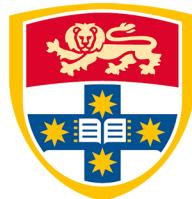


The psychology of managerial risky choice and resource allocation



THE UNIVERSITY OF
SYDNEY

Shir Dekel

The University of Sydney

A thesis submitted for the degree of

Doctor of Philosophy (Science)

2021-05-24 17:23:47

Abstract

Resource allocation decisions are critical for large organisations. The management literature mainly considers such decisions from an organisational perspective, despite the potential psychological influences. Therefore, I investigated people's cognitive limitations in the resource allocation process. I did this through three paradigms that provided participants both opportunities to make use of clear cues to moderate their decisions and also important statistical principles that would enhance their decisions even further. First, participants saw sequential risky choices without feedback. Combining the risk through *risk aggregation* allowed participants to reduce the overall potential loss. Second, participants were asked to allocate capital across a set of business projects with different attributes. The *estimate variance* associated with uncertain forecasts allowed for a better moderation of the metric used to make the choice. Third, participants saw either relevant or irrelevant anecdotal evidence that conflicted with statistical evidence for a project. The *sample distribution* of similarity to the target project from the sample of anecdotes allowed participants to clarify this conflicting evidence. The results showed that people struggled to use these statistical principles, unless in the case of risk aggregation when it was depicted visually. Otherwise, participants did appropriately use information about forecast metric reliability when this was expressed verbally, rather than numerically. Similarly, despite ignoring sample distribution information, participants moderated their use of anecdotes by the similarity of the anecdote to the target project. Further, participants tended to integrate conflicting information about the projects through a trade-off. These results show that people's resource allocation decisions are bounded by an understanding of certain statistical principles, but that they are capable of more nuanced choice when given more salient information.

Contents

| | |
|--|-------------|
| List of Figures | viii |
| List of Tables | xv |
| 1 Introduction | 1 |
| 1.1 Resource allocation in hierarchical organisations | 3 |
| 1.2 Psychology of resource allocation | 7 |
| 1.2.1 Risky choice | 9 |
| 1.2.2 Project similarity | 12 |
| 1.2.3 Reasoning from past cases | 14 |
| 1.3 Chapter overview | 16 |
| References | 18 |
| 2 Effect of choice bracketing on risk aggregation in repeated-play gambles with no feedback | 24 |
| 2.1 Introduction | 25 |
| 2.1.1 Multi-play gambles | 26 |
| 2.1.2 Repeated-play gambles | 28 |
| 2.1.3 Choice bracketing | 30 |
| 2.1.4 Internal market capital investment context | 33 |
| 2.2 Experiment 1 | 35 |
| 2.2.1 Method | 36 |
| 2.2.2 Results | 39 |
| 2.2.3 Discussion | 41 |
| 2.3 Experiment 2 | 46 |
| 2.3.1 Method | 47 |
| 2.3.2 Results | 48 |
| 2.3.3 Discussion | 50 |
| 2.4 General discussion | 53 |

Contents

| | | |
|----------------------|---|-----------|
| 2.4.1 | Theoretical implications | 54 |
| 2.4.2 | Practical implications | 57 |
| 2.4.3 | Future research | 57 |
| References | | 59 |
| 3 | Joint evaluation of multiple projects | 63 |
| 4 | Project similarity bias and variance neglect in forecast metric evaluation | 66 |
| 4.1 | Introduction | 66 |
| 4.1.1 | Experiment summary | 69 |
| 4.2 | Experiment 1 | 70 |
| 4.2.1 | Method | 71 |
| 4.2.2 | Results | 75 |
| 4.2.3 | Discussion | 77 |
| 4.3 | Experiment 2 | 77 |
| 4.3.1 | Method | 78 |
| 4.3.2 | Results | 81 |
| 4.3.3 | Discussion | 82 |
| 4.4 | Experiment 3 | 83 |
| 4.4.1 | Method | 84 |
| 4.4.2 | Results | 87 |
| 4.4.3 | Discussion | 89 |
| 4.5 | General discussion | 90 |
| References | | 93 |
| 5 | Looking for alignment in past cases | 95 |
| 6 | Anecdote similarity moderates anecdotal bias in resource allocation | 98 |
| 6.1 | Introduction | 99 |
| 6.1.1 | Anecdotal bias | 100 |
| 6.1.2 | Anecdotal bias in business | 101 |
| 6.1.3 | Effect of similarity | 102 |
| 6.1.4 | Experiment summary | 104 |
| 6.2 | Experiment 1 | 104 |
| 6.2.1 | Method | 106 |
| 6.2.2 | Results | 111 |

Contents

| | | |
|----------|------------------------------------|------------|
| 6.2.3 | Discussion | 114 |
| 6.3 | Experiment 2 | 116 |
| 6.3.1 | Method | 118 |
| 6.3.2 | Results | 120 |
| 6.3.3 | Discussion | 124 |
| 6.4 | General discussion | 126 |
| 6.4.1 | Theoretical implications | 128 |
| 6.4.2 | Practical implications | 129 |
| | References | 131 |
| 7 | Discussion | 133 |
| 7.1 | Summary of results | 134 |
| 7.2 | Theoretical implications | 136 |
| 7.3 | Practical implications | 140 |
| 7.4 | Conclusion | 142 |
| | References | 143 |
| A | Chapter 2 appendix | 146 |
| A.1 | Experiment 1 | 147 |
| A.1.1 | Method | 147 |
| A.1.2 | Results | 149 |
| A.2 | Experiment 2 | 150 |
| A.2.1 | Method | 150 |
| A.2.2 | Results | 152 |
| A.3 | Experiment 3 | 154 |
| A.3.1 | Method | 155 |
| A.3.2 | Results | 158 |
| A.3.3 | Discussion | 162 |
| A.4 | Experiment 4 | 164 |
| A.4.1 | Method | 165 |
| A.4.2 | Results | 167 |
| A.4.3 | Discussion | 173 |

Contents

| | |
|-----------------------------|------------|
| B Chapter 4 appendix | 174 |
| B.1 Experiment 1 | 175 |
| B.1.1 Method | 176 |
| B.1.2 Results | 176 |
| B.1.3 Discussion | 184 |
| B.2 Experiment 2 | 185 |
| B.2.1 Method | 185 |
| B.2.2 Results | 195 |
| B.3 Experiment 3 | 197 |
| B.3.1 Method | 197 |
| B.3.2 Results | 200 |
| B.4 Experiment 4 | 201 |
| B.4.1 Method | 203 |
| B.4.2 Results | 204 |
| B.4.3 Discussion | 207 |
| B.5 Experiment 5 | 208 |
| B.5.1 Method | 208 |
| B.5.2 Results | 209 |
| B.5.3 Discussion | 210 |
| B.6 Experiment 6 | 212 |
| B.6.1 Method | 212 |
| B.6.2 Results | 213 |
| B.6.3 Discussion | 214 |
| B.7 Experiment 7 | 215 |
| B.7.1 Method | 215 |
| B.7.2 Results | 217 |
| B.7.3 Discussion | 217 |
| B.8 Experiment 8 | 218 |
| B.8.1 Method | 219 |
| B.8.2 Results | 220 |
| B.8.3 Discussion | 225 |

Contents

| | |
|-----------------------------|------------|
| C Chapter 6 appendix | 227 |
| C.1 Experiment 1 | 227 |
| C.1.1 Method | 228 |
| C.1.2 Results | 231 |
| C.2 Experiment 2 | 236 |
| C.2.1 Method | 239 |
| C.2.2 Results | 241 |
| References | 247 |

List of Figures

| | | |
|-----|--|----|
| 1.1 | The value function in Prospect Theory (adapted from Kahneman & Tversky, 1979). The values for the exponent representing the extent diminishing returns (α) and the loss aversion coefficient (λ) were 0.88 and 2.25, respectively. These were the median values calculated by Tversky and Kahneman (1992) from experimental data. | 11 |
| 2.1 | The outcome probability distribution of the Samuelson (1963) gamble (0.5, 200; 0.5, -100) played 10 times. | 29 |
| 2.2 | Example of a project choice display in Experiment 1. Border added for clarity. | 38 |
| 2.3 | The display seen by those in the aware condition of Experiment 1. Border added for clarity. | 38 |
| 2.4 | Mean proportions of decisions to invest in each set of 10 projects, by awareness conditions. Error bars represent 95% confidence intervals. | 40 |
| 2.5 | Mean proportions of decisions to invest in each set of 10 projects, by presentation conditions. Error bars represent 95% confidence intervals. | 41 |
| 2.6 | Mean proportions of decisions to invest in each set of 10 projects, by similarity and presentation conditions. Error bars represent 95% confidence intervals. | 42 |
| 2.7 | Proportion of project acceptance in the separate presentation condition, by trial, and awareness conditions. LOESS is used for smoothing over trials, and the shading represents 95% confidence intervals. | 43 |
| 2.8 | An example of a display seen by those in the separate distribution-present condition of Experiment 2. | 49 |
| 2.9 | Mean project acceptance 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. | 50 |

List of Figures

| | | |
|------|---|-----|
| 2.10 | Mean proportion of project acceptance 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. | 51 |
| 2.11 | Mean project acceptance for separate presentation, distribution absent condition, by awareness and trial. | 52 |
| 4.1 | An example of a low alignment display in Experiment 1. Border added for clarity. | 74 |
| 4.2 | An example of a high alignment display in Experiment 1. | 75 |
| 4.3 | The allocation task. | 75 |
| 4.4 | Mean allocation. | 76 |
| 4.5 | An example of a low alignment low reliability display in Experiment 2. Border added for clarity. | 80 |
| 4.6 | An example of a high alignment high reliability display in Experiment 2. | 81 |
| 4.7 | Mean allocation. | 82 |
| 4.8 | An example of a low alignment, low verbal reliability display in Experiment 3. | 86 |
| 4.9 | An example of a high alignment, high numerical reliability display in Experiment 3. | 87 |
| 4.10 | Mean project allocation for all conditions. | 89 |
| 6.1 | Project display for the anecdote only condition in Experiment 1. . . | 108 |
| 6.2 | Project display for the statistics only, anecdote + statistics, and anecdote + enhanced statistics conditions in Experiment 1. | 109 |
| 6.3 | Anecdote display for those in the high alignment condition in Experiment 1. | 111 |
| 6.4 | Anecdote display for those in the low alignment condition in Experiment 1. | 112 |
| 6.5 | Mean allocation to Project A (the target project). Error bars represent 95% confidence intervals. | 114 |
| 6.6 | Mean allocations to the target project, by specific relevance rating and similarity condition. | 115 |
| 6.7 | An example of the anecdote display in the negative valence, low similarity condition of Experiment 2. | 120 |

List of Figures

| | |
|---|-----|
| 6.8 An example of the target projects in the negative valence, low similarity condition of Experiment 2. | 121 |
| 6.9 An example of an anecdote display in the positive valence, high similarity condition of Experiment 2. | 122 |
| 6.10 An example of the target projects in the positive valence, high similarity condition of Experiment 2. | 123 |
| 6.11 Mean allocation in Experiment 2 for the positive valence condition. | 125 |
| 6.12 Mean allocation in Experiment 2 for the positive valence condition. | 125 |
| A.1 Experiment 1 instructions. | 147 |
| A.2 The outcome distribution of the 10 gambles used in Experiment 1. | 148 |
| 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. | 149 |
| A.4 Experiment 2 project number question. | 150 |
| A.5 Experiment 2 binary portfolio question. | 151 |
| A.6 Experiment 2 numerical portfolio question. | 152 |
| A.7 Number of projects participants reported seeing, by condition. | 153 |
| 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. | 154 |
| 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. | 155 |
| A.10 Mean project acceptance for the 10 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value). | 156 |
| A.11 Mean project acceptance for the gambles' expected value, positive probability, and positive outcome. | 156 |
| A.12 An example of a project display in Experiment 3. | 157 |
| A.13 Experiment 3 project expectation question. | 158 |
| A.14 Mean project acceptance for the similarity effect. | 159 |
| A.15 Mean proportion of project acceptance for the similarity effect. | 159 |
| A.16 Mean project acceptance by similarity and trial. | 160 |

List of Figures

| | |
|---|-----|
| A.17 Number of projects participants expected to see, by similarity. | 161 |
| A.18 Number of projects participants reported seeing, by similarity. | 161 |
| A.19 Mean choice of investing in all 10 projects for the similarity effect. . | 162 |
| A.20 Mean number of projects chosen in the follow-up for the similarity effect. | 163 |
| A.21 Mean project acceptance for the 10 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value). . | 163 |
| A.22 Mean project acceptance for the gambles' expected value, positive probability, and positive outcome. | 164 |
| A.23 Instructions for those in the naive condition of Experiment 4. | 166 |
| A.24 Instructions for those in the aware condition of Experiment 4. | 166 |
| A.25 An example of a project display in Experiment 4. | 167 |
| A.26 Mean project acceptance for the awareness effect. | 168 |
| A.27 Mean proportion of project acceptance for the awareness effect. . | 168 |
| A.28 Mean project acceptance by awareness and trial. | 169 |
| A.29 Number of projects participants expected to see, by awareness. | 170 |
| A.30 Number of projects participants reported seeing, by awareness. | 171 |
| A.31 Mean choice of investing in all 20 projects for the awareness effect. . | 171 |
| A.32 Mean number of projects chosen in the follow-up for the awareness effect. | 172 |
| A.33 Mean project acceptance for the 20 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value). . | 172 |
| A.34 Mean project acceptance for the gambles' expected value, positive probability, and positive outcome. | 173 |
| B.1 Experiment 1 low reliability instructions. Border added for clarity. . | 177 |
| B.2 Experiment 1 high reliability instructions. Border added for clarity. . | 178 |
| B.3 The instructions for the no NPV condition in Experiment 1. Border added for clarity. | 179 |
| B.4 The forecasting task. | 179 |
| B.5 The ranking task. | 180 |
| B.6 The confidence task. | 180 |
| B.7 The justification task. | 181 |
| B.8 Mean ranking. | 181 |
| B.9 Mean confidence. | 182 |
| B.10 Mean forecasts. | 183 |

List of Figures

| | |
|---|-----|
| B.11 Mean forecast SD. | 184 |
| B.12 Experiment 2 instructions. Border added for clarity. | 185 |
| B.13 Experiment 2 NPV test. Border added for clarity. | 187 |
| B.14 Experiment 2 NPV test answers. Border added for clarity. | 188 |
| B.15 Experiment 2 NPV knowledge rating task. Border added for clarity. | 188 |
| B.16 Mean ranking. | 196 |
| B.17 Mean confidence. | 196 |
| B.18 Mean NPV knowledge rating. | 197 |
| B.19 Experiment 3 predicted data | 198 |
| B.20 Alignment Experiment 3 power curve. Labels indicate lowest sample size above 80% power. | 199 |
| B.21 Experiment 3 verbal reliability instructions. Border added for clarity. | 201 |
| B.22 Experiment 3 numerical reliability instructions. Border added for clarity. | 202 |
| B.23 An example of an interstitial display in Experiment 3. Border added for clarity. | 202 |
| B.24 Experiment 4 low reliability instructions. Border added for clarity. | 204 |
| B.25 Experiment 4 high reliability instructions. Border added for clarity. | 205 |
| B.26 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. | 206 |
| B.27 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. | 207 |
| B.28 An example of the forecasting task in Experiment 5. | 210 |
| B.29 Mean forecasts. | 211 |
| B.30 Mean forecast SD. | 211 |
| B.31 Mean forecasts. | 213 |
| B.32 Mean forecast SD. | 214 |
| B.33 The projects display. | 216 |
| B.34 Mean allocation. | 217 |
| B.35 Mean ranking. | 218 |
| B.36 Experiment 8 low alignment, verbal reliability display. Cropped for space (full display has eight projects). | 221 |

List of Figures

| | |
|---|-----|
| B.37 Experiment 8 low alignment, numerical reliability display. Cropped for space (full display has eight projects). | 222 |
| B.38 Experiment 8 high alignment, high reliability display. Cropped for space (full display has eight projects). | 223 |
| B.39 Experiment 8 high alignment, numerical reliability display. Cropped for space (full display has eight projects). | 224 |
| B.40 Mean project allocation, for the low alignment condition. Error bars represent 95% confidence intervals. | 225 |
| B.41 Mean project allocation, for the high alignment condition. Error bars represent 95% confidence intervals. | 226 |
| C.1 Power curves for the similarity and anecdote effects. | 230 |
| C.2 Experiment 1 general instructions. Note, the two boxes were split between two separate web-pages. | 230 |
| C.3 Experiment 1 specific instructions for those in the anecdotes only condition. | 230 |
| C.4 Experiment 1 specific instructions for those in the anecdote + statistics condition. | 231 |
| C.5 Experiment 1 specific instructions for those in the anecdote + enhanced statistics condition. | 232 |
| C.6 Experiment 1 specific instructions for those in the statistics only condition. | 233 |
| C.7 Follow-up questions in Experiment 1. | 233 |
| C.8 Mean similarity rating of Project A (the target project) to the anecdote. Error bars represent 95% confidence intervals. | 234 |
| C.9 Mean rating of how relevant participants thought the anecdote was to Project A (the target project). Error bars represent 95% confidence intervals. | 235 |
| C.10 Mean rating of how relevant participants thought the anecdote was to other oil projects. Error bars represent 95% confidence intervals. | 236 |
| C.11 Mean allocations to the target project by similarity rating | 237 |
| C.12 Rating of how relevant participants considered the anecdote to the target project, by similarity rating | 237 |
| C.13 Anecdotes Experiment 2 predicted data for the negative valence condition | 238 |
| C.14 Anecdotes Experiment 2 predicted data for the positive valence condition | 238 |

List of Figures

| | |
|---|-----|
| C.15 Anecdotes Experiment 2 power curve. Labels indicate lowest sample size above 80% power. | 240 |
| C.16 General instructions for Experiment 1. Border added for clarity. | 241 |
| C.17 Experiment 2 specific instructions for those in the anecdotes only condition. | 242 |
| C.18 Experiment 2 specific instructions for those in the combined condition. | 243 |
| C.19 Experiment 2 specific instructions for those in the statistics only condition. | 243 |
| C.20 An example of one of the follow-up question displays in Experiment 2. | 244 |
| C.21 An example of an interstitial display in Experiment 2. | 244 |
| C.22 Mean similarity rating of Project A (the target project) to the anecdote. Error bars represent 95% confidence intervals. | 245 |

List of Tables

| | | |
|-----|--|-----|
| 2.1 | Experiment 1 group allocation. | 36 |
| 2.2 | Experiment 2 group allocation. | 47 |
| 4.1 | Experiment 1 group allocation. | 72 |
| 4.2 | Experiment 3 group allocation. | 85 |
| 6.1 | Experiment 1 group allocation. | 106 |
| 6.2 | Experiment 2 group allocation. | 118 |
| A.1 | Experiment 3 group allocation. | 157 |
| A.2 | Logistic regression table of project acceptance by similarity and trial. | 160 |
| A.3 | Experiment 4 group allocation. | 165 |
| A.4 | Logistic regression table of project acceptance by awareness and trial. | 169 |
| B.1 | Experiment 4 group allocation. | 203 |
| B.2 | Experiment 5 group allocation. | 209 |
| B.3 | Experiment 6 group allocation. | 212 |
| B.4 | Experiment 7 group allocation. | 215 |

*An executive is a man who can make quick decisions
and is sometimes right.*

—Elbert Hubbard

1

Introduction

Contents

| | | |
|------------|--|-----------|
| 1.1 | Resource allocation in hierarchical organisations | 3 |
| 1.2 | Psychology of resource allocation | 7 |
| 1.2.1 | Risky choice | 9 |
| 1.2.2 | Project similarity | 12 |
| 1.2.3 | Reasoning from past cases | 14 |
| 1.3 | Chapter overview | 16 |
| | References | 18 |

Much of modern life depends on large organisations. General Electric (GE) makes the engines that go in our aircrafts, Johnson & Johnson make 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 such organisations 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,

1. Introduction

with a corporate management team and subsidiary divisions. Each division is usually 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 certain regions.

Such organisations therefore need to make resource allocation decisions. That is, given a limited amount of resources (e.g., financial, organisational), how best to invest in the multiple divisions? Equally? Pick a winner? Resource 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.

Managers of large organisations often appear as 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. However, the fact that much of an organisation's future (and by extension often many more components of the economy) depends on the decisions of a few individuals may be concerning for three reasons. First, it is unclear what the role of survivorship bias was in obtaining the manager's role, because the number of managers that used the same strategy and failed is unknown. Second, decades of work have shown that people's decision-making is often fallible and that job experience does not always alleviate this. Third, managers of large organisations often face uncertain environments, which strengthens psychological biases.

One class of biases has not been well studied: resource allocation biases. When making capital allocation decisions, there are elements of the decision making environment that can be deceiving for managers. In this thesis I examine how the framing of a series of business projects affects people's decisions about those

1. Introduction

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. I find that although people in general make sensible decisions, they fail to moderate them appropriately when presented with critical information. 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. In short, people overall tend to make sound decisions, but fail to moderate them in situations that have more subtle statistical implications.

Next, I will explain how the resource allocation process functions in hierarchical organisations, and why it is necessary to analyse such a process with a psychological approach. I will then review the literature on decision-making biases and how these may apply to resource allocation decisions. Finally, I will summarise the rest of the thesis chapters.

1.1 Resource 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. Reasons for this to work include reduced transaction costs (Coase, 1937; Teece, 1980, 1982; Williamson, 1981), shared resources (Barney, 1991; Wernerfelt, 1984), increased competitive advantage (Porter, 1985/1998a, 1980/1998b), 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 an hierarchical structure. Bower (1970) identified three levels of the typical management hierarchy: business, division, and corporate. These are equivalent to front-line (or

1. Introduction

bottom), middle, and top level managers (Noda & Bower, 1996). Early theorists suggested that the organisation's strategy originates completely in the top managers, with the rest of the organisation simply enacting 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. Subsequent work proposed and found evidence for an iterated process in which a strategic context may be set by top managers and projects advanced by lower level managers contribute to driving the strategy of the organisation (Bower, 1970; Burgelman, 1983; Noda & Bower, 1996).

The way that resources are allocated in an organisation is very important to its growth and longevity. Analysing the performance of organisations using a resource-based view has led to a substantial field of research that relates to many other approaches in the management literature (Barney, 1991; Mahoney & Pandian, 1992). The resource allocation process itself is one of the most important drivers of the strategic outcomes of an organisation (Bower, 1970; Bower & Gilbert, 2005). 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). In the resource allocation process, business-level managers 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.

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, that 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 (IRR), Return on Investment (ROI), Cost-Benefit (CB), and Pay-Back Period (PBP). In this thesis I will focus on NPV, since it is one of the most frequently used metrics for project evaluation (Graham & Harvey, 2001; Remer et al., 1993). NPV is the difference

1. Introduction

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 seen in Equation (1.1):

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

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 evaluating manager's decision.

However, there can be other influences on a 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 source of influence on managerial decision-making are the organisational structures and incentives that are in place both externally (Kokkinis, 2019), and internally to the organisation (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 financial 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 organisational influence on firm performance 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) suggested that it is not useful to refer to a “diversified” firm or the “extent of diversification.” Instead:

1. Introduction

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.

Therefore, the research on diversification is quite varied in the ways that diversification is operationalised. Some 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 & Jammie, 1988). Some of this discrepancy 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 the increase of divisions, rather than an understanding of the possible synergies will lead to the organisation actually being worth less than the sum of its parts (diversification discount). In fact, many organisations that acquired other businesses to diversify subsequently end up divesting them (Porter, 1987).

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. The influences discussed above often assume that the manager making the decisions acts rationally, as per neoclassical economic theory. However, research in psychology has shown that this is often not the case. In fact, the field of behavioural strategy now exists to study the effects

1. Introduction

of psychological biases on managerial strategy (Powell et al., 2011). I discuss this and the relevant implications for the thesis in the following section.

1.2 Psychology of resource allocation

Managers of large organisations are generally assumed to have a superior decision-making capability. However, managerial decision-making involves many of the same processes that have been shown to be affected by psychological biases (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 attempted 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.

Psychological research has shown that people tend to make decisions that are inconsistent with neoclassical economic theory. Such economic theory was based on Expected Utility Theory (EUT; Friedman & Savage, 1948; von Neumann et al., 1944), which made strong assumptions about the kind of information that people have 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 (Gigerenzer, 2008; Haselton et al., 2009). However, there are many situations in which such inconsistency with economic theory leads to poor outcomes.

Research has shown many ways in which the allocation of resources 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 funds, regardless of their composition. This bias was also found in the capital allocation of hierarchical firms (Bardolet et al., 2011). Managers allocate resources equally across the available divisions in the firm, regardless of performance. This behaviour may be damaging to firm performance because it

1. Introduction

means that lower performing business units may be getting subsidised by higher performing units, which are not operating at their full potential.

Relatedly, people tend to continue expending resources 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).

The way that information is presented can often 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 (E. J. Langer, 1975; Mannes & Moore, 2013; Puri & Robinson, 2007; Soll & Klayman, 2004), and experts often do the same (McKenzie et al., 2008), as do managers (Barone Adesi et al., 2013; Kahneman & Lovallo, 1993; Lovallo & Kahneman, 2003). This is important for managers that evaluate project proposals, both because the metrics that rely on estimates may be biased by the over-confidence of the manager that created the proposal, and because the evaluating manager may in turn be over-confident about the prospects of the proposal 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 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

1. Introduction

on past experiences that further the manager's existing agenda. 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 resource allocation. In this thesis I focus on three essential processes within the resource allocation process. When evaluating projects in hierarchical organisations, there is often 1. an element of risky choice, 2. a comparison between diversified businesses, and 3. a comparison between the target project and prior experience. Each of these is prone to separate biases, that are also interrelated. I review the literature for these processes in the subsequent subsections.

1.2.1 Risky choice

One of the most important finding of the risk choice literature is that losses have a stronger psychological impact than equivalent gains. This disparity is one of the most settled and consistent findings in psychology and economics. The fact that losses loom more than equivalent gains for the vast majority of people is referred to as loss aversion (Kahneman & Tversky, 1979). Loss aversion is a fundamental component of prospect theory, which is a theory of choice, and 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 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.

1. Introduction

Such findings are different to the prescriptions of rational choice theory. At its simplest, this theory suggests 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), expected value theory would suggest that option A should be chosen.

This basic principle was extended by Bernoulli (1738/1954), who suggested that a persons' subjective value of money differs depending on their current wealth. This concept of 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, \$1000, than for a person that initially only has \$10. This is called *diminishing returns* and is calculated by an exponential function (with a fractional exponent). Further, Prospect Theory (Kahneman & Tversky, 1979) suggests that people's subjective value of money does not depend on a person's state of wealth—it depends on a person's 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. Loss aversion showcases this idea. Namely, this is the way that people react to gains and losses relative to a reference point. Figure 1.1 shows loss aversion as the value function being steeper in the domain of losses than the domain of gains.

1. Introduction

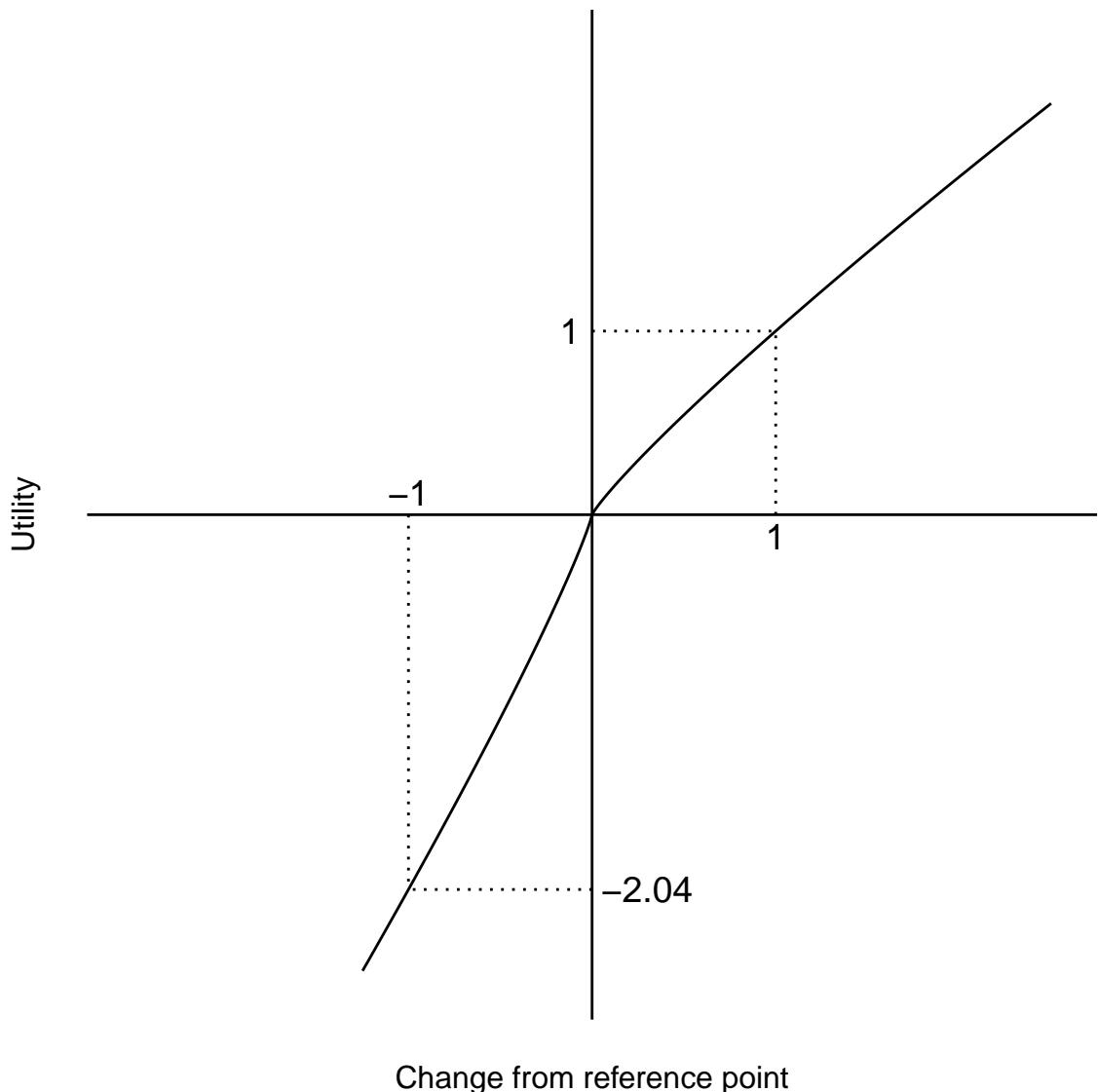


Figure 1.1: The value function in Prospect Theory (adapted from Kahneman & Tversky, 1979). The values for the exponent representing the extent diminishing returns (α) and the loss aversion coefficient (λ) were 0.88 and 2.25, respectively. These were the median values calculated by Tversky and Kahneman (1992) from experimental data.

In plain language, 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 of 2.25 of loss aversion. This means that people respond to losses 2.25 times more than equivalent gains. Figure 1.1

1. Introduction

shows this disparity. 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) brings about an increase of value (+y). However, a decrease in the same amount of wealth (-x) brings about a decrease in value 2.25 times the value of the equivalent gain (-2.25y). The straight lines adjacent to the curves showcase that the function is steeper in the loss domain than it is in the gain domain. This difference of slope is another indicator of loss aversion.

This research is relevant to resource allocation because managers often need to evaluate project proposals that 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' risk aversion to accepting projects that may jeopardise their job. However, the top 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). 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 inter-trial feedback. 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 of diversified firms. Organisations are not only varied by the number of divisions which they possess,

1. Introduction

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 such as business project similarity has not been investigated. This approach is important because the relative similarity of projects has implications about how difficult the evaluation process becomes and therefore what kind of financial metrics are used. 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 (SMT; Gentner, 1983; Gentner & Markman, 1997) provides a model of comparison that psychologically distinguishes similar and dissimilar allocation tasks. SMT models comparison as a process of bringing conceptual structures into alignment, which when possible, puts shared component 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 project success and ranking

1. Introduction

their priority more difficult. In Chapter 4, I experimentally examined business project comparisons and how project alignment affects resource 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 inherent such variance in their decisions? This is especially important to investigate given the above discussion. In cases of non-alignable comparison managers might be relying on a potentially unreliable metric. On the other hand, in an alignable comparison, managers might have the option to moderate their choice based 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).

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; Vivaldi & Coville, 2018). 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. In Chapter 4, I investigate whether people 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

1. Introduction

the target project to other cases. For instance, people tend to limit the sample size of the set of comparison cases to a target problem to a small number. Often only a handful of cases, or even one. Doing this might mean only considering surface similarity to the current situation and not an alignment of 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 might 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 MBAs 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?

1. Introduction

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? In Chapter 6, I investigate the extent to which people moderate their reliance on anecdotes or aggregated data by the relevance of the anecdote to the target project during resource allocation.

1.3 Chapter overview

In sum the potential consequences for a diversified hierarchical structure are that 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 moderate their decisions when it comes to more subtle statistical principles such as aggregation, variance, and sampling.

In the previous section I identified three resource allocation processes that are currently under-studied and so are important to investigate further. First, the evaluation of individual project proposals might 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

1. Introduction

multiple projects, the structural alignment literature suggests that managers in diversified firms will struggle to allocate more than those in more integrated firms. Further, these managers might not be sensitive to the variance inherent in the financial metrics they rely on. Third, a difficulty to compare across existing projects might instead mean a reliance on prior case studies from personal or external experience. Research on anecdotal bias suggests that managers might rely more on such case studies than on aggregated data, but it is unclear whether they will use relevance to moderate their decisions.

Therefore, I investigate the psychology of resource allocation decisions in three chapters. In Chapter 2 I investigate the effects of choice bracketing on risk aggregation without feedback. In Chapter 3, I discuss the difference between evaluating project proposals with inherent budget estimates and the process of allocating an existing budget top-down. In Chapter 4 I investigate the effects of alignment and reliability type—verbal or numerical—on allocations. In Chapter 5, I discuss the trade-offs that people make when using information to evaluate options such as project proposals. Finally, in Chapter 6 I investigate the effects of anecdote similarity on the anecdotal bias.

1. Introduction

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>
- 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>
- 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*. Retrieved February 24, 2021, from <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)
- 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.
- 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.n97>
- 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>
- Cavagnac, M. (2005). Strategic managerial incentives under adverse selection. *Managerial and Decision Economics*, 26(8), 499–512. <https://doi.org/10/cwdntp>

1. Introduction

- 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/c84dpx>
- 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/bx3wpk>
- 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>
- 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>
- 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>
- 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>
- 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>

1. Introduction

- 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., Frankenhuus, W. E., & Moore, T. (2009). Adaptive Rationality: An Evolutionary Perspective on Cognitive Bias. *Social Cognition*, 27(5), 733–763. <https://doi.org/10/c49fjq>
- 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 [magazine]. *McKinsey Quarterly*, (4), 15–17. Retrieved April 20, 2021, from <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>
- Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2), 311–328. <https://doi.org/10/bhghpg>
- 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>
- Mahoney, J. T., & Pandian, J. R. (1992). The resource-based view within the conversation of strategic management. *Strategic Management Journal*, 13(5), 363–380. <https://doi.org/10/cvv4cp>
- 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>

1. Introduction

- 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>
- 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. (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>
- Porter, M. E. (1998a). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press. (Original work published 1985)
- Porter, M. E. (1998b, June 1). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press. (Original work published 1980)
- 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>
- 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>
- 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>

1. Introduction

- 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/fdqjtjg>
- 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>
- Vivaldi, E., & Coville, A. (2018). *How Do Policymakers Update?*
http://barrett.dyson.cornell.edu/NEUDC/paper_573.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* (60th Anniversary Commemorative Edition). 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>

1. Introduction

- 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>

Fortune sides with him who dares.

—Virgil

2

Effect of choice bracketing on risk aggregation in repeated-play gambles with no feedback

Contents

| | | |
|------------|--|-----------|
| 2.1 | Introduction | 25 |
| 2.1.1 | Multi-play gambles | 26 |
| 2.1.2 | Repeated-play gambles | 28 |
| 2.1.3 | Choice bracketing | 30 |
| 2.1.4 | Internal market capital investment context | 33 |
| 2.2 | Experiment 1 | 35 |
| 2.2.1 | Method | 36 |
| 2.2.2 | Results | 39 |
| 2.2.3 | Discussion | 41 |
| 2.3 | Experiment 2 | 46 |
| 2.3.1 | Method | 47 |
| 2.3.2 | Results | 48 |
| 2.3.3 | Discussion | 50 |
| 2.4 | General discussion | 53 |
| 2.4.1 | Theoretical implications | 54 |
| 2.4.2 | Practical implications | 57 |
| 2.4.3 | Future research | 57 |
| | References | 59 |

2. Effect of choice bracketing on risk aggregation

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. As such, 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 I investigate in this chapter: sequences of large risky choices without immediate outcomes.

Risk aggregation is the combination of probability and/or variance information associated with certain outcomes for the purpose of understanding that information more comprehensively (Bjørnson & Aven, 2019). However, the psychological literature suggests that this process may be difficult for people. 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). As such, 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

2. Effect of choice bracketing on risk aggregation

individual gambles together facilitates risk aggregation. As such, the current work provides two primary contributions. First, I am the first to investigate the effect of choice bracketing on risk aggregation in independent gambles evaluated without immediate returns. Second, I introduce 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, and/or 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. As such, I 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, my 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 Expected Value (EV). That is, $200 \cdot 0.5 - 100 \cdot 0.5 = +50$.

2. Effect of choice bracketing on risk aggregation

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 (1963) argued that this choice is irrational, but other work suggests that it is consistent with expected utility theory when certain assumptions are clarified (e.g., Aloysius, 2007; Ross, 1999). However, a normative discussion is out of the scope of the present work. 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 identical gambles (a *multi-play* gamble).¹

Wedell and Böckenholt (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).

However, multi-play effects are likely robust, since there is also evidence that such gambles reduce a variety of cognitive biases, such as 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

¹I use the terminology from Bristow (2011) and Camilleri and Newell (2013) for consistency.

2. Effect of choice bracketing on risk aggregation

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).

Multi-play gambles that are displayed with an aggregated outcome distribution (that presents the probabilities of all the different possible outcomes) of those gambles are accepted more than multi-play gambles without these distributions (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) because they very clearly show the rarity of a loss. Note that this does not seem 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. These kinds of outcome distributions 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 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 not ever simply asked: “Here are 10 identical investments to consider; do you want all or none of them?”

2. Effect of choice bracketing on risk aggregation

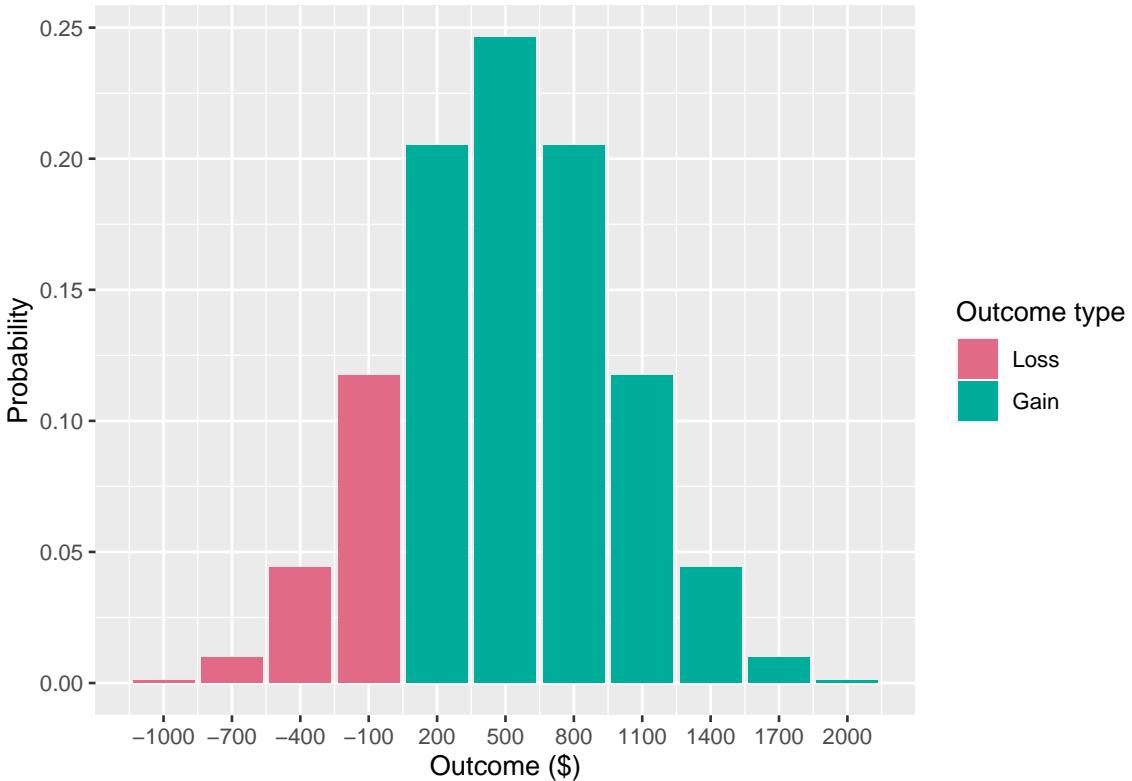


Figure 2.1: The outcome probability distribution of the Samuelson (1963) gamble (0.5, 200; 0.5, -100) played 10 times.

In *repeated-play* (rather than multi-play) gamble paradigms, participants make decisions about a series of different gambles. Research using this paradigm found that people are less risk averse when outcomes for a series of gambles are evaluated and 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 or are able to sample from a distribution 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

2. Effect of choice bracketing on risk aggregation

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, so therefore 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, thereby maximising EV, when they considered them jointly. 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.

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

2. Effect of choice bracketing on risk aggregation

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 essentially through 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.

The goal of this chapter is to facilitate risk aggregation without the experimental artefact of immediate feedback. Previous research suggests some ways of doing this. Sokol-Hessner et al. (2009) and Sokol-Hessner et al. (2012) found that lengthy instructions with an elaborate explanation to “think like a trader” – and to consider all the repeated-play gambles as a portfolio, as opposed to considering them individually – reduced risk aversion. Further, even 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. I am interested in testing whether that same intuition can be elicited and applied across sets of unique bets. What are the minimal conditions required to encourage aggregation? Both this explicit instruction effect and the multi-play gamble work suggest that participants can engage in a more intuitive form of aggregation when provided with the right contextual cues. It might be the case that people possess such an intuitive understanding, and that simply being encouraged to consider a set of choices together in more subtle ways than explicit instruction will also facilitate aggregation. This could involve simply making participants aware that they are going to be making a series of choices. This kind of contextual cue may be sufficient to encourage risk aggregation. 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.

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

2. Effect of choice bracketing on risk aggregation

multi-business firm makes decisions across multiple distinct industries. Of course, the similarity of decision contexts does not change the maths of risk aggregation, and 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. As such, people may behave similarly when considering a set of dissimilar choices, insofar as they consider them not 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 (A. B. Markman & Loewenstein, 2010; A. B. 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 that people will become 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, General Electric 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

2. Effect of choice bracketing on risk aggregation

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 those set of choices. As such, 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 market capital 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

2. Effect of choice bracketing on risk aggregation

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 principle, 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 for gains and 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 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) show 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. As such, 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. Only three managers said they would accept the investment (Thaler, 1999). 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.

In this chapter I investigate 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-utilised (Lovallo et al., 2020). The participants in my studies were taken from a general population that does not have extensive managerial experience. However, in the general population a lack of risk aggregation is most likely more common, and the variables I am researching here

2. Effect of choice bracketing on risk aggregation

are readily applicable to the financial decisions that lay people 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, because my task concerns managerial decision-making, I can investigate contextual cues to choice bracketing, while eliminating potential differences between participants in their prior experience with the specific decision-context. Future research will focus on managers with context-specific experience to investigate the effects of that experience.

2.2 Experiment 1

In Experiment 1, I investigated the effect of three choice bracketing manipulations on risky choice in hypothetical resource 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, meaning that participants had to make a choice about whether to accept a number of different hypothetical investments, but were not provided with feedback about their choices. I manipulated the similarity of the choices, whether they 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 accept.

Previous research suggests that people will exhibit more risky choice when explicitly told to bracket their choices (Sokol-Hessner et al., 2012; Sokol-Hessner et al., 2009), when choices are presented jointly (e.g., Hsee et al., 1999; Moher & Koehler, 2010), and when choices are similar (e.g., DeKay & Kim, 2005; A. B. Markman & Medin, 1995). Therefore, I tested the following hypotheses:

2. Effect of choice bracketing on risk aggregation

Table 2.1: Experiment 1 group allocation.

