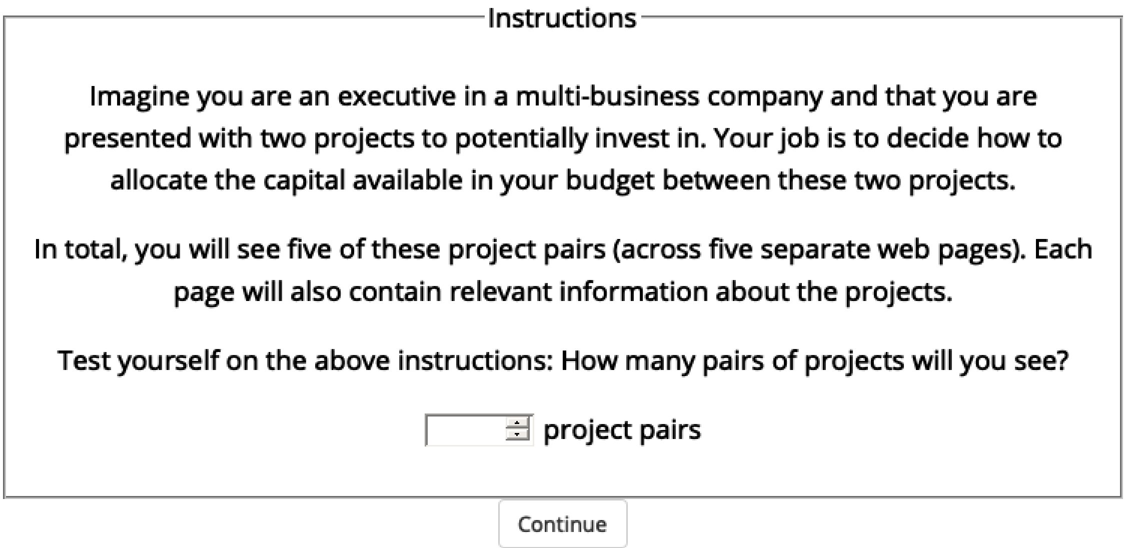


Figure C.16: General instructions for Experiment 2.

C.2.1.2 Materials

C.2.1.2.1 Instructions Figure C.16 shows the general instructions all participants received, and Figures C.17, C.18, and C.19 show the condition-specific instructions.

C.2.1.2.2 Allocation Task The following were counterbalanced: (a) project variation (five latin square variations), which is the association of each display content with each within-subject condition; and (b) anecdote variation (two variations), which is the association of each project display and being either the target or comparison project. Table column order and project display order were randomised.

C.2.1.2.3 Follow-up Questions Figure C.20 shows an example of the follow-up questions.

C.2.1.2.4 Interstitial Display Figure C.21 shows an example of one of the interstitial displays.

Instructions

Managers often find it useful to consult with previous case studies before making important decisions. As well as seeing the two target projects, you will also be provided with an example of a failed project with some information that was available just before the company decided to invest in it. This project was randomly chosen from a pool of thousands of projects. Others rated the similarity of all the case studies to the below target project based on dimensions such as the overall money invested, the quality of the proposal, the experience of the managers that proposed it, and the specific operations that were required. This case study was found to be, on average, as similar to the target as the others in the sample. Further, you are also provided with an analysis of this investment decision after it became clear that the project will not meet its expected return on investment.

Figure C.17: Experiment 2 specific instructions for those in the anecdotes only condition.

C.2.2 Results

C.2.2.1 Allocation

C.2.2.1.1 Similarity Manipulation Check The similarity manipulation worked as expected, with the negative anecdote only low similarity condition being allocated significantly more than those in the high similarity condition, $\Delta M = 26.98$, 95% CI $[18.12,\ 35.84]$, $t(186.55) = 6.01$, $p < .001$. For positive anecdotes, participants allocated more to the high similarity condition than those in the low similarity condition, $\Delta M = -22.62$, 95% CI $[-31.48,\ -13.77]$, $t(186.55) = -5.04$, $p < .001$

C.2.2.2 Ratings

C.2.2.2.1 Similarity Manipulation Check Evidence for the similarity manipulation working was also seen in the rating data. Participants rated anecdotes in the high similarity condition as more similar to the target than those in the low similarity condition, $F(1, 94) = 48.36$, $p < .001$, $\hat{\eta}^2_p = .340$.

C.2.2.2.2 Allocation is Influenced by Perceived Similarity As hypothesised, allocation was influenced by perceived similarity. That is, in the negative valence condition, there was a negative correlation between allocation and similarity

Instructions

Managers often find it useful to consult with previous case studies before making important decisions. As well as seeing the two target projects, you will also be provided with an example of a failed project with some information that was available just before the company decided to invest in it. This project was randomly chosen from a pool of thousands of projects. Others rated the similarity of all the case studies to the below target project based on dimensions such as the overall money invested, the quality of the proposal, the experience of the managers that proposed it, and the specific operations that were required. This case study was found to be, on average, as similar to the target as the others in the sample. Further, you are also provided with an analysis of this investment decision after it became clear that the project will not meet its expected return on investment.

As a part of the relevant information that will be provided for each target project, you will be provided with measures of overall reliability and Net Present Value (NPV). The NPV is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. The higher the NPV, the better the expectations for each project. Both these measures were collected as part of a research study conducted by an international consulting company that aggregated data from thousands of other projects in relevant industries.

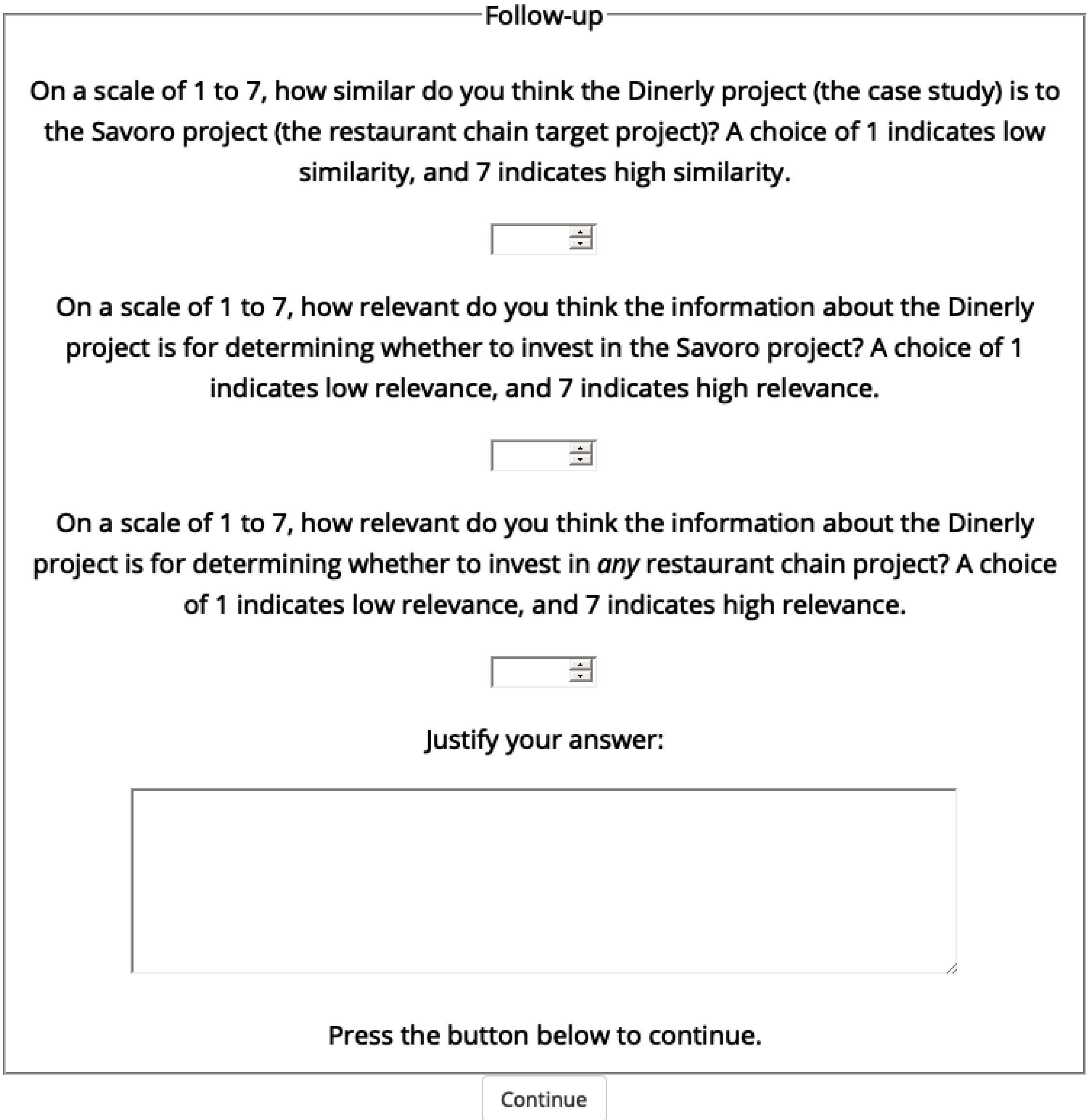
Note that the project in the case study was included in the research study, so its features are subsumed in the aggregated data.