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

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

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

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

2.2.1 Method

Participants

One hundred and ninety-eight (82 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 32.52 ($SD = 11.42$, $min = 18$, $max = 69$). Participants reported an average of 7.01 ($SD = 9.1$, $min = 0$, $max = 42$) years of work in a business setting, and an average of 1.7 ($SD = 2.85$, $min = 0$, $max = 20$) years of business education. The mean completion time was 12.04 ($SD = 11.29$, $min = 3.1$, $max = 112.4$) minutes. Table 2.1 shows the between-subjects condition allocation.

Materials

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. Appendix A.1.1 shows a screenshot of these instructions.

2. Effect of choice bracketing on risk aggregation

Risky investment task In the 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 they forecast the project to make, and probabilities for these forecasts. I constructed these projects to appear attractive when aggregated, and unattractive when segregated (see T. Langer & Weber, 2001). Project 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 ($P(\text{loss})$). Further, there was a constraint on the gambles' loss aversion coefficient (λ), which is the ratio of potential gains over the potential losses. The constraints were:

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

As such, each project cannot be considered to be a loss under expected value theory, 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 $P(\text{loss})$ was chosen to make investment in the complete set of projects seem attractive. The actual probability of a loss given the outcome distribution I used was 0.09. This was calculated by summing all probabilities in the Poisson binomial distribution whose outcomes were less than zero. For comparison, $P(\text{loss}) = 0.17$ for 10 plays of the Samuelson (1963) gamble. 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

2. Effect of choice bracketing on risk aggregation

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. Border added for clarity.

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. Border added for clarity.

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 proceeded without this message. Note, 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 are succinct, and the decisions in the task are 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 in fact frame their options as likelihoods of certain monetary outcomes.

2. Effect of choice bracketing on risk aggregation

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. Appendix A.1.1 shows this graph. After collecting data I discovered that there was a coding error in the generation of gambles, which meant that I could not make use of the outcome distribution decision data. Therefore, the effect of outcome distribution will not be discussed until [Experiment 2](#), in which I fixed this issue. Appendix A.1.2 presents an analysis of this data, and describes the coding error and its implications.

Follow-up gambles Participants were shown four further sets of gambles (11 total) that functioned to check participant attention and replicate the gambles from Samuelson (1963) and Redelmeier and Tversky (1992). See Appendix A.1.1 for details.

Procedure

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

2.2.2 Results

Project choice

I conducted a three-way ANOVA 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 to 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 to 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 seems stronger in the high similarity condition,

2. Effect of choice bracketing on risk aggregation

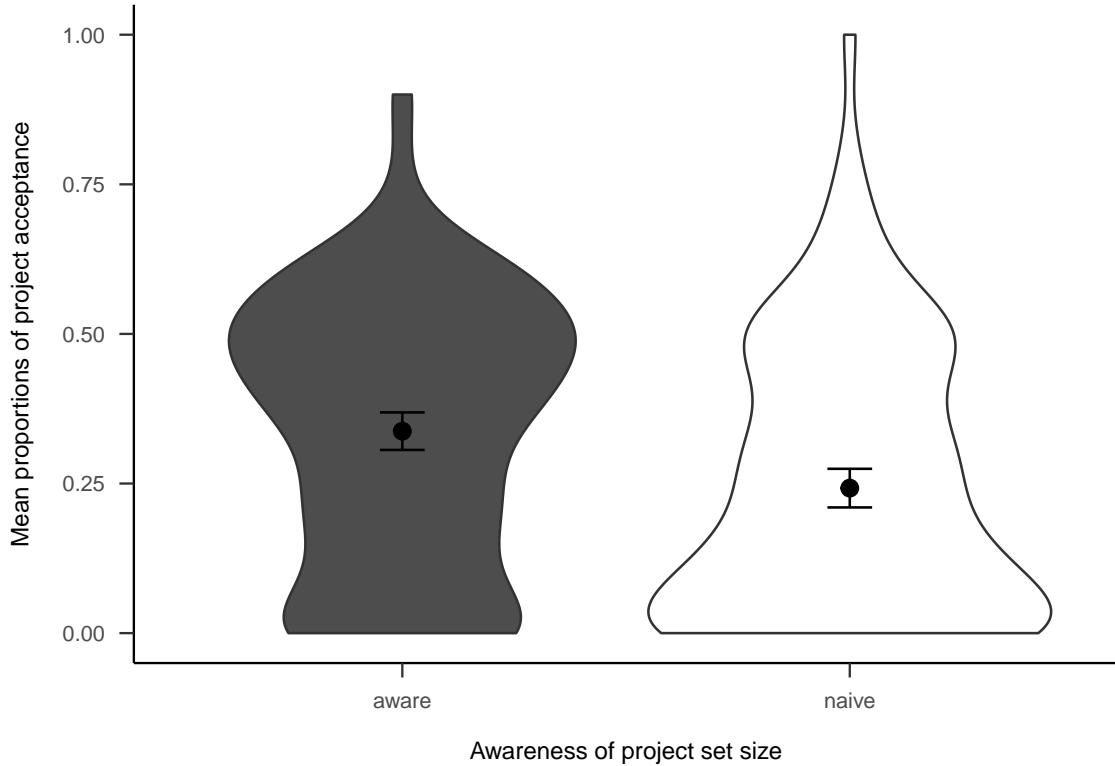


Figure 2.4: Mean proportions of decisions to invest in each set of 10 projects, by awareness conditions. Error bars represent 95% confidence intervals.

$\Delta M = 0.07$, 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], $t(194) = 2.06$, $p = .041$.

Trial-by-trial analysis

I conducted exploratory analyses into the possible effects of the manipulations on a trial-by trial basis. Appendix A.1.2 shows the data for all conditions. The key findings is 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 significant in the aware condition, $b = 0.11$, 95% CI [0.06, 0.16], $z = 4.54$, $p < .001$; but not in the naive condition, $b = 0.03$, 95% CI [-0.03, 0.08], $z = 1.01$, $p = .311$.

2. Effect of choice bracketing on risk aggregation

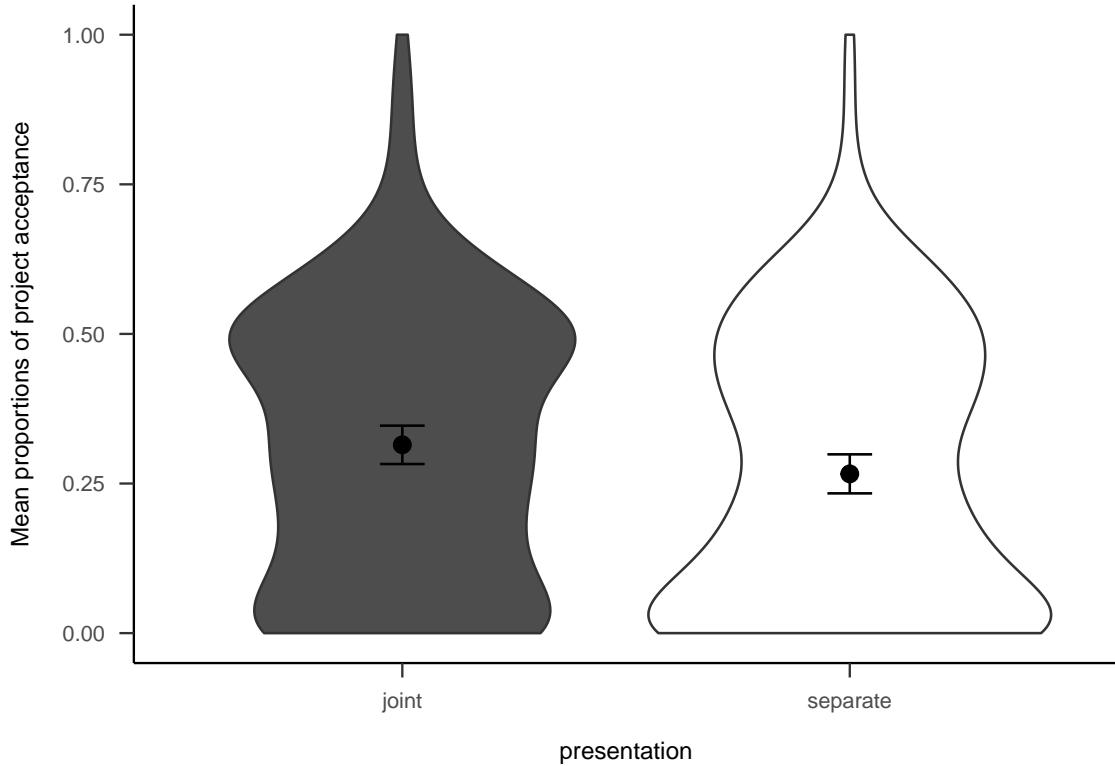


Figure 2.5: Mean proportions of decisions to invest in each set of 10 projects, by presentation conditions. Error bars represent 95% confidence intervals.

2.2.3 Discussion

I found evidence for most of the Experiment 1 hypotheses. Specifically, I found that people make more risky choices when considering those choices jointly on the same page, compared to on separate pages; and when they know how many choices were in the set. Further, I found an interaction between project similarity and presentation, and found tentative evidence that people are less risk averse when risky choices are aggregated for them. Exploratory analyses showed that participants' risk aversion seemed to decrease as they proceeded through the trials, but only when participants were aware of the number of projects.

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

2. Effect of choice bracketing on risk aggregation

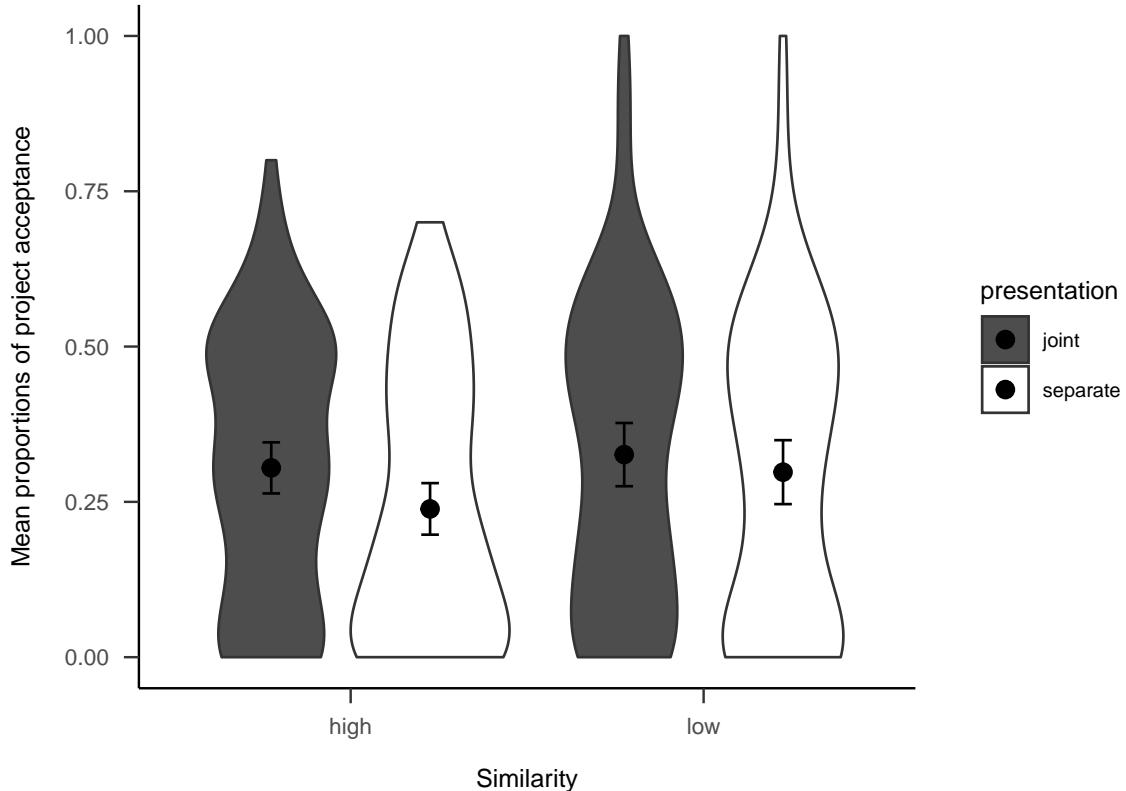


Figure 2.6: Mean proportions of decisions to invest in each set of 10 projects, by similarity and presentation conditions. Error bars represent 95% confidence intervals.

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. It may be the case that participants were encouraged by the joint display to consider the set of projects together, and perhaps conclude that ultimately investing in a higher number of gambles is potentially valuable due to the possibility that one project might pay off the projects that showed losses.

Awareness effect

I found an awareness effect such that participants that viewed the projects separately, were more likely to invest in the projects as the trials went on, regardless of the actual gambles. It might be the case that having an awareness of the total

2. Effect of choice bracketing on risk aggregation

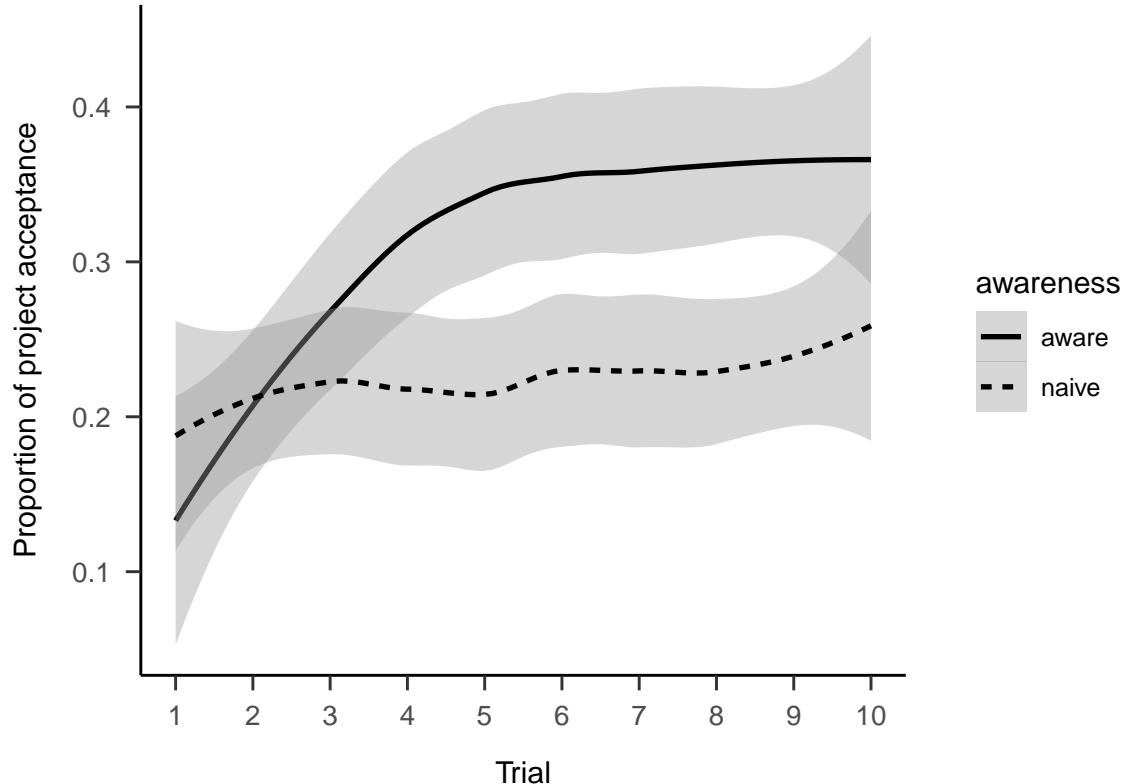


Figure 2.7: Proportion of project acceptance in the separate presentation condition, by trial, and awareness conditions. LOESS is used for smoothing over trials, and the shading represents 95% confidence intervals.

number of projects in the set would increase the likelihood that participants would naively aggregate. Specifically, knowing the number of total projects might make the idea that the gains of some projects will offset the losses of others will be more salient, because it reinforces a focus on the entire set.

Otherwise, this result may also be due to the Gambler's fallacy/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 HTTHHTTT might anticipate that the likelihood of "heads" is higher than that of a "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,

2. Effect of choice bracketing on risk aggregation

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 a Gambler's fallacy-type effect 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.

Similarity effect

I did not find a main effect of similarity in the individual choice data as predicted in Hypothesis 2.3. I found that choice similarity interacted with the presentation condition. This interaction is harder to explain since I did not originally hypothesise it. In fact, the results seem to suggest the opposite to what was originally expected. Initially, I 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 set. I thought that more similarity would act as a broad bracket, and therefore increase aggregation. That is, I would have 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 (but not statistically) higher in the low than in the high similarity condition when projects were presented separately, averaging over awareness conditions.

There was no significant difference between similarity conditions regardless of presentation condition. However, joint presentation condition allocated significantly more than those in the separate condition for both high and low similarity. The interaction seems to have been found due to this difference being larger in the high similarity condition. As such, the presentation effect can be interpreted as having benefited from a joint presentation more when projects were all from the same industry than when they were from different industries. Perhaps the ability to aggregate risk when projects are presented together is more made more salient when projects are similar.

2. Effect of choice bracketing on risk aggregation

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, it may be the case that participants were engaging 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 to the extent that people diversify, they seem to only diversify naively, meaning that they neglect co-variation when diversifying (e.g., Hedesstrom et al., 2006). That is, they could be said 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 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, it might be the case that those in the separate condition 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 would already known 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.

Limitations

This experiment had two major limitations. First, proper counterbalancing was not used in project domains, 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](#), 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. Experiment 2 replicated the effects from this experiment in the high similarity condition, so

2. Effect of choice bracketing on risk aggregation

addressed the second issue. Experiment 2 addressed the counterbalancing issue by adding more project domains.

2.3 Experiment 2

Experiment 2 investigated the effect of presentation, awareness, and distribution on project choice. For the distribution manipulation, I presented half of the sample 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 repeated-play choice literature, each choice was presented without feedback. Further, in contrast to Experiment 1, I displayed this distribution alongside each gamble, as opposed to only at the very end. This is an important manipulation because finding out whether it is effective will 1. add to the understanding of the conditions necessary for mathematical aggregation (beyond a mere intuitive sense of aggregation), and 2. 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. To my knowledge, 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, I addressed one of the limitations of Experiment 1 detailed above by manipulating all the main variables between-subjects. By manipulating presentation between-subjects, I take out the potentially confounding factor of reduced risk aversion over time.

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

Hypothesis 2.4. 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.

2. Effect of choice bracketing on risk aggregation

Table 2.2: Experiment 2 group allocation.

| Awareness | Distribution | Presentation | N |
|-----------|--------------|--------------|-----|
| Aware | Absent | Separate | 41 |
| Naive | Absent | Joint | 41 |
| Naive | Absent | Separate | 41 |
| Naive | Present | Separate | 42 |
| Total | - | - | 165 |

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, I tested the following hypothesis:

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

2.3.1 Method

Participants

One hundred and sixty-five (52 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 26.33 ($SD = 8.64$, $min = 16$, $max = 72$). Participants reported an average of 2.54 ($SD = 5.33$, $min = 0$, $max = 43$) years of work in a business setting, and an average of 1.67 ($SD = 2.93$, $min = 0$, $max = 20$) years of business education. The mean completion time was 6.51 ($SD = 5.13$, $min = 1.18$, $max = 39.93$) minutes. Table 2.2 shows the between-subjects condition allocation.

Materials

Instructions Participants were shown the same instructions as in [Experiment 1](#).

Risky investment task Participants saw a similar display to the one in [Experiment 1](#), but with new gamble values, in order to fix the mistake in the Experiment 1 gamble value calculation (detailed [above](#)).

2. Effect of choice bracketing on risk aggregation

The presentation and awareness manipulations were as in Experiment 1. However, for the distribution variable, participants either saw the project descriptions as is, or saw an outcome probability distribution of all the projects alongside the description (see Figure 2.8).

Follow-up I asked participants how many projects they think 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. See Appendix A.2.1 for screenshots.

Procedure

Participants responded to demographic questions, 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.

2.3.2 Results

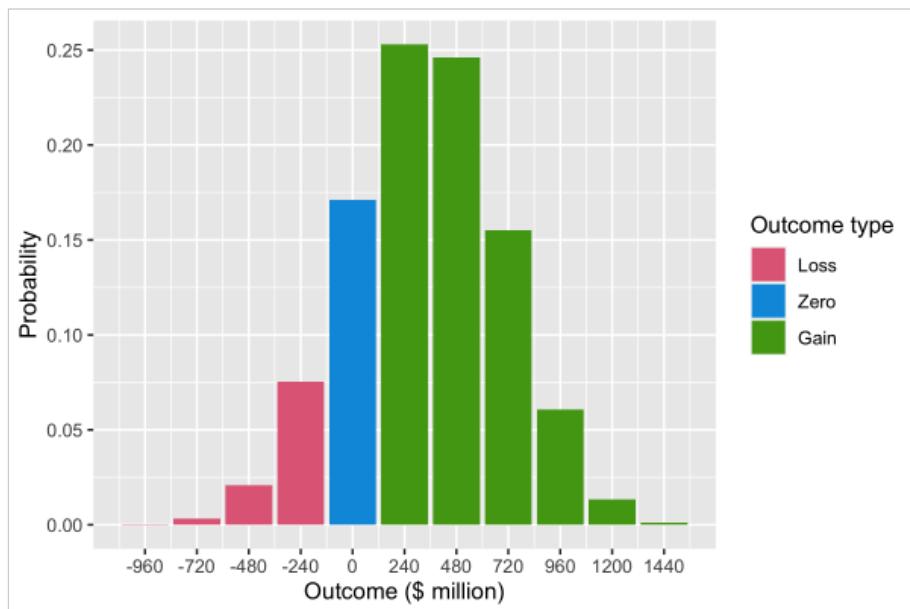
Project investment

The project investment data was analysed in two ways: as binary choice per trial (using logistic regression), and as proportions of choice per participant (using t-test). In each case I compared the relevant comparison condition to the same control condition (separate naive distribution absent). Figures 2.9 and 2.10 show the choice and proportion data, respectively. The difference between presentation conditions was not significant, both in the logistic regression $\hat{\beta} = -0.02$, 95% CI [-0.57, 0.53], $z = -0.08$, $p = .937$, and in the t-test, $d_s = 0.00$, 95% CI [-0.43, 0.43], $t(80) = 0.00$, $p = 1.000$. Similarly, the difference between awareness conditions was not significant, both in the logistic regression $\hat{\beta} = -0.25$, 95% CI [-0.86, 0.36], $z = -0.79$, $p = .428$, and in the t-test, $d_s = 0.17$, 95% CI [-0.26, 0.61], $t(80) = 0.78$, $p = .438$. However, those that saw a distribution tended to choose to invest significantly more (51.19%) than those that did not see a distribution (39.02%), seen both in the logistic regression, $OR = 1.88$, 95% CI [1.06, 3.35], $z = 2.15$, $p = .032$, and the t-test $d_s = 0.47$, 95% CI [0.03, 0.90], $t(81) = 2.11$, $p = .038$.

2. Effect of choice bracketing on risk aggregation

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.

2. Effect of choice bracketing on risk aggregation

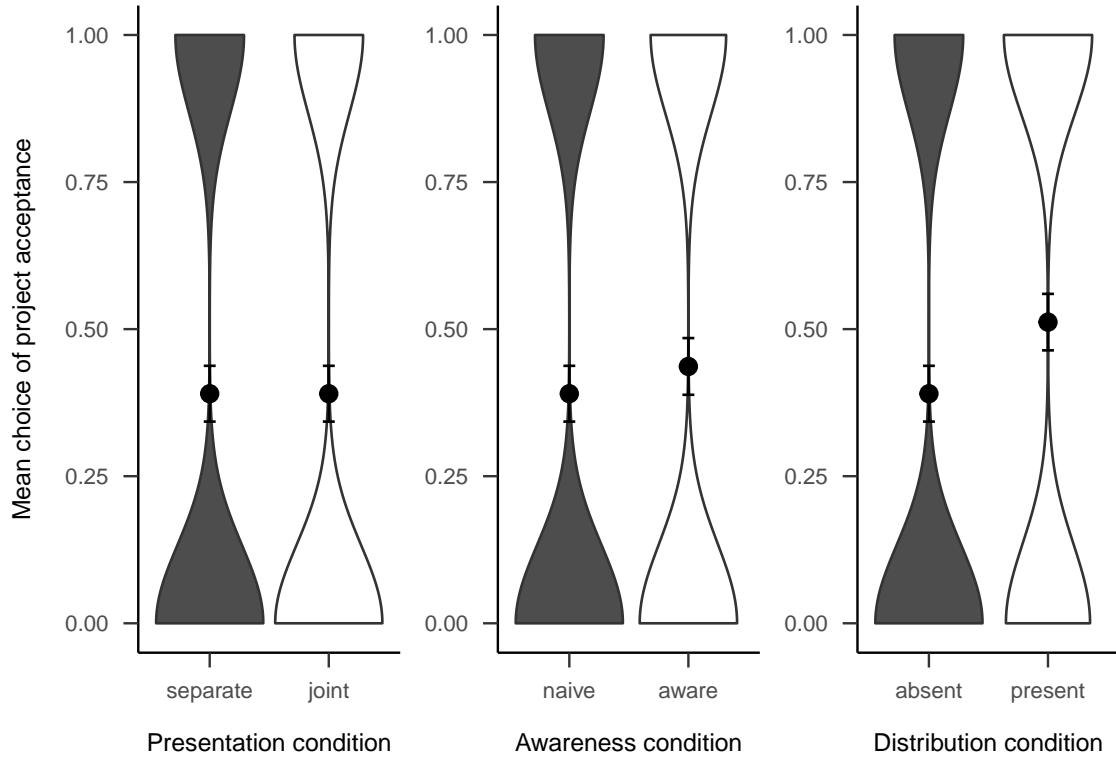


Figure 2.9: Mean project acceptance for the presentation, awareness, and distribution effects. Note, the condition on the left of each effect is the reference condition (separate presentation, naive awareness, distribution absent). As such, it is identical for the three effects.

Further, as Figure 2.11 shows, it doesn't seem as if the previous awareness by trial effect was replicated.

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

I found support for one of the Experiment 2 hypotheses. Seeing an outcome distribution of the projects 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

2. Effect of choice bracketing on risk aggregation

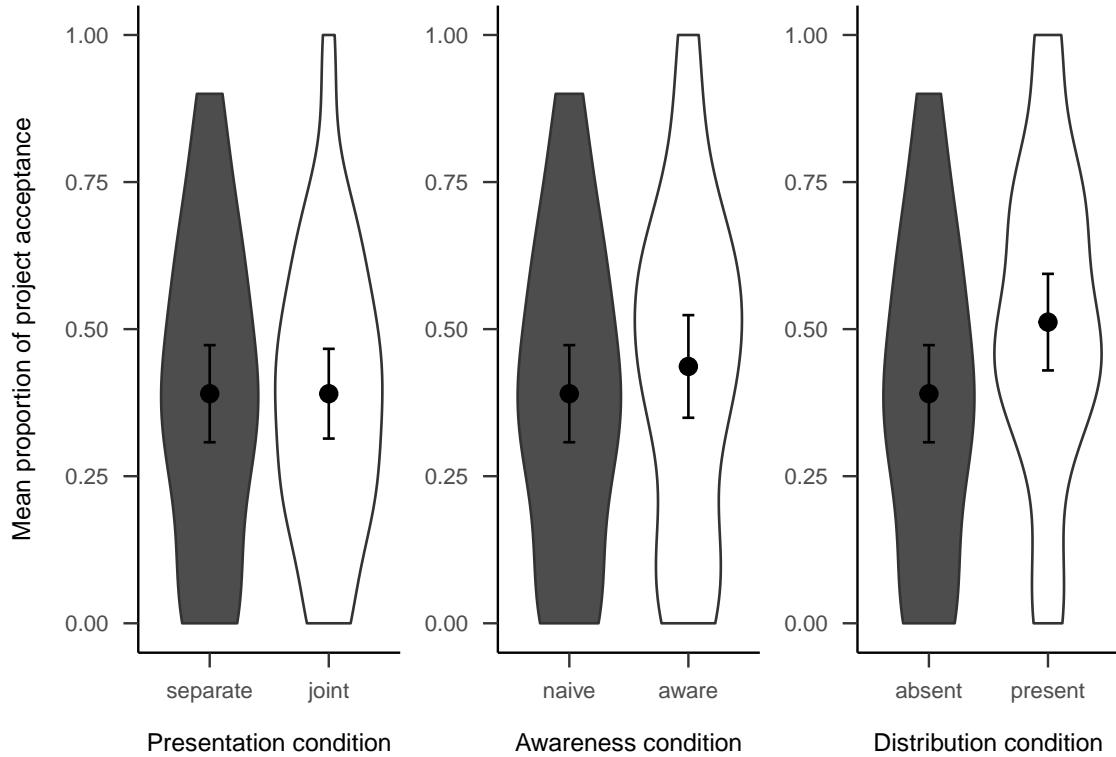


Figure 2.10: Mean proportion of project acceptance for the presentation, awareness, and distribution effects. Note, the condition on the left of each effect is the reference condition (separate presentation, naive awareness, distribution absent). As such, it is identical for the three effects.

and presentation effects I found in [Experiment 1](#) did not replicate.

The finding that seeing an outcome distribution of a set of gambles reduces risk aversion for that set of gambles provides evidence for choice bracketing. That is, people do seem to be primarily considering gambles one at a time. Specifically in this case, this finding suggests that that the main bottleneck for appropriately aggregating a set of gambles is a mathematical/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

2. Effect of choice bracketing on risk aggregation

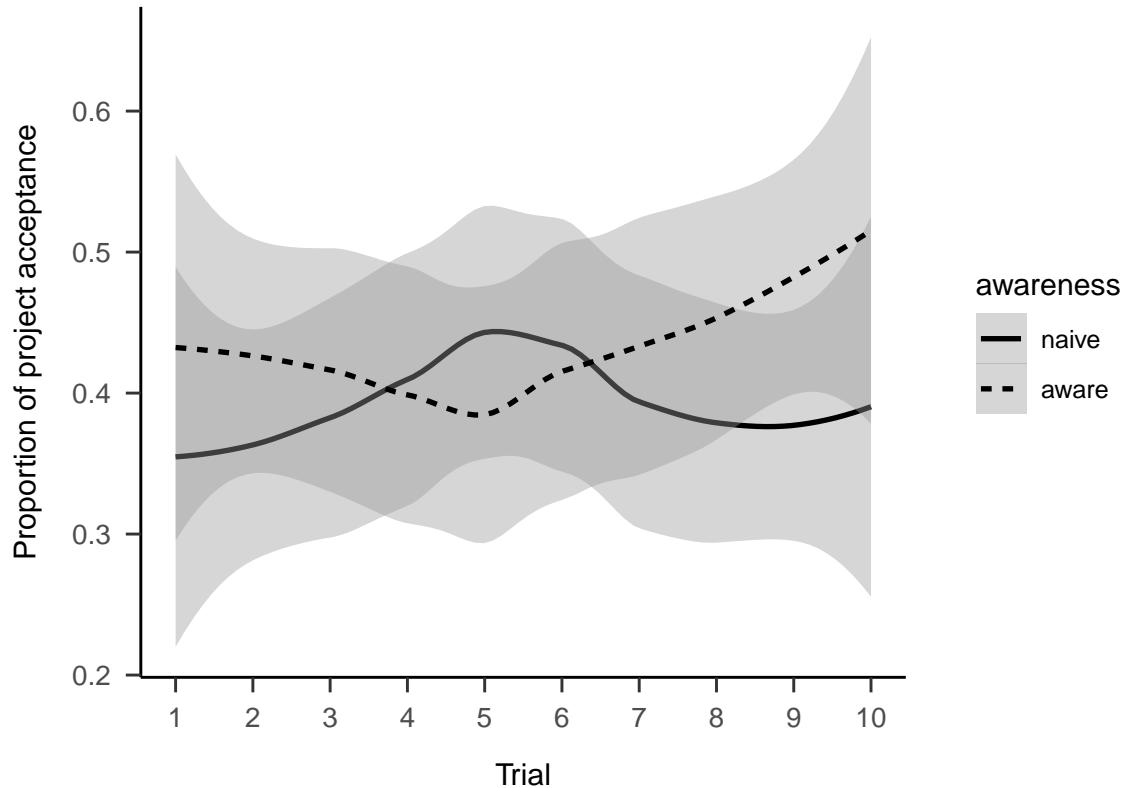


Figure 2.11: Mean project acceptance for separate presentation, distribution absent condition, by awareness and trial.

be grouped/bracketed together, then the effects from Experiment 1 should have replicated. However, 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. I did this in order to attempt to replicate the relevant effects from Experiment 1. However, it may be the case that there was 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. Effect of choice bracketing on risk aggregation

2.4 General discussion

I found that 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, my paradigm allowed for the investigation of choice bracketing in a way that is 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 (presented in Appendices [A.3](#) and [A.4](#)) again tested the similarity and awareness effects. I found evidence for naive diversification (an advantage for low similarity) when considering them 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, I found that choice bracketing facilitates risk aggregation. As per Hypothesis [2.5](#), I found that showing a distribution of outcome probabilities without inter-trial feedback reduced risk aversion. Further, I found 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. I found only minimal evidence for Hypotheses [2.1](#) and [2.2](#), suggesting that viewing projects together and awareness of the number of projects are not sufficient to encourage aggregation. Altogether, it seems that subtle contextual

2. Effect of choice bracketing on risk aggregation

cues are often not sufficient to encourage risk aggregation, and that people need risk to be aggregated for them 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 in multi-play paradigms and used immediate feedback for repeated-play paradigms (e.g., Benartzi & Thaler, 1999). As I 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. I am the first to provide evidence for this aggregation effect with non-identical gambles without feedback.

My other choice bracketing findings 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. I found that 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 seemed 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 specific elements of the experiment.

Research on the effect of option similarity on choice (e.g., A. B. 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

2. Effect of choice bracketing on risk aggregation

options as fungible (e.g., DeKay & Kim, 2005). As such, I predicted that a set of investments that involve the same type of investment would be seen as more similar and fungible. As per Hypothesis 2.3, I expected that this would facilitate a broad bracketing, and therefore more risk aggregation.

Instead, I found 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 my 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, in a way, aggregation of a portfolio already assumes that the portfolio is somewhat diversified (at least that the gambles aren’t perfectly correlated).

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

2. Effect of choice bracketing on risk aggregation

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. I proposed two explanations: 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).

2. Effect of choice bracketing on risk aggregation

2.4.2 Practical implications

This research implies some prescriptions for resource 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.

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 resource 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 in order 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.

2. Effect of choice bracketing on risk aggregation

Further, the more subtle choice bracketing manipulations should be tested with managers since it is possible that they have a greater sense of naive aggregation and therefore 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. 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 in order to see whether the effects of the manipulations replicate in a more realistic environment.

2. Effect of choice bracketing on risk aggregation

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>
- 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.
- 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.
- 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.
- 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>

2. Effect of choice bracketing on risk aggregation

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

2. Effect of choice bracketing on risk aggregation

- 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.
- 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.
- 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.
- 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.

2. Effect of choice bracketing on risk aggregation

- 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. (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>
- 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.
- Wedell, D. H., & Böckenholdt, 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., & Böckenholdt, 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 facilitated to do so intuitively. People instead tend to consider projects one at a time and only become aware of the benefits of aggregation when given an explicit visualisation of what it entails. In real-life resource allocation scenarios, when managers evaluate projects sequentially, presenting an aggregated distribution is not possible since the value of future projects is not yet known. Managers in this situation would need to aggregate intuitively by considering that for positive EV projects, some gains will win out overall. However, this is insufficient if their supervising manager is unaware of this principle. 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. Therefore, changing the organisational policy to encourage considering projects jointly is important so that the risk can be concurrently aggregated. Considering projects jointly is useful both because it allows for an actual aggregation of the risks, and also because it frames the projects as portfolios, meaning 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.

3. Joint evaluation of multiple projects

Business projects might not always be either accepted or rejected. Instead, an organisation might 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, if NPV does 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 resources 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 resources 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.

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 across litres of hydrocarbons 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 insofar 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 tested people’s sensitivity to forecast estimate variance information. That is, will people use NPV more

3. Joint evaluation of multiple projects

when the variance information suggests that it is a reliable measure, than when the information suggests that it is unreliable?

In the previous chapter I 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 considered the situation in which people are already presented with choices together and asked to evaluate the projects by allocating a hypothetical budget.

Further, I wanted to focus on identifying 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. Therefore, risk aversion was accounted for by making it clear that no losses are possible. This was achieved by using only positive NPVs, which implies that the project is not forecasted to lose money.

I also manipulated how easy the project attributes are to compare in order to identify the ways that decision-making in a diversified organisation might be different to that of a more integrated organisation. In the previous chapter I 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. I assumed that participants understood that the dissimilar project set implied that the projects were not alignable. That is, even though the underlying specific project features were not shown, I assumed that participants would realise that two projects from the same industry would be easier to compare using their intrinsic underlying features. However, this may not have been the case and may explain the equivocal similarity effect. Therefore, Chapter 4 will show the project features explicitly and manipulate alignment, so that the difficulty of the comparison is more obvious.

4

Project similarity bias and variance neglect in forecast metric evaluation

Contents

| | | |
|-------------------|---------------------------|-----------|
| 4.1 | Introduction | 66 |
| 4.1.1 | Experiment summary | 69 |
| 4.2 | Experiment 1 | 70 |
| 4.2.1 | Method | 71 |
| 4.2.2 | Results | 75 |
| 4.2.3 | Discussion | 77 |
| 4.3 | Experiment 2 | 77 |
| 4.3.1 | Method | 78 |
| 4.3.2 | Results | 81 |
| 4.3.3 | Discussion | 82 |
| 4.4 | Experiment 3 | 83 |
| 4.4.1 | Method | 84 |
| 4.4.2 | Results | 87 |
| 4.4.3 | Discussion | 89 |
| 4.5 | General discussion | 90 |
| References | | 93 |

4.1 Introduction

One of the most important tasks that executives face is allocating resources within their company. This requires ranking different projects based on their importance

4. Project similarity bias and variance neglect in forecast metric evaluation

and predicted success, and allocating limited resources respectively (not unlike a scientific funding agency). Ranking projects requires comparing them across a number of component dimensions. For example, an executive in an oil company might have multiple proposals on his or her desk for where to explore for oil next, and how to do so. Figuring out what makes one oil discovery project better than another one is relatively easy. However, consider a different scenario in which the executive has to allocate resources between an oil discovery project and an oil refinery project. The dimensions of an oil refinery project that distinguishes good from bad projects may be totally different from the dimensions of oil discovery projects that do the same. Think of a funding agency giving out fellowships, and deciding between two cognitive scientists, or a cognitive scientist and a physicist. What makes a physics proposal better for the field of physics than a cognitive science proposal is for cognitive science?

Structure-mapping theory (SMT; Gentner, 1983; Gentner & Markman, 1997) provides a model of comparison that psychologically distinguishes these two kinds of allocation tasks. SMT models comparison as a process of bringing conceptual structures into alignment, which when possible, puts shared component 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

4. Project similarity bias and variance neglect in forecast metric evaluation

to make meaningful comparisons, and so would make predicting project success and ranking their priority more difficult. In this chapter, I experimentally examined project comparisons, and how that affects resource allocation decisions. My working hypothesis was that comparisons with more alignable differences will make project predictions more precise, and project rankings easier and more informative, than a comparison with non-alignable differences.

However, what happens when the two domains are too disparate for a decision-maker to align them, but the task demands that they be aligned? This is actually a bit of uncharted territory experimentally. The prediction is that when forced to, people will grab at any piece of information that they can and then try to infer and abstract as much as seems reasonable to ease the alignment. This is in fact what occurs very frequently in business settings. Corporate enterprises continue to embrace diversification strategies in their investments, and so constantly have to make resource allocation choices that involve very disparate domains. To overcome these difficult comparisons executives rely on various financial measures that in theory can apply to any project or business proposal. These financial measures work well to ease the burden of the difficult comparison by abstracting away from the complexities of the individual projects, and just focus on financial information such as total costs, projected profits, etc. Initially hard-to-compare projects can therefore be more easily evaluated by comparing values on individual numerical measures.

The most common financial measure that is used by executives in order to value projects is Net Present Value (NPV; Graham & Harvey, 2001). 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 seen in Equation (1.1).

NPV is commonly used to decide about resource allocation and investment. The simple rule is that if a project's NPV is positive, then it is financially viable, and if NPV is negative, then it is not. However, the use of NPV has been criticised, both by academics and practitioners (Fox, 2008; Willigers et al., 2017). The main criticism is that there is a lot of underlying variance within the NPV measure that is not made apparent by the final form of the measure: a single numerical value.

4. Project similarity bias and variance neglect in forecast metric evaluation

For instance, NPV is dependent on the cash inflows that are projected for each year of the project. However, financial forecasting is known to often be 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. The duration of the project and the discount rate are further sources of variance that are hidden by the single numerical value of NPV.

As such, the secondary goal of this research was to investigate the extent to which people are sensitive to variance information (from financial forecasting) when making resource allocation decisions. This consideration is especially important in the resource allocation situations illustrated above, in which executives need to compare between projects from disparate domains and therefore have to rely on NPV. This matters because the NPV calculated from different domains may have different underlying forecasting error, which may compromise the utility of using NPV as the basis of the comparison. Do executives sufficiently account for the inherent variance in the measure that they rely on so much? Research showed that people are good at extracting variance information from numerical sequences (Rosenbaum et al., 2020). However, people struggle to use variance information when it is represented numerically (Batteux et al., 2020; Galesic & Garcia-Retamero, 2010; Konold et al., 1993; Vivaldi & Coville, 2018).

4.1.1 Experiment summary

In [Experiment 1](#) I investigated the effect of alignment on the decision-making of naive participants' resource allocation to a set of fictional projects. I manipulated NPV reliability by directly stating whether it is a reliable measure because I did not assume the naive participants would have the requisite knowledge to be sensitive to reliability information otherwise. I predicted that when projects are alignable, participants use NPV if they are told that it is a reliable measure, but do not use it if they are told that it is not reliable. However, when projects are not comparable, I predicted that participants will use NPV regardless of how reliable they are told NPV is.

4. Project similarity bias and variance neglect in forecast metric evaluation

In [Experiment 2](#), I investigated the decision-making of management students in almost the same situation as Experiment 1. The main difference was that instead of telling participants the reliability of NPV, I manipulated the level of *numerical* reliability that is associated with it. That is, the width of numerical ranges around an NPV. I predicted that participants will rely more on NPV in the non-alignable projects than in the alignable projects. However, I predicted there will be no effect of numerical reliability, since there is very little evidence that people are sensitive to variance information when represented numerically.

In [Experiment 3](#) I again tested the alignment and reliability effects in a non-business population, but manipulated both verbal and numerical reliability in the same experiment to allow for direct comparisons. I predicted an effect of reliability in the verbal reliability condition, but not in the numerical reliability condition. Further, I 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 alignment and explicit NPV reliability information on financial resource allocation decisions. The structural alignment literature suggests that people weigh alignable differences more heavily than non-alignable differences. I expected that people would use NPV more than any other product attributes (because it applied to every product). However, I expected this effect to vary with how reliable participants thought the value was. That is, if other project dimensions were alignable, then it should be easier to adjust the use of NPV based on its reliability. However, I predicted that in a low alignment condition there will be a reliance on NPV, as the sole alignable difference, regardless of its stated reliability. I used a no NPV condition in order to get a better understanding of the baseline responding to the materials without NPV. To measure these effects I considered the linear trend of how NPV across projects relates to the money allocated to the projects. Critically, participants saw both NPV and

4. Project similarity bias and variance neglect in forecast metric evaluation

other features intrinsic to each project domain, and these were inversely related. Therefore, a positive NPV amount trend indicates a reliance on NPV, whereas a negative trend indicates a reliance on the intrinsic project features. Specifically, I tested the following hypotheses:

Hypothesis 4.1. The alignment \times reliability amount \times NPV amount interaction will be significant.

Hypothesis 4.2. The linear NPV amount trend will be higher, on average, in the low alignment condition compared to the high alignment condition.

Hypothesis 4.3. The effects of alignment, reliability amount, and NPV amount on allocations will interact, such that the NPV amount \times reliability interaction will be stronger in the high alignment than in the low alignment condition.

Hypothesis 4.4. In the high alignment condition, the linear NPV amount trend will be higher in the high reliability condition than in the low reliability condition

Hypothesis 4.5. In the low alignment condition, the linear NPV amount trend will not differ significantly between reliability conditions

Hypothesis 4.6. In the high alignment condition, the linear NPV amount trend will only be higher than the no NPV reliability condition in the high reliability condition.

Hypothesis 4.7. In the low alignment condition, the linear NPV amount trend will be higher than the no NPV reliability condition in both the low and high reliability conditions.

4.2.1 Method

Participants

One hundred and eighteen (55 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 29.42 ($SD = 9.25$, $min = 18$, $max = 73$). Table 4.1 shows the between-subjects condition allocation. NPV amount was varied within subjects.

4. Project similarity bias and variance neglect in forecast metric evaluation

Table 4.1: Experiment 1 group allocation.

| Alignment | Reliability amount | N |
|-----------|--------------------|-----|
| High | High | 26 |
| High | Low | 17 |
| High | No npv | 17 |
| Low | High | 21 |
| Low | Low | 16 |
| Low | No npv | 21 |
| Total | - | 118 |

Materials

Instructions The instructions page explained to the participants, who did not necessarily have business experience, about the task and NPV. However, I also used this page to manipulate whether they were told that NPV was reliable or unreliable as a financial measure. Participants in the low NPV reliability condition were told that NPV is an unreliable metric, whereas those in the high NPV reliability condition were told that NPV is a reliable metric. Those in the no NPV condition saw instructions that did not include the NPV explanation. Critically, participants were told to invest in products with a high objective value (because a better quality product is not always better in a consumer goods market). Since this might still not use this instruction when directly viewing the projects, Experiment 3 used projects with attributes that more inherently conveyed quality. Appendix B.1.1 shows screenshots of these web-pages.

Project display Participants read sets of fictional business projects to potentially allocate resources to. The high alignment display was a table that listed various attributes for five projects (see Figure 4.2). The low alignment display had each project expressed as individual paragraphs that described the relevant attributes through full sentences (see Figure 4.1). For the high alignment display, the product was the same, and the concrete attributes were consistent, while in the low alignment display, each project was a different product with concrete attributes relevant to that project. For both alignment conditions each project description included an NPV.

4. Project similarity bias and variance neglect in forecast metric evaluation

This alignment manipulation was reinforced by visual presentation. The table format is better afforded by the high alignment condition because all the dimensions are shared. However, this confounds alignment with presentation style. I address this in [Experiment 3](#) by equating the table format across both alignment conditions.

Critically, the value of the concrete attributes were always in conflict with the NPV. For instance, Project 4 always had the lowest value for each of the concrete attributes for laptops, but always had the highest NPV. This was done in order to be able to use participants' allocations as a proxy measure for an individual's degree of dependence on NPV.

Allocation Participants completed a resource allocation task (see Figure 4.3), adapted from Bardolet et al. (2011) in which they were asked to allocate their hypothetical yearly budget across the given five projects.

Additional measures As well as measuring allocation, I included other measures, whose presentation and analyses are reported in the Appendix. Specifically, participants were asked to forecast the future returns of the projects (see Appendix B.1.1), rank the projects (see Appendix B.1.1), indicate their confidence in their decisions (see Appendix B.1.1), and justify their decisions (see Appendix B.1.1).

Procedure

Following ethics and demographics web-pages, participants were introduced to the study and the concept of NPV through the instructions page. They then completed the forecasting task directly after each project display in the low alignment condition, and directly under all projects for the high alignment condition. Participants then ranked the projects, and subsequently answered the allocation, confidence, and justification questions.

4. Project similarity bias and variance neglect in forecast metric evaluation

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 4.1: An example of a low alignment display in Experiment 1. Border added for clarity.

4. Project similarity bias and variance neglect in forecast metric evaluation

| | 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 4.2: An example of a high alignment display in Experiment 1.

How will you allocate your budget across these five projects (percentages must sum to 100%)?

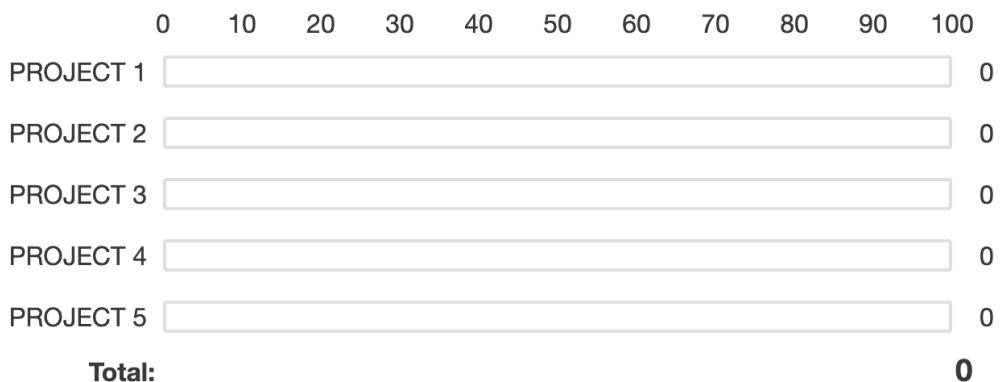


Figure 4.3: The allocation task.

4.2.2 Results

A mixed factorial ANOVA was conducted to investigate the effects of alignment and verbally-instructed NPV reliability on participants' allocations. As seen in Figure 4.4, the alignment \times reliability amount \times NPV amount interaction was significant, $F(6.57, 367.76) = 2.18, p = .039, \eta_p^2 = .038$. 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 the low alignment condition, the linear NPV trend was significantly lower in the no NPV condition than both the low

4. Project similarity bias and variance neglect in forecast metric evaluation

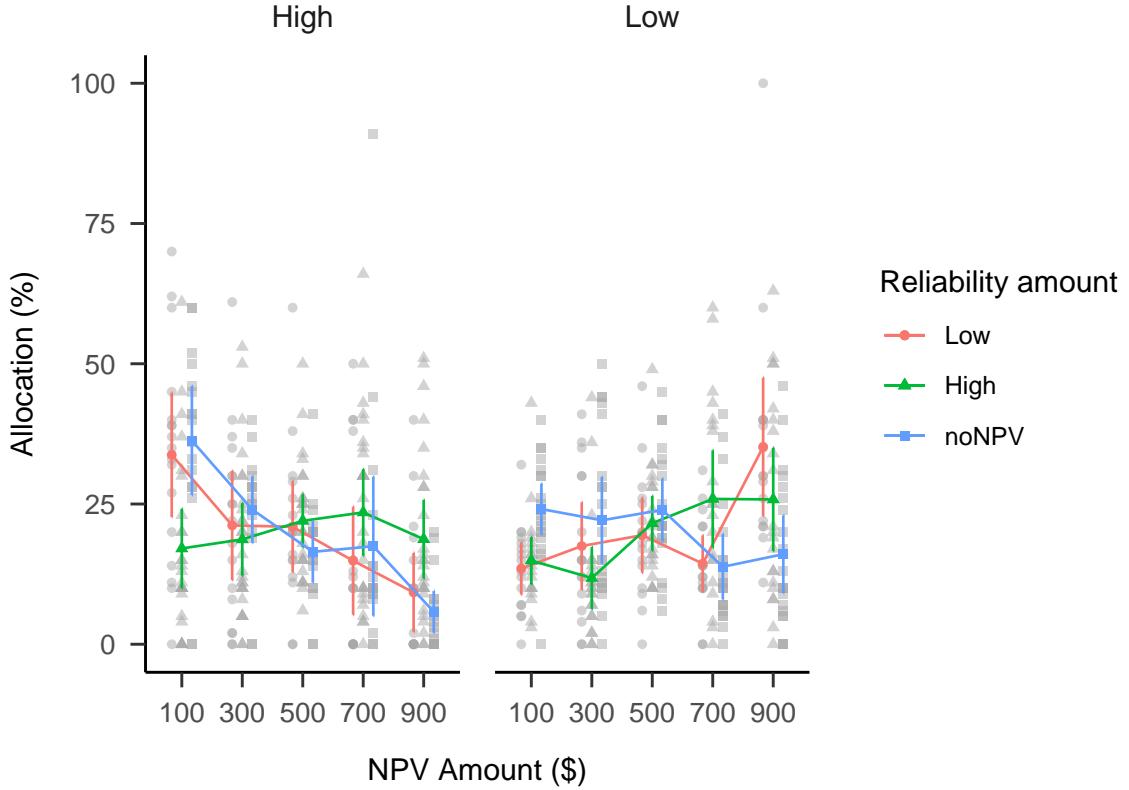


Figure 4.4: Mean allocation.

reliability condition, $M = 64.63$, 95% CI [25.76, 103.50], $t(112) = 3.29$, $p = .001$, and the high reliability condition, $M = 60.29$, 95% CI [24.14, 96.43], $t(112) = 3.30$, $p = .001$. However, in the high alignment condition, the linear NPV trend was only significantly lower in the no NPV condition than the high reliability condition, $M = 75.70$, 95% CI [39.17, 112.24], $t(112) = 4.11$, $p < .001$, and not the low reliability condition, $M = 12.24$, 95% CI [-27.94, 52.41], $t(112) = 0.60$, $p = .547$.

The ranking, confidence, and forecast mean data were all largely congruent with the allocation findings (see Appendix B.1.2). I also found that the forecasts of those in the low alignment condition had higher standard deviations than those in the high alignment condition (see Appendix B.1.2). However, this did not replicate in subsequent experiments (see Appendix B.5.2 and B.6.2).

4. Project similarity bias and variance neglect in forecast metric evaluation

4.2.3 Discussion

Experiment 1 found evidence for the effect of alignment on laypeople's decision-making in resource allocation scenarios. Specifically, I found that when projects are comparable, people use NPV when they are told it is reliable, but not when they are told it is unreliable. However, they use NPV regardless of reliability when it is the only shared dimension across products.

In Experiment 1 I manipulated *verbal* NPV reliability. That is, participants were explicitly told whether NPV is 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 and estimate. In Experiment 2 I will attempt to replicate the alignment effects, but while manipulating the numerical reliability associated with each project, rather than the verbal reliability as I have previously used. Further, it might be the case that only those with sufficient experience with financial theory and analysis will be able to successfully draw inferences from such information. Therefore, in Experiment 2 I used a sample of Masters of Management students, instead of the laypeople in Experiment 1.

4.3 Experiment 2

Experiment 2 further investigated the effects of alignment and numerically-expressed variance information on financial resource allocation decisions. In [Experiment 1](#) the information about the variance inherent in the NPVs was communicated explicitly as the conclusion, e.g., NPV is unreliable. In Experiment 2 however, only the actual variance information itself was communicated without the conclusion about its reliability. Here, I simply communicated variance by showing the range of predicted values (akin to a confidence interval). Therefore, while previously I manipulated *verbal* reliability, in Experiment 2 I manipulated *numerical* reliability. Further, in this experiment, I studied participants with more business experience than the previously used laypeople samples. The main two interests in this experiment were first, whether the previous findings of an alignment effect will replicate with

4. Project similarity bias and variance neglect in forecast metric evaluation

people with more business experience; and second, whether this population is sensitive to variance in forecasts. As well as again testing Hypothesis 4.2, I tested the following hypothesis:

Hypothesis 4.8. The NPV amount \times reliability amount interaction will not be significant in either alignment condition.

I was also interested in the potential to quickly change participants' understanding, if they did not initially use numerical reliability for their allocations. Therefore, I presented participants with a short lecture on the importance of paying attention to variance in financial decision-making. However, results were inconclusive, so see but see Appendix B.2, for a more detailed discussion.

4.3.1 Method

Participants

Fifty-four (28 female) people were recruited from a Masters of Management course at an Australian university. Age information was not recorded. Both the reliability amount conditions (low and high) and alignment conditions (low and high) were presented within subjects and their order of presentation was counterbalanced.

Materials

Instructions Participants were shown similar instructions to Experiment 1. However, here they were given more information about NPV (about the discount rate and initial investment). See Appendix B.2.1 for screenshots of the instructions.

NPV test Participants were presented with a short and simple test of their understanding of NPV (see Appendix B.2.1).

Project display As seen in Figures 4.5 and 4.6, the project display was as in Experiment 1, except for the addition of another set of projects that differed in semantic content. That is, the projects were about different types of projects (to allow for a within-subjects manipulation). Along with the single NPV, the

4. Project similarity bias and variance neglect in forecast metric evaluation

forecasted ranges of cash flow that the NPV was calculated from were presented. In the low numerical reliability condition, the ranges were $\pm 85\%$ around the mean (e.g., \$150-\$1850, where the forecast mean is \$1000); whereas in the high numerical reliability condition, the ranges were $\pm 5\%$ around the mean (e.g., \$950-\$1050, where the forecast mean is \$1000). Wide ranges should indicate low reliability in the measure, while narrow ranges should indicate high reliability. Between the four displays, participants were told to treat each new set of projects as independent to the other sets.

NPV knowledge ratings I was interested if the participants would be overconfident in their knowledge of NPV, so asked them to rate this at multiple points in the experiment (see Appendix [B.2.1](#)).

Variance lecture Participants were shown a short lecture on the importance of paying attention to variance information, in an attempt to facilitate a subsequent increased use of the numerical reliability information in their allocations (see Appendix [B.2.1](#) for more details and the lecture slides).

Procedure

Participants saw the instructions, NPV explanation, and completed a simple test to demonstrate their understanding of NPV. They then completed four counterbalanced resource allocation trials (equivalent to each condition combination), and saw a brief presentation about the importance of paying attention to variance in financial decision-making. Subsequently, participants completed a further two resource allocation trials, of two of the trials that they subsequently saw earlier. They were shown the allocations they provided earlier, and given the opportunity to change them. With the knowledge ratings, participants rated once before the NPV test, re-rated after the test, and after the four project displays. They were then asked to rate their knowledge of NPV as they believe it had been both before and after the variance presentation.

4. Project similarity bias and variance neglect in forecast metric evaluation

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. Border added for clarity.

4. Project similarity bias and variance neglect in forecast metric evaluation

| | 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.

4.3.2 Results

A mixed factorial ANOVA was conducted to investigate the effects of NPV amount, alignment, and numerical NPV reliability on participants' project allocations. Figure 4.7 shows these data. The alignment \times reliability amount \times NPV amount 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 amount trend and reliability amount, $\Delta M = -42.28, 95\% \text{ CI } [-76.96, -7.59], t(53) = -3.14, p = .011$, even after applying a Šidák correction. The same interaction but with the linear NPV trend was not significant, $\Delta M = -6.13, 95\% \text{ CI } [-31.50, 19.25], t(53) = -0.62, p = .954$. Further, the linear NPV trend did not differ between reliability amount condition in neither the low alignment condition, $\Delta M = -3.19, 95\% \text{ CI } [-18.77, 12.40], t(53) = -0.41, p = .683$, nor the high alignment condition, $\Delta M = 2.94, 95\% \text{ CI } [-12.63, 18.52], t(53) = 0.38, p = .706$. However, averaging over reliability amount, the linear NPV trend was higher in the low alignment condition than in the high alignment condition, $\Delta M = 28.19, 95\% \text{ CI } [5.57, 50.81], t(53) = 2.50, p = .016$.

4. Project similarity bias and variance neglect in forecast metric evaluation

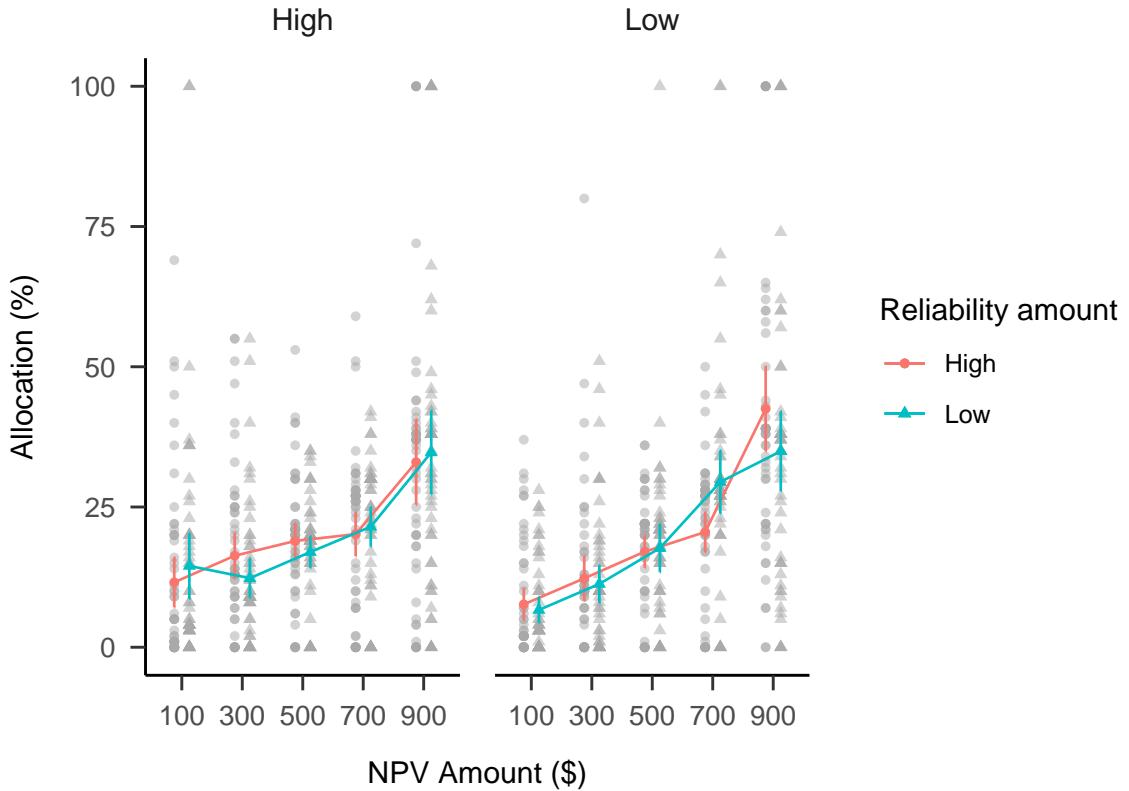


Figure 4.7: Mean allocation.

The ranking data were congruent with these results, while the confidence data was less so. Further, the NPV knowledge data did not replicate the effect from Long et al. (2018, Study 1). These analyses are reported in Appendix B.2.2.

4.3.3 Discussion

Experiment 2 replicated the alignment effect found in Experiment 1 with participants with real-world business experience. I found evidence for Hypothesis 4.2, as people seemed to rely more on NPV when faced with a set of dissimilar project than when projects were all from the same domain. Further, I found evidence for Hypothesis 4.8, with no significant differences between the numerical reliability amount conditions. While I did not replicate the interaction found in Experiment 1, it should be emphasised that these are two different effects. In Experiment 1, I explicitly told participants about the reliability of the NPV measure, whereas in Experiment 2, I provided them with variance information that merely implies NPV reliability. Thus, I showed that business students are affected by comparability

4. Project similarity bias and variance neglect in forecast metric evaluation

of project sets, but not by numerical NPV reliability information. Specifically, participants appeared to only focus on the NPV itself, and not on any variations in the underlying noisiness of the measure for a specific project.

However, it is unclear whether laypeople would also display this variance neglect. Further, in Experiments 1 and 2, 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, with various features displayed. Consumer projects were chosen to be more easily accessible to participants without business experience. However, with these projects, individual features do not necessarily indicate a project's profitability. For instance, a laptop with a low capacity can be more profitable than a laptop with a high capacity due to the existence of consumer goods markets. In Experiment 3, I addressed these limitations.

Another limitation of the last two experiments was the potential presentation style confound. 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 in paragraphs, whereas the information in the high alignment condition was presented in tables. While it is likely that each of these data types would be presented in this way in a business setting, it is important to rule out that this manipulation did not also 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 alignment, reliability type, NPV amount, and reliability amount on allocations. Experiment 1 manipulated reliability amount by using *verbal* prompts. That is, participants were told explicitly whether or not NPV was reliable for a certain project industry. Experiment 2 investigated whether people were able to extract this same kind of reliability information from *numerical* prompts. That is, participants saw NPVs with either wide or narrow

4. Project similarity bias and variance neglect in forecast metric evaluation

ranges, indicating either low or high reliability of the metric, respectively. However, Experiment 1 sampled laypeople, whereas Experiment 2 sampled a Masters of Management course. Therefore, I was not able to compare the two reliability types (verbal and numerical) without ruling out the potential confound of population type. As such, in Experiment 3 I manipulated NPV amount, alignment, and reliability amount, but also added a reliability type factor. Further, presentation style was a possible confound in the alignment manipulation of the previous experiments. That is, the business projects in the high alignment condition were always displayed as a part of a table, whereas the projects in the low alignment condition were displayed in prose, as paragraphs. Experiment 3 fixed this by using the same presentation style across alignment condition.

As well as again testing Hypotheses 4.2, 4.3, 4.4, and 4.5, I tested the following hypothesis (see Figure B.19 for a plot of a simulation of all hypothesised effects):

Hypothesis 4.9. The alignment \times reliability amount \times reliability type \times NPV amount interaction will be significant, such that the effects hypothesised above will be seen in the verbal reliability condition, but none of these effects will be seen in the numerical reliability condition.

4.4.1 Method

Participants

Four hundred and forty-eight (176 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 41.65 ($SD = 10.3$, $min = 29$, $max = 78$). Participants reported an average of 6.94 ($SD = 8.23$, $min = 0$, $max = 43$) years of work in a business setting, and an average of 3.73 ($SD = 6.27$, $min = 0$, $max = 45$) years of business education. The mean completion time was 11.35 ($SD = 10.79$, $min = 1.92$, $max = 183.7$) minutes. Table 4.2 shows the between-subjects condition allocation. The two reliability amount conditions (low and high) were presented within subjects and the order of their presentation was randomised. As before, NPV amount was

4. Project similarity bias and variance neglect in forecast metric evaluation

Table 4.2: Experiment 3 group allocation.

| Alignment | Reliability type | N |
|-----------|------------------|-----|
| High | Explicit | 112 |
| High | Implicit | 112 |
| Low | Explicit | 112 |
| Low | Implicit | 112 |
| Total | - | 448 |

varied within subjects. Therefore, each participant saw two separate project displays. Appendix B.3.1 describes the power analysis conducted to arrive at this sample size.

Materials

Instructions Participants saw similar instructions to the previous experiments, with an added explanation of the NPV reliability information for each reliability type condition (see Appendix B.3.1). Further, they completed a test of basic NPV understanding. This test was added also as an attention check as, although it was required to answer, the response should only be one of two letters.

Project display The project displays were similar to the previous experiments. However, here participants saw the same presentation style in both alignment conditions. Each display had a table describing the projects in the set, with ranking and allocation inputs. The project details were presented as dot 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: 1. the association of reliability amount and project set (two variations), 2. the association of business name with NPV (five latin square variations), and 3. project variation (five variations per alignment condition), which for high alignment meant the project type. For low alignment this meant the intrinsic feature variant for the relevant project type. Table column order and project display order were both randomised.

4. Project similarity bias and variance neglect in forecast metric evaluation

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: <input type="text"/> | Ranking: <input type="text"/> | Ranking: <input type="text"/> | Ranking: <input type="text"/> | Ranking: <input type="text"/> |
| Project allocation (%) | Allocation: <input type="text"/> | Allocation: <input type="text"/> | Allocation: <input type="text"/> | Allocation: <input type="text"/> | Allocation: <input type="text"/> |
| Business name | Pressbloom | Cweb | Pharmacore | Erectic | Railmont |
| Project type | <u>national newspaper</u> | <u>software</u> | <u>pharmaceutical</u> | <u>high-rise construction</u> | <u>railway</u> |
| Predicted project features | <ul style="list-style-type: none"> Newspapers printed: 50,000 a day Number of weekly advertisers: 80 Ink that is not discarded due to impurities: 5,000L a day | <ul style="list-style-type: none"> Code written: 1,000 lines a day Security rating: 60% Number of potential customers in first year: 3 million | <ul style="list-style-type: none"> Pills pressed: 300,000 an hour Shelf life: 20 months Probability of symptom reduction after a week: 90% | <ul style="list-style-type: none"> 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 | <ul style="list-style-type: none"> 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.

4. Project similarity bias and variance neglect in forecast metric evaluation

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: <input type="text"/> | Ranking: <input type="text"/> | Ranking: <input type="text"/> | Ranking: <input type="text"/> | Ranking: <input type="text"/> |
| Project allocation (%) | Allocation: <input type="text"/> | Allocation: <input type="text"/> | Allocation: <input type="text"/> | Allocation: <input type="text"/> | Allocation: <input type="text"/> |
| Business name | Liquid Pipeline | Enfuel | Petroyield | Refinera | Oilpier |
| Project type | <u>oil well</u> | <u>oil well</u> | <u>oil well</u> | <u>oil well</u> | <u>oil well</u> |
| Predicted project features | <ul style="list-style-type: none"> • Oil extracted: 3,400L an hour • Time the machinery lasts before requiring maintenance: 11 years • Probability of finding oil: 96% | <ul style="list-style-type: none"> • Oil extracted: 2,000L an hour • Time the machinery lasts before requiring maintenance: 7 years • Probability of finding oil: 90% | <ul style="list-style-type: none"> • Oil extracted: 3,870L an hour • Time the machinery lasts before requiring maintenance: 13 years • Probability of finding oil: 99% | <ul style="list-style-type: none"> • Oil extracted: 2,470L an hour • Time the machinery lasts before requiring maintenance: 8 years • Probability of finding oil: 92% | <ul style="list-style-type: none"> • 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.

Interstitial Before each project display, participants saw an “interstitial” page, whose role was to 1. introduce the next display, and 2. check the participant’s attention (not required to answer, so can be skipped if the interstitial text isn’t read). See Appendix B.3.1 for an example.

4.4.2 Results

A mixed factorial ANOVA was conducted to investigate the effects of NPV amount, alignment, NPV reliability amount, and NPV reliability type on participants’ project allocations. Figure 4.10 shows these data. Only the main results are reported here,

4. Project similarity bias and variance neglect in forecast metric evaluation

while the rest of the hypothesised allocation effects are reported in Appendix B.3.2.

The four-way interaction (alignment \times reliability amount \times NPV amount \times reliability type) was not significant, $F(3.20, 1, 420.19) = 0.71, p = .555, \hat{\eta}_p^2 = .002$. Further, the three-way interaction (alignment \times reliability amount \times NPV amount) in the verbal reliability condition was not significant, $\Delta M = 13.42, 95\% \text{ CI } [-1.27, 28.11], t(444) = 1.80, p = .073$. This is most likely due to the size of the two-way interactions in each alignment condition being relatively similar, despite the expectation of no interaction in the low alignment condition (as was seen in [Experiment 1](#)). In high alignment, this interaction (between the linear NPV amount trend and NPV reliability amount) was significant, $\Delta M = -36.63, 95\% \text{ CI } [-47.02, -26.25], t(444) = -6.93, p < .001$. Specifically, the trend was stronger in the high reliability amount condition, $\Delta M = 27.26, 95\% \text{ CI } [17.69, 36.83], t(444) = 5.60, p < .001$, than in the low reliability amount condition, $\Delta M = -9.38, 95\% \text{ CI } [-18.86, 0.11], t(444) = -1.94, p = .053$. The same interaction was significant in the low alignment condition, $\Delta M = -23.21, 95\% \text{ CI } [-33.60, -12.83], t(444) = -4.39, p < .001$. Despite the lack of a three-way interaction in the verbal reliability condition (indicating a difference in the two alignment two-way interactions), the linear NPV amount trend was stronger in the low alignment condition than in the high alignment condition when averaging over reliability amount, $\Delta M = 28.97, 95\% \text{ CI } [17.68, 40.26], t(444) = 5.04, p < .001$.

In the numerical reliability condition, the linear NPV amount trend was not significantly “equivalent” between those in the low and high alignment conditions (averaging over reliability amount), $\Delta M = 15.19, 95\% \text{ CI } [3.90, 26.48], t(444) = 2.64, p = .996$. In fact, a post-hoc comparison suggests that the low alignment trend was stronger (with Bonferroni adjustment), $\Delta M = 15.19, 95\% \text{ CI } [0.78, 29.60], t(444) = 2.64, p = .034$. Further, the difference between the linear NPV trend was not significantly different between reliability amount conditions in both the low alignment condition, $\Delta M = 1.64, 95\% \text{ CI } [-11.61, 14.90], t(444) = 0.31, p > .999$, and high alignment condition, $\Delta M = -1.21, 95\% \text{ CI } [-14.46, 12.05]$,

4. Project similarity bias and variance neglect in forecast metric evaluation

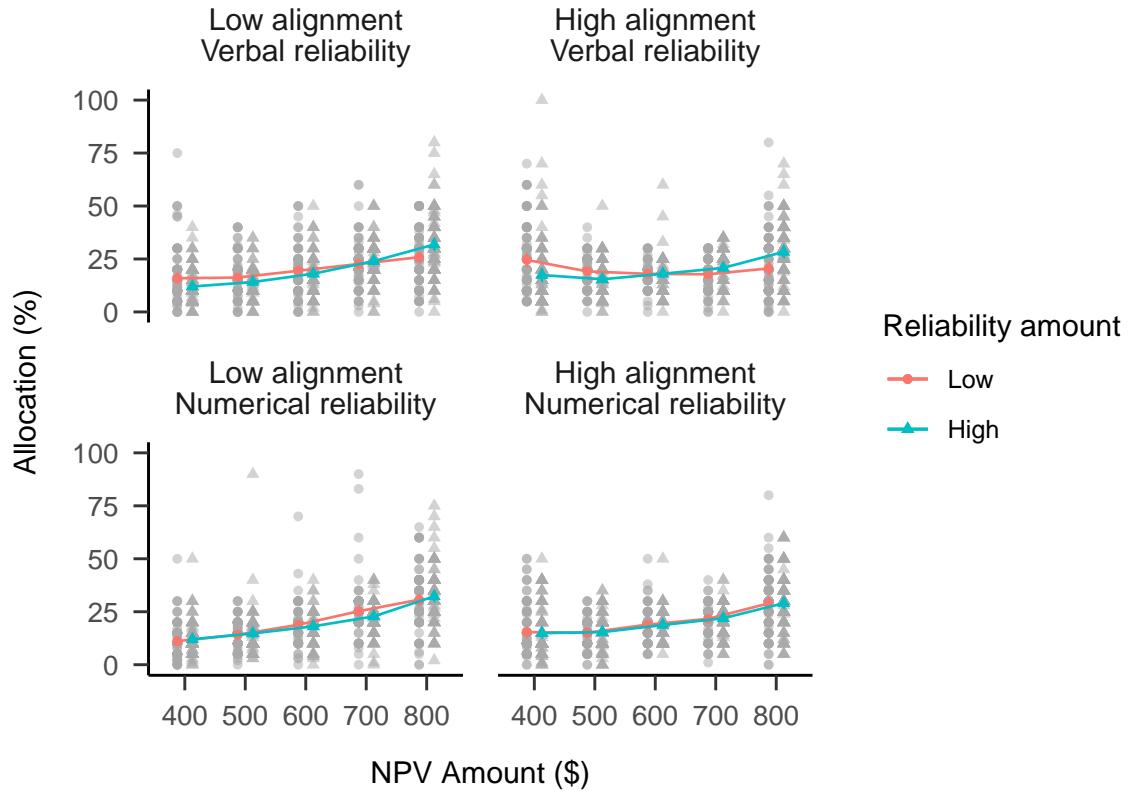


Figure 4.10: Mean project allocation for all conditions.

$t(444) = -0.23, p > .999$. This replicates the finding from [Experiment 2](#) that NPV is used more in low alignment, but numerical reliability is ignored.

4.4.3 Discussion

Hypotheses [4.2](#), [4.3](#), and [4.4](#) were supported. This shows that while overall participants prefer to use NPV as a proxy for project quality in their allocations, they still use verbal reliability information. Specifically, when projects are similar, people use NPV when they are told that it is reliable, and use alternative metrics when told that it is not reliable. However, Experiment 3 did not find evidence for Hypothesis [4.5](#). Instead, I found that even in the low alignment condition, participants were still able to use NPV more when told that it was reliable.

Further, I did not find support for Hypothesis [4.9](#). The hypothesis was constructed in response to the results of a pilot study (documented in [Appendix B.8](#)) that replicated the results of Experiment 1 in the verbal reliability condition, but

4. Project similarity bias and variance neglect in forecast metric evaluation

did not replicate the results of Experiment 2 in the numerical reliability condition. That is, I found that when faced with numerical ranges as variance information, people did not seem to even use the midpoint in their decisions. In Experiment 3, on the other hand, I replicated the finding of Experiment 2 in the numerical reliability condition. Specifically, I found that people used NPV more when projects were dissimilar, but critically, that they did not use numerical range information to moderate their allocations.

4.5 General discussion

Across three experiments there were two main findings: 1. NPV is used more when options are hard to compare; and 2. people do not consider numerical variance information, despite this being important to the reliability of the NPV forecasts. I found these effects consistently across both naive and experienced populations, which indicates their persistence. When finding themselves in a situation that requires comparison across disparate options, people make use of metrics with alignable differences. However, they do not sufficiently moderate their use of such metrics even when they have alternative attributes to use.

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

Alignable differences have been shown to be important to decision making in many settings (A. B. Markman & Loewenstein, 2010; A. B. Markman & Medin, 1995). The experiments in this chapter are novel in the investigation of alignment effects in a resource allocation paradigm. Further, I considered the extent to which

4. Project similarity bias and variance neglect in forecast metric evaluation

the reliability of an alignable measure (NPV) affects its use in choice. I found that this is dependent on the availability of other alignable differences in the choice set. If other alignable differences are available, then participants are willing to reduce their use of a supposedly unreliable alignable measure (and use it when told that it is reliable). However, when no other alignable differences are available, then the unreliable, but alignable, measure is used less. I found this result in Experiment 1 and 3, as well as a pilot study (although the alignment effects appeared weaker perhaps due to interference due to a within-subjects manipulation; see Appendix B.4).

Financial measure such as NPV are useful because of their alignability. That is, they act as an alignable difference regardless of the inherent similarity of 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 theoretical account of structural alignment does not directly speak to the real-world implications when there is a need for non-alignable comparisons. NPV is a type of abstraction that allows comparison between different aspects of a company. For instance, comparing an oil field project with a refinery project might be made easier by using NPV. On the other hand, this increased alignment might actually hide important information because it does not consider the finer complexities inherent within each business unit. The forecasts that are the basis of NPV are based on different indicators for each unit, there are differences in variance between the estimates from each unit, etc. As such, one can imagine a continuum of similarity comparisons in which usefulness of comparison increases with the level of alignability, but is moderated by the level of abstraction that is required to make the alignment.

The finding that people, even with some business experience, do not sufficiently consider variance information is surprising, but understandable. It is surprising because so much of financial decision making depends on considering different sources of variance, e.g., risk, volatility, and uncertainty. However, it is understandable because research from psychology and statistics education shows that statistics

4. Project similarity bias and variance neglect in forecast metric evaluation

students and people in general have a poor ability to make statistical inferences (Galesic & Garcia-Retamero, 2010; Konold et al., 1993). Future research should investigate the conditions under which people's sensitivity to variance information can be facilitated. For instance, it is unclear whether it is merely salience that is lacking, and that therefore visual aids could be useful, or whether further explicit explanation of the statistical inference is necessary. Pilot experiments suggest that participants still struggle to use numerical reliability information, even when given very explicit instructions (see Appendix B.7).

A possible limitation of these experiments was the use of NPV as the only financial metric. In the business world there are many metrics that serve similar functions and I predict would be used as a tool to deal with non-alignable options, as NPV was in the current study. Therefore, future research should attempt to replicate the current findings with different financial measures.

Future research should also investigate the boundary conditions of the reliability effect. That is, people seem to be responding to explicit reliability information, but not to variance information that implies reliability. As such, it would be interesting to find out what is the minimal kind of information about variance that participants need in order to understand the relevant implications about reliability. It may be the case that participants simply do not notice the variability information, or that they find it irrelevant. For instance, future research could test participants in a condition in which the variability information is more salient.

4. Project similarity bias and variance neglect in forecast metric evaluation

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>
- 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). A brief critical history of NPV. *BAA Conference*, 16.
- 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>
- 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>
- 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>
- Vivaldi, E., & Coville, A. (2018). *How Do Policymakers Update?*
http://barrett.dyson.cornell.edu/NEUDC/paper_573.pdf

4. Project similarity bias and variance neglect in forecast metric evaluation

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

Looking for alignment in past cases

Chapter 4 found that people do not sufficiently weigh the importance of numerical variance information in resource allocation. This is important for when projects are dissimilar because the results showed that people rely more on NPV in low alignment. Not paying attention to the underlying variance obscures uncertainty inherent in the measure. When projects are similar this means that managers may miss out on an opportunity to use different, potentially more reliable measures. Such measures are more likely to be available in alignable comparisons. This again many mean another unnecessary use of an unreliable measure.

An evaluation of a non-alignable set of projects can therefore lead to many potential pitfalls. Such a situation is likely to occur in most hierarchical organisations and be more common the more the organisation is diversified. In Chapter 3 I discussed that an organisation that may realise the benefits of evaluating projects concurrently can change the frequency of project evaluation meetings to facilitate a portfolio approach. However, in the case of a diversified organisation, increasing the alignment between the projects being evaluated is likely to involve significantly more difficult structural changes in the organisation. For instance, this may mean divesting certain divisions of the organisation, as General Electric has been doing over the last few years. Therefore, without enough points of reference for a project

5. Looking for alignment in past cases

proposal within an organisation, managers may look to evidence from similar projects from outside the organisation.

Such evidence may include an individual case study from another organisation, or research report that describes a statistical result. Case studies are especially important in managerial decision-making since they are used extensively in business school teaching materials. Therefore, managers are likely to look to case studies to inform their decisions. But would they think that a single case study is more useful than statistical data? The literature on anecdotal bias suggests that they might. Therefore, Chapter 6 considers the influence of an anecdote on project allocation when it conflicts with statistical evidence.

Previous work showed that people often do not always give evidence appropriate weighting in their decisions (Griffin & Tversky, 1992). Therefore I investigated the role of anecdotal and statistical evidence, which are potentially conflicting sources of evidence. This is possible since statistical estimates commonly refer to the mean value of a distribution, whereas individual cases may be sampled, for instance, from either of the tails of the distribution. This comparison would give the appearance of conflicting information, especially if the distribution is skewed. In the same way that in Chapter 4 the intrinsic project features conflicted with the abstract financial metric, in Chapter 6 the description of an anecdote conflicted with the financial metrics of the target project.

Chapter 6 also considered how people dealt with such conflicting information. That is, would they focus on one metric or use a trade-off? In the previous chapter, people did not seem to predominantly use one cue or another. The fact that those in the low alignment condition relied on NPV more than those in the high alignment condition means that those in the high alignment condition were still referring to the intrinsic project features to some extent. Specifically, the different measures' influence may have been integrated in a form 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, I set up the conditions

5. Looking for alignment in past cases

such that it was 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

6

Anecdote similarity moderates anecdotal bias in resource allocation

Contents

| | |
|--|------------|
| 6.1 Introduction | 99 |
| 6.1.1 Anecdotal bias | 100 |
| 6.1.2 Anecdotal bias in business | 101 |
| 6.1.3 Effect of similarity | 102 |
| 6.1.4 Experiment summary | 104 |
| 6.2 Experiment 1 | 104 |
| 6.2.1 Method | 106 |
| 6.2.2 Results | 111 |
| 6.2.3 Discussion | 114 |
| 6.3 Experiment 2 | 116 |
| 6.3.1 Method | 118 |
| 6.3.2 Results | 120 |
| 6.3.3 Discussion | 124 |
| 6.4 General discussion | 126 |
| 6.4.1 Theoretical implications | 128 |
| 6.4.2 Practical implications | 129 |
| References | 131 |

6. Anecdote similarity moderates anecdotal bias in resource allocation

6.1 Introduction

A good story is often more persuasive than data. While usually harmless in daily settings, poor judgement due to a bias towards anecdotal evidence can lead to larger-scale negative consequences. Perhaps the most prominent example of such an error in judgement is 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, usually called *case studies*, as a part of strategic decision-making. Case studies are examples of previous situations that are considered similar by the decision-maker and are used to draw inferences about a target problem. When comparing such examples with aggregated data these are called anecdotes.

Many businesses use case studies to inform their decisions, but often struggle to use them successfully (Gavetti & Rivkin, 2005). This is likely because of the high salience of companies that have ended up either very successful or very unsuccessful. That is, people are often uninterested in average outcomes, but are captivated by both positive or negative extreme outcomes. This has two related impacts: increased salience of an anecdote may increase its influence above statistical data and may shift attention away from structural similarities in favour of more surface similarities. The extent to which statistical data is consulted and the decision-maker's judgement of the anecdote's similarity are two issues that 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. Research has shown that people sometimes prefer anecdotal evidence over statistical data (e.g., 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

6. Anecdote similarity moderates anecdotal bias in resource allocation

problem. Research on similarity distinguishes between surface and relational similarity (Gentner, 1983); the consensus of this research is that the more conceptual structures two cases share, the more useful they are in decision-making (Lassaline, 1996; A. B. Markman & Medin, 1995). As such, case studies that are similar to a target problem on a merely surface level are less useful than those that are related by shared conceptual structure.

Previous research has considered the role of similarity and analogical reasoning in business-related decision-making (Gavetti et al., 2005), but it is unclear to what extent the similarity of a anecdote to the target will affect the strength of the anecdotal bias.

6.1.1 Anecdotal bias

Anecdotal bias is the finding that anecdotal evidence influences people's beliefs more than statistical evidence. Journalists are well aware of the power of anecdotes. For example, an analysis of approximately 29,000 New York Times editorials showing a reliance on anecdotes to drive arguments (Al Khatib et al., 2017). While some research concluded that statistics are more persuasive than anecdotes (e.g., Allen & Preiss, 1997; Hoeken, 2001; Hornikx, 2005) and others were equivocal (Winterbottom et al., 2008), a number of studies 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) suggested that this disparity in findings might be due to statistics having an effect on beliefs and attitudes, and anecdotes affecting intention. A more recent meta-analysis of 61 studies showed that overall people find statistical evidence more persuasive than anecdotal evidence (Freling et al., 2020). However, even if statistical evidence is overall more persuasive across studies, anecdotes than add no additional information to co-presented statistics still influence 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, the decisions are clearly relevant to the decision-maker.

6. Anecdote similarity moderates anecdotal bias in resource allocation

6.1.2 Anecdotal bias in business

It is important to investigate anecdotal bias in business because of the implications this might have on managers' use of case studies. There are many cases of managers successfully using analogies from anecdotal cases, but also of failures to analyse properly (Gavetti et al., 2005; Gavetti & Rivkin, 2005). There is very little research on anecdote bias in business, but the existing work finds clear evidence of the effect. In fact, the recent meta-analysis by Freling et al. (2020) only included the work in 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 that was based on their audit deficiencies for various clients. The experiment was designed in a way that the statistical evidence favoured one firm and the anecdotal evidence favoured the other firm. Participants either viewed an *anecdote only* condition in which they were shown examples of firm deficiencies; an *anecdote + statistics* condition in which they were shown the same as in the anecdote condition, but also the number of deficiencies and clients (but they were not explicitly provided the proportions); a *statistics only* condition in which the proportions and clients without deficiencies are made explicit; an *anecdote + enhanced statistics* condition that added the anecdotes to the statistics only condition; and an *anecdote + enhanced statistics – judgment orientation* condition, which emphasised the importance of proportions and keeping absolute numbers in their relevant context.

Wainberg et al. (2013) found equivalent proportions of participants choosing the firm favoured by the statistical data by those viewing just the anecdotal information and both the anecdotal information and the statistics. Further, participants chose the statistically favoured firm less when seeing both types of information than when seeing just statistics even when the underlying proportions were made explicit (anecdote + enhanced statistics condition). This provided evidence of anecdotal bias, as participants seemed to ignore the contradictory statistical data. Further, having found no difference between the anecdote + statistics condition and the anecdote only condition implies that the anecdotal bias effect is "complete," because

6. Anecdote similarity moderates anecdotal bias in resource allocation

it shows that the statistics displayed played no role in influencing choice. A “partial” effect would have been if the anecdote + statistics condition had been chosen more than the anecdote only condition. This would mean that statistics still played some role in influencing choice. Highlighting relevant statistical features and providing some explanation of statistical inference reduced this anecdotal bias.

Wainberg (2018) conducted a similar study to Wainberg et al. (2013), but with a capital budgeting task. Participants had to choose between purchasing three production-line machines for a mid-sized company that prints circuit boards. The statistical data was constructed such that Machine A was better than Machine B, which was better than Machine C. Participants were either given just this information, or were also provided with an anecdote. This anecdote was in the form of an email correspondence from a colleague from another company that recommended against Machine A (the best option). As in Wainberg et al. (2013), participants were assigned to *anecdote + statistics* and *statistics only* conditions. As well as these, in the *judgment orientation I & II* conditions, participants were told to “think like a scientist” and either received a short or long explanation of what this means (essentially explaining the importance of statistical inference). Wainberg (2018) found that including a contradictory anecdote alongside the statistics (the anecdote + statistics condition) reduced the proportion of choosing Machine A. This indicated evidence for anecdotal bias. The addition of instructions that emphasise scientific thinking further reduced this bias.