Figure C.18: Experiment 2 specific instructions for those in the combined condition.

Instructions

As a part of the relevant information that will be provided for each target project, you will be provided with measures of overall reliability and Net Present Value (NPV). The NPV is the company's estimation of the future returns of the project. An NPV that is greater than 0 (zero) indicates that there is an expectation of profit. The higher the NPV, the better the expectations for each project. Both these measures were collected as part of a research study conducted by an international consulting company that aggregated data from thousands of other projects in relevant industries.

Figure C.19: Experiment 2 specific instructions for those in the statistics only condition.



Follow-up

On a scale of 1 to 7, how similar do you think the Dinerly project (the case study) is to the Savoro project (the restaurant chain target project)? A choice of 1 indicates low similarity, and 7 indicates high similarity.

On a scale of 1 to 7, how relevant do you think the information about the Dinerly project is for determining whether to invest in the Savoro project? A choice of 1 indicates low relevance, and 7 indicates high relevance.

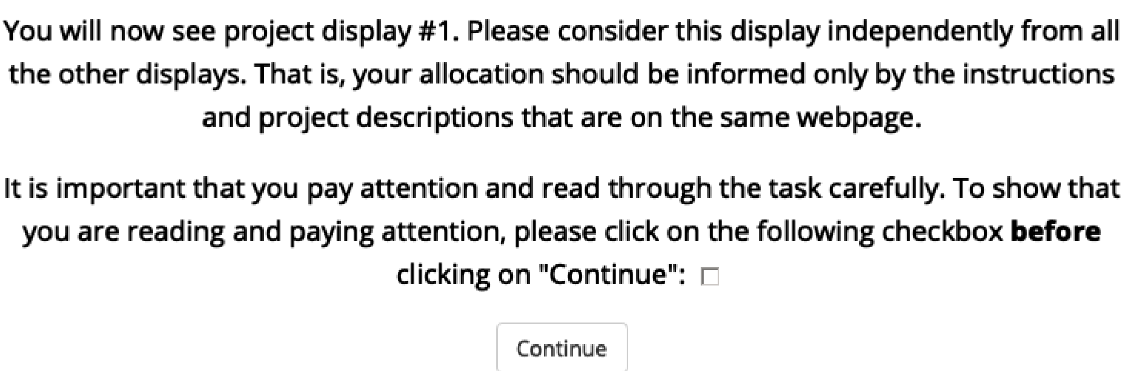
On a scale of 1 to 7, how relevant do you think the information about the Dinerly project is for determining whether to invest in *any* restaurant chain project? A choice of 1 indicates low relevance, and 7 indicates high relevance.

Justify your answer:

Press the button below to continue.

Continue

Figure C.20: An example of one of the follow-up question displays in Experiment 2.



You will now see project display #1. Please consider this display independently from all the other displays. That is, your allocation should be informed only by the instructions and project descriptions that are on the same webpage.

It is important that you pay attention and read through the task carefully. To show that you are reading and paying attention, please click on the following checkbox **before** clicking on "Continue":

Continue

Figure C.21: An example of an interstitial display in Experiment 2.

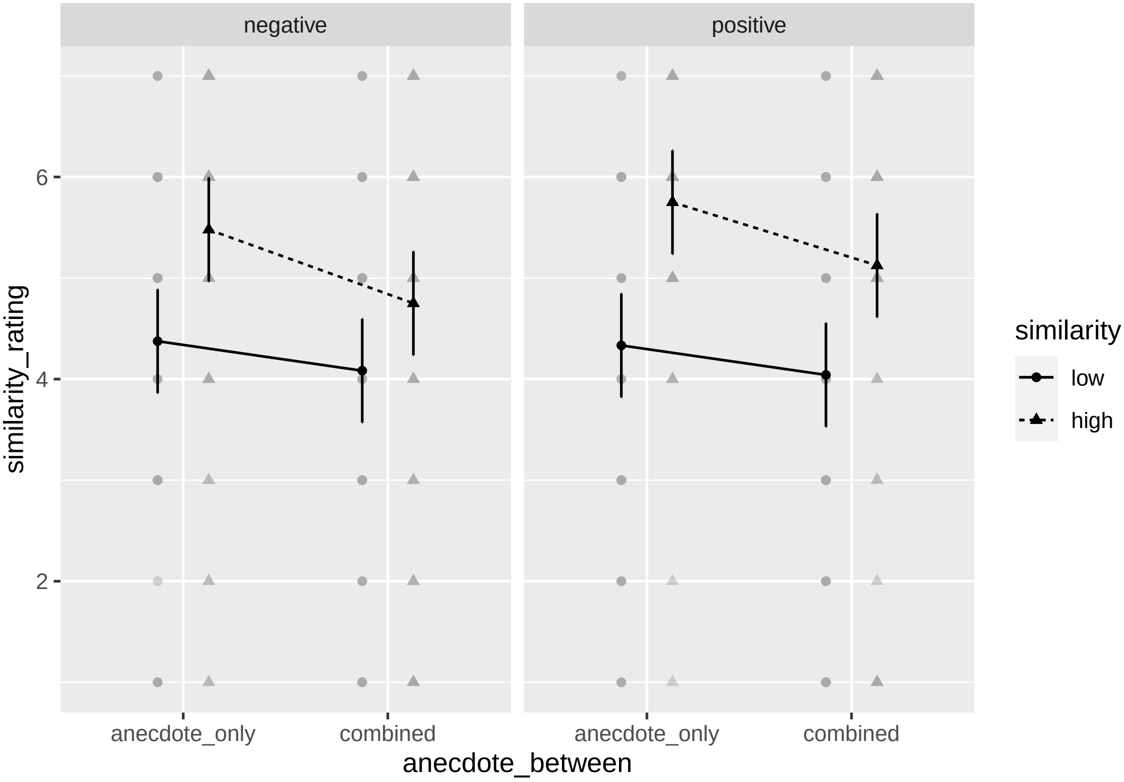


Figure C.22: Mean similarity rating of Project A (the target project) to the anecdote. Error bars represent 95% confidence intervals.

rating, $\Delta M = 0.34$, 95% CI $[-3.72,\ 4.39]$, $t(376) = 0.16$, $p = .870$. However, in the positive valence condition, there was a positive correlation between allocation and similarity rating, $\Delta M = 2.86$, 95% CI $[-1.47,\ 7.18]$, $t(376) = 1.30$, $p = .195$.

C.2.2.2.3 The Relationship Between Allocation and Specific-Relevance Depends on Similarity In the negative valence condition, there was no significant difference between the slopes of the high and low similarity conditions, $M = -2.02$, 95% CI $[-6.44,\ 2.41]$, $t(376) = -0.90$, $p = .371$. In the low similarity condition, allocation and specific-relevance rating were not correlated, $\Delta M = 1.01$, 95% CI $[-1.21,\ 3.22]$, $t(376) = 0.90$, $p = .371$, as in the low similarity condition, $\Delta M = -1.01$, 95% CI $[-3.22,\ 1.21]$, $t(376) = -0.90$, $p = .371$.

In the positive valence condition, there was no significant difference between the slopes of the high and low similarity conditions, $M = 4.25$, 95% CI $[-0.20,\ 8.70]$, $t(376) = 1.88$, $p = .061$. In the low similarity condition, allocation and specific-relevance rating were not correlated, $\Delta M = -2.12$, 95% CI $[-4.35,\ 0.10]$, $t(376) = -1.88$, $p = .061$, as in the low similarity condition, $\Delta M = 2.12$, 95% CI $[-0.10,\ 4.35]$, $t(376) = 1.88$, $p = .061$.

C.2.2.2.4 People do not Consider General-Relevance in Their Allocation There were no significant correlations between allocation and general-relevance rating.

References

Al Khatib, K., Wachsmuth, H., Hagen, M., & Stein, B. (2017). Patterns of Argumentation Strategies across Topics. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 1351–1357. https://doi.org/10/gjscsq

Allen, M., & Preiss, R. W. (1997). Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports*, *14*(2), 125–131. https://doi.org/10/djqrp7

Aloysius, J. A. (2007). Decision making in the short and long run: Repeated gambles and rationality. *British Journal of Mathematical and Statistical Psychology*, *60*(1), 61–69. https://doi.org/10/czgkh3

Alpert, M., & Raiffa, H. (1982). A progress report on the training of probability assessors. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty* (pp. 294–305). Cambridge University Press. https://doi.org/10.1017/CBO9780511809477.022

Arrfelt, M., Wiseman, R. M., McNamara, G., & Hult, G. T. M. (2015). Examining a key corporate role: The influence of capital allocation competency on business unit performance. *Strategic Management Journal*, *36*(7), 1017–1034. https://doi.org/10/gkhhc5

Aust, F. (2020). *Rmdfiltr: 'lua'-filters for r markdown* [R package version 0.1.3]. https://CRAN.R-project.org/package=rmdfiltr

Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R Markdown* [R package version 0.1.0.9997]. https://github.com/crsh/papaja

Bache, S. M., & Wickham, H. (2020). *Magrittr: A forward-pipe operator for r* [R package version 2.0.1]. https://CRAN.R-project.org/package=magrittr

Bardolet, D., Brown, A., & Lovallo, D. (2017). The Effects of Relative Size, Profitability, and Growth on Corporate Capital Allocations. *Journal of Management*, *43*(8), 2469–2496. https://doi.org/10/gjvc4v

Bardolet, D., Fox, C. R., & Lovallo, D. (2011). Corporate capital allocation: A behavioral perspective. *Strategic Management Journal*, *32*(13), 1465–1483. https://doi.org/10/cn6xsb

Bardolet, D., Lovallo, D., & Rumelt, R. (2010). The hand of corporate management in capital allocations: Patterns of investment in multi- and single-business firms. *Industrial and Corporate Change*, *19*(2), 591–612. https://doi.org/10/ctxs4t

Barney, J. B. (1988). Returns to bidding firms in mergers and acquisitions: Reconsidering the relatedness hypothesis. *Strategic Management Journal*, *9*(S1), 71–78. https://doi.org/10/dmpq6x

Barney, J. B. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10/gpm

Barone-Adesi, G., Mancini, L., & Shefrin, H. (2013). *A Tale of Two Investors: Estimating Optimism and Overconfidence*. http://www.ssrn.com/abstract=2319260

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*(3), 215–233. https://doi.org/10/d3jsr8

Barron, G., & Leider, S. (2010). The role of experience in the Gambler's Fallacy. *Journal of Behavioral Decision Making*, *23*(1), 117–129. https://doi.org/10/d3p92r

Barth, M. (2021). *tinylabels: Lightweight variable labels* [R package version 0.2.1]. https://github.com/mariusbarth/tinylabels

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bates, D., & Maechler, M. (2021). *Matrix: Sparse and dense matrix classes and methods* [R package version 1.3-2]. https://CRAN.R-project.org/package=Matrix

Batteux, E., Bilovich, A., Johnson, S., & Tuckett, D. (2020, May 7). *Impressed by Numbers: The Extent to Which Novice Investors Favor Precise Numerical Information in a Context of Uncertainty* (SSRN Scholarly Paper No. ID 3595409). Social Science Research Network. Rochester, NY. https://doi.org/10.2139/ssrn.3595409

Bednar, M. K., Boivie, S., & Prince, N. R. (2013). Burr Under the Saddle: How Media Coverage Influences Strategic Change. *Organization Science*, *24*(3), 910–925. https://doi.org/10/gf33wh

Bellemare, C., Krause, M., Kröger, S., & Zhang, C. (2005). Myopic loss aversion: Information feedback vs. investment flexibility. *Economics Letters*, *87*(3), 319–324. https://doi.org/10/dfjq7n

Benartzi, S., & Thaler, R. H. (1999). Risk Aversion or Myopia? Choices in Repeated Gambles and Retirement Investments. *Management Science*, *45*(3), 364–381. https://doi.org/10/cjhp6t

Benartzi, S., & Thaler, R. H. (2001). Naive Diversification Strategies in Defined Contribution Saving Plans. *American Economic Review*, *91*(1), 79–98. https://doi.org/10/cfd4pb

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica*, *22*(1), 23–36. https://doi.org/10/cw729v (Original work published 1738)

Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2016). Does Aggregated Returns Disclosure Increase Portfolio Risk Taking? *The Review of Financial Studies*, *30*(6), 1971–2005. https://doi.org/10/gjscs7

Bettis, R. A. (2017). Organizationally Intractable Decision Problems and the Intellectual Virtues of Heuristics. *Journal of Management*, *43*(8), 2620–2637. https://doi.org/10/gf2r6w

Birshan, M., Engel, M., & Sibony, O. (2013). Avoiding the quicksand: Ten techniques for more agile corporate resource allocation. *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/avoiding-the-quicksand

Bjørnsen, K., & Aven, T. (2019). Risk aggregation: What does it really mean? *Reliability Engineering & System Safety*, *191*, 106524. https://doi.org/10/gjscst

Bolker, B., & Robinson, D. (2020). *Broom.mixed: Tidying methods for mixed models* [R package version 0.2.6]. https://CRAN.R-project.org/package=broom.mixed

Bostrom, A., Anselin, L., & Farris, J. (2008). Visualizing Seismic Risk and Uncertainty: A Review of Related Research. *Annals of the New York Academy of Sciences*, *1128*(1), 29–40. https://doi.org/10/c32d9k

Bower, J. L. (1970). *Managing the resource allocation process: A study of corporate planning and investment*. Division of Research, Graduate School of Business Administration, Harvard University.

Bower, J. L., & Gilbert, C. G. (Eds.). (2005). *From resource allocation to strategy*. Oxford University Press.

Boyd, T. (2021, June 3). *Wesfarmers CEO proves his worth*. Australian Financial Review. https://www.afr.com/chanticleer/wesfarmers-ceo-proves-his-worth-20210603-p57xmq

Bradley, D. (2013). Decision Theory, Philosophical Perspectives. In H. E. Pashler (Ed.), *Encyclopedia of the mind*. SAGE Publications, Inc. https://doi.org/10.4135/9781452257044.n81

Bristow, R. E. (2011). *There is more to Life than Expected Values: Results of Free Distributions of Multiple-Play Gambles* (Masters thesis). Miami University. Oxford, Ohio. https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=miami1304352729

Brodlie, K., Allendes Osorio, R., & Lopes, A. (2012). A Review of Uncertainty in Data Visualization. In J. Dill, R. Earnshaw, D. Kasik, J. Vince, & P. C. Wong (Eds.), *Expanding the Frontiers of Visual Analytics and Visualization* (pp. 81–109). Springer London. https://doi.org/10.1007/978-1-4471-2804-5_6

Buchanan, E. M., Gillenwaters, A., Scofield, J. E., & Valentine, K. (2019). *MOTE: Measure of the Effect: Package to assist in effect size calculations and their confidence intervals* [R package version 1.0.2]. http://github.com/doomlab/MOTE

Bukszar, E., & Connolly, T. (1988). Hindsight Bias and Strategic Choice: Some Problems in Learning From Experience. *Academy of Management Journal*, *31*(3), 628–641. https://doi.org/10/gjscqv

Burgelman, R. A. (1983). A Process Model of Internal Corporate Venturing in the Diversified Major Firm. *Administrative Science Quarterly*, *28*(2), 223–244. https://doi.org/10/dsvjgf

Camerer, C., & Lovallo, D. (1999). Overconfidence and Excess Entry: An Experimental Approach. *American Economic Review*, *89*(1), 306–318. https://doi.org/10/bnkbkb

Camilleri, A. R., & Newell, B. R. (2011). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review*, *18*(2), 377–384. https://doi.org/10/cz55xk

Camilleri, A. R., & Newell, B. R. (2013). The long and short of it: Closing the description-experience “gap” by taking the long-run view. *Cognition*, *126*(1), 54–71. https://doi.org/10/f4gq3w

Carvalho, P. F., Chen, C.-h., & Yu, C. (2021). The distributional properties of exemplars affect category learning and generalization. *Scientific Reports*, *11*(1), 11263. https://doi.org/10.1038/s41598-021-90743-0

Cavagnac, M. (2005). Strategic managerial incentives under adverse selection. *Managerial and Decision Economics*, *26*(8), 499–512. https://doi.org/10/cwdbtp

Champely, S. (2020a). *Pwr: Basic functions for power analysis*. manual. https://github.com/heliosdrm/pwr

Champely, S. (2020b). *Pwr: Basic functions for power analysis* [R package version 1.3-0]. https://CRAN.R-project.org/package=pwr

Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006). How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior. *Journal of Political Economy*, *114*(3), 517–537. https://doi.org/10/fr7jw7

Christensen-Szalanski, J. J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*(1), 147–168. https://doi.org/10/fhnf22

Clark, A., & Karmiloff-Smith, A. (1993). The cognizer's innards. *Mind & Language*, *8*(4), 487–519. https://doi.org/10/csfpck

Clarke, E., & Sherrill-Mix, S. (2017). *Ggbeeswarm: Categorical scatter (violin point) plots* [R package version 0.6.0]. https://CRAN.R-project.org/package=ggbeeswarm

Coase, R. H. (1937). The Nature of the Firm. *Economica*, *4*(16), 386–405. https://doi.org/10/fnm876

Cohen, L. J. (1981). Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, *4*(3), 317–331. https://doi.org/10/fn9rpc

Coombs, C. H., & Bowen, J. N. (1971). A test of VE-theories of risk and the effect of the central limit theorem. *Acta Psychologica*, *35*(1), 15–28. https://doi.org/10/dm5gbv

Courtney, H., Kirkland, J., & Viguerie, P. (1997). Strategy under uncertainty. *Harvard Business Review*, *75*(6), 67–79. https://hbr.org/1997/11/strategy-under-uncertainty

Courtney, H., Lovallo, D., & Clarke, C. (2013). Deciding How to Decide. *Harvard Business Review*, *91*(11), 62–70.