6.1.3 Effect of similarity

The extent of one’s reliance on an anecdote should arguably be moderated by its similarity to the target problem. Previous work has discussed the importance of weighting previous cases by their similarity to the present situation (Gilboa & Schmeidler, 1995; Lovallo et al., 2012). For instance, on the one hand, an individual’s decision of whether to take a certain medical treatment should be informed more by a large-scale aggregated study showing 99% efficacy than by a story of someone getting sick as a side-effect of the treatment that they read on

6. Anecdote similarity moderates anecdotal bias in resource allocation

social media. On the other hand, the individual might have reason to be concerned if the person who got sick was the individual's identical twin. The inference that the individual may therefore also need to be cautious about the treatment arises from a specific causal model based on the two cases' shared genetics.

There have been mixed results regarding the effect of similarity of anecdote on the extent of anecdotal bias. Hoeken and Hustinx (2009, Study 3) found evidence for an effect of similarity on anecdotal bias with laypeople for a variety of claims. As well as manipulating whether participants received a claim supported by an anecdote 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. Hoeken (2001) did not find evidence for an effect of similarity about a local government proposal with a student sample. Similarly, Hornikx (2018) considered the effect of similarity on anecdote bias in local government policy decision-making. The researchers did not find an effect of similarity, or an effect of anecdotes. However, they measured persuasiveness and perhaps requiring to make more concrete decisions will be a better test for a more real-life scenario.

Apart from the need to clarify 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 made the distinction between simple surface similarity and deeper relational similarity (Gentner, 1983). As mentioned above, one's use of an anecdote should be moderated by the extent to which it is associated by an underlying causal mechanism or mere surface similarity. Imagine a manager of a multi-divisional company that is deciding how to allocate capital between an oil well project and a technology project. Would hearing of a recent failed oil well project at a different company influence the manager's allocation decision? If so, would it influence the decision because of the fact that the anecdote and one of the target projects are from the same industry (surface similarity)? Or would the manager look to the underlying reason of why the anecdote failed and first identify if this mechanism is relevant to the target oil project? The experiments in this chapter investigate whether the

6. Anecdote similarity moderates anecdotal bias in resource allocation

anecdotal bias effect is due to causal inductive reasoning or merely the association between the valence of the anecdote and surface similarity with the target.

6.1.4 Experiment summary

Experiment 1 investigated whether anecdotal bias in a resource allocation paradigm was moderated by similarity of the anecdotes. Further, I tested whether giving extra information about statistical thinking would encourage participants to consider the statistics over the anecdote. Experiment 1 used a negative anecdote, as this 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 of a bias in positive anecdotes, and Wainberg (2018) did not consider these at all. Therefore, Experiment 2 attempted to replicate the effect of similarity on anecdotal bias with a positive valence anecdote.

6.2 Experiment 1

In Experiment 1 I investigated the effects of similarity and anecdotal bias on resource allocation. I asked participants to allocate a hypothetical budget between two business projects. I also presented participants with a case study that was either similar or dissimilar to the target project (but still from the same industry). Further, participants were allocated to the same conditions as in Wainberg (2018), except that of the two judgement orientation conditions, I only used the *judgment orientation II* condition. Further, I included an anecdote only condition. For the conditions with statistical evidence, participants also saw aggregated information about the success of similar projects in the form of Net Present Value (NPV) and a reliability measure. As in Wainberg (2018), one project (Project A) was clearly better than the other (Project B) when considering the statistical data, but the anecdotal evidence suggested the opposite.

Previous research found that in resource allocation scenarios, people are more persuaded by negative anecdotes than by positive statistical data (Wainberg, 2018).

6. Anecdote similarity moderates anecdotal bias in resource allocation

While other work has shown that similar anecdotes are more persuasive than dissimilar anecdotes (Hoeken & Hustinx, 2009, Study 3), it is unclear how similarity changes the anecdotal bias effect. As such, my main question is whether anecdotal bias will be greater when the anecdote is similar, compared to when it is dissimilar. Evidence of anecdotal bias is given when a statistics only condition is different to an anecdote + statistics condition. Therefore, I tested the following hypothesis:

Hypothesis 6.1—Anecdotal bias is moderated by the similarity of negative anecdotes. The target project is supported by the statistics, but is inconsistent with the anecdotes, thus: Allocations to the target project will be higher when only the statistics are presented and when the statistics are accompanied by a low similarity anecdote, in comparison to when the statistics are accompanied by a high similarity anecdote. In addition, allocations are predicted to not be affected by the low similarity anecdote at all. That is, the statistics only condition should not differ from the low similarity anecdote + statistics condition.

I predicted that the anecdotal bias effect will be complete, as in Wainberg et al. (2013), such that the participants presented with the high similarity anecdote along with the statistics will not use any statistical information. Testing the high similarity condition will allow for an equivalent test to Wainberg et al. (2013). Therefore, I tested the following hypothesis:

Hypothesis 6.2—Effect of statistics for negative anecdotes. Participants will allocate resources equivalently to the target project when in the high similarity anecdote + statistics condition and when in the high similarity anecdote only condition without enhancing the statistics explanation.

Participants with additional information explaining the importance of “scientific thinking” and statistical data may be less affected by anecdotes, as in the *judgment orientation II* condition in Wainberg (2018), here called “the enhanced statistics condition.” Here, I test whether this effect protecting against anecdotal bias would replicate in a resource allocation scenario. Therefore, I tested the following hypothesis:

6. Anecdote similarity moderates anecdotal bias in resource allocation

Table 6.1: Experiment 1 group allocation.

| Anecdote | Alignment | N |
|------------|-----------|-----|
| Anecdote | High | 41 |
| Anecdote | Low | 41 |
| Combined | High | 41 |
| Combined | Low | 41 |
| Enhanced | High | 41 |
| Enhanced | Low | 41 |
| Statistics | NA | 41 |
| Total | - | 287 |

Hypothesis 6.3—Effect of enhanced statistics for negative anecdotes.

Participants will allocate more resources to the target project when in the high similarity anecdote + enhanced statistics condition than when in the high similarity anecdote + statistics condition.

6.2.1 Method

Participants

Two hundred and eighty-seven (199 female) people were recruited from a Psychology undergraduate sample at The University of Sydney. Participants were compensated with course credit. The average age was 20.79 ($SD = 4.93$, $min = 16$, $max = 58$). Participants reported an average of 1.66 ($SD = 3.61$, $min = 0$, $max = 32$) years of work in a business setting, and an average of 0.8 ($SD = 1.57$, $min = 0$, $max = 12$) years of business education. The mean completion time was 22.11 ($SD = 96.95$, $min = 1.67$, $max = 1,101.48$) minutes. Table 6.1 shows the between-subjects condition allocation. Appendix C.1.1 describes the power analysis conducted to arrive at this sample size.

Materials

Instructions All participants initially saw general instructions that explained the task. The subsequent instructions that participants saw depended on their experimental condition. Those in the anecdote only condition were told that they

6. Anecdote similarity moderates anecdotal bias in resource allocation

will see a case study of a failed project and an analysis of why it failed. Those in the statistics only condition were told that they will see NPV and reliability information as a part of their target project descriptions, and were explained that these values were sourced from a study with a large sample. Those in the anecdote + statistics condition were given both of these instructions, and were also told that the information in the anecdote is subsumed in the data of the aggregated study. Those in the anecdote + enhanced statistics condition saw the same as those in the anecdote + statistics condition, but were subsequently given the explanation of scientific thinking that Wainberg (2018) used. Appendix C.1.1 shows screenshots of of these instructions.

Allocation task In the allocation task, participants allocated a percentage of a hypothetical budget between two target projects that come from different businesses within a company. Participants were presented with information about each business' name, location, integration (vertical or horizontal), and organisational structure (centralised or decentralised). See Appendix C.1.1 for an explanation of these terms. Further, participants were presented with information about features of each project that they were told were available to managers before the time of investment. Participants in the anecdote only condition saw just this information (see Figure 6.1), while those in the statistics conditions saw this information along with measures of NPV and “Overall reliability rating” (see Figure 6.2). Participants entered their allocation data underneath this table, in two textboxes labelled *Project A allocation* and *Project B allocation*, respectively.

Anecdote Participants that were presented with an anecdote (those in either the anecdote only, anecdote + statistics , or anecdote + enhanced statistics conditions) saw a description of a business project and an accompanying “analysis.” Figures 6.3 and 6.4 show the anecdote display for those in the high and low similarity conditions, respectively. The project description had a similar layout to the target projects. That is, it contained information about the business name, location, integration,

6. Anecdote similarity moderates anecdotal bias in resource allocation

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 | <ul style="list-style-type: none"> • Oil extracted: 2200L an hour • Time the machinery lasts before requiring maintenance: 8 years • Probability of finding oil: 88% • Type of well: onshore | <ul style="list-style-type: none"> • 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: Project display for the anecdote only condition in Experiment 1.

6. Anecdote similarity moderates anecdotal bias in resource allocation

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 | <ul style="list-style-type: none"> • Oil extracted: 2200L an hour • Time the machinery lasts before requiring maintenance: 8 years • Probability of finding oil: 88% • Type of well: onshore | <ul style="list-style-type: none"> • 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: Project display for the statistics only, anecdote + statistics, and anecdote + enhanced statistics conditions in Experiment 1.

6. Anecdote similarity moderates anecdotal bias in resource allocation

and organisational structure of the business, and detailed predicted features of the project. Underneath this description was a paragraph of text that participants were told was an analysis of why the project failed. Therefore, this text references each of the features in the description in order to justify the project failing.

Those in the high similarity condition saw a description of a project from a business with the same type of investment. All categorical attributes were identical to the relevant target project (Project A), and the numerical attributes were all made to be lower than those in Project A. In the analysis, the numerical attributes were explained to have failed because they were not as high as certain cut-offs. Critically, these cut-offs were made to all be higher than the relevant values in Project A. This was done to make sure that the numerical attributes in the anecdote seem more relevant to those in Project A. For instance, for Project A, oil extraction was 2200L an hour, for the anecdote it was 2000L an hour, and the cutoff was 3000L an hour. As such, a failure of the anecdote because of an insufficient oil extraction rated will seem more relevant since they both share the state of being lower than the cut-off in the analysis. Note, however, that there was uncertainty about the generalisability of these cut-off values because the participants did not receive a explicit indication about whether these values were meant to generalise to other cases.

Follow-up questions Participants that saw the anecdote were subsequently presented with follow-up questions. They were asked how similar they believe the anecdote was to the target project, how relevant it was for their allocations, and how relevant it would be for judgements about other projects of that type. See Appendix C.1.1 for a screenshot of the display.

Procedure

Following ethics and demographics web-pages, participants were introduced to the study through the general instructions and the specific instructions relevant to the condition. They then saw the allocation task, which included the anecdote analysis and description (for those not in the statistics condition), and the target

6. Anecdote similarity moderates anecdotal bias in resource allocation

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 display for those in the high alignment condition in Experiment 1.

projects description. Those that saw an anecdote were subsequently shown the follow-up questions.

6.2.2 Results

The effect of similarity on anecdotal bias

Figure 6.5 shows the results for the allocation data. I tested anecdotal bias by comparing the statistics only condition with the high and low similarity anecdote + (not enhanced) statistics conditions. The omnibus one-way ANOVA test of these three conditions was significant, $F(2, 120) = 4.40, p = .014, \hat{\eta}_p^2 = .068$. Planned comparisons revealed that participants allocated more to the target project when seeing only statistics than when seeing the high similarity anecdote and statistics,

6. Anecdote similarity moderates anecdotal bias in resource allocation

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 display for those in the low alignment condition in Experiment 1.

$\Delta M = -12.78$, 95% CI $[-21.90, -3.66]$, $t(120) = -2.77$, $p = .006$; but not compared to seeing the low similarity anecdote and statistics, $\Delta M = -2.17$, 95% CI $[-11.29, 6.95]$, $t(120) = -0.47$, $p = .638$. Therefore, I found evidence of anecdotal bias only in the high similarity condition.

The effect of enhanced statistics

To investigate the effect of enhanced statistics I compared the conditions in which participants saw both an anecdote and statistics to the conditions in which they saw the same, but with enhanced statistics. The two-way interaction between similarity and the two anecdote + statistics conditions was not significant, $M = 4.12$, 95% CI $[-8.56, 16.80]$, $t(240) = 0.64$, $p = .523$, as was the main effect of anecdote + statistics

6. Anecdote similarity moderates anecdotal bias in resource allocation

condition (averaging over similarity), $\Delta M = -0.23$, 95% CI [-6.57, 6.11], $t(240) = -0.07$, $p = .943$. Therefore, I did not find evidence that providing participants with instructions with how to think statistically facilitated a focus on statistics.

The effect of statistics

A two-way ANOVA was conducted to investigate the interaction of similarity (low and high) and anecdote conditions (anecdote only, statistics + anecdote, excluding anecdote + enhanced statistics). This allowed me to identify the role of statistics. The interaction between anecdote condition and similarity, excluding the enhanced condition was significant, $M = -13.71$, 95% CI [-26.39, -1.03], $t(240) = -2.13$, $p = .034$. Specifically, the difference between allocations when only seeing an anecdote and seeing the anecdote + statistics was greater when the anecdote was similar, $\Delta M = -21.56$, 95% CI [-32.30, -10.83], $t(240) = -4.74$, $p < .001$; compared to when it was dissimilar, $\Delta M = -7.85$, 95% CI [-18.59, 2.88], $t(240) = -1.73$, $p = .198$. Therefore, I found evidence of “partial” anecdotal bias in the high similarity condition, since the anecdote + statistics condition was lower than the statistics only condition, but higher than the anecdote only condition.

Relevance ratings

I conducted regression analyses to determine the relationship between allocations and the follow-up relevance ratings. As seen in Figure 6.6 the specific relevance ratings interact with similarity condition, $b = -2.81$, 95% CI [-4.76, -0.86], $t(242) = -2.83$, $p = .005$. It appears that specific relevance ratings are related to allocations, but only in the high similarity condition. Further, there were no significant associations with the general relevance ratings. This suggests that people are reasoning about the connection between the anecdote and target, as opposed to simply reacting to the failed project and associating that with the project industry.

6. Anecdote similarity moderates anecdotal bias in resource allocation

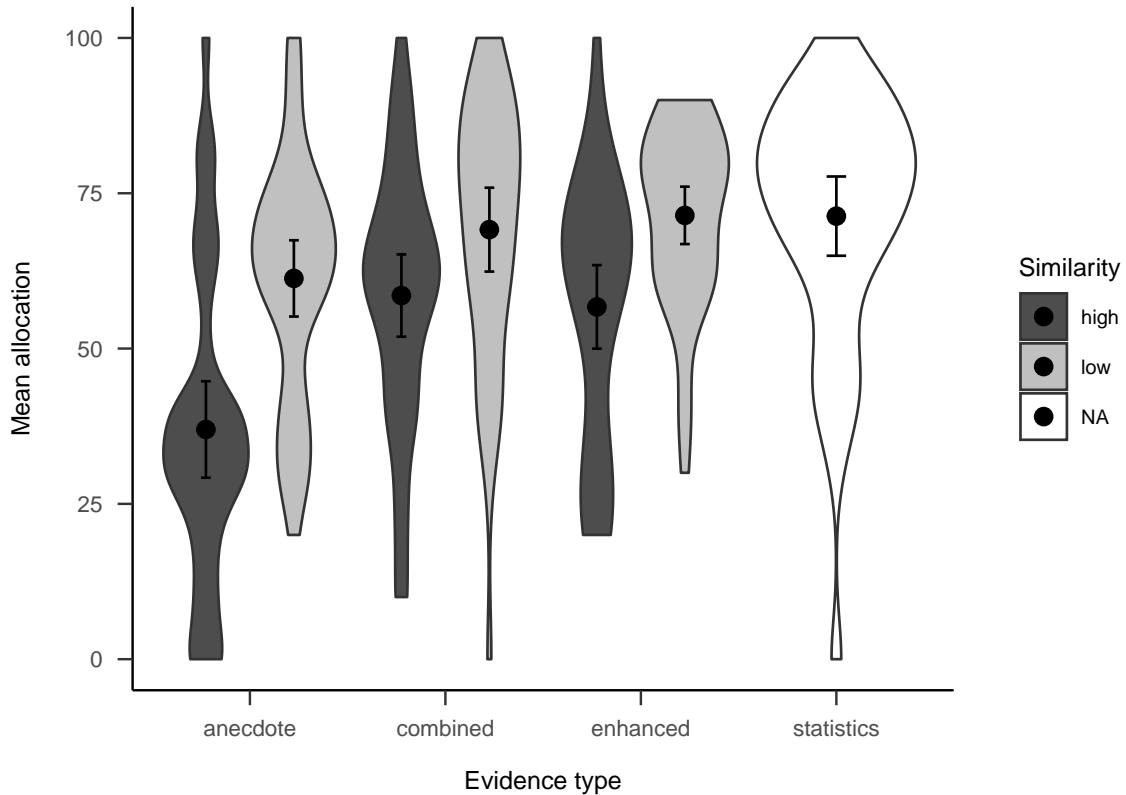


Figure 6.5: Mean allocation to Project A (the target project). Error bars represent 95% confidence intervals.

6.2.3 Discussion

I found support for Hypothesis 6.1, as participants allocated less resources when seeing both the anecdote and statistics, than when just seeing statistics, in the high similarity condition, but not in the low similarity condition. This shows that while anecdotal bias exists when the anecdote is similar, participants are not influenced when the causal mechanisms do not match. Contrary to Hypothesis 6.2, I found that while participants were influenced by the anecdote, they still made some use of the statistics. This is different from Wainberg et al. (2013), who found no difference between an anecdote only and a anecdote + statistics condition, indicating a “complete” effect of anecdotal bias. Hypothesis 6.3 was also not supported, as the added enhanced statistical language used to encourage participants to use the statistical information did not contribute to reducing participants’ reliance on anecdotes.

6. Anecdote similarity moderates anecdotal bias in resource allocation

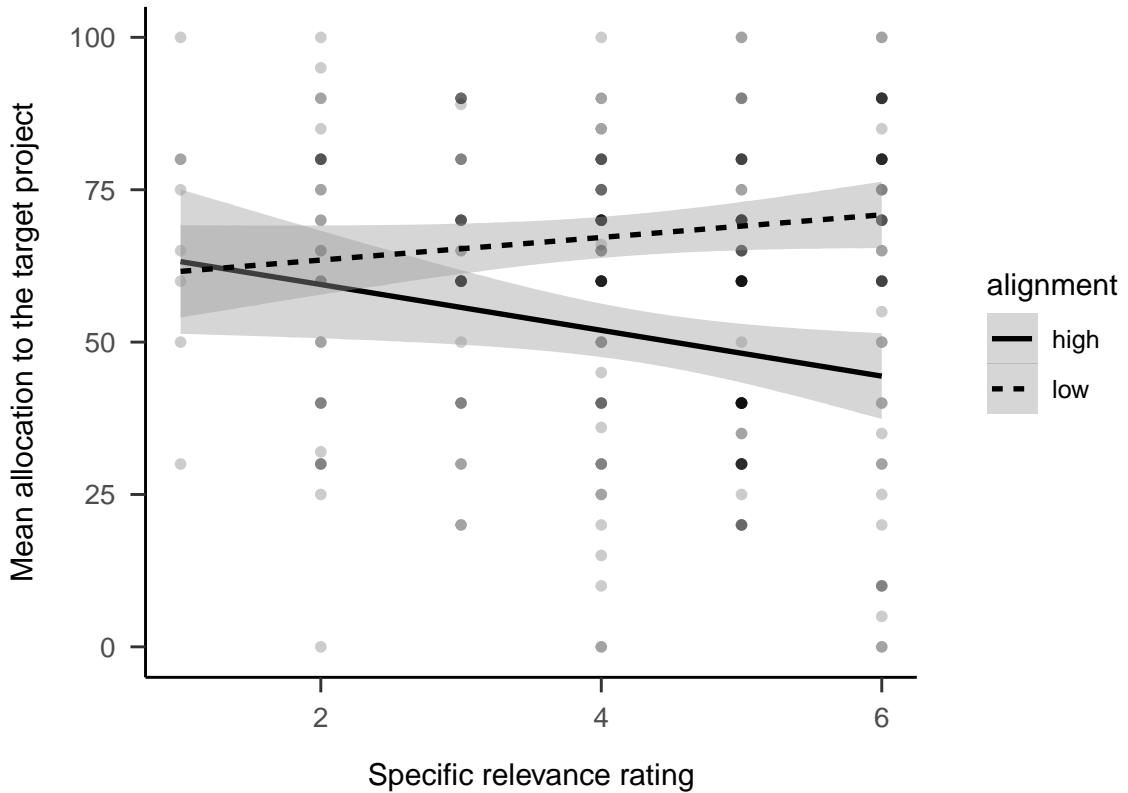


Figure 6.6: Mean allocations to the target project, by specific relevance rating and similarity condition.

Experiment 1 was limited because it only considered an anecdote with a *negative* valence. That is, the case study was of a project that failed. However, in real life these case studies are often ones with *positive* valence. That is, a story of a successful company. In fact, it may be the case that in business, the anecdotes that are used are more likely to be positive, because of survivorship bias. Jaramillo et al. (2019) found an anecdotal bias effect in negative anecdotes, but not in positive anecdotes. This may be because the work was done in medical decision-making and in this domain, a loss of health may be felt stronger than an equivalent gain of health. Therefore, in Experiment 2 I added a condition with a positive anecdote, in order to investigate whether anecdote valence will impact the anecdotal bias effect.

Further, it was unclear if the effects found in Experiment 1 were related to participants' perceptions of the type of sampling used in selecting the anecdotes. The instructions in Experiment 1 do not explain how the anecdote that is displayed to participants was chosen. Intentional or random sampling has been shown to

6. Anecdote similarity moderates anecdotal bias in resource allocation

affect people's decision-making (e.g., Hayes et al., 2019). In the case of the current experiments, the sampling assumption changes the extent to which it is rational to use the anecdote or not. It may be considered rational to choose the anecdote over the aggregated data if 1. the anecdote was not sampled randomly from the pool of anecdotes, and 2. the anecdote is more similar to the target project than any of the other anecdotes in the pool in relevant ways. That is, if the anecdote was 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 might have held these beliefs. In order to control for these assumptions, in Experiment 2 I added text to the instructions that clarified that the anecdote 1. was sampled randomly from the pool of anecdotes, and 2. is 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. That is, people use an anecdote more when presented with conflicting statistics than when anecdote alone and less than when statistics alone. However, anecdote similarity moderated this effect, such that anecdotal bias is stronger when the anecdote is similar to the current task, than when it is dissimilar. Experiment 1 only used a negative anecdote because previous research found anecdotal bias for negative, but not for positive anecdotes (Jaramillo et al., 2019). However, Jaramillo et al. (2019) considered medical decision-making, so this effect of anecdote valence may be different in a business scenario. When considering medicine, it is likely that negative anecdotes are more salient than positive ones. Positive health can generally simply mean a lack of change of one's current state, whereas a negative health outcome almost necessarily involves a change for the worse. Further, Experiment 1 did not clarify certain assumptions about the way the displayed anecdote was sampled from the pool of anecdotes.

Therefore, in Experiment 2 I added a within-subjects anecdote valence manipulation and manipulated anecdote similarity within-subjects, in order to increase

6. Anecdote similarity moderates anecdotal bias in resource allocation

the experiment's power. Further, I removed the anecdote + enhanced statistics manipulation as Experiment 1 did not find evidence for its efficacy. All participants saw the statistics only condition, as it did not contain an anecdote, and therefore did not need to be manipulated between-subjects. Each participant therefore saw five displays, with one statistics only condition, and four displays for either the anecdote only condition, or the anecdote + statistics condition. These four anecdote displays consisted of the similarity (low and high) \times valence (negative and positive) conditions.

I expected to replicate the effects of Experiment 1, as well as seeing the reverse effect in the positive valence condition. Also, I expected that statistics will have an effect similar to Experiment 1, contrary to Hypothesis 6.2. Therefore, as well as again testing Hypothesis 6.1, I tested the following hypotheses (see Appendix C.2 for a plot of a simulation of all hypothesised effects):

Hypothesis 6.4—Overall effect. Three-way interaction of similarity \times valence \times anecdote, excluding statistics-only

Hypothesis 6.5—Anecdotal bias moderated by similarity for positive anecdotes. When the anecdote is positive, allocations will be higher in the statistics-only condition than in both the anecdote + statistics conditions (high and low similarity). Within these two anecdote + statistics conditions, allocations will be higher when the anecdote is similar than when it is dissimilar.

After not replicating the lack of a statistics effect as in Wainberg et al. (2013), in Experiment 2 I expected to replicate the finding in Experiment 1 that participants do somewhat integrate statistics in their decisions. Therefore, I tested the following hypotheses:

Hypothesis 6.6—Effect of statistics for negative anecdotes. In the negative valence condition, allocations will be higher for the high similarity anecdote + statistics condition than the high similarity anecdote only condition.

6. Anecdote similarity moderates anecdotal bias in resource allocation

Table 6.2: Experiment 2 group allocation.

| Anecdote between | N |
|------------------|----|
| Anecdote only | 48 |
| Combined | 48 |
| Total | 96 |

Hypothesis 6.7—Effect of statistics for positive anecdotes. In the positive valence condition, allocations will be higher for the high similarity anecdote only condition than the high similarity statistics + anecdote condition.

6.3.1 Method

Participants

Ninety-six (50 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 41.69 ($SD = 11.29$, $min = 27$, $max = 74$). Participants reported an average of 7.19 ($SD = 8.34$, $min = 0$, $max = 43$) years of work in a business setting, and an average of 3.91 ($SD = 7.66$, $min = 0$, $max = 50$) years of business education. The mean completion time was 15.01 ($SD = 8.73$, $min = 2.57$, $max = 58.71$) minutes. Table 6.2 shows the between-subjects condition allocation. Similarity and valence were manipulated within-subjects. Therefore, each participant was in one of two between-subjects anecdote conditions, and saw five displays (statistics only, and one for each similarity/valence combination). Appendix C.2.1 describes the power analysis conducted to arrive at this sample size.

Materials

Instructions Participants were shown similar instructions to [Experiment 1](#). I also included a test of basic instructions understanding that also functioned as an attention check. As in Experiment 1, participants also saw 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 target projects. One important difference

6. Anecdote similarity moderates anecdotal bias in resource allocation

from Experiment 1 was that I clarified in the instructions text both that the anecdote was sampled randomly and that the anecdotes in the pool were all equally similar to the target project. Appendix C.2.1 shows screenshots of these web-pages.

Allocation task As in Experiment 1, the allocation task included a description and analysis of an anecdote (except for those in the anecdote only condition) and a project display with a table describing the two target projects. Figures 6.7 and 6.8 show the anecdote and target projects for the negative valence low similarity condition, respectively. Figures 6.9 and 6.10 show the anecdote and target projects for the positive valence high similarity conditions, respectively. In the statistics only condition, participants only saw the target projects display. Appendix C.2.1 details the counterbalancing and randomisation that I used.

Interstitial Before each display, participants saw an “interstitial” page, whose role was 1. to introduce the next display, and 2. to provide an attention check (not required to answer, so can be skipped if the interstitial text isn’t read). See Appendix C.2.1.

Follow-up questions Participants were shown similar follow-up questions as in Experiment 1, except that here the rating scales were 1-7, instead of 1-6. See Appendix C.2.1 for an example of a follow-up display.

Procedure

Following ethics and demographics web-pages, participants were introduced to the study through the general instructions. They then saw five sets of two web-pages (in randomised order). Each “set” contained two web-pages: the allocation task and a follow-up questions page (except for the anecdotes only condition, in which participants did not see 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 target projects description. The only exception was the statistics only display, for which there was no anecdote description or analysis.

6. Anecdote similarity moderates anecdotal bias in resource allocation

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

6.3.2 Results

I analysed the allocation data that was relevant to the Experiment 2 hypotheses. See Appendix C.2.2 for manipulation check analyses, and analyses of the follow-up rating data.

Overall effect of manipulations

As seen in Figures 6.11 and 6.12, the similarity \times valence \times anecdote interaction (excluding the statistics-only condition) was not significant, $F(1, 94) = 3.42, p = .067, \hat{\eta}_p^2 = .035$. However, the similarity \times valence interaction was significant,

6. Anecdote similarity moderates anecdotal bias in resource allocation

| 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 | <ul style="list-style-type: none"> • Code written: 1000 lines a day • Security rating: 85% • Number of potential customers in first year: 3 million • Target users: ordinary consumers | <ul style="list-style-type: none"> • 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: <input type="text" value=""/> | Allocation: <input type="text" value=""/> |
| Overall reliability rating (%) | 91 | 90 |
| NPV (\$) | 901 | 100 |

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

6. Anecdote similarity moderates anecdotal bias in resource allocation

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 6.9: 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 anecdote \times valence interaction,
 $F(1, 94) = 10.11, p = .002, \hat{\eta}_p^2 = .097$.

Anecdotal bias moderated by similarity

To investigate whether anecdotal bias was moderated by similarity, I compared allocations between the anecdote + statistics high and low similarity conditions and the statistics-only condition. I found that in the negative valence condition, participants allocated more in the statistics-only condition than in the anecdote + statistics high similarity condition, $\Delta M = -19.25, 95\% \text{ CI } [-29.66, -8.84]$, $t(141) = -3.65, p < .001$. Participants in the anecdote + statistics low similarity

6. Anecdote similarity moderates anecdotal bias in resource allocation

| 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 | Solistics | Altchip |
| Project type | shipping logistics | microchip |
| Location | Kuala Lumpur, Malaysia | Toronto, Canada |
| Integration | vertical | horizontal |
| Structure | centralised | decentralised |
| Predicted project features | <ul style="list-style-type: none"> • 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 | <ul style="list-style-type: none"> • 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: <input type="text"/> | Allocation: <input type="text"/> |
| Overall reliability rating (%) | 93 | 90 |
| NPV (\$) | 905 | 105 |

Figure 6.10: An example of the target projects in the positive valence, high similarity condition of Experiment 2.

6. Anecdote similarity moderates anecdotal bias in resource allocation

allocated more than when they saw the anecdote + statistics high similarity display, $\Delta M = -18.17$, 95% CI [-28.58, -7.75], $t(141) = -3.45$, $p = .001$. However, the difference between responses to the statistics-only projects and the anecdote + statistics low similarity projects was not significant, $\Delta M = -1.08$, 95% CI [-11.50, 9.33], $t(141) = -0.21$, $p = .837$. This provides evidence for the moderation of anecdotal bias by similarity for negative anecdotes.

In the positive valence condition, allocations were higher in the statistics-only condition than in the anecdote + statistics low similarity condition, $\Delta M = -31.73$, 95% CI [-41.86, -21.60], $t(141) = -6.19$, $p < .001$. Allocations were also higher in the anecdote + statistics high similarity condition than in the anecdote + statistics low similarity condition, $\Delta M = 14.10$, 95% CI [3.97, 24.23], $t(141) = 2.75$, $p = .007$. Further, allocations were higher in the statistics-only condition than in the anecdote + statistics high similarity condition, $\Delta M = -17.62$, 95% CI [-27.75, -7.50], $t(141) = -3.44$, $p = .001$. This provides evidence for the moderation of anecdotal bias by similarity for positive anecdotes.

Effect of statistics

As in Experiment 1, I investigated the extent to which the statistical information influenced participants' allocations. When in the negative valence condition, participants allocated more to the high similarity anecdote + statistics project than those in the high similarity anecdote-only condition, $\Delta M = -12.67$, 95% CI [-23.53, -1.81], $t(336.36) = -2.29$, $p = .022$. When in the positive valence condition, they allocated more to the high similarity anecdote-only condition 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.

6.3.3 Discussion

I found support for Hypothesis 6.1 and Hypothesis 6.5, as participants showed a stronger anecdotal bias effect when the anecdote was more similar to the target

6. Anecdote similarity moderates anecdotal bias in resource allocation

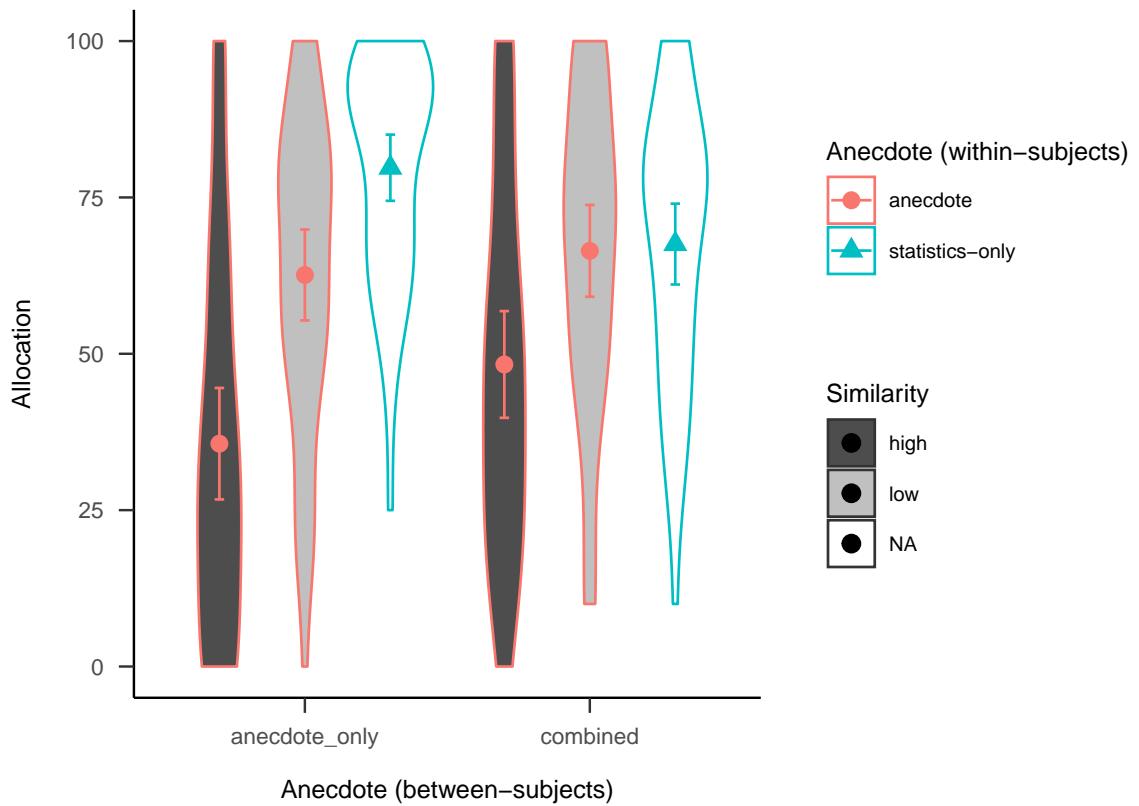


Figure 6.11: Mean allocation in Experiment 2 for the positive valence condition.

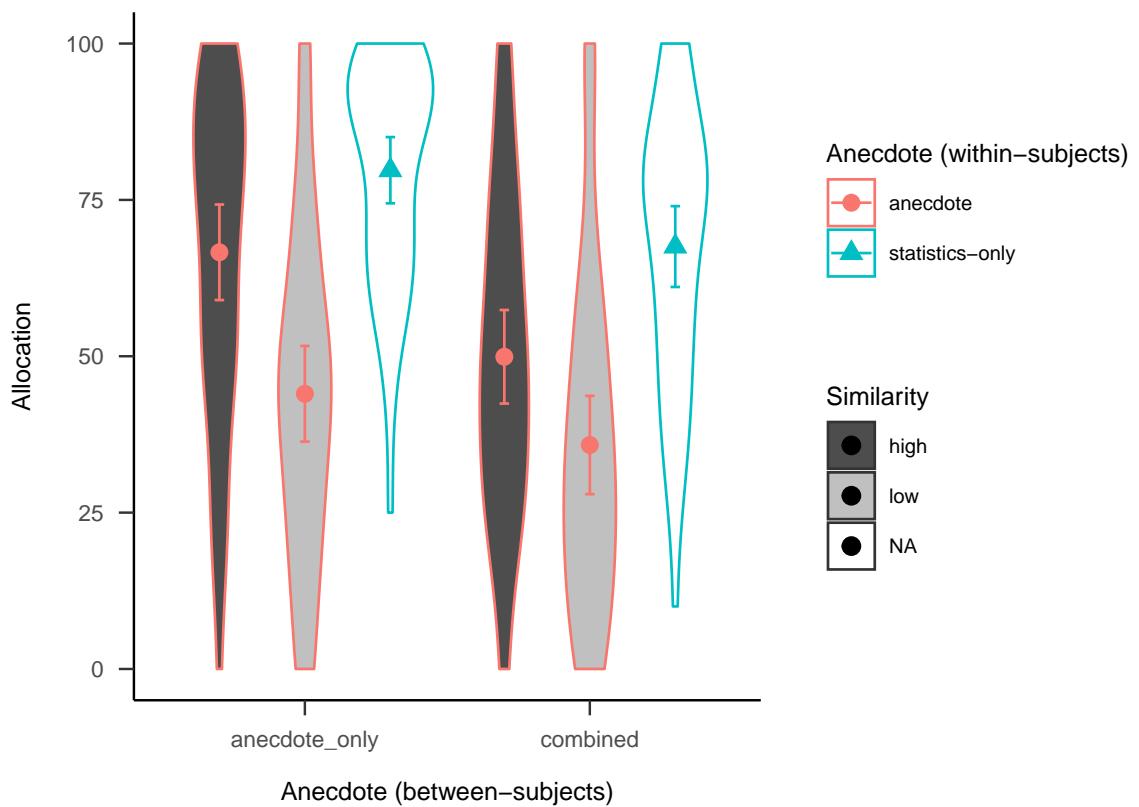


Figure 6.12: Mean allocation in Experiment 2 for the positive valence condition.

6. Anecdote similarity moderates anecdotal bias in resource allocation

project, both for positive and negative anecdotes. Further, as per Hypothesis 6.6 and Hypothesis 6.7, I found that participants seemed to incorporate the statistical information into their judgements, in both negative and positive anecdotes.

Experiment 2 therefore found that, unlike in the medical domain, the effect of anecdotes in financial decision-making does not seem to depend on anecdote valence. Further, as in Experiment 1, and unlike Wainberg et al. (2013), the anecdotal bias effect does not seem 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. I found that people's decisions are influenced by anecdotes even when aggregated data is available in a resource-allocation context. I also made three novel findings: 1. I found that the anecdotal bias effect is only seen when participants consider the anecdote as sufficiently relevant to the target at hand, 2. participants still seem to integrate statistics in their decisions, and 3. these effects are seen in both negative and positive anecdotes.

The first novel finding from these experiments was that participants appeared to moderate their use of anecdotal evidence. Specifically, when the anecdote appeared to be causally relevant, participants used it in their decisions. However, when it appeared irrelevant, participants relied on statistics almost entirely. The findings in the high similarity condition are largely congruent with findings from other work investigating anecdotal bias in business decision-making. As in Wainberg et al. (2013) and Wainberg (2018), I found that people allocate less resources to a project that is successful according to statistical evidence, but is displayed alongside a contradictory similar anecdote, than to a project with just the statistics.

It seems that participants made the distinction between the low and high similarity conditions based on underlying structure of the anecdote. The low similarity condition always consisted of the same project type, for each domain, as in the high similarity condition. For instance, in one variation, both the high and

6. Anecdote similarity moderates anecdotal bias in resource allocation

low similarity anecdotes were of oil well projects. This means that participants were sensitive to the specific information presented to them 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, while they considered that the anecdote was relevant in when considering the target project, they did not consider it as relevant to other such projects. In other words, participants did not appear to carelessly use anecdotal evidence in their decisions, but instead appeared to carefully consider the anecdote based on its particular causal structure.

The second novel finding from these experiments was that participants that saw both statistics and anecdotal evidence did not completely disregard the statistical measures. Wainberg et al. (2013) found a “complete” effect of anecdotal bias, because in their study, these two conditions were equivalent. This meant that the statistics they provided had a negligible effect on participants’ decisions. These experiments, on the other hand, showed a “partial” anecdotal bias effect, with allocations being different between when participants saw only an anecdote and when they also saw the statistics. It seems as if participants integrated the anecdote with the statistical information. As mentioned above, this suggests that people’s evaluation of evidence might be more sensitive than previously thought.

This discrepancy might be a result of the sampled population. Since Freling et al. (2020) found a stronger effect of anecdote when decisions were more personally relevant, the manager sample in Wainberg et al. (2013) may have simply been more personally invested in the task than the laypeople in the experiments in this chapter. Similarly, Yang et al. (2015) found that anxiety increases anecdotal bias in risky choice. It might also be due to the anecdote + statistics condition not being equivalent between Wainberg et al. (2013) and the present work. Specifically, the statistics shown in the anecdote + statistics condition in Wainberg et al. (2013) were not the same ones that were 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

6. Anecdote similarity moderates anecdotal bias in resource allocation

the statistics only condition. This suggests that people integrate statistics when they are sufficiently clear and no further interpretation is required.

The third novel finding from these experiments was that anecdotal bias was seen in both negative and positive anecdotes. Most studies in the literature considered anecdotes that involve an example with negative consequences (a *negative* anecdote). For instance, a medication that leads to an adverse reaction in a patient. However, there is not much work in the literature that involves an anecdote with positive consequences (a *positive* anecdote). Jaramillo et al. (2019) found an asymmetry in the effect of the anecdote, such that the effect was stronger when a person in a description did not get better after a medication (negative), compared to when they did get better (positive). In the present experiments I found a more symmetrical effect, such that both the effects of the moderated anecdotal bias and the influence of statistics were found in both valence conditions. The difference between this and the finding from Jaramillo et al. (2019) might be due to the domain of the task in that people are more focused on negative outcomes in the medical domain.

6.4.1 Theoretical implications

The current work adds to the current understanding of the way people use different forms of evidence in their decision-making. Previous work mostly investigated the relative influence of statistics and anecdotes by comparing anecdote and statistics conditions. The current work shows that comparing a joint anecdote + statistics condition to both an anecdote only and statistics only condition allows for a more specific representation of participants' anecdotal bias. The influence of anecdote can be seen in the comparison between statistics only and the anecdote + statistics condition, but the effect of statistics can be seen in the comparison between the combined condition and the anecdote only condition. These two effects allow to determine the independent influence of anecdote and statistics, respectively. Further use of such a design in future research might help to further understand the conditions under which these types of evidence are used.

6. Anecdote similarity moderates anecdotal bias in resource allocation

It seems that in some of the anecdotal bias literature there is an assumption that using anecdotal evidence over statistical evidence is necessarily irrational. This likely arises from examples from 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 a rational use of anecdote based on the similarity of the anecdote to the target. For instance, there are times in which an anecdote is so similar to the target situation (e.g., the identical twin example in the [Introduction](#)) that it would be unwise not to consider the anecdote. That is, the use of anecdote should depend on both 1. the strength of the underlying causal structure between it and the target problem, and 2. the distribution of similarity across cases in the sample on which the statistics are based. People should use an anecdote when causal structure is significantly more relevant than other cases in available data. It is also important to note that there could also be misleading similarity. For instance, if someone is highly similar, but not along some key hidden dimension that is the real causal thing to care about, then using the anecdote may be the wrong thing to do. What seems to be important is a sensitivity to relational, rather than surface level similarity. Future research should further investigate how varying the assumptions that people have about sampling from a data set of anecdotes influences their anecdotal bias. Such assumptions can include the size of the sample and the shape of the distribution.

6.4.2 Practical implications

The current work can contribute to managerial decision making by suggesting insights into how managers make decisions about business case studies and statistical information about certain industries. Managers of large companies are often in a difficult position; they have incomplete information and an uncertain environment. Despite this, different biases and responses to those biases can be anticipated for different levels of uncertainty. For instance, a manager may be in a position in which they are presented with both a convincing case study that suggests a certain course

6. Anecdote similarity moderates anecdotal bias in resource allocation

of action, and also be in the possession of 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: 1. the quality of the aggregated data (determined by factors such as the sample size), 2. the relative similarity of the cases in the data pool to the situation at hand, and 3. the similarity of the anecdote to the situation at hand. For instance, if the anecdote is uniquely similar to the target situation, and it is significantly more similar than the rest of the cases in the data set, then it should have more weight than an anecdote that comes from a pool of cases that are all just as similar to the target.

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

6. Anecdote similarity moderates anecdotal bias in resource allocation

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>
- 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.
- 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/gjscrif>
- 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>

6. Anecdote similarity moderates anecdotal bias in resource allocation

- 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/gjscrn>
- 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. United States – New York. Retrieved April 6, 2021, from <http://search.proquest.com/docview/304937594/abstract/4B8166E9517494FPQ/1>
- 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 Piercy, 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

| | | |
|-----|------------------------------------|-----|
| 7.1 | Summary of results | 134 |
| 7.2 | Theoretical implications | 136 |
| 7.3 | Practical implications | 140 |
| 7.4 | Conclusion | 142 |
| | References | 143 |

This thesis investigated the psychology of resource allocation decisions. The influence of psychological factors on such decisions has not been sufficiently considered in the literature despite its 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. Each of the empirical chapters investigated distinct but related processes that are relevant to the resource allocation process. By doing this I was able to investigate whether people are 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

7. Discussion

conflicts with statistical evidence (Chapter 6). I will first summarise the results of the empirical chapters and then discuss their theoretical and practical implications.

7.1 Summary of results

Chapter 2 investigated the choice of risky business projects that are displayed sequentially and without feedback. This design modelled the real-life situation that managers are faced with 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 encourage aggregation.

Chapter 4 investigated resource allocation when projects were evaluated jointly and capital was allocated as a proportion of the budget, rather than a binary choice. I manipulated whether all the project attributes were alignable, or only the abstract financial metric (NPV) was alignable. I also manipulated whether NPV was to be considered a reliable metric or not, and whether this information was expressed as explicit verbal instruction or as numerical ranges. The results showed that when reliability information was presented verbally, participants used it 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 regardless of the reliability information. However, when reliability information was presented

7. Discussion

numerically, there was no moderation of allocation based on the ranges—participants used NPV even when they had an opportunity to use the intrinsic features of the project. Overall, however, participants tended to rely on NPV more in the low alignment condition than in the high alignment condition.

Chapter 6 investigated the effect of anecdote similarity on allocations when anecdote conflicted with statistical data. Participants were asked to allocate a hypothetical budget between two projects. One of the projects 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. This anecdotal project was always in the same industry as one of the target projects. However, the anecdote description either contained substantive connections to the target or not. I also manipulated whether the anecdote conflicted with the statistical measures because it was 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. I found this for both positive and negative anecdotes. However, 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 decision and still showed an anecdotal bias effect. Therefore, people seem to appropriately moderate their use of anecdotes based on the anecdotes' relevance, but do not understand the implications of more complex statistical principles.

Together, these results show the bounds of people's decision-making capability in resource allocation. The participants in these experiments in general behaved rationally but struggled to incorporate more advanced statistical principles into their decisions. Further, when confronted with multi-attribute choice, participants tended to allocate resources 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 were able to moderate their allocations when the moderating factors were sufficiently clear (as in the verbal

7. Discussion

reliability condition in Chapter 4). However, participants struggled to do this when the moderating factor involved an understanding of a more advanced statistical principle. Each empirical chapter included such a principle: risk aggregation in Chapter 2, metric variance in Chapter 4, and sample distribution in Chapter 4. However, a formal understanding of such principles does not seem to be necessary if they are expressed explicitly (as in the aggregated distribution in Chapter 2).

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 resource allocation scenarios. People made good decisions most of the time, but sometimes do not take into account important moderating factors 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 resource allocation decisions are bounded. To this end, in each experiment, participants were given resource 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 principles in their decisions, but that they are also capable of making nuanced trade-offs and can be assisted by decision aides.

7. Discussion

In Chapter 2 I presented participants with a resource 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 of all the gambles some 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 principle 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 numerical variance would have allowed participants to moderate their allocations 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: 1. noticing that in the low numerical reliability condition the ranges were all overlapping, and 2. noticing the difference in the width of the ranges between the two within-subjects reliability amount 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 moderated allocation when reliability was expressed explicitly as words, but not when it was expressed numerically.

7. Discussion

In Chapter 6 participants did not make use of important information about the 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 was negatively 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.

While people struggled to understand and use difficult statistical principles they still seemed to be able to integrate multiple cues and create trade-offs. As discussed in Chapter 5, in both Chapters 4 and 6 participants were provided 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 have some relevance. Similar to 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 using 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 explain individual differences.

7. Discussion

While trade-offs allow people to integrate multiple cues, decision aides allow for people to use statistical principles for more complex moderation. In Chapter 2 I 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, in a follow-up experiment to Chapter 4 (detailed in Appendix B.7), I 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 resource 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). A Hypothetical Outcome Plot (HOP; Hullman et al., 2015; Kale et al., 2019) is one method that is likely to be beneficial to people's understanding of ranges as used in this thesis. HOPs express variance information as dynamic plots. 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 might also be necessary to facilitate understanding of the relevance of the sample distribution. Ultimately, it seems that visualisations of the effects of complex statistical concepts are often necessary for people to use them appropriately.

Future research should also investigate the potential expertise effects that may influence the findings in the thesis. This is important because of the potential downstream effects that biased managerial decision-making might have. 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 perform better

7. Discussion

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, there are reasons to believe that these effects may actually be stronger in managers. 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 managerial one. Further, Chapter 4 found evidence of variance neglect for both laypeople and Masters of Management students. In the case of the work in Chapter 6, it might be possible that business managers often prefer anecdotal accounts 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 according to Freling et al. (2020) would suggest 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, in this section I will assume that the findings of this 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

7. Discussion

& Sibony, 2014; Sibony et al., 2017), with only some specifically focused on resource allocation decisions (Birshan et al., 2013). I will now 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 way that business projects are framed 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 make it easier for managers to 1. group projects together, and 2. aggregate a portfolio of projects for them. This suggests implementing organisational changes that will facilitate the resource 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 underlying abstract financial measures. Further, translating such numerical variance estimates into clear verbal clear 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.

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

7. Discussion

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; A. Markman & Gentner, 1993). Therefore, managers should take care to first identify such relational structures first before making subsequent inferences.

In each of these cases, addressing these psychological effects will help eliminate some of the biases in the resource 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, which can in turn be addressed (Flyvbjerg et al., 2018).

7.4 Conclusion

Resource allocation decisions can be consequential for large organisations. Therefore, this thesis tested the conditions under which people behave rationally or are fallible when allocating resources. The experiments found that people struggle to incorporate concepts such as risk aggregation, estimate variance, and sample distribution into their decisions. People 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 people were capable of integrating conflicting information in their decisions. Identifying such cognitive bounds helped to both better understand how people evaluate multiple choices and future research can use this work to develop methods to facilitate better decisions.

7. Discussion

References

- Birshan, M., Engel, M., & Sibony, O. (2013). Avoiding the quicksand: Ten techniques for more agile corporate resource allocation [magazine]. *McKinsey Quarterly*, 6. <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
- 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. Retrieved July 24, 2020, from <http://search.ebscohost.com/login.aspx?direct=true&db=bsu&AN=91571432&zsite=bsi-live>
- 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](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>
- Hall, S., Lovallo, D., & Musters, R. (2012). How to put your money where your strategy is [magazine]. *McKinsey Quarterly*. <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/how-to-put-your-money-where-your-strategy-is>
- 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>

7. Discussion

- 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 [magazine]. *McKinsey Quarterly*, (4), 15–17. Retrieved April 20, 2021, from <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/overcoming-a-bias-against-risk>
- Kox, E. (2018). *Evaluating the effectiveness of uncertainty visualizations: A user-centered approach* (Masters thesis). University of Utrecht.
- 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.
- Lapinski, A.-L. S. (2009, October 1). *A Strategy for Uncertainty Visualization Design*. DEFENCE RESEARCH AND DEVELOPMENT ATLANTIC DARTMOUTH (CANADA). Retrieved May 4, 2021, from <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? [magazine]. *McKinsey Quarterly*, 3. <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/gjscc9>
- Markman, A., & 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

7. Discussion

- 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>
- 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. Retrieved May 4, 2021, from <http://eprints.cs.univie.ac.at/4598/>

A

Chapter 2 appendix

Contents

| | |
|-----------------------------------|------------|
| A.1 Experiment 1 | 147 |
| A.1.1 Method | 147 |
| A.1.2 Results | 149 |
| A.2 Experiment 2 | 150 |
| A.2.1 Method | 150 |
| A.2.2 Results | 152 |
| A.3 Experiment 3 | 154 |
| A.3.1 Method | 155 |
| A.3.2 Results | 158 |
| A.3.3 Discussion | 162 |
| A.4 Experiment 4 | 164 |
| A.4.1 Method | 165 |
| A.4.2 Results | 167 |
| A.4.3 Discussion | 173 |

This appendix contains supplementary materials and analyses for the two experiments reported in Chapter 2. In addition, I 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

A. Chapter 2 appendix

Imagine that you are an executive in a large company composed of many individual businesses.

You will see various projects from these businesses and have to decide whether you would like to invest in them.

Imagine that making good investment decisions will result in you receiving a generous bonus and a potential promotion, and that doing poorly will result in you receiving a large pay cut and a potential demotion.

We want to know what choices you would actually make in these scenarios.

Continue

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

Materials

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

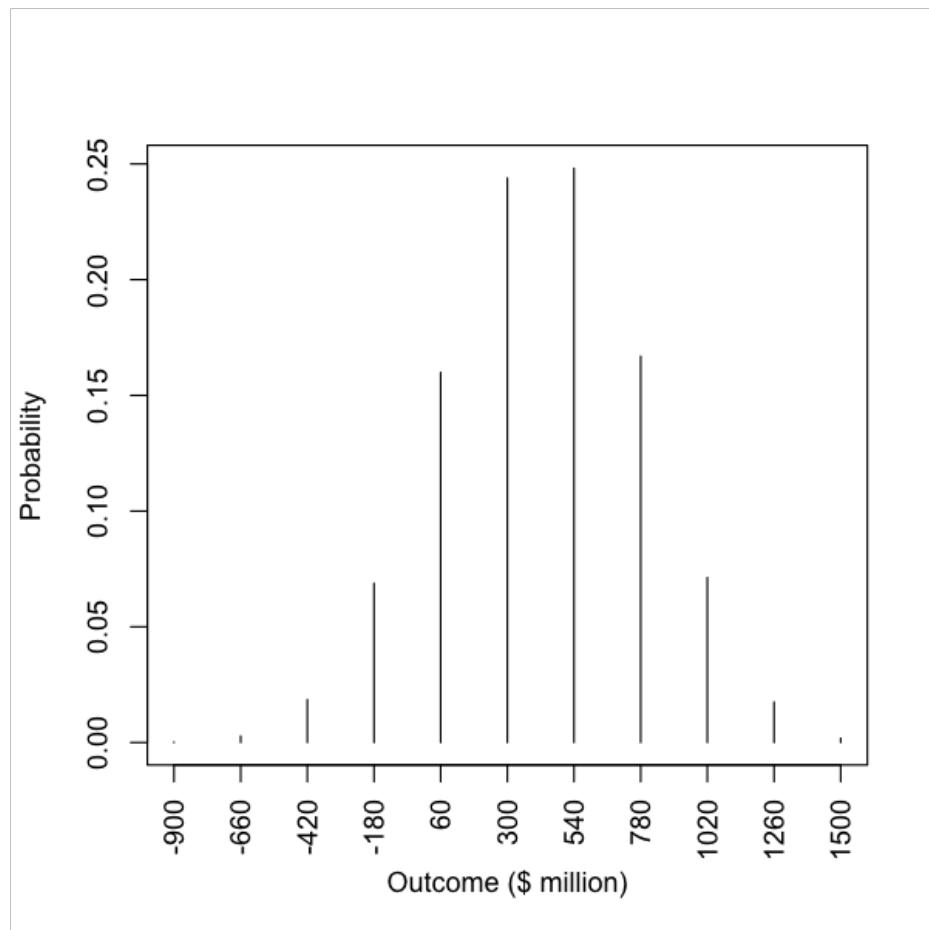
Outcome distribution decision Figure A.2 shows the outcome distribution display that participants saw in Experiment 1.

Follow-up gambles

Negative EV gambles I wanted 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, I presented participants two project decisions that had a negative EV. Out of the 396 negative EV gambles included (two per participant), all but four were rejected.

A. Chapter 2 appendix

Below is the distribution of final outcomes after the last 10 investments you just saw. That is, each possible outcome is shown on the x-axis, and the probability of each outcome is shown on the y-axis. Regardless of what you decided previously, would you invest in all 10 of those investments, given the below distribution?



*

Yes No

[Continue](#)

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

A. Chapter 2 appendix

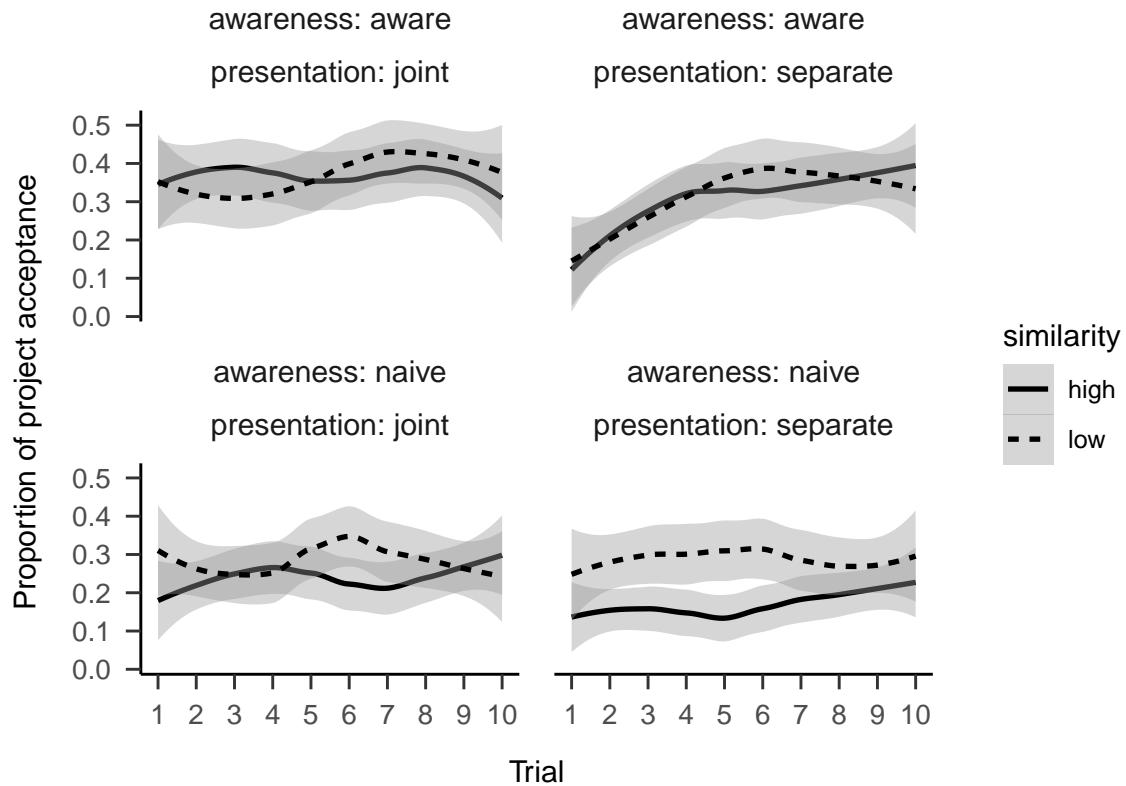


Figure A.3: Proportion of project acceptance by trial, similarity, awareness, and presentation conditions. LOESS is used for smoothing over trials, and the shading represents 95% confidence intervals.

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

Redelmeier and Tversky (1992) gambles I then showed participants 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

Trial-by-trial analysis

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

A. Chapter 2 appendix

In total, how many projects did you just see?

A screenshot of a survey interface. At the top, there is a question: "In total, how many projects did you just see?". Below the question is a numeric input field containing the number "10". To the right of the input field is the word "projects". Below the input field is a "Continue" button.

Figure A.4: Experiment 2 project number question.

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, $t(197) = 5.48, p < .001, d_z = 0.50, 95\% \text{ CI } [0.31, 0.68]$; and in the joint presentation phase, $t(197) = 4.17, p < .001, d_z = 0.37, 95\% \text{ CI } [0.19, 0.56]$.

However, I 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 I found an effect of distribution, it is 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

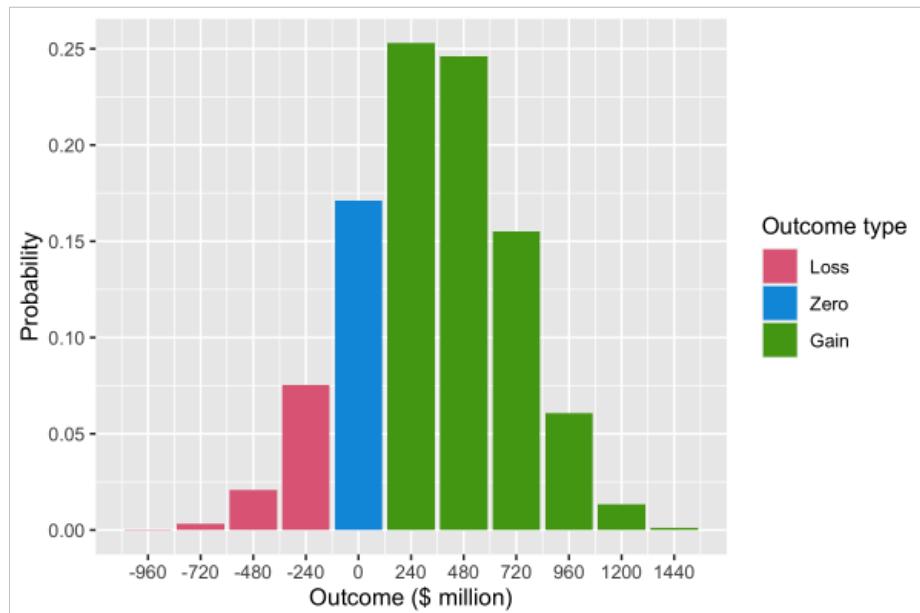
Materials

Follow-up Figure A.4 shows the project number question (maximum value 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 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. Chapter 2 appendix

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:

Consider all the projects you saw. If you had to choose between investing in all of them, or investing in none of them, which would you choose?

Figure A.5: Experiment 2 binary portfolio question.

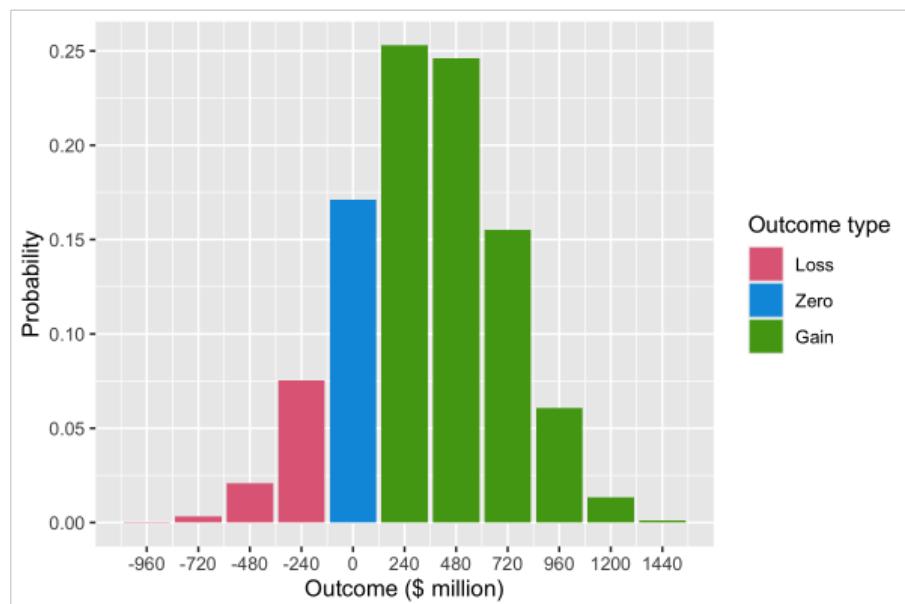
Procedure

Participants responded to demographic questions, 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.

A. Chapter 2 appendix

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:

The total number of projects you were shown is 10 . If you could choose to invest in a certain number of those 10 projects, how many would you invest in?

projects

Figure A.6: Experiment 2 numerical portfolio question.

A.2.2 Results

Follow-up

Project number We asked participants how many projects they think they saw.

Figure A.7 shows that overall people to correctly estimate the number of projects, with more accuracy for those in the aware condition.

A. Chapter 2 appendix

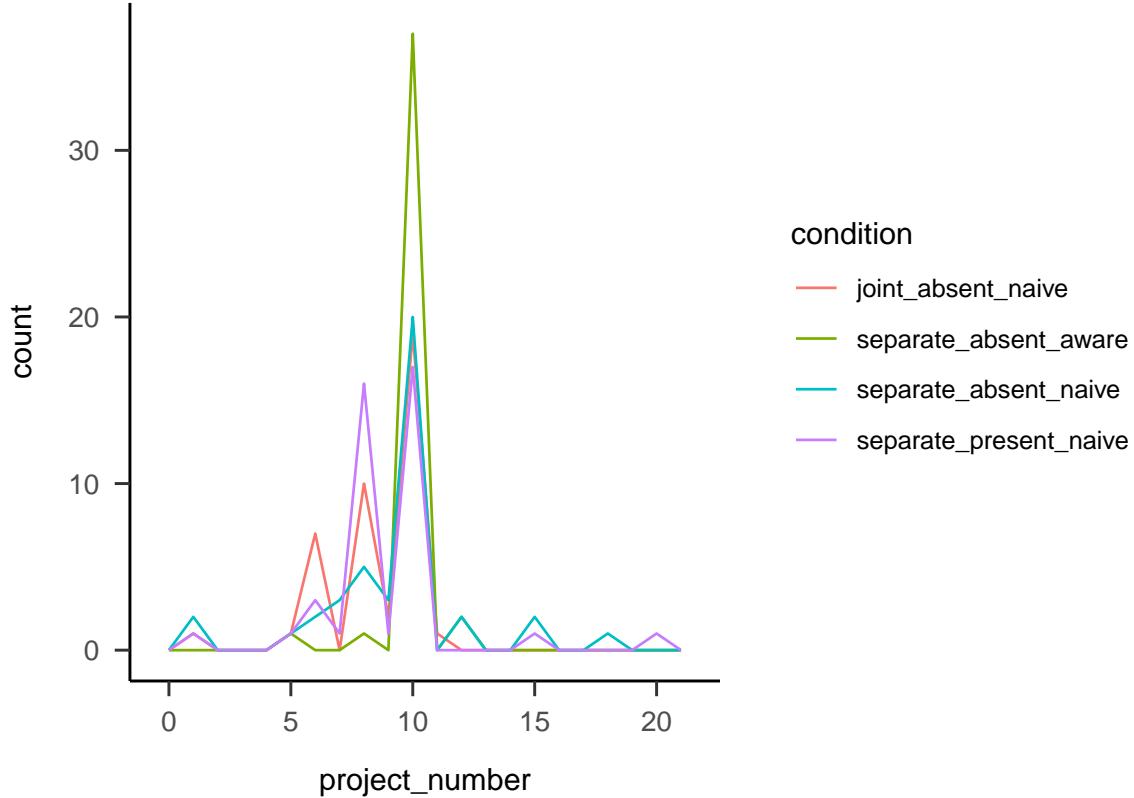


Figure A.7: Number of projects participants reported seeing, by condition.

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, $\hat{\beta} = -0.30$, 95% CI $[-1.19, 0.58]$, $z = -0.67$, $p = .500$. The awareness effect was also not significant, $\hat{\beta} = -0.60$, 95% CI $[-1.48, 0.28]$, $z = -1.33$, $p = .184$. However, those 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%), $OR = 4.33$, 95% CI $[1.76, 11.24]$, $z = 3.11$, $p = .002$.

Portfolio choice - number Subsequently, we asked participants 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, $d_s = 0.13$, 95% CI $[-0.31, 0.56]$, $t(80) = 0.57$, $p = .570$. However, those 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 $[0.15, 1.03]$, $t(81) = 2.70$, $p = .009$.

A. Chapter 2 appendix

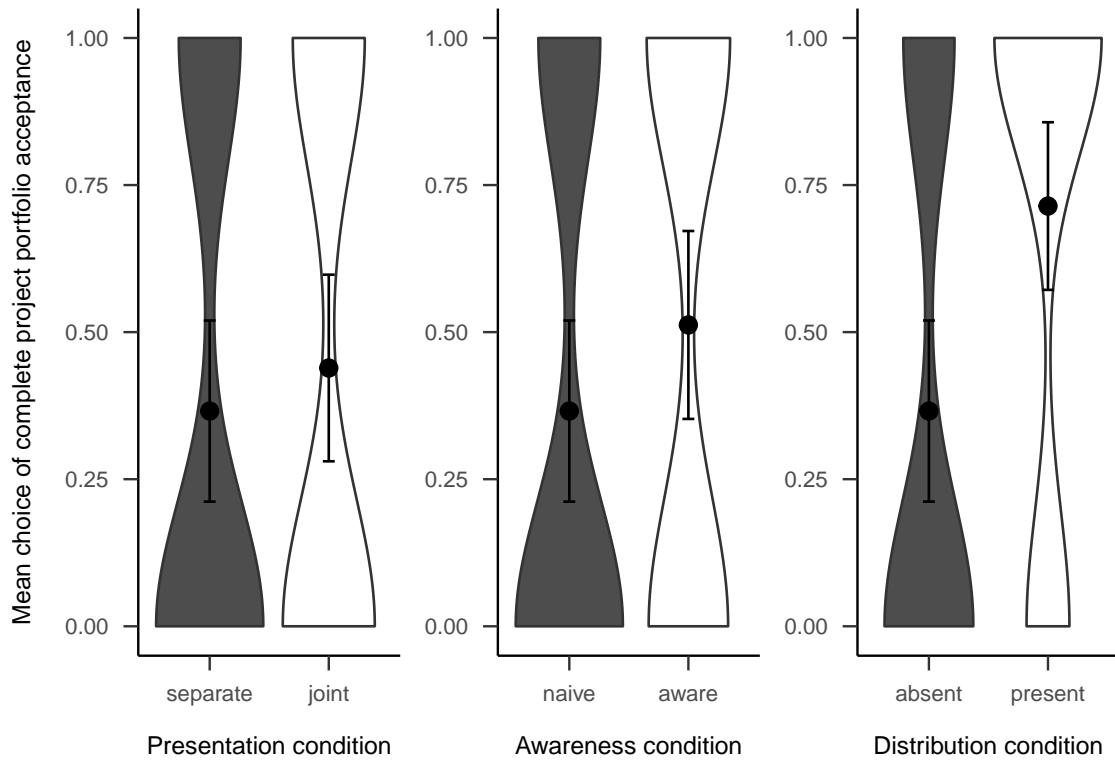


Figure A.8: Mean choice of investing in all 10 projects for the presentation, awareness, and distribution effects. Note, the condition on the left of each effect is the reference condition (separate presentation, naive awareness, distribution absent). As such, it is identical for the three effects.

Gambles

Figures A.10 and A.11 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 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. In Experiment 3, I 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. I again tested Hypothesis 2.3.

A. Chapter 2 appendix

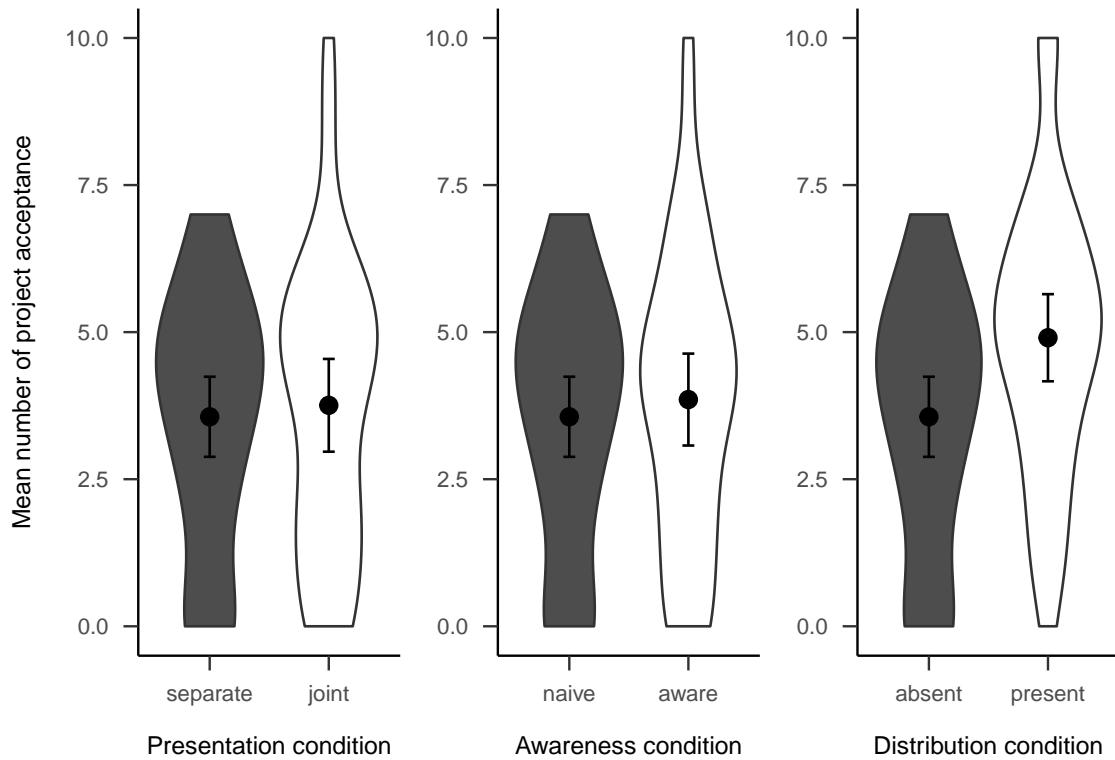


Figure A.9: Mean number of projects chosen in the follow-up for the presentation, awareness, and distribution effects. Note, the condition on the left of each effect is the reference condition (separate presentation, naive awareness, distribution absent). As such, it is identical for the three effects.

A.3.1 Method

Participants

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

Materials

Instructions Participants were shown the same instructions as in [Experiment 1](#).

A. Chapter 2 appendix

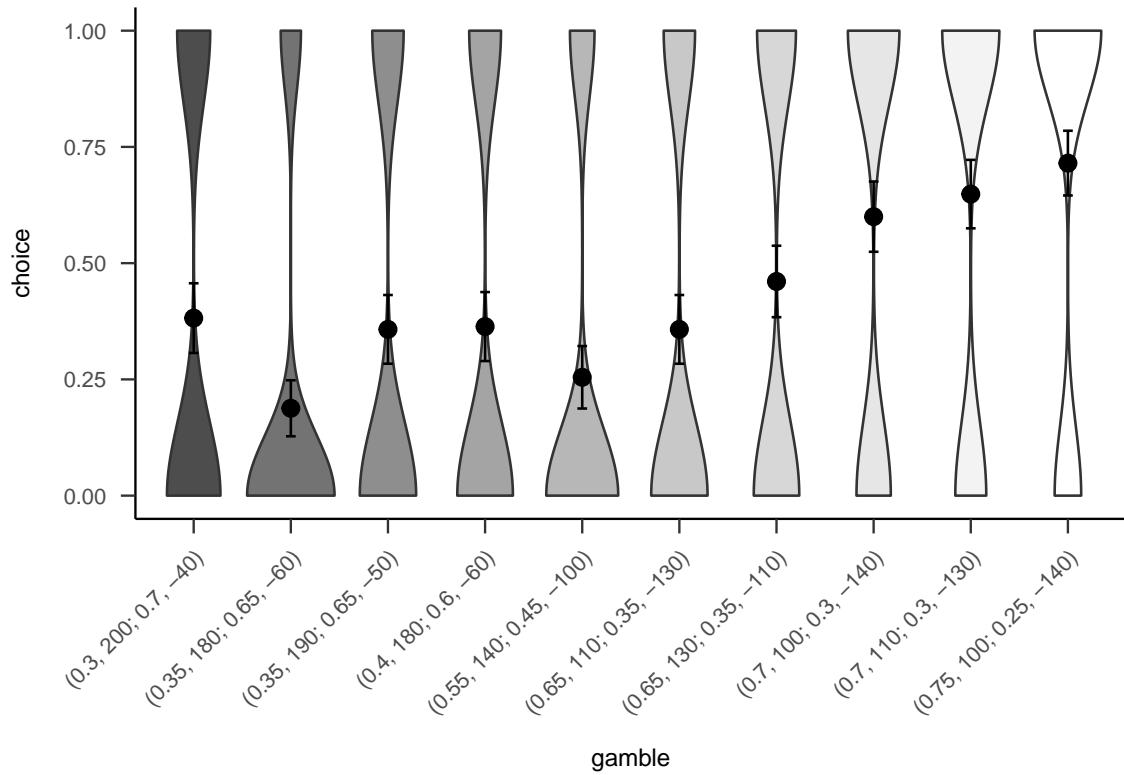


Figure A.10: Mean project acceptance for the 10 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value).

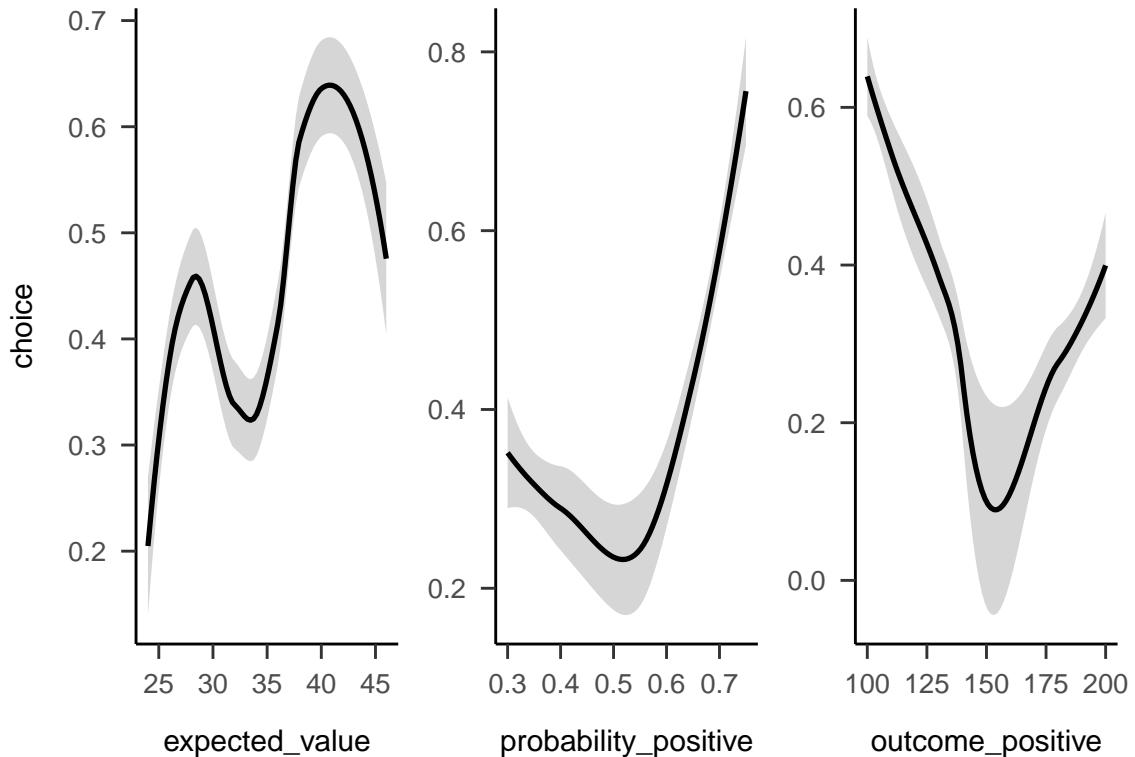


Figure A.11: Mean project acceptance for the gambles' expected value, positive probability, and positive outcome.

A. Chapter 2 appendix

Table A.1: Experiment 3 group allocation.

| Similarity | N |
|------------|-----|
| High | 133 |
| Low | 133 |
| Total | 266 |

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.

Risky investment task Participants saw displays with the same gamble values as those in [Experiment 2](#), but with some changes in wording and sentence structure. I kept the same gamble information, but added extra prose describing the projects and randomised the order of the sentences, 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, here I varied project domain so that in the high similarity condition participants saw one of the ten project domains.

A. Chapter 2 appendix

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

project(s)

Figure A.13: Experiment 3 project expectation question.

Follow-up The follow-up questions were similar to those in [Experiment 2](#), except in the portfolio number question I added the number of projects that they saw (10). Further, I added a question asking how many projects participants were expecting to see at the beginning of the experiment (see [Figure A.13](#)).

Procedure

Participants responded to demographic questions, 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

Project investment

The project investment data was analysed as in [Experiment 2](#). Figures [A.14](#) and [A.15](#) show the choice and proportion data, respectively. The difference between similarity conditions was not significant, both in the logistic regression $b = 0.01$, 95% CI $[-0.34, 0.36]$, $z = 0.04$, $p = .966$, 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.

Follow-up

Project expectation We asked participants 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. Chapter 2 appendix

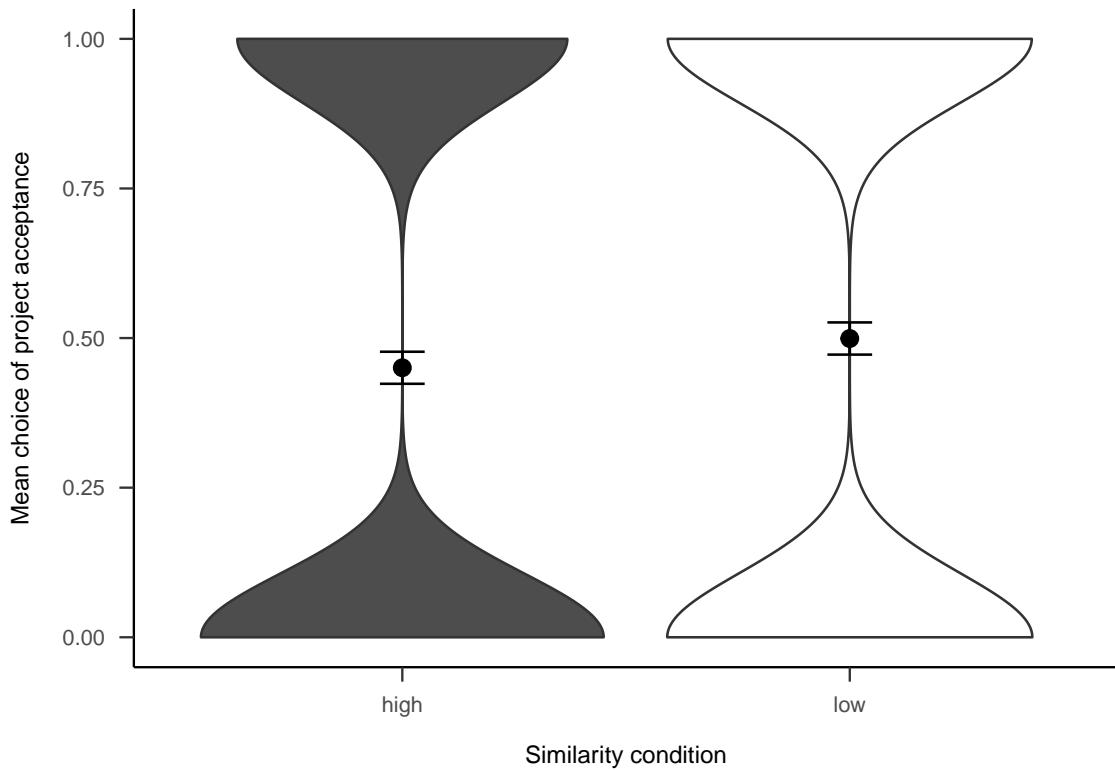


Figure A.14: Mean project acceptance for the similarity effect.

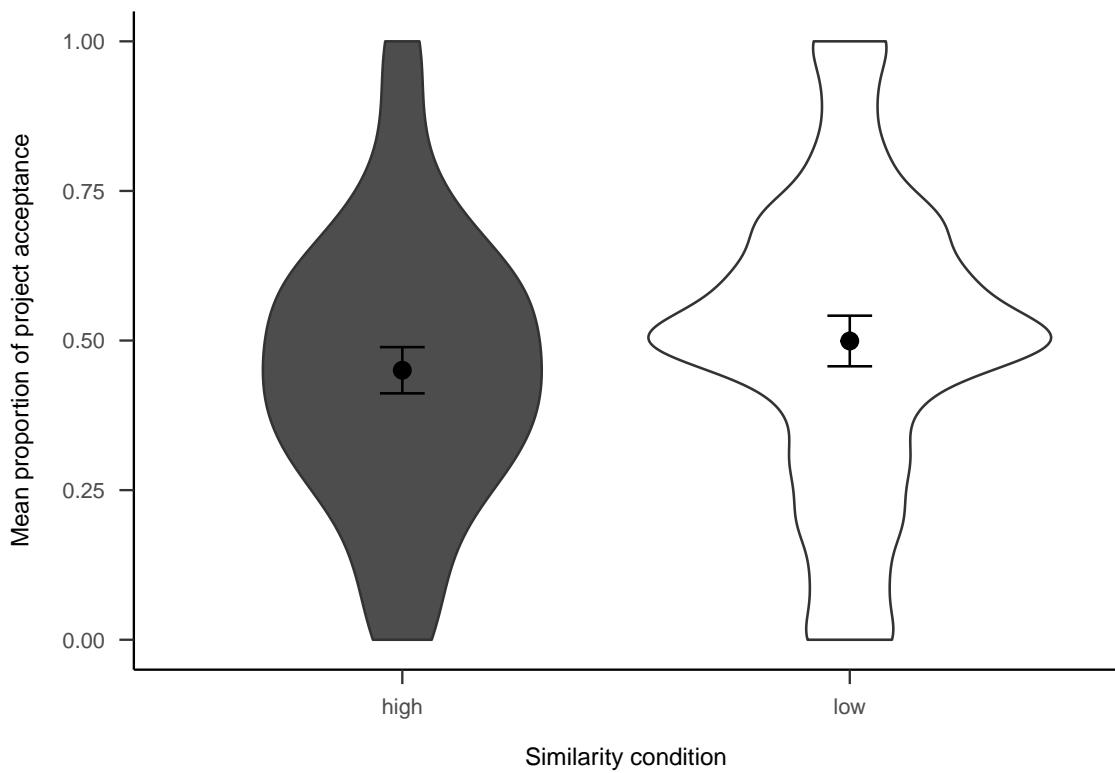


Figure A.15: Mean proportion of project acceptance for the similarity effect.

A. Chapter 2 appendix

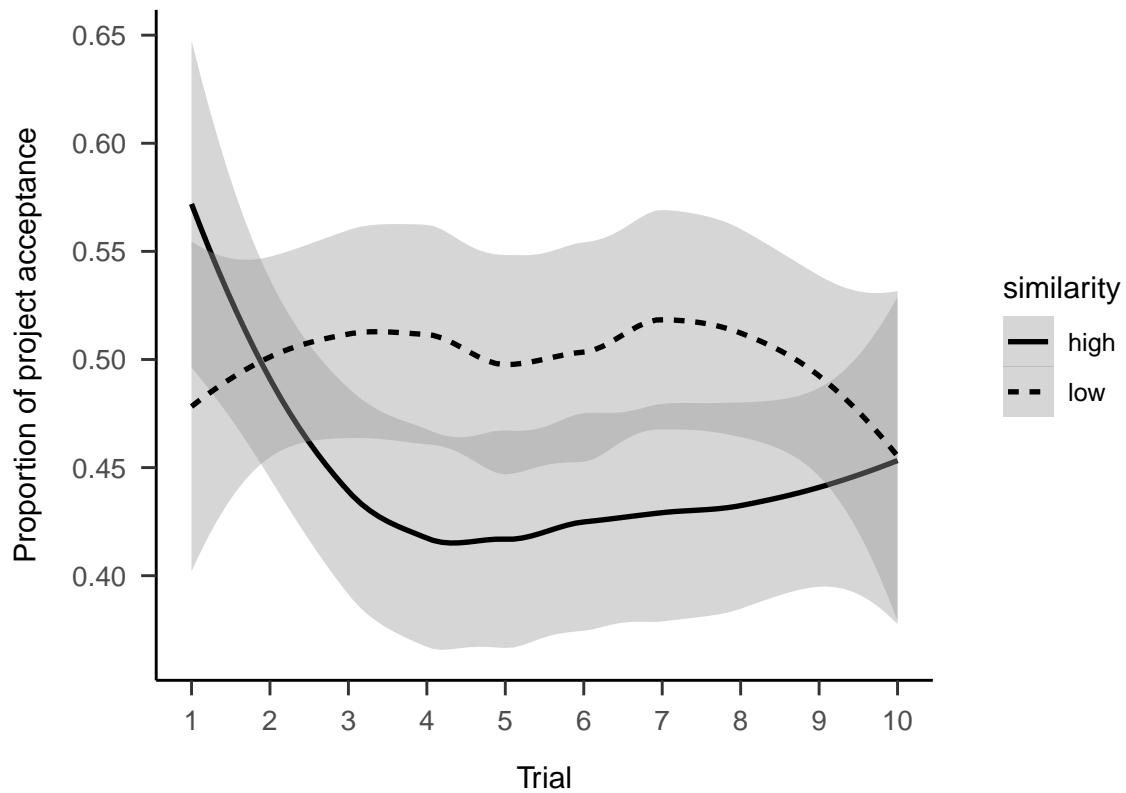


Figure A.16: Mean project acceptance by similarity and trial.