Cousineau, D., & O'Brien, F. (2014). Error bars in within-subject designs: A comment on Baguley (2012). *Behavior Research Methods*, *46*(4), 1149–1151. https://doi.org/10/f6vdsw

Cyert, R. M., Simon, H. A., & Trow, D. B. (1956). Observation of a Business Decision. *The Journal of Business*, *29*(4), 237–248. https://doi.org/10/fm8x9b

Das, T. K., & Teng, B.-S. (1999). Cognitive Biases and Strategic Decision Processes: An Integrative Perspective. *Journal of Management Studies*, *36*(6), 757–778. https://doi.org/10/c84dxp

Davis, P. S., Robinson, R. B., Pearce, J. A., & Park, S. H. (1992). Business unit relatedness and performance: A look at the pulp and paper industry. *Strategic Management Journal*, *13*(5), 349–361. https://doi.org/10/bx3wpx

Davis, T. J., & Keller, C. P. (1997). Modelling and visualizing multiple spatial uncertainties. *Computers & Geosciences*, *23*(4), 397–408. https://doi.org/10.1016/S0098-3004(97)00012-5

DeKay, M. L. (2011). Are Medical Outcomes Fungible? A Survey of Voters, Medical Administrators, and Physicians. *Medical Decision Making*, *31*(2), 338–353. https://doi.org/10/b539tb

DeKay, M. L., Hershey, J. C., Spranca, M. D., Ubel, P. A., & Asch, D. A. (2006). Are medical treatments for individuals and groups like single-play and multiple-play gambles? *Judgment and Decision Making*, *1*(2), 12. http://journal.sjdm.org/jdm06133.pdf

DeKay, M. L., & Kim, T. G. (2005). When things don't add up: The role of perceived fungibility in repeated-play decisions. *Psychological Science*, *16*(9), 667–672. https://doi.org/10/ddgt5v

Dekel, S. (2021a). *Aggregation1: Aggregation 1 experiment* [R package version 1.0]. https://github.com/shirdekel/aggregation1

Dekel, S. (2021b). *Aggregation2: Aggregation 2 experiment* [R package version 1.0]. https://github.com/shirdekel/aggregation2

Dekel, S. (2021c). *Aggregation3: Aggregation 3 experiment* [R package version 1.0]. https://github.com/shirdekel/aggregation3

Dekel, S. (2021d). *Aggregation4: Aggregation 4 experiment* [R package version 1.0]. https://github.com/shirdekel/aggregation4

Dekel, S. (2021e). *Alignment1: Alignment 1 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment1

Dekel, S. (2021f). *Alignment2: Alignment 2 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment2

Dekel, S. (2021g). *Alignment3: Alignment 3 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment3

Dekel, S. (2021h). *Alignment4: Alignment 4 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment4

Dekel, S. (2021i). *Alignment5: Alignment 5 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment5

Dekel, S. (2021j). *Alignment6: Alignment 6 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment6

Dekel, S. (2021k). *Alignment7: Alignment 7 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment7

Dekel, S. (2021l). *Alignment8: Alignment 8 experiment* [R package version 1.0]. https://github.com/shirdekel/alignment8

Dekel, S. (2021m). *Anecdotes1: Anecdotes 1 experiment* [R package version 1.0]. https://github.com/shirdekel/anecdotes1

Dekel, S. (2021n). *Anecdotes2: Anecdotes 2 experiment* [R package version 1.0]. https://github.com/shirdekel/anecdotes2

Dekel, S. (2021o). *Shir: Shir's thesis functions* [R package version 0.0.0.9000].

Dekel, S. (2021p). *Shirthesis: Thesis functions* [R package version 0.0.0.9000]. https://github.com/shirdekel/shirthesis

Doumas, L. A. A., & Hummel, J. E. (2013). Comparison and Mapping Facilitate Relation Discovery and Predication (O. Sporns, Ed.). *PLoS ONE*, *8*(6), 1–8. https://doi.org/10/gjscsn

Einhorn, H. J., & Hogarth, R. M. (1987). Decision making: Going forward in reverse. *Harvard Business Review*, *65*(1), 66–70. https://hbr.org/1987/01/decision-making-going-forward-in-reverse

Epictetus. (1995). *The Art of Living: The Classical Manual on Virtue, Happiness, and Effectiveness* (S. Lebell, Trans.). HarperCollins. (Original work published ca. 125)

Ert, E., & Erev, I. (2013). On the descriptive value of loss aversion in decisions under risk: Six clarifications. *Judgment and Decision Making*, *8*(3), 22. http://journal.sjdm.org/12/12712/jdm12712.pdf

Firke, S. (2021). *Janitor: Simple tools for examining and cleaning dirty data* [R package version 2.1.0]. https://CRAN.R-project.org/package=janitor

Flyvbjerg, B., Ansar, A., Budzier, A., Buhl, S., Cantarelli, C., Garbuio, M., Glenting, C., Holm, M. S., Lovallo, D., Lunn, D., Molin, E., Rønnest, A., Stewart, A., & van Wee, B. (2018). Five things you should know about cost overrun. *Transportation Research Part A: Policy and Practice*, *118*, 174–190. https://doi.org/10/ghdgv4

Fox, R. (2008–April 3). A brief critical history of NPV. *BAA Conference*, 16. http://usir.salford.ac.uk/id/eprint/9291/2/NPV_paper.pdf

Freling, T. H., Yang, Z., Saini, R., Itani, O. S., & Rashad Abualsamh, R. (2020). When poignant stories outweigh cold hard facts: A meta-analysis of the anecdotal bias. *Organizational Behavior and Human Decision Processes*, *160*, 51–67. https://doi.org/10/gg4t2f

Friedman, M., & Savage, L. J. (1948). The Utility Analysis of Choices Involving Risk. *Journal of Political Economy*, *56*(4), 279–304. https://doi.org/10.1086/256692

Galesic, M., & Garcia-Retamero, R. (2010). Statistical Numeracy for Health: A Cross-cultural Comparison With Probabilistic National Samples. *Arch Intern Med*, *170*(5), 462–468. https://doi.org/10/fmj7q3

Garbuio, M., & Lovallo, D. (2017). Does organizational politics kill company growth? *Review of International Business and Strategy*, *27*(4), 410–433. https://doi.org/10/gjscsw

Gaughan, P. A. (Ed.). (2012a). Horizontal Integration and M&A. In *Maximizing Corporate Value through Mergers and Acquisitions* (pp. 117–157). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119204374.ch5

Gaughan, P. A. (Ed.). (2012b). Vertical Integration. In *Maximizing Corporate Value through Mergers and Acquisitions* (pp. 159–178). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119204374.ch6

Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*(8), 691–712. https://doi.org/10/b64gsr

Gavetti, G., & Rivkin, J. W. (2005). How Strategists Really Think. *Harvard Business Review*, *83*(4), 54–63.