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

| Term | $\hat{\beta}$ | 95% CI | z | p |
|--------------------------------------|---------------|---------------|-------|------|
| Intercept | -0.02 | [-0.31, 0.28] | -0.11 | .916 |
| Similaritylow | 0.05 | [-0.37, 0.46] | 0.22 | .826 |
| Project order | -0.04 | [-0.08, 0.00] | -1.83 | .067 |
| Similaritylow \times Project order | 0.03 | [-0.03, 0.09] | 1.07 | .284 |

Project number We asked participants how many projects they think they saw.

Figure A.18 shows that overall people correctly estimate the number of projects.

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, $b = 0.52$, 95% CI [0.04, 1.02], $z = 2.10$, $p = .036$.

A. Chapter 2 appendix

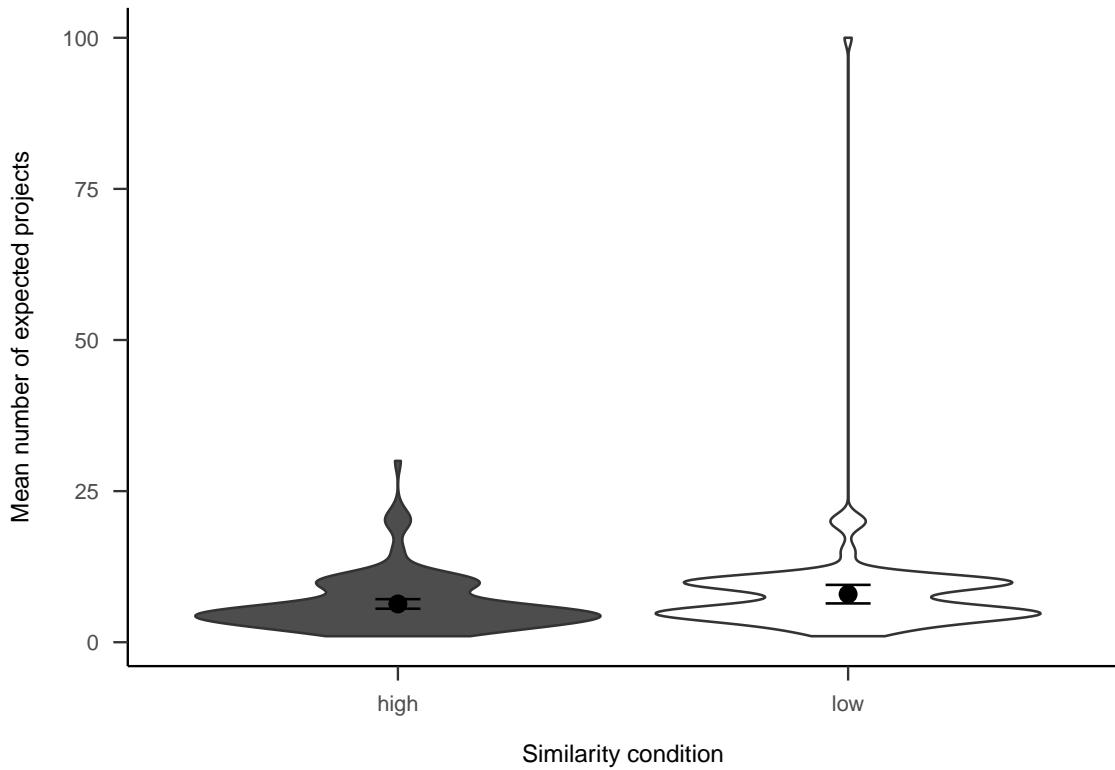


Figure A.17: Number of projects participants expected to see, by similarity.

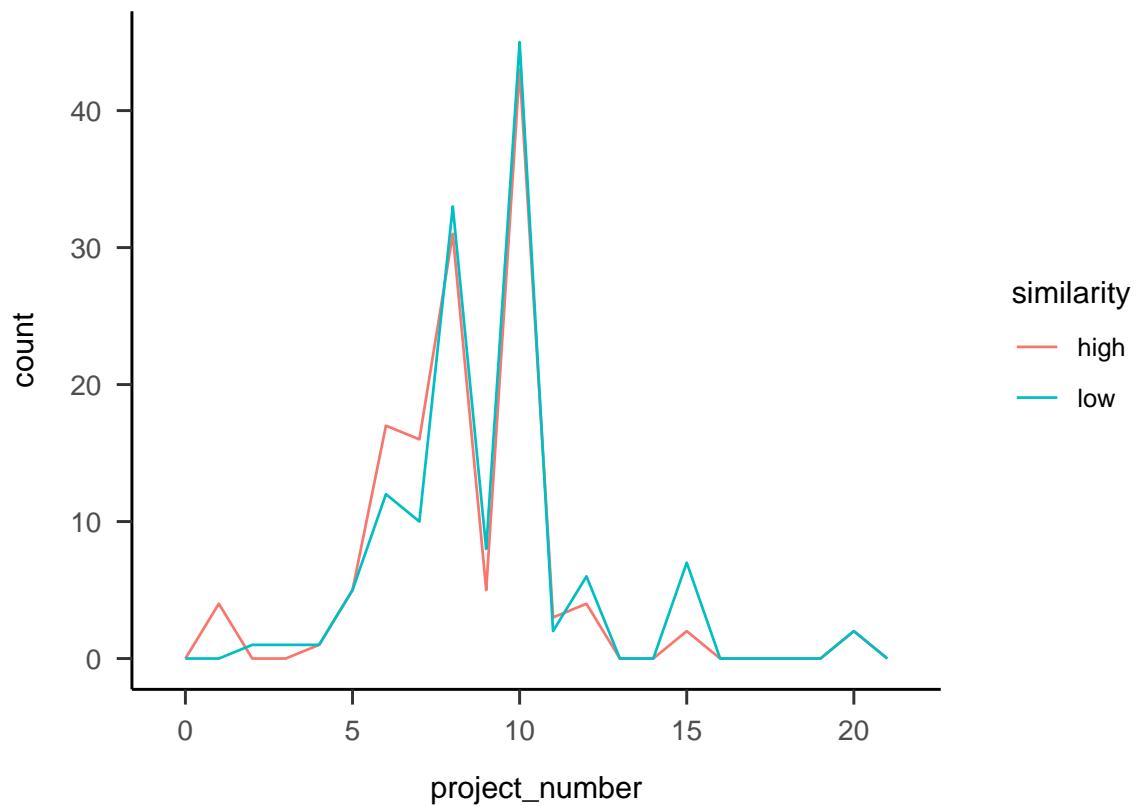


Figure A.18: Number of projects participants reported seeing, by similarity.

A. Chapter 2 appendix

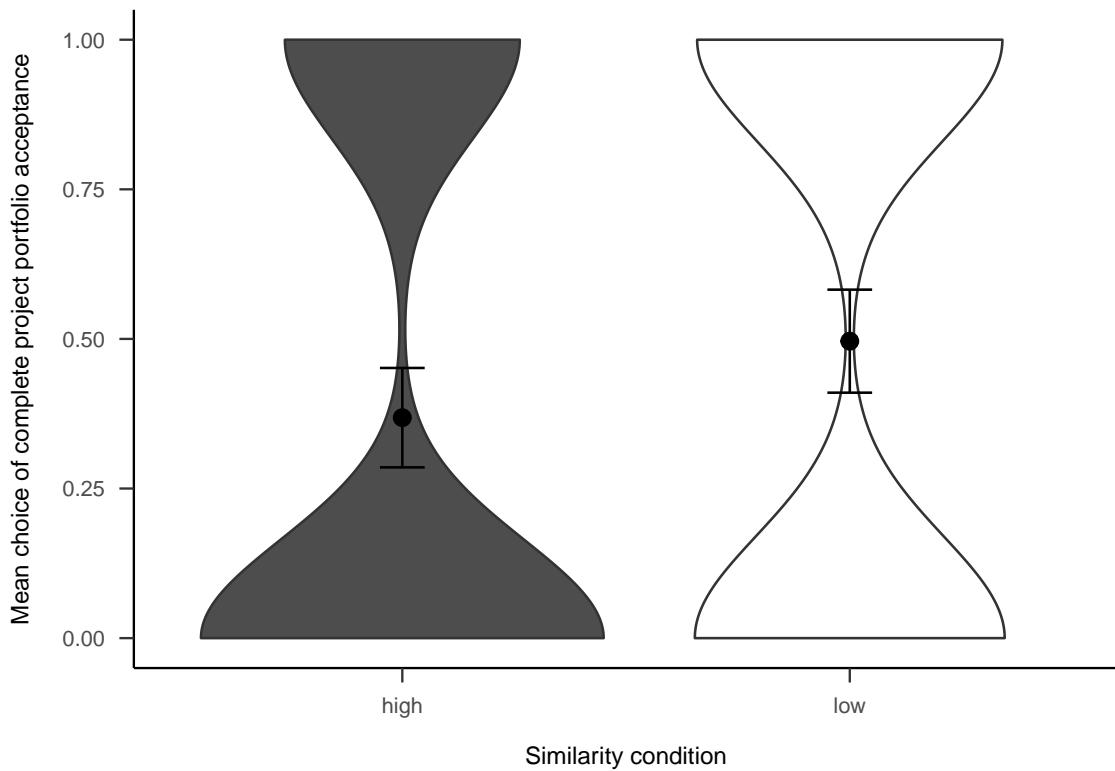


Figure A.19: Mean choice of investing in all 10 projects for the similarity effect.

Portfolio choice - number Subsequently, we asked participants how many projects they would invest in out of the 10 that they saw. As Figure A.20 shows, the difference between similarity conditions was not significant, $d_s = -0.14$, 95% CI [-0.38, 0.10], $t(264) = -1.12$, $p = .264$.

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

I found some evidence for the effect of similarity on project choice, but it was in the opposite direction to the one hypothesised. Specifically, I found that when considering projects individually, participants' risk aversion did not differ between

A. Chapter 2 appendix

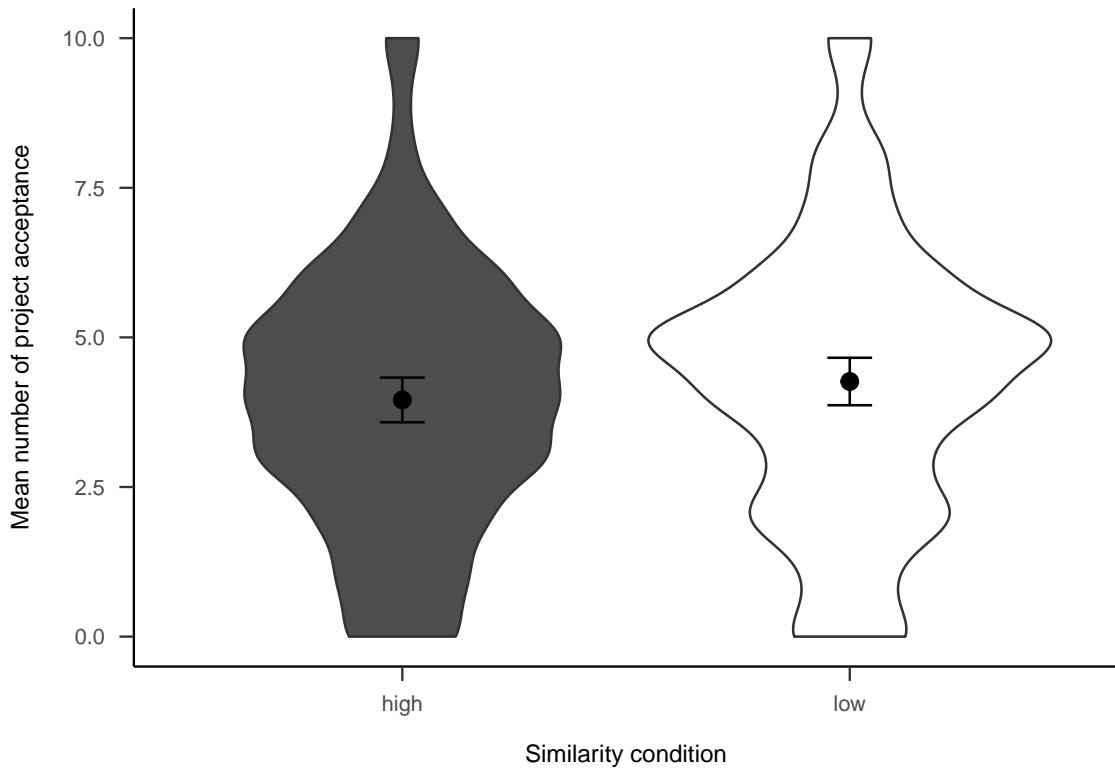


Figure A.20: Mean number of projects chosen in the follow-up for the similarity effect.

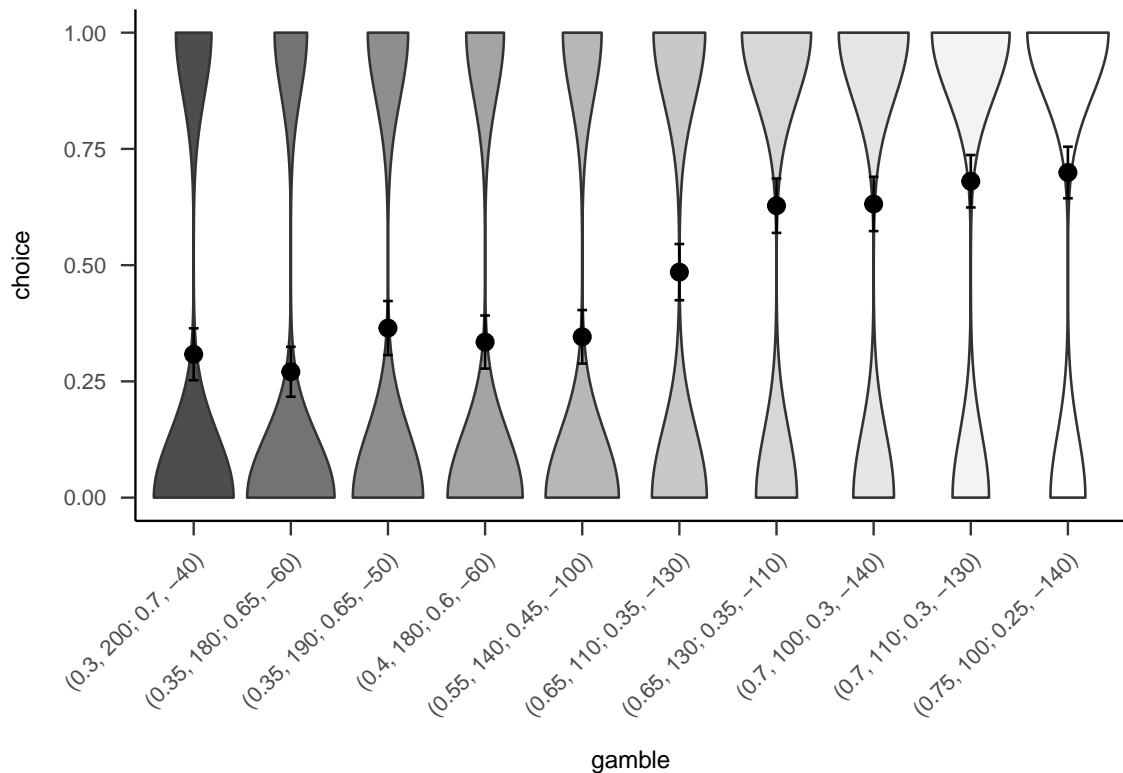


Figure A.21: Mean project acceptance for the 10 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value).

A. Chapter 2 appendix

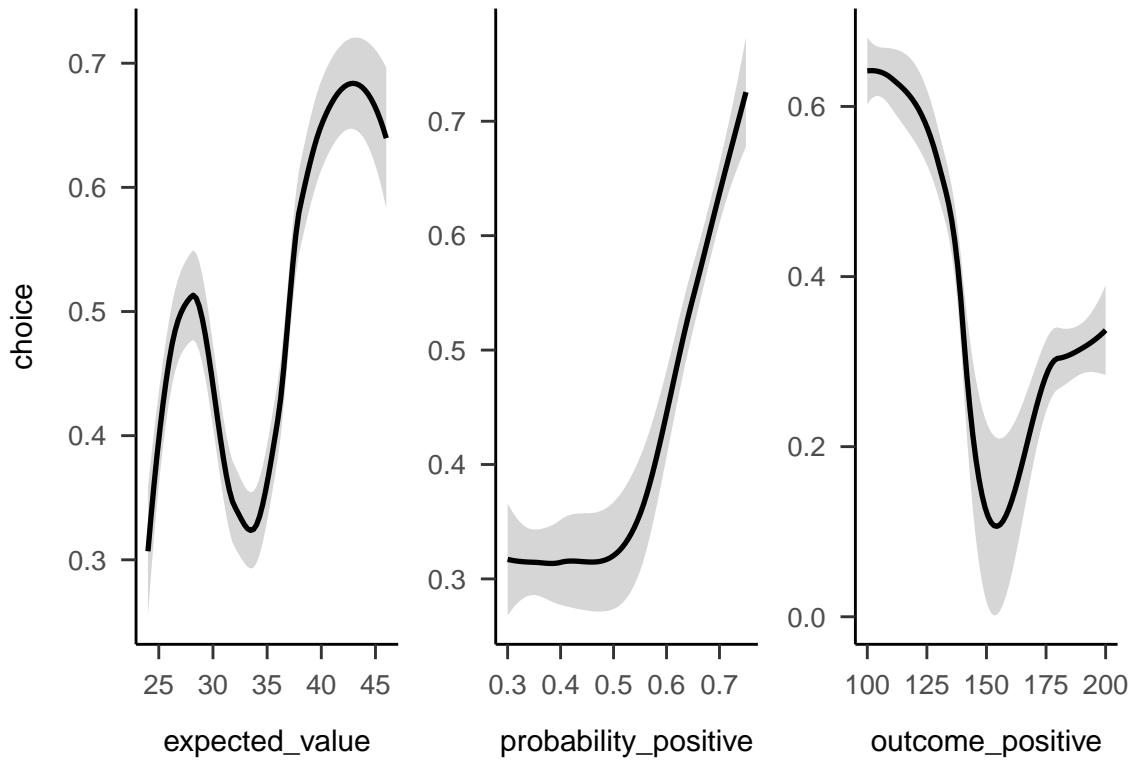


Figure A.22: Mean project acceptance for the gambles' expected value, positive probability, and positive outcome.

similarity conditions, but when offered a portfolio of the projects, those that saw the dissimilar projects were more likely to invest.

These results provide evidence for the naive diversification account expressed above. Specifically, it may be the case that participants really are naively diversifying, but only when they are explicitly given an opportunity to do so. This is similar to the multi-play effects in that the question itself provides a sort of choice bracketing. That is, the gambles are grouped together as a package/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. In Experiment 1 we found an effect of awareness in the trial-by-trial data that was not replicated

A. Chapter 2 appendix

Table A.3: Experiment 4 group allocation.

| Awareness | N |
|-----------|-----|
| Aware | 133 |
| Naive | 133 |
| Total | 266 |

in Experiment 2. Previously, I explained this effect using the law of small numbers; people may have been anticipating less risky gambles towards the end of the set. As such, it might be the case that the effect will be seen with more trials. In Experiment 4 we 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 say how many projects there were. The “aware” condition attempted to encourage participants to think of all 20 projects (by saying the total number in the beginning, and notifying participants where they were in the project order). I again tested Hypothesis 2.4.

A.4.1 Method

Participants

Two hundred and sixty-six (110 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 40.62 ($SD = 9.59$, $min = 25$, $max = 74$). Participants reported an average of 7.45 ($SD = 7.8$, $min = 0$, $max = 47$) years of work in a business setting, and an average of 5.52 ($SD = 7.27$, $min = 0$, $max = 48$) years of business education. The mean completion time was 12.66 ($SD = 8.26$, $min = 1.48$, $max = 53.47$) minutes. Table A.3 shows the between-subjects condition allocation.

Materials

Instructions Participants were shown similar instructions as in Experiment 1, except that the awareness manipulation was incorporated into the text. Participants

A. Chapter 2 appendix

Imagine that you are an executive in a large company composed of many individual businesses. You need to make decisions about projects that come across your desk.

**As the executive, your pay will be determined by the performance of each investment.
We want to know what choices you would actually make.**

[< Previous](#) [Next >](#)

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

Imagine that you are an executive in a large company composed of many individual businesses. You need to make decisions about projects that come across your desk.

**As the executive, your pay will be determined by the performance of your investments.
We want to know what choices you would actually make.**

There will be 20 projects that you will decide on this quarter.

[< Previous](#) [Next >](#)

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

in the naive condition saw the instructions in Figure A.23, and those in the aware condition saw the instructions in Figure A.24.

Risky investment task Participants saw similar displays to those in Experiment 3. However, here participants viewed 20 projects, so while the gamble constraints 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.

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

A. Chapter 2 appendix

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.

Procedure

Participants responded to demographic questions, 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

Project investment

The project investment data was analysed as in [Experiment 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 $b = 0.09$, 95% CI $[-0.25, 0.44]$, $z = 0.53$, $p = .595$, 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.

A. Chapter 2 appendix

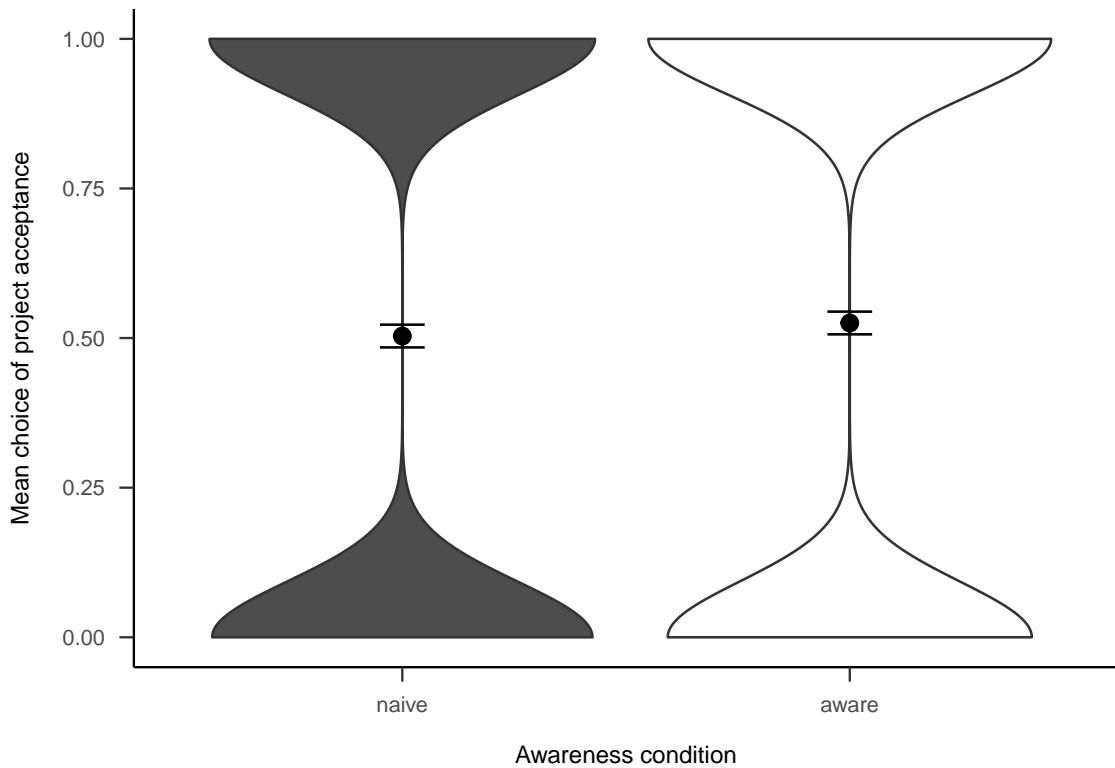


Figure A.26: Mean project acceptance for the awareness effect.

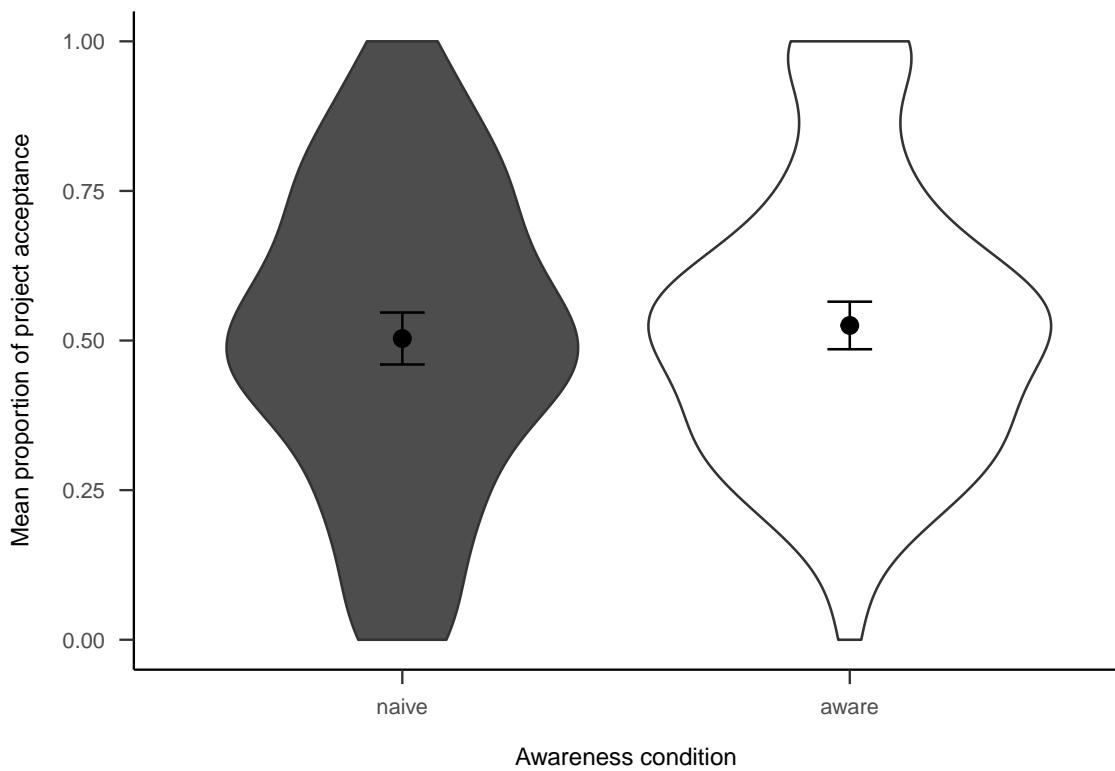


Figure A.27: Mean proportion of project acceptance for the awareness effect.

A. Chapter 2 appendix

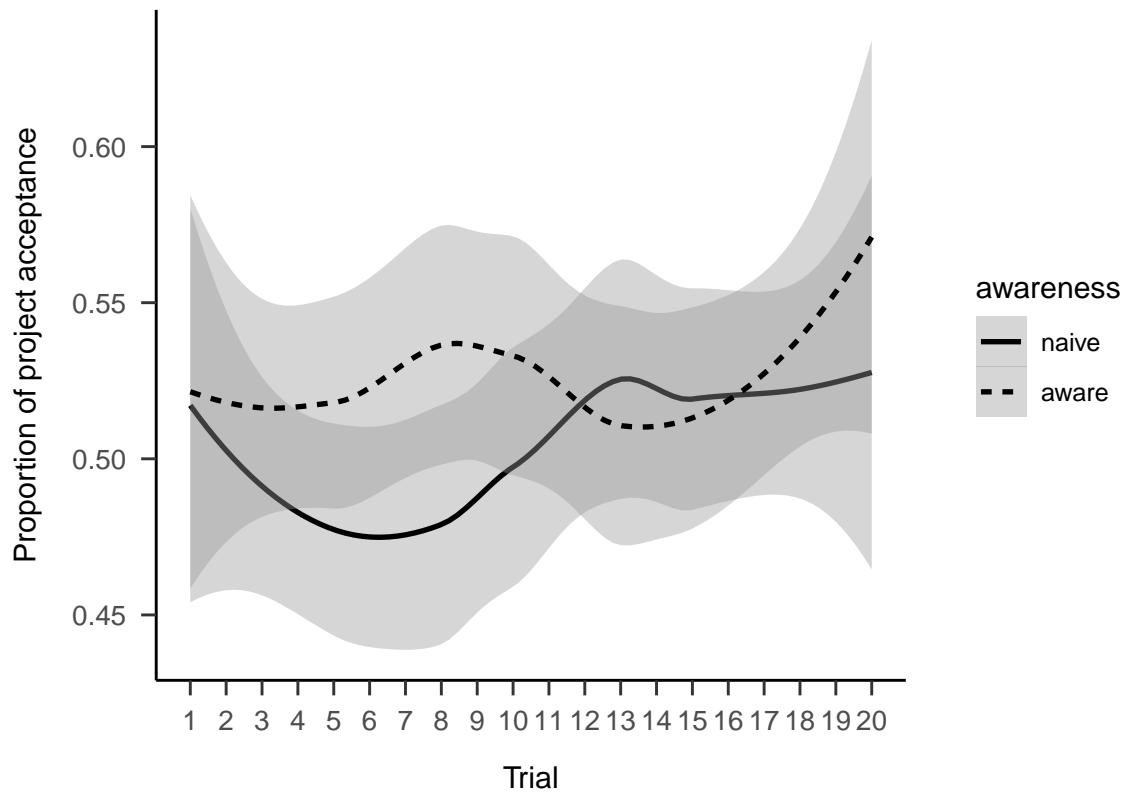


Figure A.28: Mean project acceptance by awareness and trial.

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

| Term | $\hat{\beta}$ | 95% CI | z | p |
|---------------------------------------|---------------|---------------|-------|------|
| Intercept | -0.11 | [-0.37, 0.15] | -0.82 | .410 |
| Awarenessaware | 0.20 | [-0.17, 0.57] | 1.05 | .293 |
| Project order | 0.01 | [0.00, 0.03] | 1.37 | .170 |
| Awarenessaware \times Project order | 0.00 | [-0.03, 0.02] | -0.29 | .775 |

Follow-up

Project expectation We asked participants 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.

Project number We asked participants how many projects they think they saw. Figure A.30 shows that overall people correctly estimate the number of projects,

A. Chapter 2 appendix

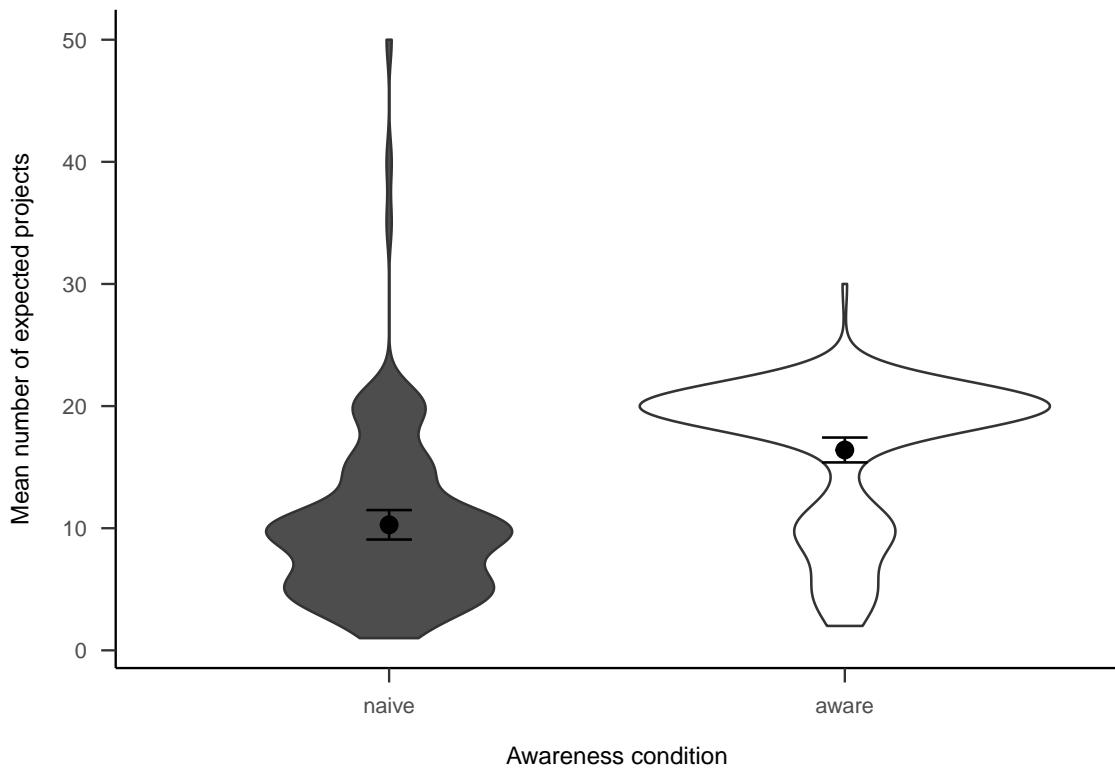


Figure A.29: Number of projects participants expected to see, by awareness.

with higher accuracy for those in the aware condition.

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, $b = 0.18$, 95% CI $[-0.30, 0.67]$, $z = 0.74$, $p = .460$.

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 $[-0.36, 0.12]$, $t(264) = -0.97$, $p = .334$.

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. Chapter 2 appendix

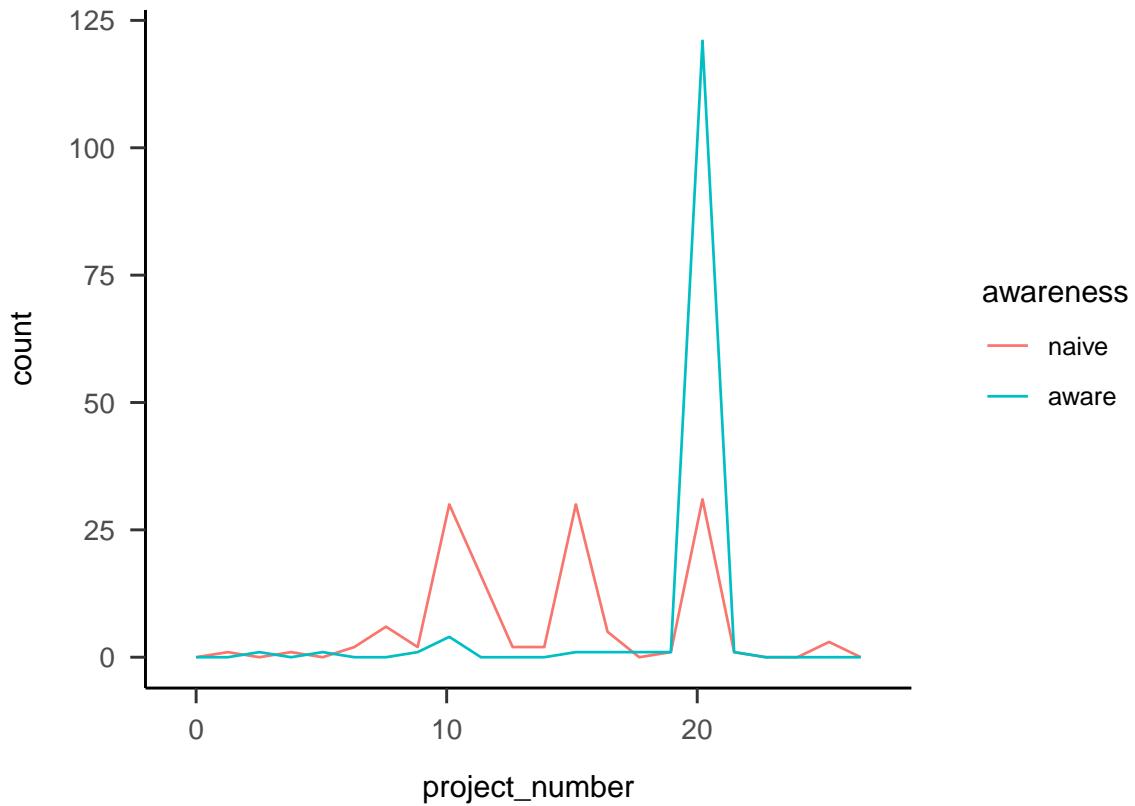


Figure A.30: Number of projects participants reported seeing, by awareness.

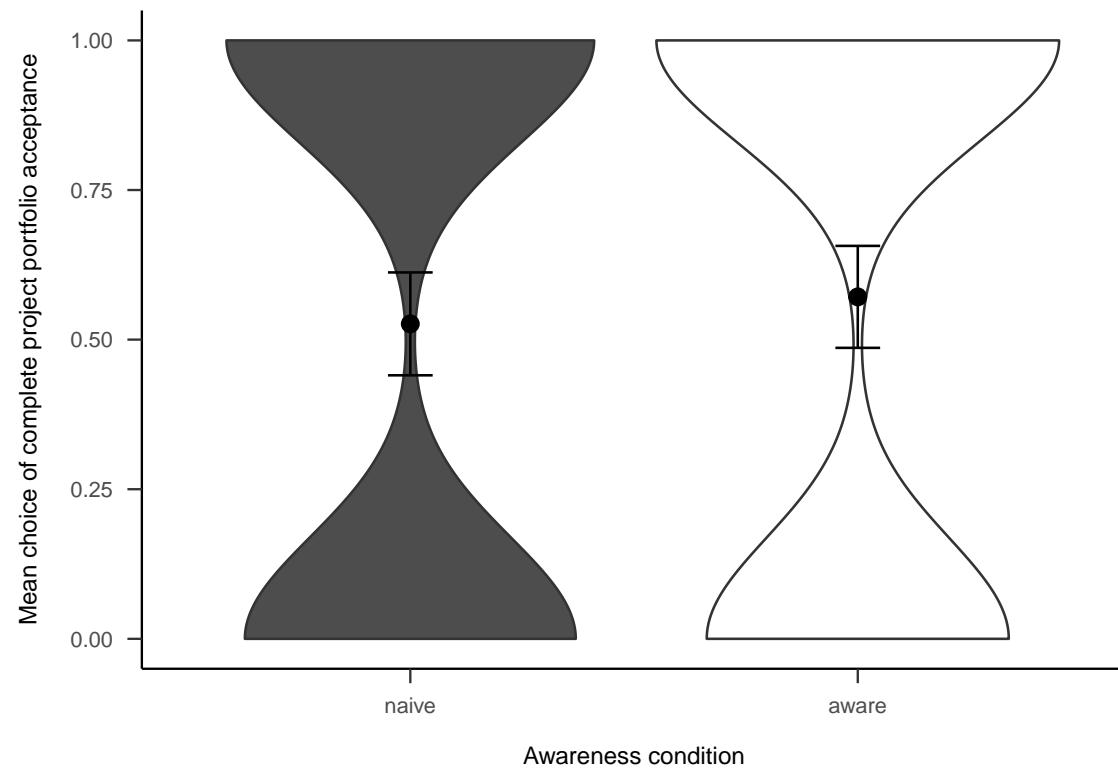


Figure A.31: Mean choice of investing in all 20 projects for the awareness effect.

A. Chapter 2 appendix

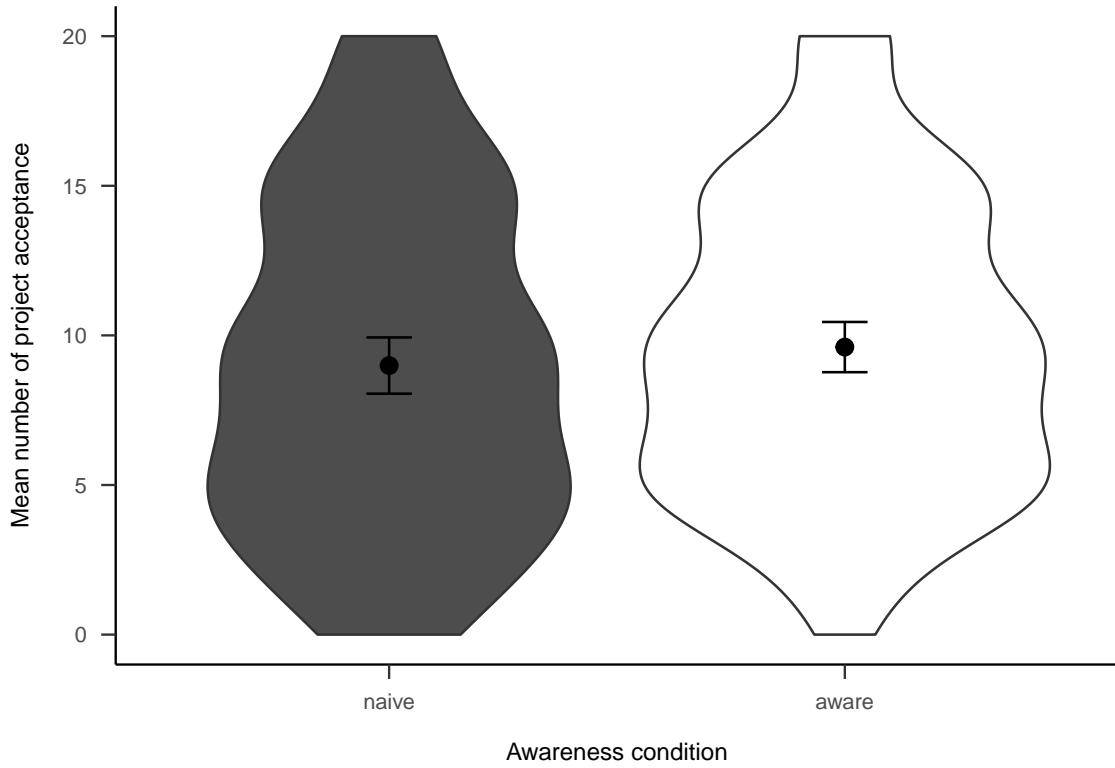


Figure A.32: Mean number of projects chosen in the follow-up for the awareness effect.

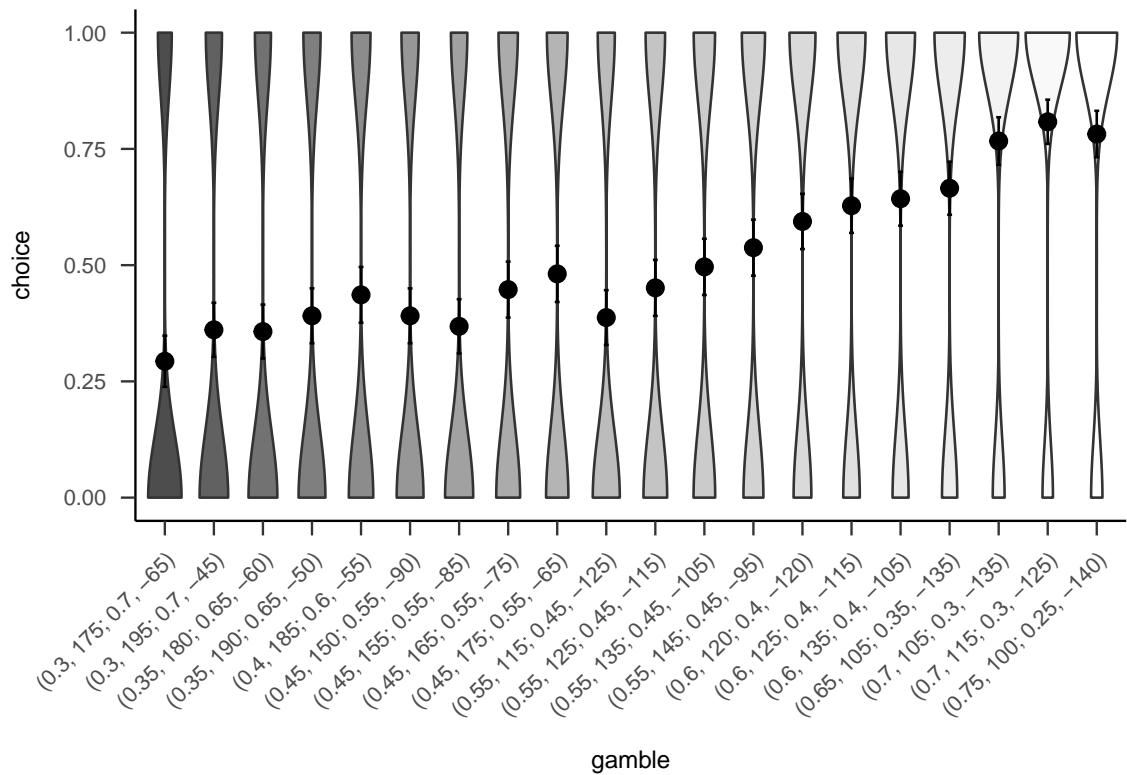


Figure A.33: Mean project acceptance for the 20 gambles. The format of the labels indicate: (gain probability, gain value; loss probability, loss value).