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, *7*(2), 155–170. https://doi.org/10/dw52z8

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, *52*(1), 45–56. https://doi.org/10/fm4rrb

Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The Roles of Similarity in Transfer: Separating Retrievability From Inferential Soundness. *Cognitive Psychology*, *25*(4), 524–575. https://doi.org/10/b7tzks

Gertner, R. H., Scharfstein, D. S., & Stein, J. C. (1994). Internal Versus External Capital Markets. *The Quarterly Journal of Economics*, *109*(4), 1211–1230. https://doi.org/10.2307/2118361

Gigerenzer, G. (2008). Why Heuristics Work. *Perspectives on Psychological Science*, *3*(1), 20–29. https://doi.org/10/d8h4qv

Gilboa, I., & Schmeidler, D. (1995). Case-Based Decision Theory. *The Quarterly Journal of Economics*, *110*(3), 605–639. https://doi.org/10/c7tz7x

Glöckner, A., Hilbig, B. E., & Jekel, M. (2014). What is adaptive about adaptive decision making? A parallel constraint satisfaction account. *Cognition*, *133*(3), 641–666. https://doi.org/10/f6q9fj

Gneezy, U., & Potters, J. (1997). An Experiment on Risk Taking and Evaluation Periods. *The Quarterly Journal of Economics*, *112*(2), 631–645. https://doi.org/10/bpkbhz

Graham, J. R., & Harvey, C. R. (2001). The theory and practice of corporate finance: Evidence from the field. *Journal of Financial Economics*, *60*(2), 187–243. https://doi.org/10/fpdzrj

Graham, J. R., Harvey, C. R., & Puri, M. (2015). Capital allocation and delegation of decision-making authority within firms. *Journal of Financial Economics*, *115*(3), 449–470. https://doi.org/10/gfvz8d

Grant, R. M., & Jammine, A. P. (1988). Performance differences between the wrigley/rumelt strategic categories. *Strategic Management Journal*, *9*(4), 333–346. https://doi.org/10/fdh8xm

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, *24*(3), 411–435. https://doi.org/10/frw7xm

Grosser, M. (2019). *Snakecase: Convert strings into any case* [R package version 0.11.0]. https://CRAN.R-project.org/package=snakecase

Haigh, M. S., & List, J. A. (2005). Do Professional Traders Exhibit Myopic Loss Aversion? An Experimental Analysis. *The Journal of Finance*, *60*(1), 523–534. https://doi.org/10/c7jn9k

Haisley, E., Mostafa, R., & Loewenstein, G. (2008). Myopic risk-seeking: The impact of narrow decision bracketing on lottery play. *Journal of Risk and Uncertainty*, *37*(1), 57–75. https://doi.org/10/czj8x7

Hall, S., Lovallo, D., & Musters, R. (2012). How to put your money where your strategy is. *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-to-put-your-money-where-your-strategy-is

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2001). Are adults better behaved than children? Age, experience, and the endowment effect. *Economics Letters*, *70*(2), 175–181. https://doi.org/10/db826t

Harrison, J. S., Hall, E. H., & Nargundkar, R. (1993). Resource Allocation as an Outcropping of Strategic Consistency: Performance Implications. *Academy of Management Journal*, *36*(5), 1026–1051. https://doi.org/10/gjscqw

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & Moore, T. (2009). Adaptive Rationality: An Evolutionary Perspective on Cognitive Bias. *Social Cognition*, *27*(5), 733–763. https://doi.org/10/c49fjq

Hayes, B. K., Navarro, D. J., Stephens, R. G., Ransom, K., & Dilevski, N. (2019). The diversity effect in inductive reasoning depends on sampling assumptions. *Psychonomic Bulletin & Review*, *26*(3), 1043–1050. https://doi.org/10/gjscss

Hedesstrom, T. M., Svedsater, H., & Garling, T. (2006). Covariation neglect among novice investors. *Journal of Experimental Psychology*, *12*(3), 155–165. https://doi.org/10/ftmd77

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools* [R package version 0.3.4]. https://CRAN.R-project.org/package=purrr

Henry, L., & Wickham, H. (2021). *Rlang: Functions for base types and core r and 'tidyverse' features* [https://rlang.r-lib.org, https://github.com/r-lib/rlang].

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*(8), 534–539. https://doi.org/10/b274n8

Hoeken, H. (2001). Convincing citizens: The role of argument quality. In D. Janssen & R. Neutelings (Eds.), *Reading and writing public documents: Problems, solutions, and characteristics* (pp. 147–169). John Benjamins Publishing Company. https://doi.org/10.1075/ddcs.1.08hoe

Hoeken, H., & Hustinx, L. (2009). When is Statistical Evidence Superior to Anecdotal Evidence in Supporting Probability Claims? The Role of Argument Type. *Human Communication Research*, *35*(4), 491–510. https://doi.org/10/fgtwjd

Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis*, *59*, 41–51. https://doi.org/10/gjscrd

Horn, J. T., Lovallo, D. P., & Viguerie, S. P. (2005). Beating the odds in market entry. *McKinsey Quarterly*, (4). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/hidden-flaws-in-strategy

Hornikx, J. (2005). A review of experimental research on the relative persuasiveness of anecdotal, statistical, causal, and expert evidence. *Studies in Communication Sciences*, *5*(1), 205–216. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.725.6516&rep=rep1&type=pdf

Hornikx, J. (2018). Combining Anecdotal and Statistical Evidence in Real-Life Discourse: Comprehension and Persuasiveness. *Discourse Processes*, *55*(3), 324–336. https://doi.org/10/gjscrf

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin*, *125*(5), 576–590. https://doi.org/10/dm3fvc

Hubbard, E. (1914). *The Roycroft Dictionary, Concocted by Ali Baba and the Bunch on Rainy Days.* The Roycrofters.

Hullman, J., Resnick, P., & Adar, E. (2015). Hypothetical Outcome Plots Outperform Error Bars and Violin Plots for Inferences about Reliability of Variable Ordering. *PLOS ONE*, *10*(11), e0142444. https://doi.org/10/f3tvsd

Jaramillo, S., Horne, Z., & Goldwater, M. (2019, May 15). *The impact of anecdotal information on medical decision-making* (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/r5pmj

Jessup, R. K., Bishara, A. J., & Busemeyer, J. R. (2008). Feedback Produces Divergence From Prospect Theory in Descriptive Choice. *Psychological Science*, *19*(10), 1015–1022. https://doi.org/10/bgb3qs

Joag, S. G., Mowen, J. C., & Gentry, J. W. (1990). Risk perception in a simulated industrial purchasing task: The Effects of single versus multi-play decisions. *Journal of Behavioral Decision Making*, *3*(2), 91–108. https://doi.org/10/czwknv

Johnson, C., & Sanderson, A. (2003). A Next Step: Visualizing Errors and Uncertainty. *IEEE Computer Graphics and Applications*, *23*(5), 6–10. https://doi.org/10/df8kvd

Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *The American Economic Review*, *93*(5), 1449–1475. https://doi.org/10/b5gfrt

Kahneman, D., & Lovallo, D. (1993). Timid Choices and Bold Forecasts: A Cognitive Perspective on Risk Taking. *Management Science*, *39*(1), 17–31. https://doi.org/10/c8vntn

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10/g98

Kale, A., Nguyen, F., Kay, M., & Hullman, J. (2019). Hypothetical Outcome Plots Help Untrained Observers Judge Trends in Ambiguous Data. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 892–902. https://doi.org/10/gghfzn

Kenton, W. (2021, March 1). *How Organizational Structures Work*. Investopedia. https://www.investopedia.com/terms/o/organizational-structure.asp

Keren, G. (1991). Additional tests of utility theory under unique and repeated conditions. *Journal of Behavioral Decision Making*, *4*(4), 297–304. https://doi.org/10/bqqkt4

Keren, G., & Wagenaar, W. A. (1987). Violation of utility theory in unique and repeated gambles. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(3), 387. https://doi.org/10/dkr96j

Kinkeldey, C., MacEachren, A. M., Riveiro, M., & Schiewe, J. (2017). Evaluating the effect of visually represented geodata uncertainty on decision-making: Systematic review, lessons learned, and recommendations. *Cartography and Geographic Information Science*, *44*(1), 1–21. https://doi.org/10/f3m63m

Klos, A. (2013). Myopic loss aversion: Potential causes of replication failures. *Judgment and Decision Making*, *8*(5), 13. http://journal.sjdm.org/12/121229/jdm121229.pdf

Klos, A., Weber, E. U., & Weber, M. (2005). Investment Decisions and Time Horizon: Risk Perception and Risk Behavior in Repeated Gambles. *Management Science*, *51*(12), 1777–1790. https://doi.org/10/bbrvhd

Koehler, J. J., Gibbs, B. J., & Hogarth, R. M. (1994). Shattering the illusion of control: Multi-shot versus single-shot gambles. *Journal of Behavioral Decision Making*, *7*(3), 183–191. https://doi.org/10/fwmwjs

Kokkinis, A. (2019). Exploring the effects of the 'bonus cap' rule: The impact of remuneration structure on risk-taking by bank managers. *Journal of Corporate Law Studies*, *19*(1), 167–195. https://doi.org/10/gjspz9

Koller, T., Lovallo, D., & Williams, Z. (2012). Overcoming a bias against risk. *McKinsey Quarterly*, (4). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/overcoming-a-bias-against-risk

Koller, T., Lovallo, D., & Williams, Z. (2017). Should assessing financial similarity be part of your corporate portfolio strategy? *McKinsey on Finance*, (64). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/should-assessing-financial-similarity-be-part-of-your-corporate-portfolio-strategy

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., & Lipson, A. (1993). Inconsistencies in Students' Reasoning about Probability. *Journal for Research in Mathematics Education*, *24*(5), 392. https://doi.org/10/bq4hvm

Kox, E. (2018). *Evaluating the effectiveness of uncertainty visualizations: A user-centered approach* (Masters thesis). University of Utrecht. http://dspace.library.uu.nl/handle/1874/367380

Kurtz, K. J., Boukrina, O., & Gentner, D. (2013). Comparison promotes learning and transfer of relational categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(4), 1303. https://doi.org/10/gjvc63

Lakens, D., & Caldwell, A. R. (2019, May 28). *Simulation-Based Power-Analysis for Factorial ANOVA Designs* (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/baxsf

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10/gdj7s9

Landau, W. M. (2021a). *Tarchetypes: Archetypes for targets* [https://docs.ropensci.org/tarchetypes/, https://github.com/ropensci/tarchetypes].

Landau, W. M. (2021b). The targets R package: A dynamic Make-like function-oriented pipeline toolkit for reproducibility and high-performance computing. *Journal of Open Source Software*, *6*(57), 2959. https://doi.org/10/ghwrqs

Lang, L. H. P., & Stulz, R. M. (1994). Tobin's q, Corporate Diversification, and Firm Performance. *Journal of Political Economy*, *102*(6), 1248–1280. https://doi.org/10/dgjvg3

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32*(2), 311–328. https://doi.org/10/bhghpg

Langer, T., & Weber, M. (2001). Prospect Theory, Mental Accounting, and Differences in Aggregated and Segregated Evaluation of Lottery Portfolios. *Management Science*, *47*(5), 716–733. https://doi.org/10/fcfk69

Lapinski, A.-L. S. (2009, October 1). *A Strategy for Uncertainty Visualization Design.* Defence R&D Canada – Atlantic. https://apps.dtic.mil/sti/citations/ADA523694

Lassaline, M. E. (1996). Structural alignment in induction and similarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(3), 754–770. https://doi.org/10/fq9fww

Lenth, R. V. (2021). *Emmeans: Estimated marginal means, aka least-squares means* [R package version 1.5.4-09004]. https://github.com/rvlenth/emmeans

Li, S. (2003). The role of Expected Value illustrated in decision-making under risk: Single-play vs multiple-play. *Journal of Risk Research*, *6*(2), 113–124. https://doi.org/10/cz6phv

Liebeskind, J. P. (2000). Internal Capital Markets: Benefits, Costs, and Organizational Arrangements. *Organization Science*, *11*(1), 58–76. https://doi.org/10/dkmq6t

Lipkus, I. M., & Hollands, J. G. (1999). The Visual Communication of Risk. *JNCI Monographs*, *1999*(25), 149–163. https://doi.org/10/gd589v

Lipkus, I. M. (2007). Numeric, Verbal, and Visual Formats of Conveying Health Risks: Suggested Best Practices and Future Recommendations. *Medical Decision Making*, *27*(5), 696–713. https://doi.org/10/b8p3gf

Liu, B., & McConnell, J. J. (2013). The role of the media in corporate governance: Do the media influence managers' capital allocation decisions? *Journal of Financial Economics*, *110*(1), 1–17. https://doi.org/10/gffggq

Liu, H.-H., & Colman, A. M. (2009). Ambiguity aversion in the long run: Repeated decisions under risk and uncertainty. *Journal of Economic Psychology*, *30*(3), 277–284. https://doi.org/10/d5p9kw

Long, A. R., Fernbach, P. M., & De Langhe, B. (2018). Circle of Incompetence: Sense of Understanding as an Improper Guide to Investment Risk. *Journal of Marketing Research*, *55*(4), 474–488. https://doi.org/10/gjscr7

Lopes, L. L. (1996). When Time Is of the Essence: Averaging, Aspiration, and the Short Run. *Organizational Behavior and Human Decision Processes*, *65*(3), 179–189. https://doi.org/10/fdtw45

Lovallo, D., Clarke, C., & Camerer, C. (2012). Robust analogizing and the outside view: Two empirical tests of case-based decision making. *Strategic Management Journal*, *33*(5), 496–512. https://doi.org/10/dnkh8m

Lovallo, D., & Kahneman, D. (2003). Delusions of Success: How Optimism Undermines Executives' Decisions. *Harvard Business Review*, *81*(7).

Lovallo, D., Koller, T., Uhlaner, R., & Kahneman, D. (2020). Your Company Is Too Risk-Averse. *Harvard Business Review*, *98*(2), 104–111.

Lovallo, D., & Sibony, O. (2014). Is your budget process stuck on last year's numbers? *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/is-your-budget-process-stuck-on-last-years-numbers

Lubatkin, M., & Shrieves, R. E. (1986). Towards Reconciliation of Market Performance Measures to Strategic Management Research. *Academy of Management Review*, *11*(3), 497–512. https://doi.org/10/ftmktj

Ludvig, E. A., & Spetch, M. L. (2011). Of Black Swans and Tossed Coins: Is the Description-Experience Gap in Risky Choice Limited to Rare Events? *PLoS ONE*, *6*(6), e20262. https://doi.org/10/ds5q2k

Lyngs, U. (2019). Oxforddown: An oxford university thesis template for r markdown. https://doi.org/10.5281/zenodo.3484682

MacCrimmon, K. R., Wehrung, D. A., & Stanbury, W. T. (1986). *Taking risks: The management of uncertainty*. Free Press.

MacEachren, A. M. (1992). Visualizing Uncertain Information. *Cartographic Perspectives*, (13), 10–19. https://doi.org/10/gjscq9

Mahr, T. (2021). *Printy: Helper functions for pretty-printing numbers* [R package version 0.0.0.9003].

Mannes, A. E., & Moore, D. A. (2013). A Behavioral Demonstration of Overconfidence in Judgment. *Psychological Science*, *24*(7), 1190–1197. https://doi.org/10/f443q6

March, J. G., & Shapira, Z. (1987). Managerial Perspectives on Risk and Risk Taking. *Management Science*, *33*(11), 1404–1418. https://doi.org/10/ft2phq

Markman, A. B., & Gentner, D. (1993). Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, *25*(4), 431–467. https://doi.org/10/cqtx7q

Markman, A. B., & Loewenstein, J. (2010). Structural comparison and consumer choice. *Journal of Consumer Psychology*, *20*(2), 126–137. https://doi.org/10/d7b49c

Markman, A. B., & Medin, D. L. (1995). Similarity and Alignment in Choice. *Organizational Behavior and Human Decision Processes*, *63*(2), 117–130. https://doi.org/10/c8z7r9

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10/bhzd

Mazzolini, R. (1981). How strategic decisions are made. *Long Range Planning*, *14*(3), 85–96. https://doi.org/10/dm8852
McBain, M. (2020). *Dflow: Setup a project in the dflow style for using the drake* [R package version 0.0.0.9000].
McCray, G. E., Purvis, R. L., & McCray, C. G. (2002). Project Management under Uncertainty: The Impact of Heuristics and Biases. *Project Management Journal*, *33*(1), 49–57. https://doi.org/10/gjscsr
McKenzie, C. R., Liersch, M. J., & Yaniv, I. (2008). Overconfidence in interval estimates: What does expertise buy you? *Organizational Behavior and Human Decision Processes*, *107*(2), 179–191. https://doi.org/10/b22rw7
Meissner, P., Sibony, O., & Wulf, T. (2015). Are you ready to decide? *McKinsey Quarterly*. https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/are-you-ready-to-decide
Mintzberg, H., Raisinghani, D., & Théorêt, A. (1976). The Structure of ”Unstructured” Decision Processes. *Administrative Science Quarterly*, *21*(2), 246–275. https://doi.org/10/b2t8ks
Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal*, *6*(3), 257–272. https://doi.org/10/bs2pm3
Moher, E., & Koehler, D. J. (2010). Bracketing effects on risk tolerance: Generalizability and underlying mechanisms. *Judgment and Decision Making*, *5*(5), 8. http://journal.sjdm.org/10/10422/jdm10422.pdf
Montgomery, H., & Adelbratt, T. (1982). Gambling decisions and information about expected value. *Organizational Behavior and Human Performance*, *29*(1), 39–57. https://doi.org/10/cvgjp4
Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames* [R package version 3.1.2]. https://CRAN.R-project.org/package=tibble
Noda, T., & Bower, J. L. (1996). Strategy making as iterated processes of resource allocation. *Strategic Management Journal*, *17*(S1), 159–192. https://doi.org/10/brg23k
Ooms, J. (2021). *Magick: Advanced graphics and image-processing in r* [R package version 2.6.0]. https://CRAN.R-project.org/package=magick
Ortner, J., Velthuis, L., & Wollscheid, D. (2017). Incentive systems for risky investment decisions under unknown preferences. *Management Accounting Research*, *36*, 43–50. https://doi.org/10/gbz24d
Padilla, L. M., Creem-Regehr, S. H., Hegarty, M., & Stefanucci, J. K. (2018). Decision making with visualizations: A cognitive framework across disciplines. *Cognitive Research: Principles and Implications*, *3*(1), 29. https://doi.org/10/ggrtng
Pang, A. T., Wittenbrink, C. M., & Lodha, S. K. (1997). Approaches to uncertainty visualization. *The Visual Computer*, *13*(8), 370–390. https://doi.org/10/fdnbmw
Parayre, R. (1995). The strategic implications of sunk costs: A behavioral perspective. *Journal of Economic Behavior & Organization*, *28*(3), 417–442. https://doi.org/10/d2wg72
Pascal, B. (1999, June 10). *Pensées and other writings* (A. Levi, Ed.; H. Levi, Trans.). Oxford University Press. (Original work published 1670)
Pedersen, T. L. (2020). *Patchwork: The composer of plots* [R package version 1.1.1]. https://CRAN.R-project.org/package=patchwork
Penrose, E. (2009). *The Theory of the Growth of the Firm* (4. ed., rev. ed). Oxford University Press. (Original work published 1959)

Porter, M. E. (1980). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press.