A. Chapter 2 appendix

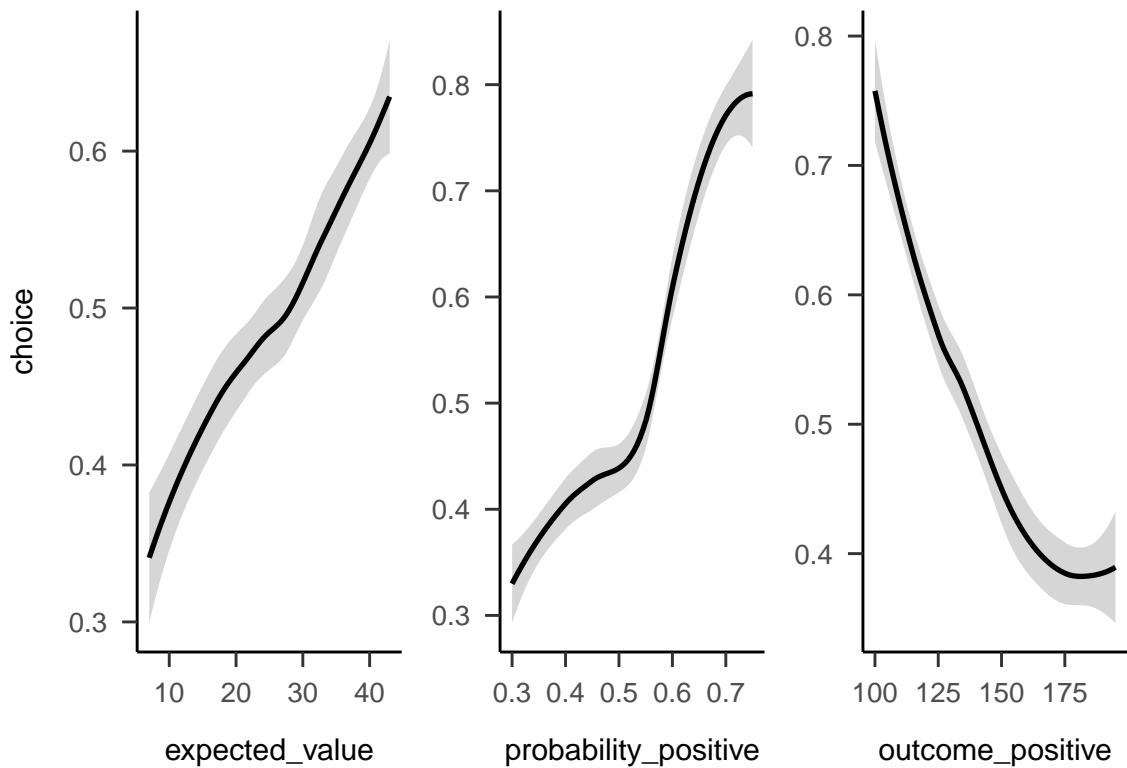


Figure A.34: Mean project acceptance for the gambles' expected value, positive probability, and positive outcome.

A.4.3 Discussion

I did not find evidence for my hypothesis. There was no significant effect of awareness on project choice by trial. I expected to see participants in the aware condition 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.

B

Chapter 4 appendix

Contents

| | |
|-----------------------------------|------------|
| B.1 Experiment 1 | 175 |
| B.1.1 Method | 176 |
| B.1.2 Results | 176 |
| B.1.3 Discussion | 184 |
| B.2 Experiment 2 | 185 |
| B.2.1 Method | 185 |
| B.2.2 Results | 195 |
| B.3 Experiment 3 | 197 |
| B.3.1 Method | 197 |
| B.3.2 Results | 200 |
| B.4 Experiment 4 | 201 |
| B.4.1 Method | 203 |
| B.4.2 Results | 204 |
| B.4.3 Discussion | 207 |
| B.5 Experiment 5 | 208 |
| B.5.1 Method | 208 |
| B.5.2 Results | 209 |
| B.5.3 Discussion | 210 |
| B.6 Experiment 6 | 212 |
| B.6.1 Method | 212 |
| B.6.2 Results | 213 |
| B.6.3 Discussion | 214 |
| B.7 Experiment 7 | 215 |
| B.7.1 Method | 215 |
| B.7.2 Results | 217 |
| B.7.3 Discussion | 217 |
| B.8 Experiment 8 | 218 |

B. Chapter 4 appendix

| | | |
|-------|----------------------|-----|
| B.8.1 | Method | 219 |
| B.8.2 | Results | 220 |
| B.8.3 | Discussion | 225 |

This appendix contains supplementary materials and analyses for the three experiments reported in Chapter 4. In addition, I report five related experiments. 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 with 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, I also asked participants to rank the projects and forecast their future returns. I included the ranking task before the allocation task in order to encourage alignment and to have another measure of participants' decision-making. I added the forecasting task (described further [below](#)) 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, I expect that more alignable differences therefore bring about more certainty about forecasting decisions, since participants will have more easily comparable information. As such, people's forecasting should be less

B. Chapter 4 appendix

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

Materials

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.

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 me to calculate each participant's forecasting mean and standard deviation (the latter as inversely proportional to forecasting precision).

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

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").

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

B.1.2 Results

Ranking

A mixed factorial ANOVA was conducted to investigate the effects of alignment and verbally-instructed NPV reliability on participants' rankings of the target

B. Chapter 4 appendix

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. Border added for clarity.

project. As seen in Figure B.8, the alignment \times reliability amount \times NPV amount interaction was significant, $F(6.62, 370.54) = 2.70$, $p = .011$, $\eta_p^2 = .046$. 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 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

B. Chapter 4 appendix

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. Border added for clarity.

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 $[-5.54, 2.13]$, $t(112) = -0.88$, $p = .380$.

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 seen in Figure B.9, the alignment \times reliability amount \times NPV amount

B. Chapter 4 appendix

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. Border added for clarity.

Imagine that you have 100 points to assign to the following options for Project 1's rate of return on investment after one month. Assign points according to how likely you think each rate of return is.

| | 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |
|---------------|---|----|----|----|----|----|----|----|----|----|----------|
| more than 9% | | | | | | | | | | | 0 |
| 7% to 9% | | | | | | | | | | | 0 |
| 5% to 7% | | | | | | | | | | | 0 |
| 3% to 5% | | | | | | | | | | | 0 |
| 1% to 3% | | | | | | | | | | | 0 |
| -1% to -3% | | | | | | | | | | | 0 |
| -5% to -7% | | | | | | | | | | | 0 |
| -7% to -9% | | | | | | | | | | | 0 |
| less than -9% | | | | | | | | | | | 0 |
| Total: | | | | | | | | | | | 0 |

Figure B.4: The forecasting task.

B. Chapter 4 appendix

Rank the projects in order of investment priority (drag and drop).

Remember that you're trying to enter a high-quality market.

- 1 PROJECT 1
- 2 PROJECT 2
- 3 PROJECT 3
- 4 PROJECT 4
- 5 PROJECT 5

Figure B.5: The ranking task.

How confident are you in each of your decisions?



Figure B.6: The confidence task.

B. Chapter 4 appendix

Justify your decision

Figure B.7: The justification task.

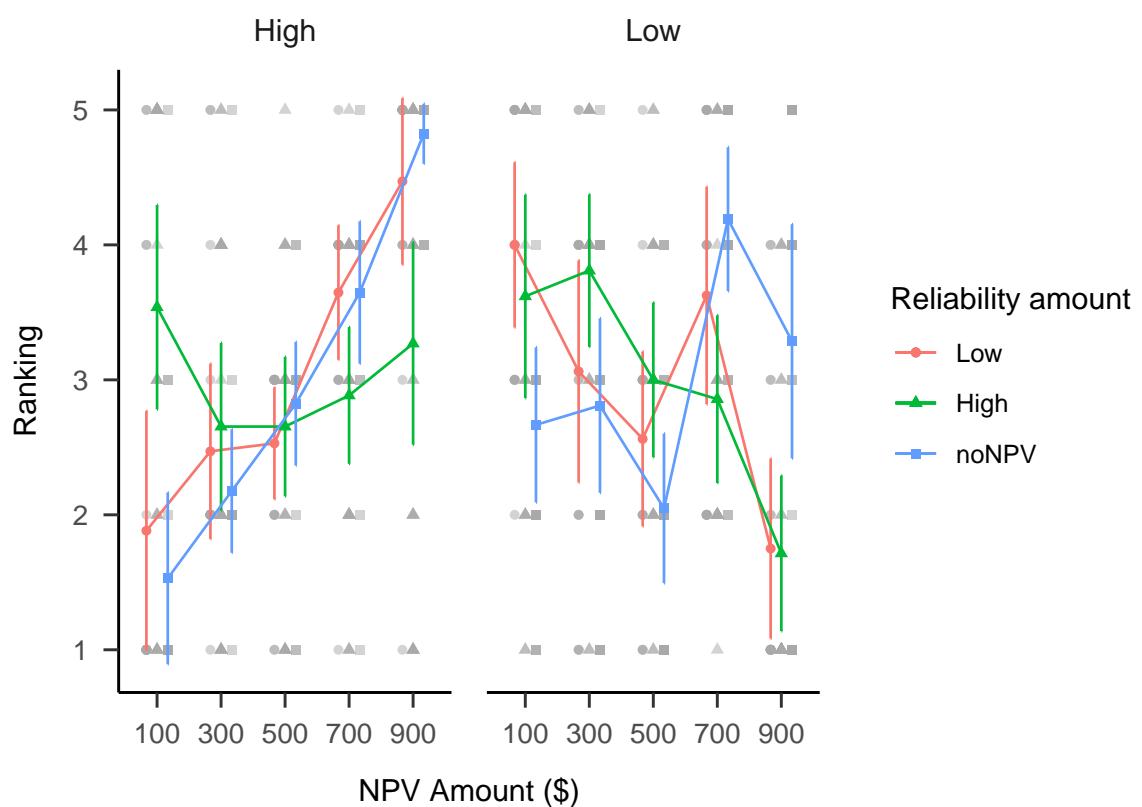


Figure B.8: Mean ranking.

B. Chapter 4 appendix

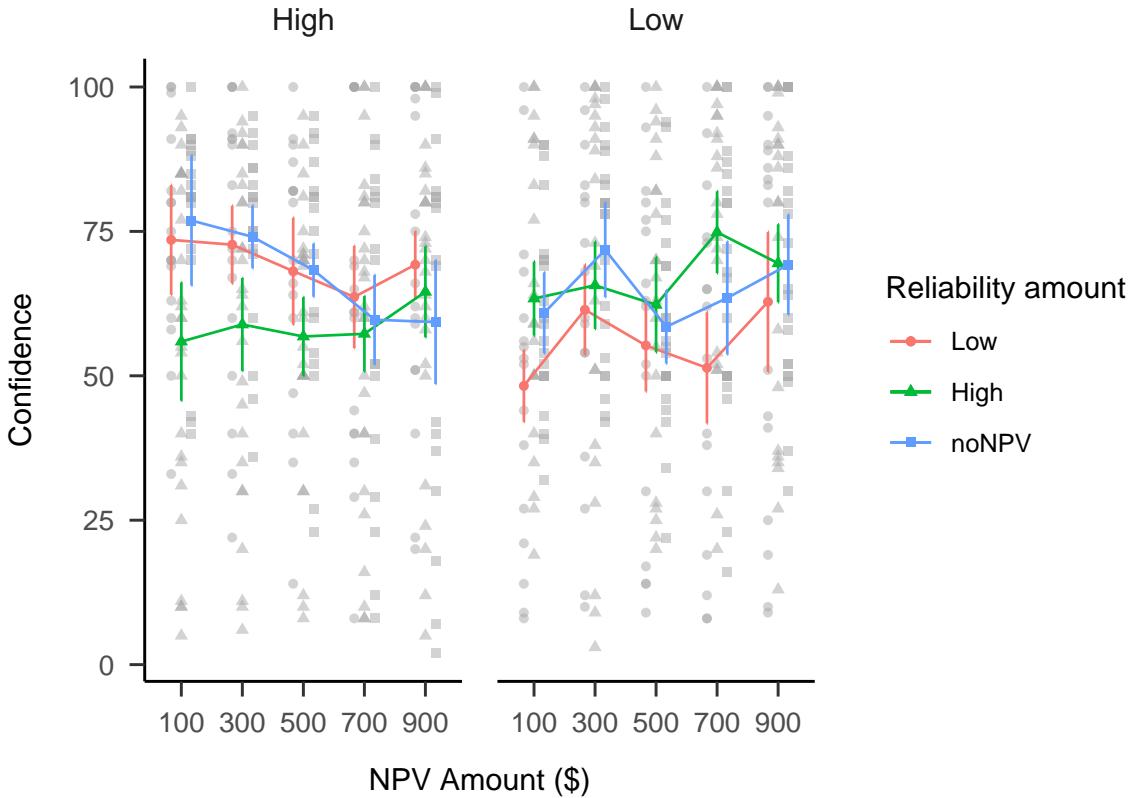


Figure B.9: Mean confidence.

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, $M = 10.73, 95\% \text{ CI } [-30.15, 51.61], t(112) = 0.52, p = .604$, nor the high reliability condition, $M = 13.05, 95\% \text{ CI } [-24.97, 51.07], t(112) = 0.68, p = .498$. 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, $M = 65.14, 95\% \text{ CI } [26.72, 103.57], t(112) = 3.36, p = .001$, and not the low reliability condition, $M = 31.88, 95\% \text{ CI } [-10.38, 74.14], t(112) = 1.49, p = .138$.

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 \times reliability amount \times NPV amount interaction was not significant, $F(5.26, 142.10) = 1.89, p = .095, \hat{\eta}_p^2 = .066$. However, the alignment \times

B. Chapter 4 appendix

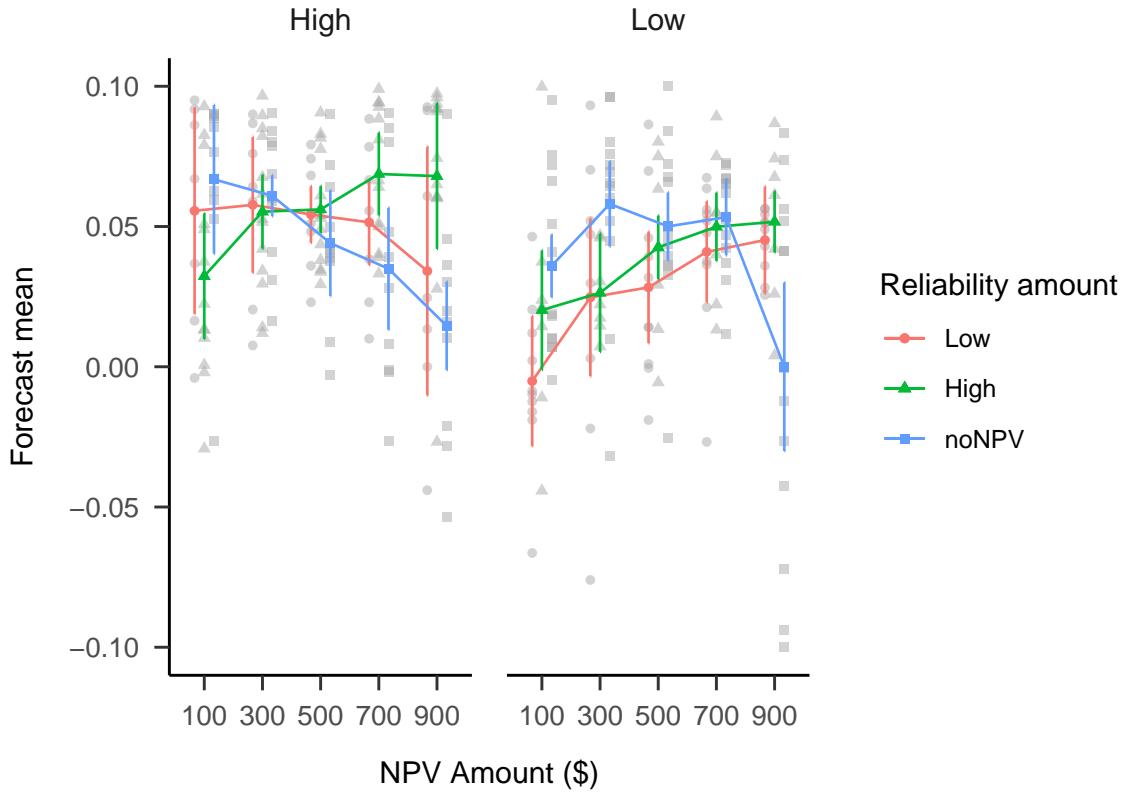


Figure B.10: Mean forecasts.

NPV amount interaction was significant, $F(2.63, 142.10) = 2.89, p = .044, \hat{\eta}_p^2 = .051$; as well as the reliability amount \times NPV amount 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, $M = 0.19, 95\% \text{ CI } [0.09, 0.30], t(54) = 3.63, p = .001$, and the high reliability condition, $M = 0.16, 95\% \text{ CI } [0.04, 0.28], t(54) = 2.75, p = .008$. However, in the high alignment condition, the linear NPV trend was only significantly lower in the no NPV condition than the high reliability condition, $M = 0.22, 95\% \text{ CI } [0.11, 0.32], t(54) = 4.04, p < .001$, and not the low reliability condition, $M = 0.08, 95\% \text{ CI } [-0.04, 0.21], t(54) = 1.30, p = .198$.

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 \times reliability amount \times NPV amount interaction was

B. Chapter 4 appendix

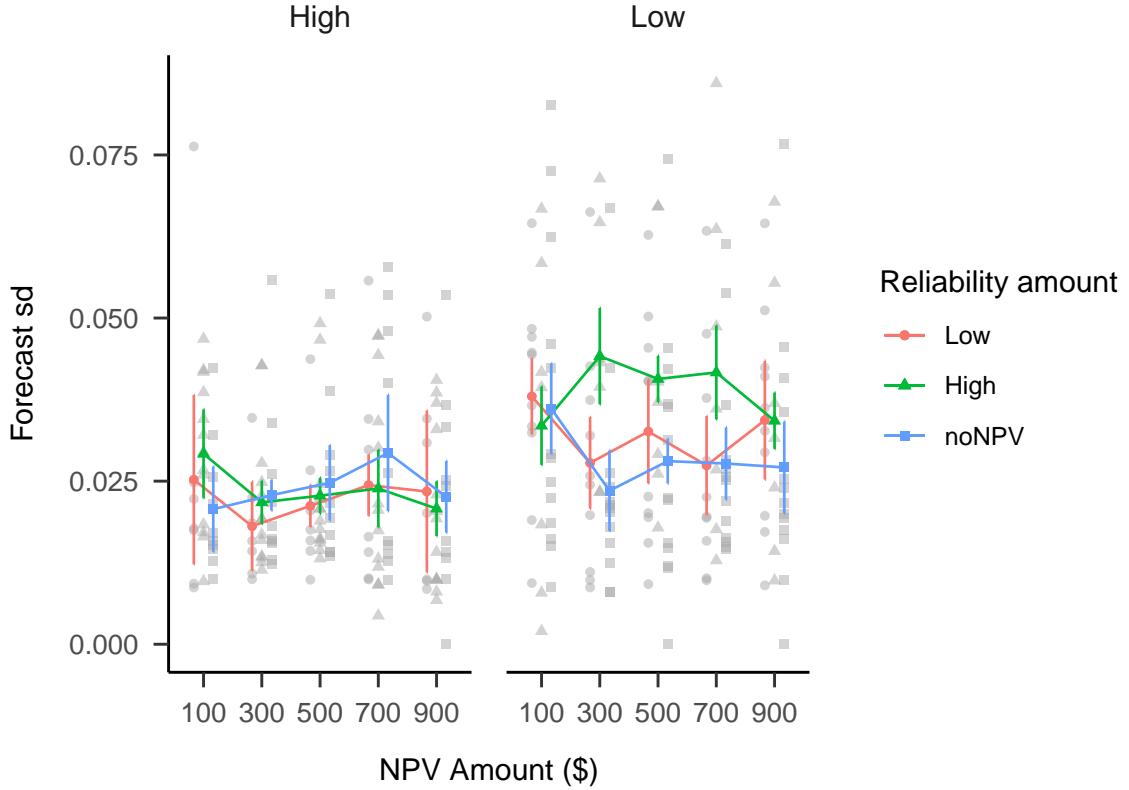


Figure B.11: Mean forecast SD.

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 I did not find 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' 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

B. Chapter 4 appendix

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. Border added for clarity.

reliance on NPV in this situation. There was limited evidence for the effect of alignment on forecast variability. As such, a future experiment will attempt to replicate this result with more participants.

B.2 Experiment 2

B.2.1 Method

Materials

Instructions Figure B.12 shows the instructions.

NPV test Participants were given more extensive information about NPV than in the previous experiment and were tested on their ability to calculate simple

B. Chapter 4 appendix

averages from given numerical ranges, as seen in Figures B.13 and B.14.

NPV knowledge ratings I used a similar design to Long et al. (2018, Study 1) to test whether this sample may be overconfident in their understanding on NPV. Therefore, I asked participants to rate their knowledge of NPV in various points in the study (see the [procedure](#)).

Variance lecture

NPV variance

B. Chapter 4 appendix

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:

$$NPV = \frac{\text{Cash inflow for year 1}}{(1 + \text{discount rate})^1} + \frac{\text{Cash inflow for year 2}}{(1 + \text{discount rate})^2} + \frac{\text{Cash inflow for year 3}}{(1 + \text{discount rate})^3} \dots - \text{Initial investment}$$

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

$$NPV = \frac{[\text{range: } 1500 - 2500]}{1.1} + \frac{[\text{range: } 750 - 1250]}{1.21} + \frac{[\text{range: } 1875 - 3125]}{1.331} - 3000$$

Please calculate the mid-points for these ranges and type them in below:

| | |
|--------------------|----------------------|
| Year 1 cash inflow | <input type="text"/> |
| Year 2 cash inflow | <input type="text"/> |
| Year 3 cash inflow | <input type="text"/> |

Figure B.13: Experiment 2 NPV test. Border added for clarity.

B. Chapter 4 appendix

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. Border added for clarity.

Please rate your knowledge of Net Present Value (NPV) on this 1-7 scale:

| | | | | | | | |
|---------|---|---|---|---------|---|---|------|
| Shallow | 2 | 3 | 4 | Partial | 5 | 6 | Deep |
| 1 | | | | | | | 7 |

NPV knowledge

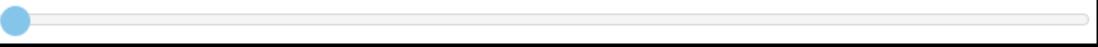
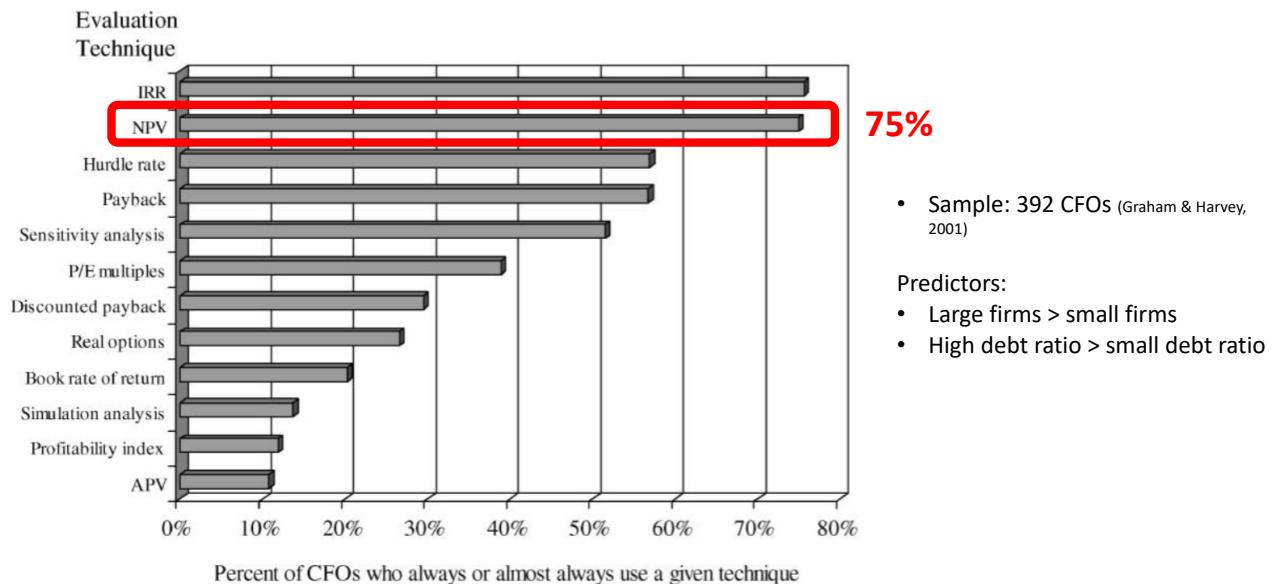


Figure B.15: Experiment 2 NPV knowledge rating task. Border added for clarity.

NPV

$$NPV = \frac{Year\ 1\ inflow}{(1+discount\ \%)^1} + \frac{Year\ 2\ inflow}{(1+discount\ \%)^2} + \frac{Year\ 3\ inflow}{(1+discount\ \%)^3} \dots - Initial\ investment$$

NPV is used very frequently



The NPV paradox

- “Although the NPV method is criticized by both practitioners and academics, the traditional NPV calculation is by far the most commonly used tool for [exploration & production] project valuation.” (Willigers et al., 2017)
- “NPV is almost always applicable but is almost always wrong” (Fox, 2008)
- “the NPV rule as governing all capital budgeting decisions may not be appropriate” (Arya et al., 1998)

Consequences

- Researchers studied 174 cases of fraudulent financial reporting
 - Fraudulent “facts” vs “forecasts”
- Forecasts based on unreasonable accounting assumptions
 - Form 40% of fraud cases
 - Account for 44% of economic losses
- Total damages by fraudulent *facts*: US\$ 27 billion
- Total damages by fraudulent *forecasts*: US\$ 23 billion

NPV

$$NPV = \frac{\text{Year 1 inflow}}{(1+discount\%)^1} + \frac{\text{Year 2 inflow}}{(1+discount\%)^2} + \frac{\text{Year 3 inflow}}{(1+discount\%)^3} \dots - \text{Initial investment}$$

Where do these cash inflows come from?

“It’s impossible to forecast most projects’ actual cash flows accurately” (Myers, 1984)

Forecasting is error-prone

- Future forecasts tend to be overly-optimistic
 - For longevity
 - For relationships
 - When dopamine is increased
 - In animal behaviour
- Executives are similarly overly-optimistic
 - In stock market returns
 - For firm earnings

Forecasting is error-prone

- CFO survey between 2001-2011
- Over the next year, I expect the annual S&P 500 return will be:
 - There is a 1-in-10 chance the actual return will be less than ____%.
 - I expect the return to be: ____%.
 - There is a 1-in-10 chance the actual return will be greater than ____%.
- 13,346 estimates

Forecasting is error-prone

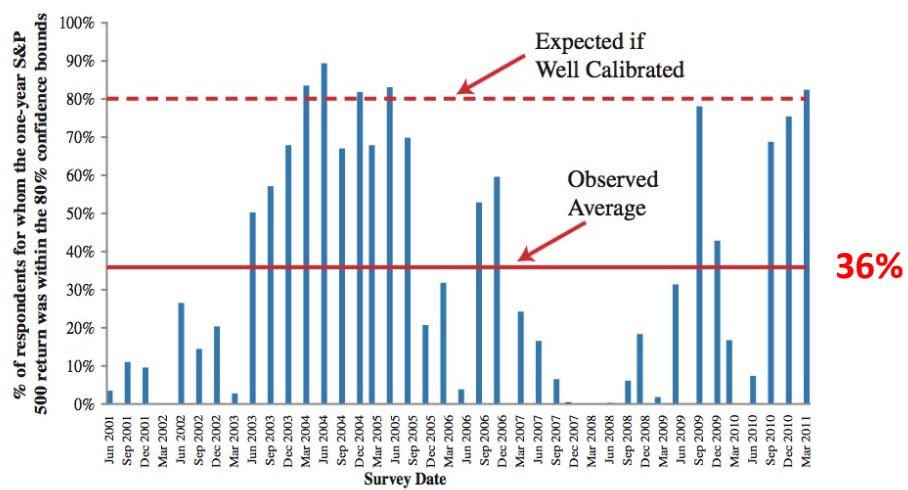
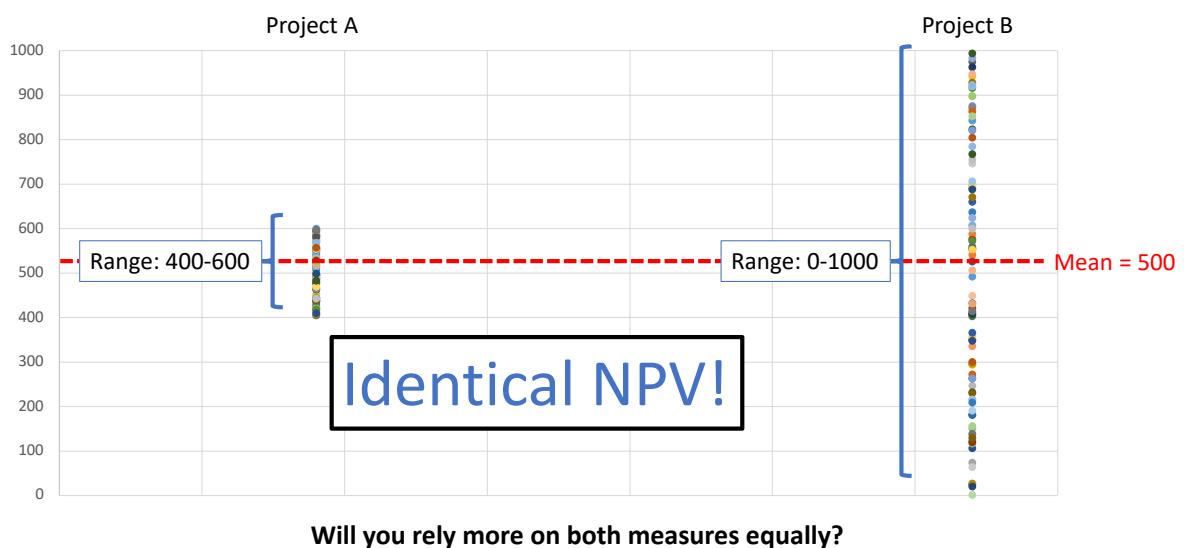


FIGURE I
Time-Series of CFO Miscalibration

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

Paying attention to variance - Example



Summary

- NPV is used a lot, but criticised by some
- The costs of poor forecasting are potentially high
- NPV relies on forecasting
- Executives may underestimate forecast variance

Bottom line

- Pay attention to cash inflow variance
- Not all NPVs are created equal
 - NPV based on more variance should be weighted less than other measures

B.2.2 Results

Ranking

A mixed factorial ANOVA was conducted to investigate the effects of NPV amount, alignment, and numerical NPV reliability on participants' project rankings. Figure B.16 shows these data. The alignment \times reliability amount \times NPV amount interaction was not significant, $F(3.00, 159.10) = 2.44, p = .066, \hat{\eta}_p^2 = .044$. However, the alignment \times NPV amount interaction was significant, $F(3.31, 370.54) = 21.00, p < .001, \hat{\eta}_p^2 = .158$; as well as the reliability amount \times NPV amount 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 not differ between reliability amount condition in neither the low alignment condition, $\Delta M = 0.43, 95\% \text{ CI } [-0.77, 1.63], t(53) = 0.71, p = .480$, nor the high alignment condition, $\Delta M = 0.46, 95\% \text{ CI } [-0.92, 1.84], t(53) = 0.67, p = .504$. However, averaging over reliability amount, the linear NPV trend was higher in the low alignment condition than in the high alignment condition, $\Delta M = -4.54, 95\% \text{ CI } [-6.39, -2.68], t(53) = -4.91, p < .001$.

Confidence

A mixed factorial ANOVA was conducted to investigate the effects of NPV amount, alignment, and numerical NPV reliability on participants' confidence ratings. Figure B.17 shows these data. Only the main effect of NPV amount was significant, $F(2.62, 139.08) = 2.97, p = .041, \hat{\eta}_p^2 = .053$.

NPV knowledge

A repeated-measures ANOVA was conducted to investigate the effects of experiment phase condition on participants' NPV knowledge rating. Figure B.18 shows these data. The main effect of phase was significant, $F(2.43, 128.59) = 7.80, p < .001, \hat{\eta}_p^2 = .128$. The post-explanation rating was significantly higher than the pre-explanation rating, $\Delta M = -0.59, 95\% \text{ CI } [-0.92, -0.26], t(53) = -5.07, p < .001$. However, there were no significant differences in rating between any of the later phases.

B. Chapter 4 appendix

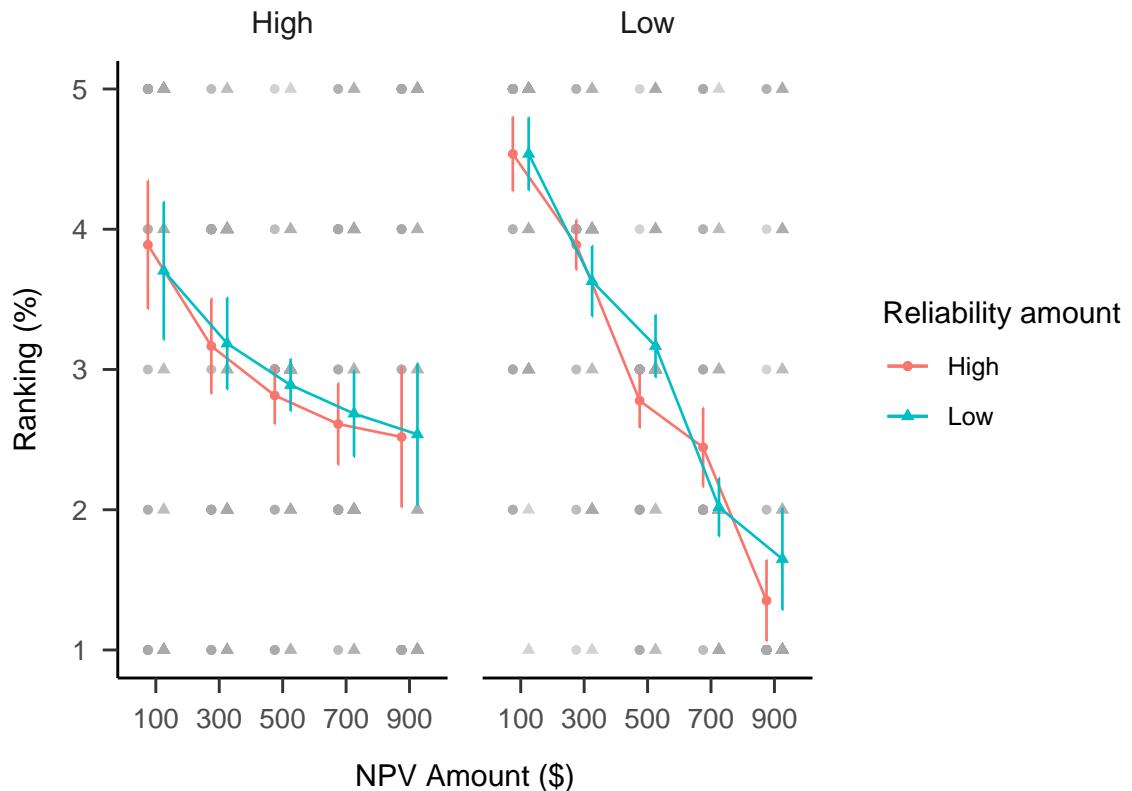
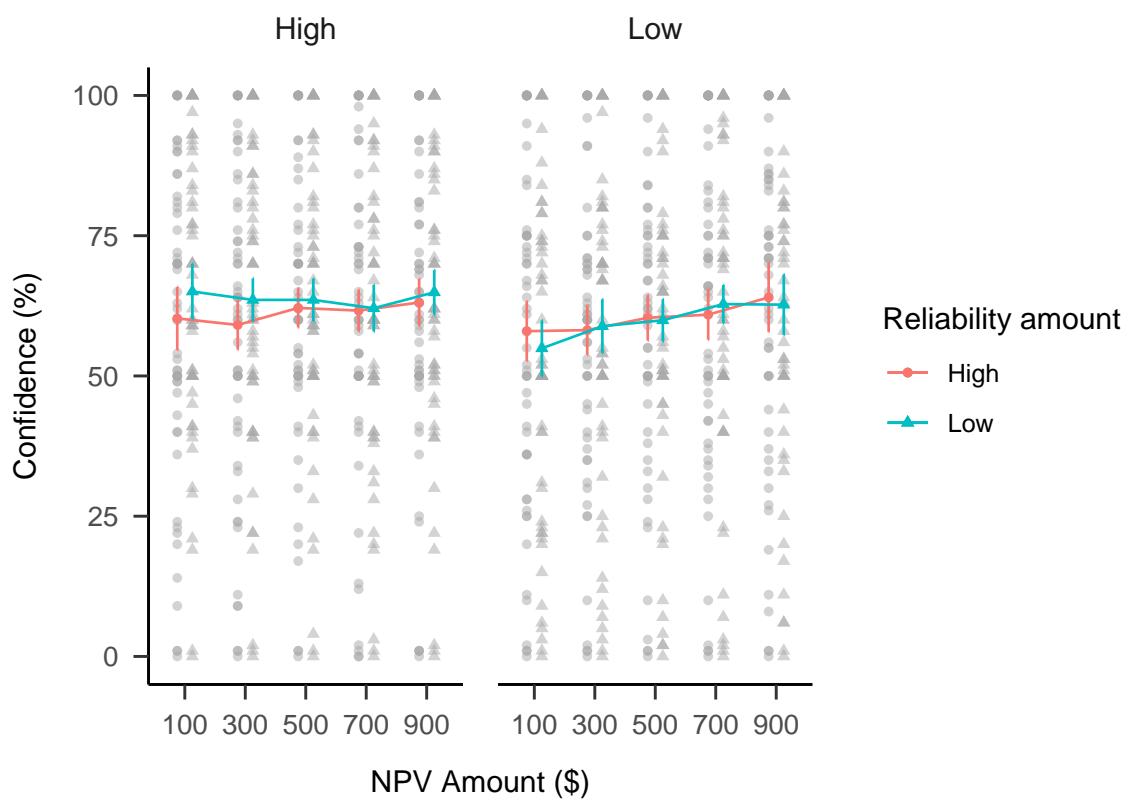


Figure B.16: Mean ranking.



B. Chapter 4 appendix

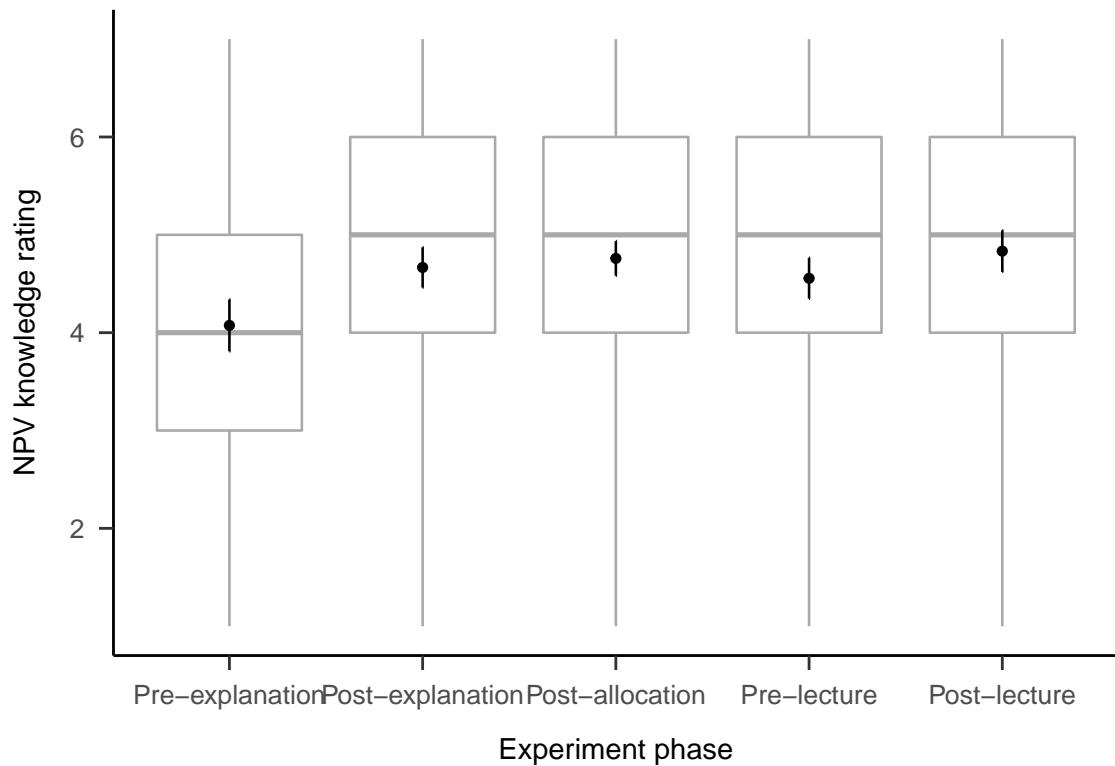


Figure B.18: Mean NPV knowledge rating.

B.3 Experiment 3

I simulated the effects I hypothesised in Experiment 3 in See Figure B.19. These effects were taken 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.

B.3.1 Method

Participants

Power analysis I conducted a power analysis through simulation of the effects hypothesised in Experiment 3 (and the simple effects implied by them). I simulated data with the same regression coefficients as Experiment 2 for the explicit condition, no effects for the implicit condition (as shown in Figure B.19), and the intercept and residual variance of Experiment 2. I analysed the null effects using the two one-sided tests (TOST) procedure, or *equivalence* testing (Lakens et al., 2018), and

B. Chapter 4 appendix

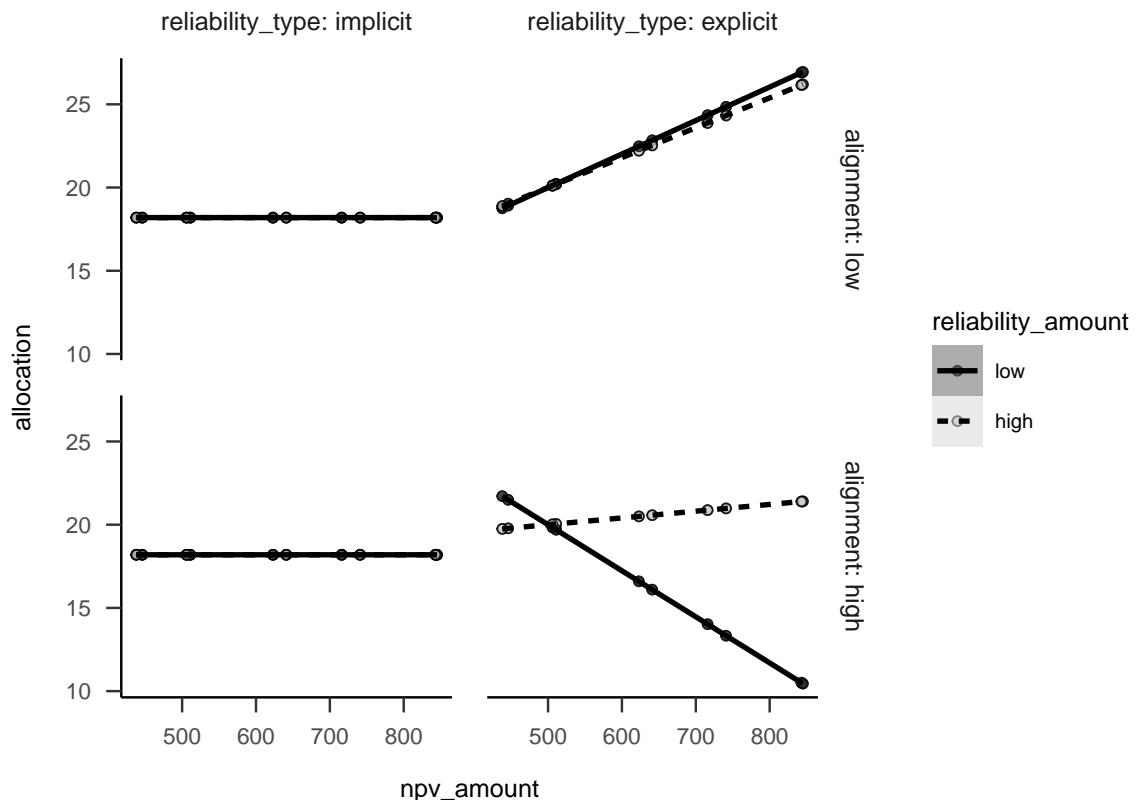


Figure B.19: Experiment 3 predicted data

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 3. Figure B.20 shows the resulting power curve. The analysis suggests a total sample size of 448 ($112 \cdot 4$).