Porter, M. E. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press.

Porter, M. E. (1987). From competitive advantage to corporate strategy. *Harvard Business Review*, *65*(3), 43–43. https://hbr.org/1987/05/from-competitive-advantage-to-corporate-strategy

Potter, K., Kirby, R. M., Xiu, D., & Johnson, C. R. (2012). Interactive visualization of probability and cumulative density functions. *International Journal for Uncertainty Quantification*, *2*(4), 397–412. https://doi.org/10/ghhdw2

Powell, T. C., Lovallo, D., & Fox, C. R. (2011). Behavioral strategy. *Strategic Management Journal*, *32*(13), 1369–1386. https://doi.org/10/d8m8pw

Puri, M., & Robinson, D. T. (2007). Optimism and economic choice. *Journal of Financial Economics*, *86*(1), 71–99. https://doi.org/10/c9839j

R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Rabin, M., & Weizsäcker, G. (2009). Narrow Bracketing and Dominated Choices. *American Economic Review*, *99*(4), 1508–1543. https://doi.org/10/fk4rmz

Rajan, R., Servaes, H., & Zingales, L. (2000). The Cost of Diversity: The Diversification Discount and Inefficient Investment. *The Journal of Finance*, *55*(1), 35–80. https://doi.org/10/bdsh6h

Ratcliff, C. L., & Sun, Y. (2020). Overcoming Resistance Through Narratives: Findings from a Meta-Analytic Review. *Human Communication Research*, *46*(4), 412–443. https://doi.org/10/gjscrm

Read, D., & Loewenstein, G. (1995). Diversification bias: Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied*, *1*(1), 34. https://doi.org/10/fcgvrw

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice Bracketing. *Journal of Risk and Uncertainty*, *19*(1), 171–197. https://doi.org/10/dh3rmv

Redelmeier, D. A., & Tversky, A. (1992). On the Framing of Multiple Prospects. *Psychological Science*, *3*(3), 191–193. https://doi.org/10/ctw2k6

Reichelstein, S. (1997). Investment Decisions and Managerial Performance Evaluation. *Review of Accounting Studies*, *2*(2), 157–180. https://doi.org/10/fp2zqc

Reinard, J. C. (1988). The Empirical Study of the Persuasive Effects of Evidence The Status After Fifty Years of Research. *Human Communication Research*, *15*(1), 3–59. https://doi.org/10/ccb67v

Reinhart, A. M. (2006). *Comparing the persuasive effects of narrative versus statistical messages: A meta -analytic review* (Ph.D.). State University of New York at Buffalo. New York, US. https://www.researchgate.net/profile/Amber-Reinhart/publication/34707525_Comparing_the_persuasive_effects_of_narrative_versus_statistical_messages_electronic_resource_A_meta-analytic_review/links/57335b6108ae9f741b26120c/Comparing-the-persuasive-effects-of-narrative-versus-statistical-messages-electronic-resource-A-meta-analytic-review.pdf

Remer, D. S., Stokdyk, S. B., & Van Driel, M. (1993). Survey of project evaluation techniques currently used in industry. *International Journal of Production Economics*, *32*(1), 103–115. https://doi.org/10/bsc6bs

Ristovski, G., Preusser, T., Hahn, H. K., & Linsen, L. (2014). Uncertainty in medical visualization: Towards a taxonomy. *Computers & Graphics*, *39*, 60–73. https://doi.org/10/f5v59d

Robinson, C. L. (1944). Getting Most Impressions Per Dollar. *Broadcasting*, *27*(1), 13. https://worldradiohistory.com/Archive-BC/BC-1944/1944-07-03-BC.pdf

Robinson, D., Hayes, A., & Couch, S. (2021). *Broom: Convert statistical objects into tidy tibbles* [https://broom.tidymodels.org/, https://github.com/tidymodels/broom].

Rosenbaum, D. M., Glickman, M., & Usher, M. (2020, July 30). *Extracting summary statistics of rapid numerical sequences* (preprint). PsyArXiv. https://doi.org/10.31234/osf.io/6scav

Ross, S. A. (1999). Adding Risks: Samuelson's Fallacy of Large Numbers Revisited. *The Journal of Financial and Quantitative Analysis*, *34*(3), 323–339. https://doi.org/10/bj6r8r

Roxburgh, C. (2003). Hidden flaws in strategy. *McKinsey Quarterly*, (2). https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/hidden-flaws-in-strategy

Ruggeri, K., Alí, S., Berge, M. L., Bertoldo, G., Bjørndal, L. D., Cortijos-Bernabeu, A., Davison, C., Demić, E., Esteban-Serna, C., Friedemann, M., Gibson, S. P., Jarke, H., Karakasheva, R., Khorrami, P. R., Kveder, J., Andersen, T. L., Lofthus, I. S., McGill, L., Nieto, A. E., … Folke, T. (2020). Replicating patterns of prospect theory for decision under risk. *Nature Human Behaviour*, *4*(6), 622–633. https://doi.org/10/ggwrdh

Rumelt, R. P. (1974). *Strategy, Structure, and Economic Performance*. Division of Research, Graduate School of Business Administration, Harvard University.

Samuelson, P. A. (1963). Risk and Uncertainty: A Fallacy of Large Numbers. *Scientia*, *57*(98), 108–113. https://www.casact.org/sites/default/files/database/forum_94sforum_94sf049.pdf

Sarama, J., & Clements, D. H. (2009). Building Blocks and Cognitive Building Blocks: Playing to Know the World Mathematically. *American Journal of Play*, *1*(3), 313–337. Retrieved June 23, 2021, from https://eric.ed.gov/?id=EJ1069014

Schaefer, R. E., & Borcherding, K. (1973). The assessment of subjective probability distributions: A training experiment. *Acta Psychologica*, *37*(2), 117–129. https://doi.org/10/dpzkfb

Schwenk, C. R. (1984). Cognitive simplification processes in strategic decision-making. *Strategic Management Journal*, *5*(2), 111–128. https://doi.org/10/b2hsgw

Scott, A. (2018). GE breakup leaves it with best and worst performers [newspaper]. *Reuters*. https://www.reuters.com/article/us-ge-divestiture-idUSKBN1JM0ZT

Sengul, M., Costa, A. A., & Gimeno, J. (2019). The Allocation of Capital within Firms. *Academy of Management Annals*, *13*(1), 43–83. https://doi.org/10/gfgvqj

Shapira, Z., & Shaver, J. M. (2014). Confounding changes in averages with marginal effects: How anchoring can destroy economic value in strategic investment assessments. *Strategic Management Journal*, *35*(10), 1414–1426. https://doi.org/10/gkhfb3

Shelton, L. M. (1988). Strategic business fits and corporate acquisition: Empirical evidence. *Strategic Management Journal*, *9*(3), 279–287. https://doi.org/10/fbhcgd

Shen, F., Sheer, V. C., & Li, R. (2015). Impact of Narratives on Persuasion in Health Communication: A Meta-Analysis. *Journal of Advertising*, *44*(2), 105–113. https://doi.org/10/gfkwj7

Sibony, O., Lovallo, D., & Powell, T. C. (2017). Behavioral Strategy and the Strategic Decision Architecture of the Firm. *California Management Review*, *59*(3), 5–21. https://doi.org/10/gcp2w3

Simon, H. A. (1955). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, *69*(1), 99. https://doi.org/10/dw3pfg

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2021). *Afex: Analysis of factorial experiments* [R package version 0.28-1]. https://CRAN.R-project.org/package=afex

Smith, J. F., & Kida, T. (1991). Heuristics and biases: Expertise and task realism in auditing. *Psychological bulletin*, *109*(3), 472–489. https://doi.org/10/fwv6z6

Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2012). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience*, *8*(3), 341–350. https://doi.org/10/fx5cn6

Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*(13), 5035–5040. https://doi.org/10/fhdrcw

Soll, J. B., & Klayman, J. (2004). Overconfidence in Interval Estimates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 299–314. https://doi.org/10/dtpsvd

Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, *333*(6048), 1393–1400. https://doi.org/10.1126/science.1191181

Staw, B. M. (1981). The Escalation of Commitment To a Course of Action. *Academy of Management Review*, *6*(4), 577–587. https://doi.org/10/fdqtjg

Stephens, J., Simonov, K., Xie, Y., Dong, Z., Wickham, H., Horner, J., reikoch, Beasley, W., O'Connor, B., & Warnes, G. R. (2020). *Yaml: Methods to convert r data to yaml and back* [R package version 2.2.1]. https://CRAN.R-project.org/package=yaml

Stutzer, M. (2013). Misperceptions of long-term investment performance Insights from an experiment. *The Journal of Behavioral Finance & Economics*, *3*(1), 1–20. http://www.aobf.org/wp-content/uploads/2020/06/1-Stutzer-and-Grant.pdf

Su, Y., Rao, L.-L., Sun, H.-Y., Du, X.-L., Li, X., & Li, S. (2013). Is making a risky choice based on a weighting and adding process? An eye-tracking investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(6), 1765–1780. https://doi.org/10/gjscr2

Swalm, R. O. (1966). Utility Theory–Insights into Risk Taking. *Harvard Business Review*, *44*(6), 123–136.

Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable*. Random House.

Teece, D. J. (1980). Economies of scope and the scope of the enterprise. *Journal of Economic Behavior & Organization*, *1*(3), 223–247. https://doi.org/10/fxcj3r

Teece, D. J. (1982). Towards an economic theory of the multiproduct firm. *Journal of Economic Behavior & Organization*, *3*(1), 39–63. https://doi.org/10/bdwvvq

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral decision making*, *12*(3), 183–206. https://doi.org/10/d4njp3

Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *The Quarterly Journal of Economics*, *112*(2), 647–661. https://doi.org/10/fcf346

Thamhain, H. J. (2014). Assessing the Effectiveness of Quantitative and Qualitative Methods for R&D Project Proposal Evaluations. *Engineering Management Journal*, *26*(3), 3–12. https://doi.org/10/gjsm2q

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The Neural Basis of Loss Aversion in Decision-Making Under Risk. *Science*, *315*(5811), 515–518. https://doi.org/10/bhtqgn

Torsney-Weir, T., Sedlmair, M., & Möller, T. (2015). Decision making in uncertainty visualization. http://eprints.cs.univie.ac.at/4598/

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10/gwh

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458. https://doi.org/10/fj3z3r

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*(4), 297–323. https://doi.org/10/cb57hk

Ullrich, M. J., & Tuttle, B. M. (2004). The Effects of Comprehensive Information Reporting Systems and Economic Incentives on Managers' Time-Planning Decisions. *Behavioral Research in Accounting*, *16*, 89–105. https://doi.org/10/b99r4n

Ushey, K. (2021). *Renv: Project environments*. manual. https://CRAN.R-project.org/package=renv

Venkatraman, S., Aloysius, J. A., & Davis, F. D. (2006). Multiple prospect framing and decision behavior: The mediational roles of perceived riskiness and perceived ambiguity. *Organizational Behavior and Human Decision Processes*, *101*(1), 59–73. https://doi.org/10/dszh5v

Vivalt, E., & Coville, A. (2021, April 23). *How Do Policy-Makers Update Their Beliefs?* http://evavivalt.com/wp-content/uploads/How-Do-Policymakers-Update.pdf

von Holstein, C.-A. S. S. (1971). Two techniques for assessment of subjective probability distributions — An experimental study. *Acta Psychologica*, *35*(6), 478–494. https://doi.org/10/fgg6jn

von Neumann, J., Morgenstern, O., & Rubinstein, A. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

Wainberg, J. S. (2018, January 1). Stories vs Statistics: The Impact of Anecdotal Data on Managerial Decision Making. *Advances in Accounting Behavioral Research* (pp. 127–141). Emerald Publishing Limited. https://doi.org/10.1108/S1475-148820180000021006

Wainberg, J. S., Kida, T., David Piercey, M., & Smith, J. F. (2013). The impact of anecdotal data in regulatory audit firm inspection reports. *Accounting, Organizations and Society*, *38*(8), 621–636. https://doi.org/10/gjscqz

Webb, E. C., & Shu, S. B. (2017). Is broad bracketing always better? How broad decision framing leads to more optimal preferences over repeated gambles. *Judgment and Decision Making*, *12*(4), 382. http://journal.sjdm.org/17/17317/jdm17317.pdf

Weber, E. U., & Hsee, C. (1998). Cross-Cultural Differences in Risk Perception, but Cross-Cultural Similarities in Attitudes Towards Perceived Risk. *Management Science*, *44*(9), 1205–1217. https://doi.org/10/c4rqqd

Wedell, D. H., & Bockenholt, U. (1994). Contemplating Single versus Multiple Encounters of a Risky Prospect. *The American Journal of Psychology*, *107*(4), 499. https://doi.org/10/b4fs2p

Wedell, D. H., & Böckenholt, U. (1990). Moderation of preference reversals in the long run. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 429–438. https://doi.org/10/bmn8hf

Wernerfelt, B. (1984). A Resource-Based View of the Firm. *Strategic Management Journal*, *5*(2), 171–180. https://doi.org/10/cm664p

Wernerfelt, B., & Montgomery, C. A. (1988). Tobin's q and the Importance of Focus in Firm Performance. *The American Economic Review*, *78*(1), 246–250.

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2019a). *Conflicted: An alternative conflict resolution strategy* [R package version 1.0.4]. https://CRAN.R-project.org/package=conflicted

Wickham, H. (2019b). *Stringr: Simple, consistent wrappers for common string operations* [R package version 1.4.0]. https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors)* [R package version 0.5.1]. https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). *Tidyr: Tidy messy data* [R package version 1.1.3]. https://CRAN.R-project.org/package=tidyr

Wickham, H., & Bryan, J. (2021). *Usethis: Automate package and project setup* [R package version 2.0.1]. https://CRAN.R-project.org/package=usethis

Wickham, H., François, R., Henry, L., & Müller, K. (2021). *Dplyr: A grammar of data manipulation* [https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr].

Wickham, H., Hester, J., & Chang, W. (2021). *Devtools: Tools to make developing r packages easier* [R package version 2.4.1]. https://CRAN.R-project.org/package=devtools

Wickham, H., & Seidel, D. (2020). *Scales: Scale functions for visualization* [R package version 1.1.1]. https://CRAN.R-project.org/package=scales

Wilke, C. O. (2020). *Cowplot: Streamlined plot theme and plot annotations for 'ggplot2'* [R package version 1.1.1]. https://CRAN.R-project.org/package=cowplot

Williamson, O. E. (1981). The Economics of Organization: The Transaction Cost Approach. *American Journal of Sociology*, *87*(3), 548–577. https://doi.org/10/bqzkpb

Willigers, B. J. A., Jones, B., & Bratvold, R. B. (2017). The Net-Present-Value Paradox: Criticized by Many, Applied by All. *SPE Economics & Management*, *9*(04), 090–102. https://doi.org/10/gjscsx

Winterbottom, A., Bekker, H. L., Conner, M., & Mooney, A. (2008). Does narrative information bias individual's decision making? A systematic review. *Social Science & Medicine*, *67*(12), 2079–2088. https://doi.org/10/cfpr4z

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (2018). A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin*, *144*(2), 140–176. https://doi.org/10/gc2s4c

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd) [ISBN 978-1498716963]. Chapman; Hall/CRC. https://yihui.org/knitr/

Xie, Y. (2016). *Bookdown: Authoring books and technical documents with R markdown*. Chapman and Hall/CRC. https://bookdown.org/yihui/bookdown

Xie, Y., Allaire, J., & Grolemund, G. (2018). *R markdown: The definitive guide*. Chapman and Hall/CRC. https://bookdown.org/yihui/rmarkdown

Yang, Z., Saini, R., & Freling, T. (2015). How Anxiety Leads to Suboptimal Decisions Under Risky Choice Situations. *Risk Analysis*, *35*(10), 1789–1800. https://doi.org/10/f7wtcx

Yates, J. F., Jagacinski, C. M., & Faber, M. D. (1978). Evaluation of partially described multiattribute options. *Organizational Behavior and Human Performance*, *21*(2), 240–251. https://doi.org/10/br3dbv

Zebregs, S., van den Putte, B., Neijens, P., & de Graaf, A. (2015). The Differential Impact of Statistical and Narrative Evidence on Beliefs, Attitude, and Intention: A Meta-Analysis. *Health Communication*, *30*(3), 282–289. https://doi.org/10/ghk97p

Zeisberger, S. (2020, May 15). *Do People Care About Loss Probabilities?* https://ssrn.com/abstract=2169394