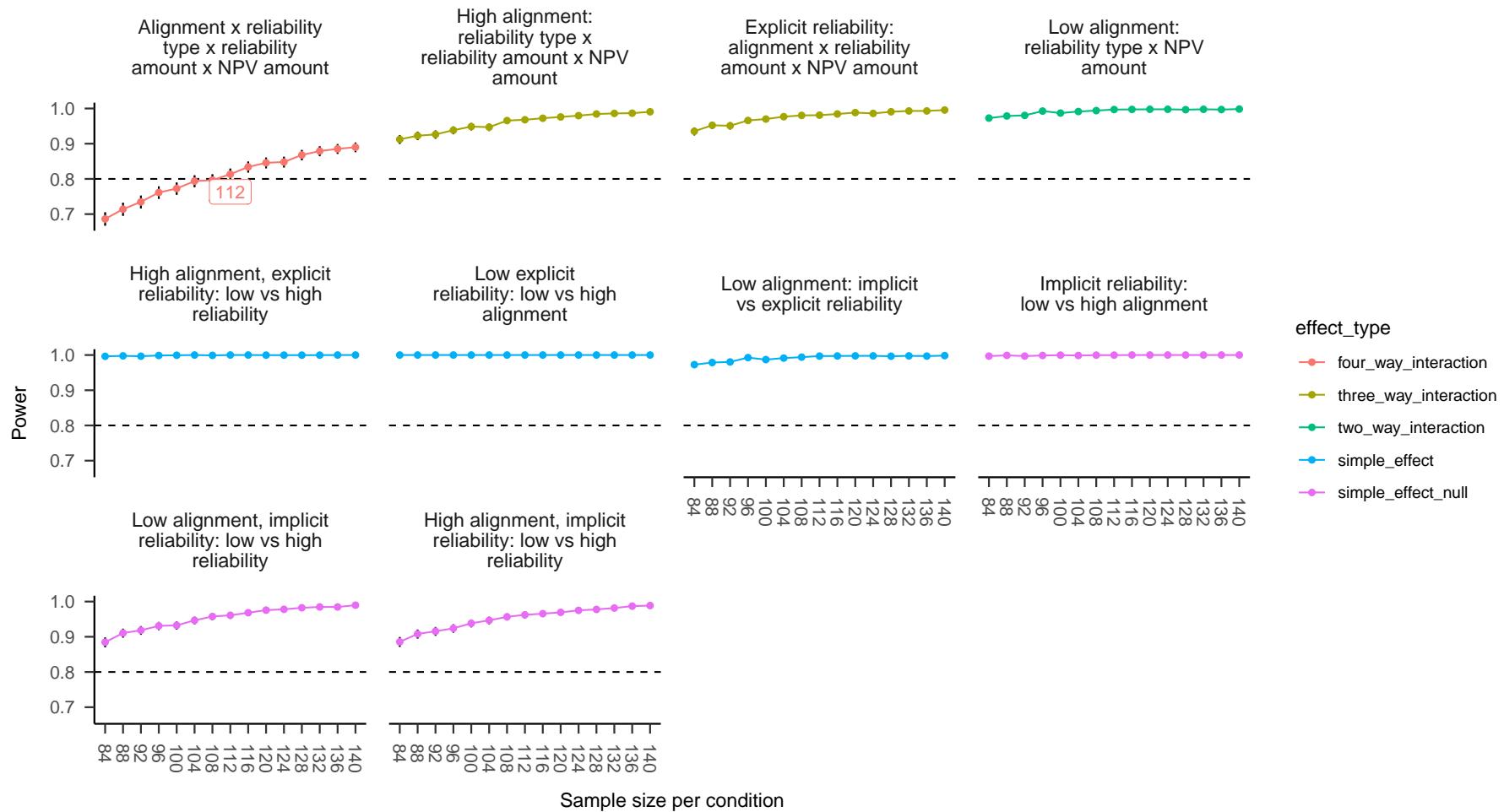


Figure B.20: Alignment Experiment 3 power curve. Labels indicate lowest sample size above 80% power.

B. Chapter 4 appendix

Materials

Instructions Figures B.21 and B.22 show the instructions for the verbal and numerical reliability conditions, respectively.

Interstitial display

B.3.2 Results

Allocation

The three-way interaction (reliability amount \times NPV amount \times 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 amount \times reliability type (averaging over reliability amount) 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 amount 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 amount 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 amount), $\Delta M = 11.48$, 95% CI [0.19, 22.77], $t(444) = 2.00$, $p = .046$. The linear NPV amount 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 I used a “lowest effect size of interest” that 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 amount as a continuous predictor), whereas the data was ultimately analysed using a multivariate linear model (treating NPV amount as a repeated measures factor).

B. Chapter 4 appendix

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 an NPV, alongside a statement of whether NPV is considered to be a reliable (or an unreliable) metric for that project. There are usually a range of plausible NPV outcomes, so when NPV is considered to be "reliable" this means that the range of possible values is relatively narrow (indicating high confidence in the estimate). Conversely, when NPV is considered to be "unreliable", this means that the range of possible values is relatively wider (indicating low confidence in the estimate).

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.21: Experiment 3 verbal reliability instructions. Border added for clarity.

B.4 Experiment 4

Experiment 4 further investigated the effects of alignment and verbal NPV reliability information on financial resource allocation decisions. I replicated the same methodology as in [Experiment 1](#), except for two main changes. First, I manipulated the alignment conditions within subjects. Second, I removed the no NPV condition to the NPV reliability variable.

I expected to replicate the results of [Experiment 1](#). Specifically, I expect 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, here I report

B. Chapter 4 appendix

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.22: Experiment 3 numerical reliability instructions. Border added for clarity.

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.23: An example of an interstitial display in Experiment 3. Border added for clarity.

B. Chapter 4 appendix

Table B.1: Experiment 4 group allocation.

| Reliability amount | N |
|--------------------|----|
| High | 34 |
| Low | 37 |
| Total | 71 |

analyses for the just for the “target project.” This is allocation of resources 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, I report the method and analyses for the confidence measure.

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

Participants

Seventy-one (44 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 33.27 ($SD = 10.21$, $min = 18$, $max = 65$). Table B.1 shows the between-subjects condition allocation. The two alignment conditions (low and high) were presented within subjects and the order of their presentation was randomised. Further, NPV amount was varied within subjects.

Materials

The project display, allocation task, and confidence task were the same as in [Experiment 1](#).

Instructions Participants were shown similar instructions to [Experiment 1](#), except for the addition of references to the multiple displays and the removal of an

B. Chapter 4 appendix

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.24: Experiment 4 low reliability instructions. Border added for clarity.

explanation about the forecasting task. Figures B.24 or B.25 show the instructions, depending on NPV reliability condition.

Procedure

The procedure was the same as in Experiment 1, except that there was 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 amount on participants' project allocations. As

B. Chapter 4 appendix

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.25: Experiment 4 high reliability instructions. Border added for clarity.

seen in Figure B.26, the alignment \times reliability amount \times NPV amount 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 amount \times NPV amount interaction was significant in the high alignment condition, $\Delta M = -64.82, 95\% \text{ CI } [-102.70, -26.93], t(69) = -3.41, p = .001$, the low alignment condition, $\Delta M = -37.74, 95\% \text{ CI } [-70.92, -4.56], t(69) = -2.27, p = .026$, 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 amount interaction was significant, $F(3.64, 250.93) = 3.19, p = .017, \hat{\eta}_p^2 = .044$, such that the linear trend of NPV amount was stronger in the low alignment, $\Delta M = 13.28, 95\% \text{ CI } [-3.31, 29.87], t(69) = 1.60, p = .115$ than in the high alignment condition, $\Delta M = -10.67, 95\% \text{ CI } [-29.62, 8.27], t(69) = -1.12, p = .265$. However, neither of these trends were individually significant.

B. Chapter 4 appendix

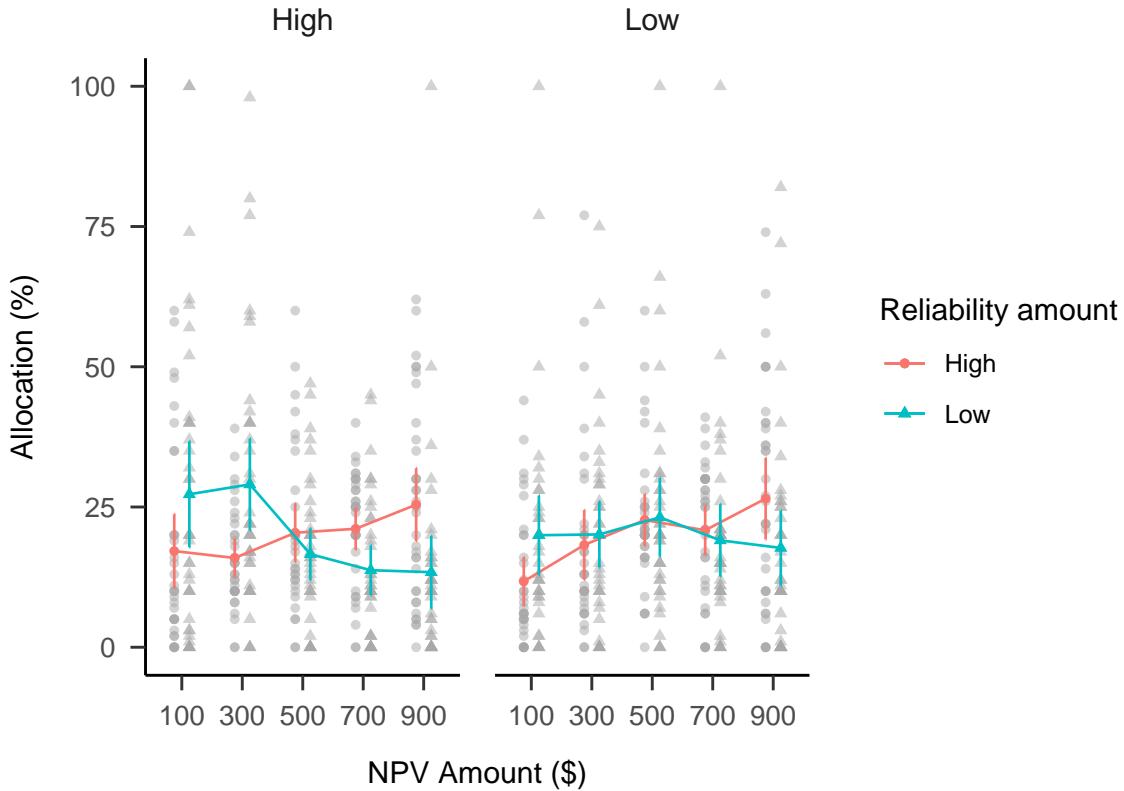


Figure B.26: Mean project allocation in Experiment 4. Error bars represent 95% confidence intervals based on the multivariate model. Note that this mixed factorial design does not allow for using confidence intervals to make inferences by “eye” across conditions.

Confidence

A mixed factorial ANOVA was conducted to investigate the effects of alignment, verbal NPV reliability, and NPV amount on participants’ confidence in their allocations. As seen in Figure B.27, the difference between alignment conditions was not significant, $F(1, 69) = 2.76, p = .101, \hat{\eta}_p^2 = .038$. However, the reliability \times alignment interaction was significant, as well as the NPV amount \times alignment interaction. I conducted an exploratory analysis 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, $\Delta M = -8.83, 95\% \text{ CI } [-17.84, 0.18], t(69) = -1.95, p = .055$ compared to the high alignment condition, $\Delta M = 2.37, 95\% \text{ CI } [-8.65, 13.40], t(69) = 0.43, p = .669$. Further, there was a

B. Chapter 4 appendix

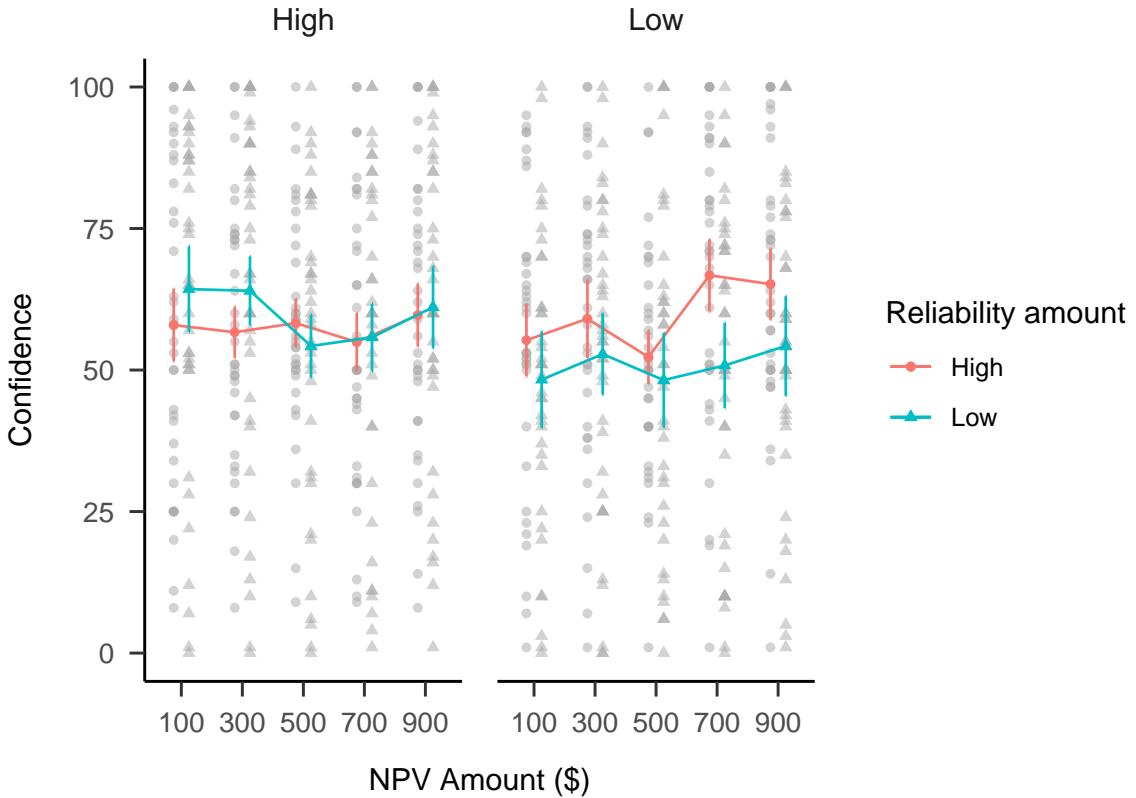


Figure B.27: Mean confidence. Error bars represent 95% confidence intervals based on the multivariate model. Note that this mixed factorial design does not allow for using confidence intervals to make inferences by “eye” across conditions.

stronger linear trend of NPV amount 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 to verbal reliability instructions in the high alignment condition. Contrary to Hypothesis 4.5, however, I found that 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, I did not find evidence that this effect is moderated by alignment condition, contrary to Hypothesis 4.3. However, I did find that the linear NPV amount trend was higher

B. Chapter 4 appendix

in the high than low alignment condition, when averaging over reliability amount, 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.

Despite these results, it is unclear to what extent the within-subjects design influenced participants' allocations, as they saw both high and low alignment conditions. Further, it is unclear how participants would have responded to the projects without NPV. As such, in Experiment 2 I used a between-subject design for all factors and added a no NPV condition.

Hypothesis B.4 was not supported, as I did not find 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' 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. I set out to replicate the forecasting results of Experiment 1, but with a sample that has investing experience. As before, I hypothesised that people's forecasting would be less variable when comparing projects with alignable differences, than when comparing projects with non-alignable differences.

B.5.1 Method

Participants

Sixty (2 female) people were recruited from Reddit. Participants were compensated with a virtual Gold Award, which gives the recipient a week of a premium

B. Chapter 4 appendix

Table B.2: Experiment 5 group allocation.

| Alignment | Reliability amount | N |
|-----------|--------------------|----|
| High | Absent | 19 |
| High | Present | 17 |
| Low | Absent | 14 |
| Low | Present | 10 |
| Total | - | 60 |

version of Reddit and 100 virtual coins. The average age was 28.17 ($SD = 8.73$, $min = 16$, $max = 61$). Table B.2 shows the between-subjects condition allocation.

Materials

Instructions

Risky investment task We only used 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.28.

Procedure

The procedure was the same as in Experiment 1, except participants only completed the forecasting task.

B.5.2 Results

Forecast mean

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecasts. As seen in Figure B.29, the alignment \times reliability amount \times NPV amount 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 $[-0.21, -0.02]$,

B. Chapter 4 appendix

Imagine that you have 100 points to assign to the following options for Project I's rate of return on investment after one year. Assign points according to how likely you think each rate of return is.

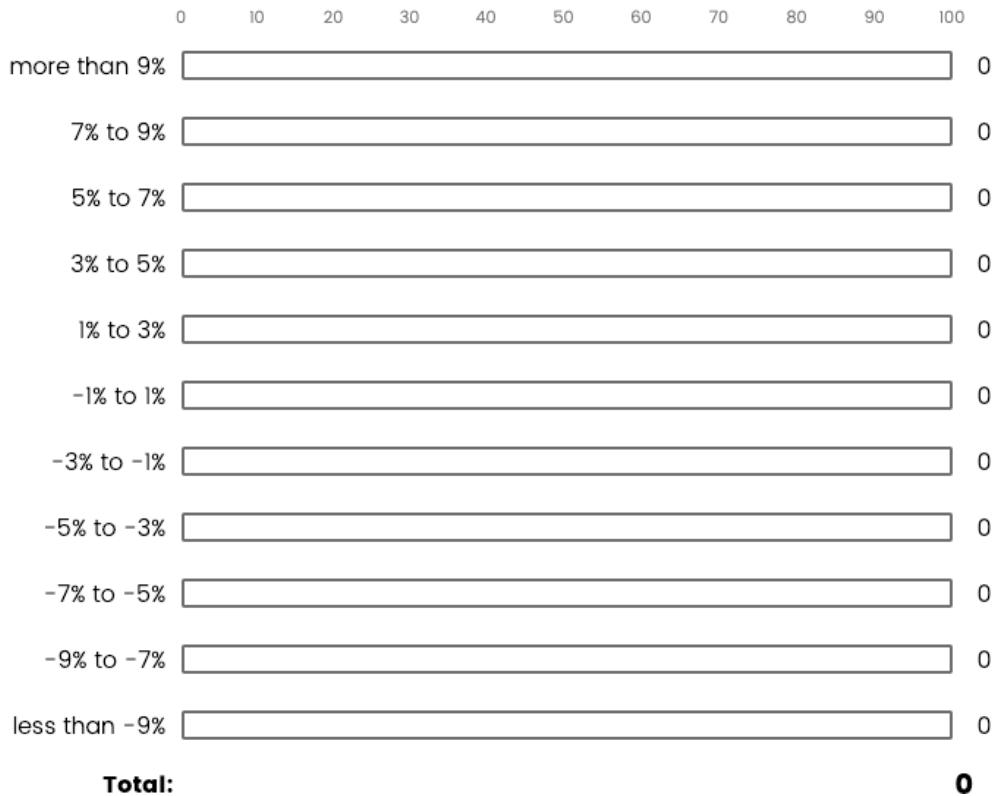


Figure B.28: An example of the forecasting task in Experiment 5.

$t(56) = -2.50$, $p = .015$, but not in the low alignment condition, $M = -0.05$, 95% CI $[-0.16, 0.07]$, $t(56) = -0.81$, $p = .424$.

Forecast SD

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecast SDs. As seen in Figure B.30, the main effect of alignment was significant, with the high alignment condition having higher SDs than the low alignment condition.

B.5.3 Discussion

Experiment 5 found that people with some investing experience respond to alignable information in the form of NPV when it is given, but do not show the same effect

B. Chapter 4 appendix

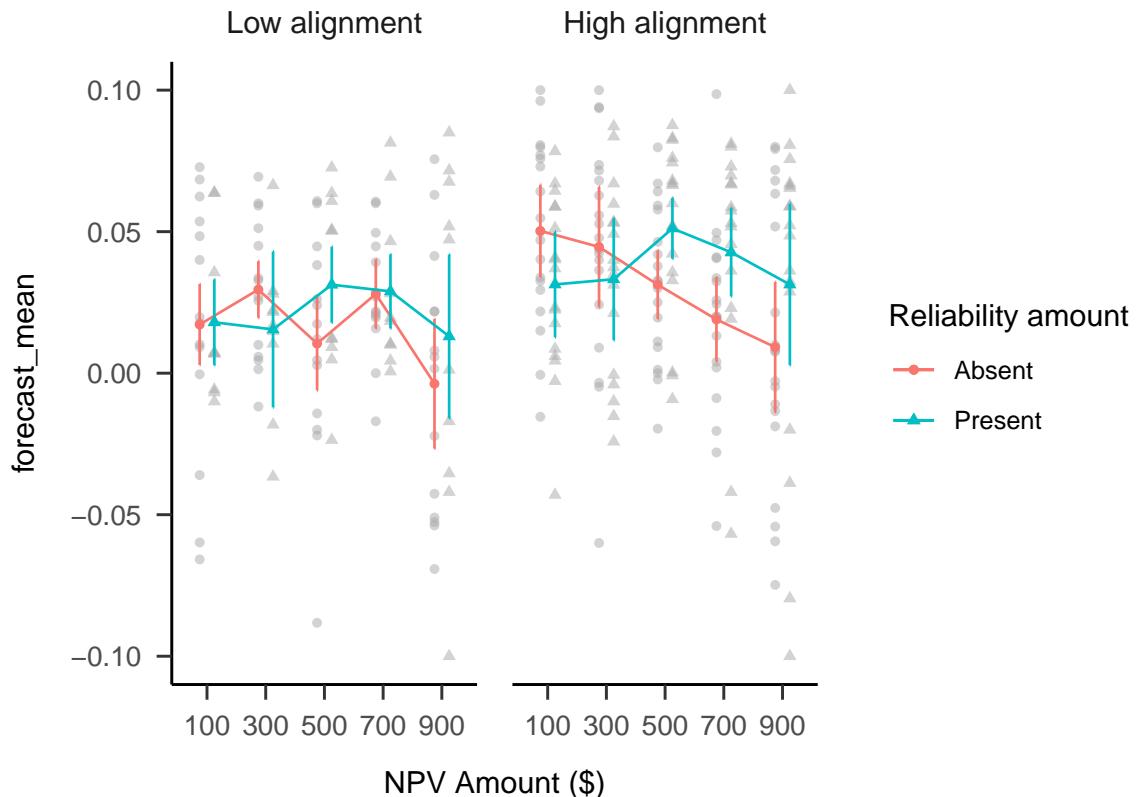


Figure B.29: Mean forecasts.

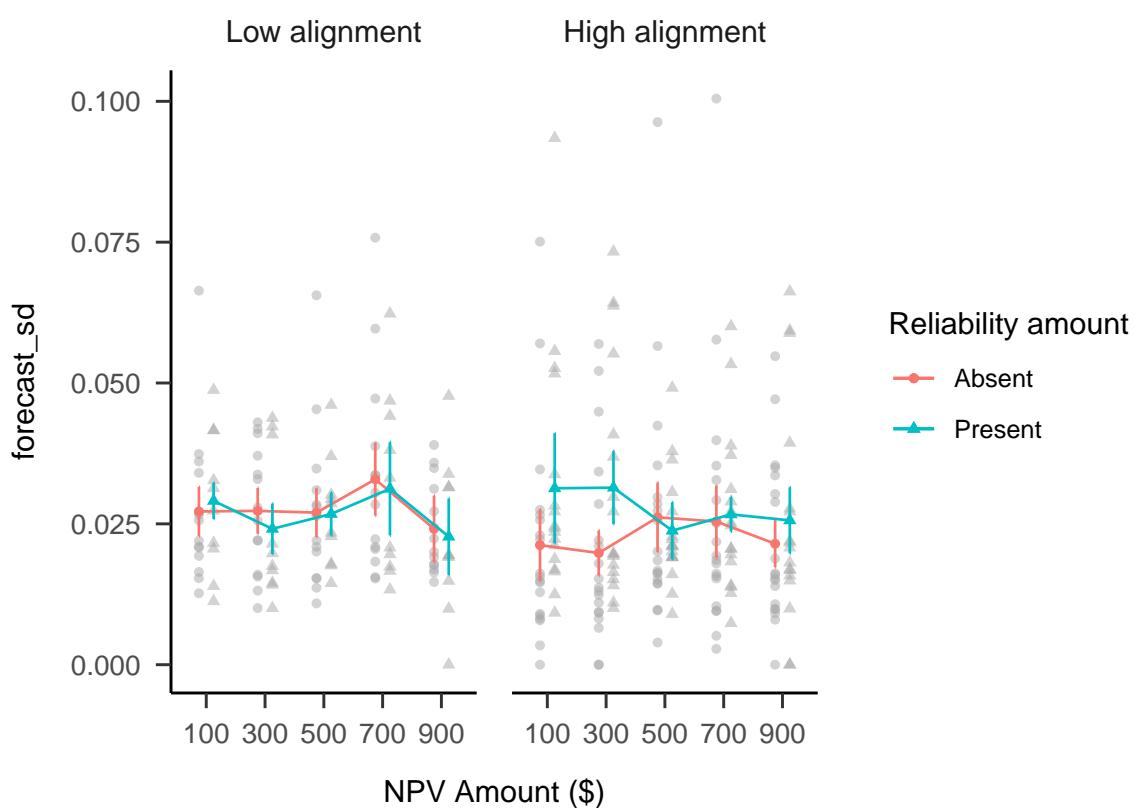


Figure B.30: Mean forecast SD.

B. Chapter 4 appendix

Table B.3: Experiment 6 group allocation.

| Alignment | Reliability amount | N |
|-----------|--------------------|-----|
| High | Absent | 97 |
| High | Present | 87 |
| Low | Absent | 101 |
| Low | Present | 104 |
| Total | - | 389 |

of alignment that was seen in Experiment 1.

B.6 Experiment 6

Experiment 6 further investigated the effects of alignment and NPV Presence information on forecasting. In Experiment 5 I did not clearly replicate the forecasting results of Experiment 1, potentially due to low power, so in this experiment I collected a much larger sample size. As before, I 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

Participants

Three hundred and eighty-nine (170 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 32.39 ($SD = 11.89$, $min = 18$, $max = 75$). Table B.3 shows the condition allocation.

Materials

The materials were the same as in Experiment 5.

Procedure

The procedure was the same as in Experiment 5.

B. Chapter 4 appendix

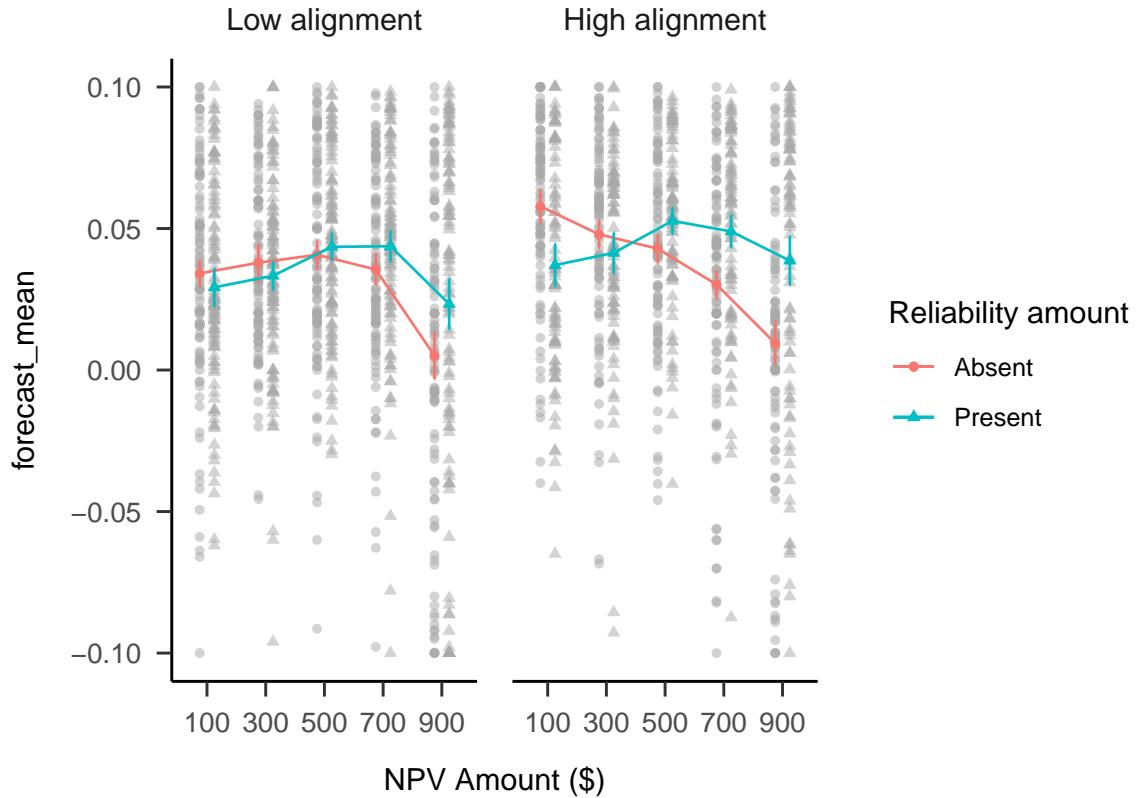


Figure B.31: Mean forecasts.

B.6.2 Results

Forecast mean

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecasts. As seen in Figure B.31, the alignment \times reliability amount \times NPV amount 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 $[-0.09, -0.02]$, $t(385) = -3.28$, $p = .001$.

Forecast sd

A mixed factorial ANOVA was conducted to investigate the effects of alignment and NPV presence on participants' forecast SDs. As seen in Figure B.32,

B. Chapter 4 appendix

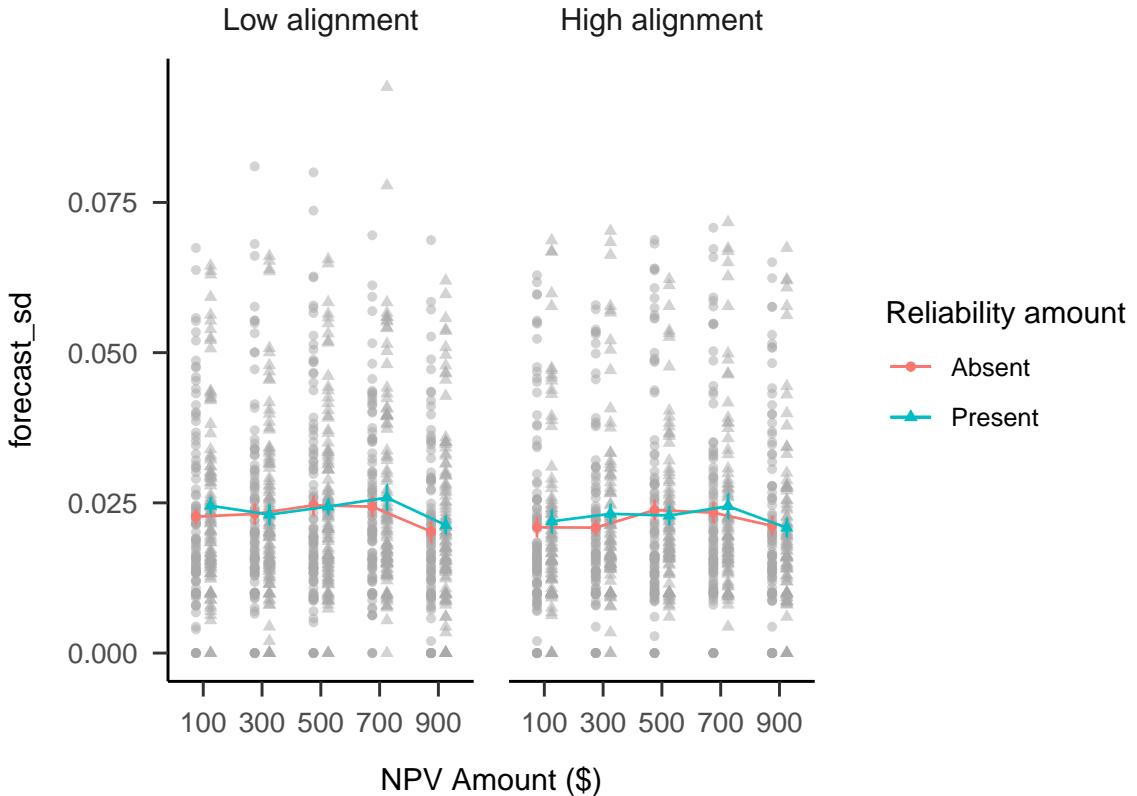


Figure B.32: Mean forecast SD.

the alignment \times reliability amount \times NPV amount interaction was significant, $F(3.08, 1, 186.45) = 3.13, p = .024, \hat{\eta}_p^2 = .008$. The interaction between the linear NPV trend and NPV presence was significant in both the high alignment condition, $M = -0.13, 95\% \text{ CI } [-0.16, -0.09], t(385) = -6.57, p < .001$, and in the low alignment condition, $M = -0.06, 95\% \text{ CI } [-0.09, -0.02], t(385) = -3.28, p = .001$.

B.6.3 Discussion

Experiment 6 did not replicate the findings of Experiment 1, finding contradictory effects of alignment on forecast variance. 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. Chapter 4 appendix

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 |

B.7 Experiment 7

Experiment 7 investigated potential ways to facilitate people’s use of variance in resource allocation. Arguably, people’s decisions should be moderated by variance, especially with a small set of projects. That is, when considering between two potential measures to use for resource allocation, underlying variance should serve as a moderator for decision making, with measures with narrow ranges being relied upon more than those with wider ranges. As such, in this experiment I presented participants with the same resource allocation scenario as in Experiment 1, but only in low numerical reliability displays. I varied both the variance associated with NPV, and the extent to which I explicitly hinted to them to use the variance information. I predicted that participants would be more likely to moderate their allocations through 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

Participants

Seventy-nine (35 female) people were recruited from the online recruitment platform Prolific. Participants were compensated at a rate of £5 an hour. The average age was 31.15 ($SD = 11.11$, $min = 16$, $max = 71$). Table B.4 shows the between-subjects condition allocation.

B. Chapter 4 appendix

| | 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.33: The projects display.

Materials

The materials were similar to Experiment 1, as the same laptop displays were used, but I added the *Cash inflow range* row to the table (see Figure B.33). Further, participants in the No Hint condition saw the same instructions as in Experiment 1, those in the Salience Only condition saw the instructions along with a sentence that draws attention to the *Cash inflow range* row, and 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.

Procedure

Participants completed a demographics page, read the instruction page as per their Hint condition, and then proceeded to complete one set of allocations.

B. Chapter 4 appendix

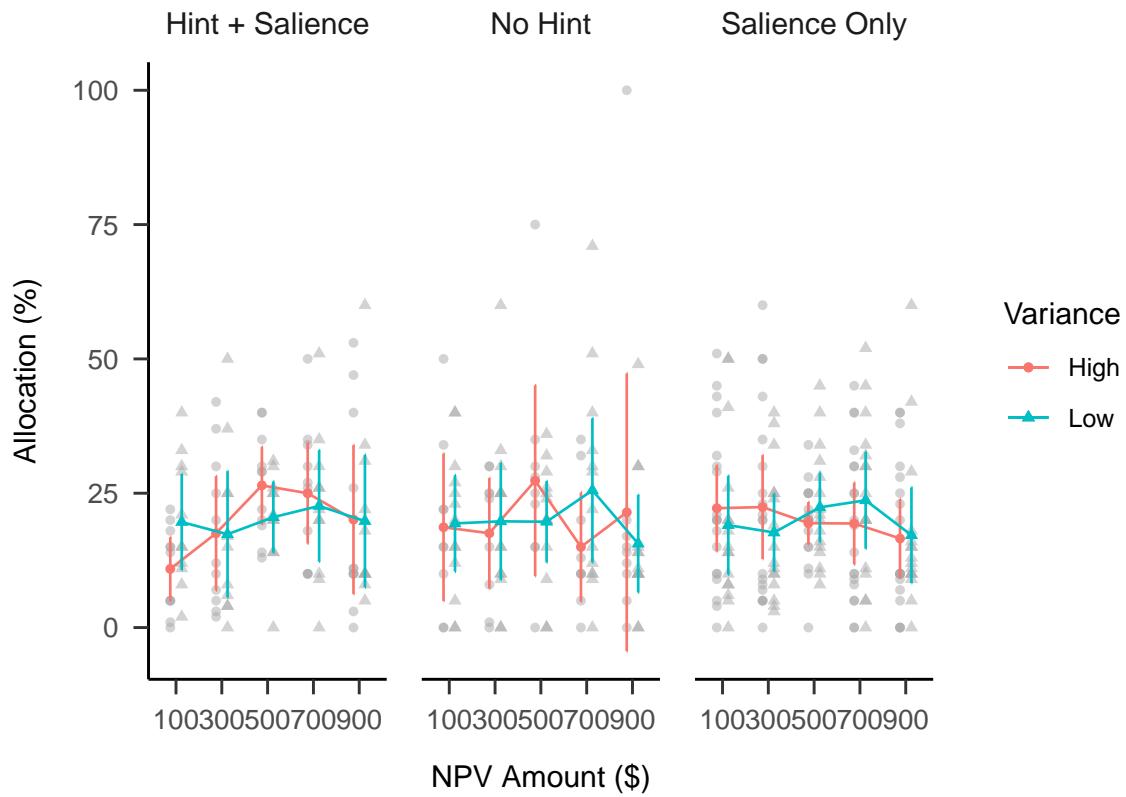


Figure B.34: Mean allocation.

B.7.2 Results

Allocation

A mixed factorial ANOVA was conducted to investigate the effects of hint and NPV variance on participants' allocations. As seen in Figure B.34, none of the interactions or main effects were significant.

Ranking

A mixed factorial ANOVA was conducted to investigate the effects of hint and NPV variance on participants' project rankings. As seen in Figure B.35, only the main effect of NPV amount was significant.

B.7.3 Discussion

Experiment 7 found that explicitly telling participants how to use variance information to moderate their allocations did not help them do so. However, there was an

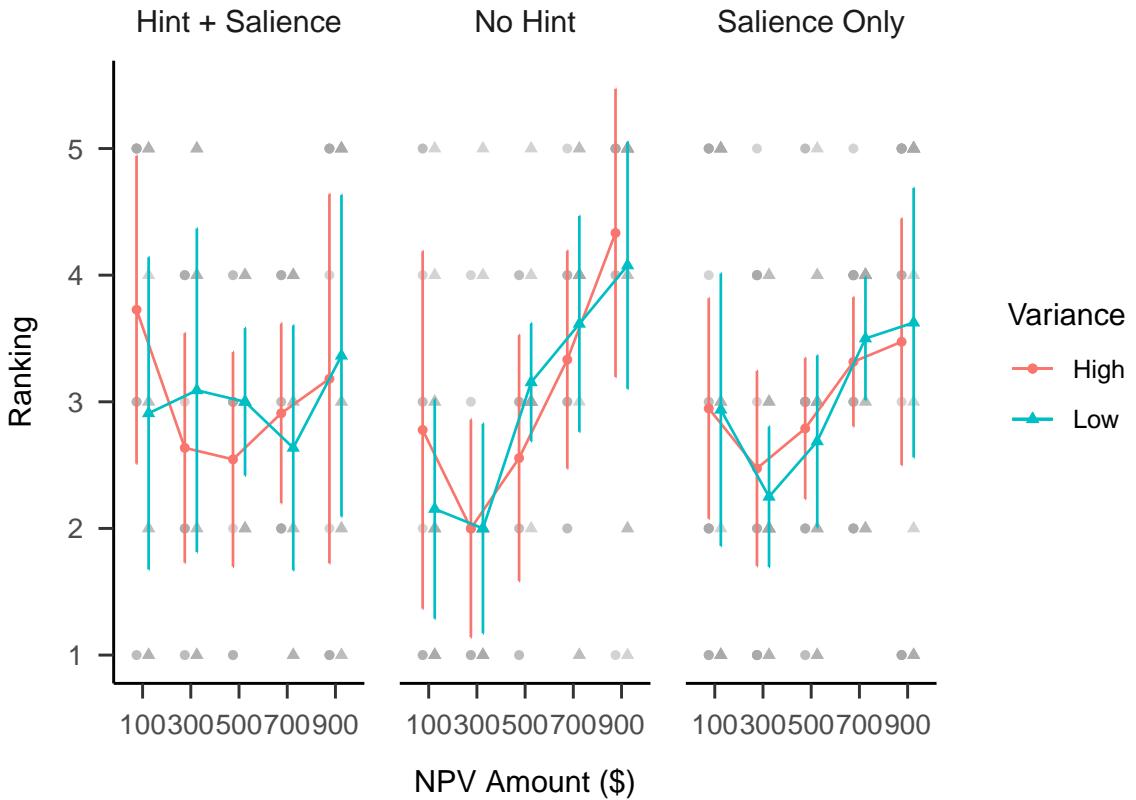


Figure B.35: Mean ranking.

increased reliance on NPV with more hints in the ranking data. This suggests that the hint manipulations potentially simply increase participants' attention to NPV.

B.8 Experiment 8

Experiment 8 tested the alignment and reliability effects found in the previous experiments addressing the limitations of the previous experiments. In Experiments 1 and 4 I found a verbal reliability effect, in which laypeople allocated more resources to a high NPV project, depending on how reliable they were told NPV was as a measure. In Experiment 2 I found a lack of a numerical reliability effect, in which business students allocated an equivalent amount of resources to projects associated with a high variance NPV, as projects with a low NPV. Testing these two effects in two different populations did not account for potential expertise effects. As such, in Experiment 8 I tested both effects with a naive sample. Further, I used projects whose features more clearly indicate their profitability,

B. Chapter 4 appendix

and included more project domains.

B.8.1 Method

Participants

Fifty-two (33 female) people were recruited from both the online recruitment platform Prolific and a Psychology undergraduate sample at The University of Sydney. Participants from Prolific were compensated at a rate of £5 an hour, and participants from the undergraduate sample were compensated with course credit. The average age was 24.46 ($SD = 7.77$, $min = 18$, $max = 68$). Participants reported an average of 2.63 ($SD = 4.16$, $min = 0$, $max = 25$) years of work in a business setting, and an average of 0.81 ($SD = 1.39$, $min = 0$, $max = 5$) years of business education. The mean completion time was 35.57 ($SD = 71.96$, $min = 7.36$, $max = 511.74$) minutes.. All conditions were presented within-subjects: alignment (low and high), NPV reliability type (numerical and verbal), NPV amount (low and high), and NPV reliability amount (low and high).

Materials

Instructions Participants saw instructions similar to the previous experiments.

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 the type of comparison (and the associated preamble text).

B. Chapter 4 appendix

Figures B.36, B.37, B.38, B.39 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 is in an industry in which NPV is considered a reliable or unreliable measure. Implicit reliability was manipulated by presenting NPVs alongside numerical ranges instead of verbal reliability information about them, and varying whether the range is high or low. Both of these were manipulated within-display, such that for four projects in each display NPV is reliable, and for four it is unreliable.

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, we inferred a reliance on NPV if participants allocated the high NPV project more resources, or a reliance on the intrinsic features if participants allocated the low NPV project more resources.

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. I was unable to conduct a direct comparison of the two alignment conditions due to the different allocation input scales, so I compared the NPV reliability amount \times NPV amount interaction separately in each alignment condition (see Figures B.40 and B.41).

B. Chapter 4 appendix

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.

| | | |
|--|--|---|
| <p>Business name: FreightCog. - Investment: <u>railway</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Railmont. - Investment: <u>railway</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Pharmacore. - Investment: <u>pharmaceutical</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Biotechly. - Investment: <u>pharmaceutical</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Pressbloom. - Investment: <u>national newspaper</u>. - Predicted project features: - Newspapers printed: 30000 a day. - Number of weekly advertisers: 48. - Ink that is not discarded due to impurities: 3000L a day</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |

Figure B.36: Experiment 8 low alignment, verbal reliability display. Cropped for space (full display has eight projects).

B. Chapter 4 appendix

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.

| | | |
|--|---------------------------------------|--|
| <p>Business name: Dinerly. - Investment: <u>restaurant chain</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Savoro. - Investment: <u>restaurant chain</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Altchip. - Investment: <u>microchip</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Microxy. - Investment: <u>microchip</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Enfuel. - Investment: <u>oil well</u>. - Predicted project features: - Oil extracted: 1200L an hour. - Time the machinery lasts before requiring maintenance: 4 years. - Probability of finding oil: 50%</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |

Figure B.37: Experiment 8 low alignment, numerical reliability display. Cropped for space (full display has eight projects).

B. Chapter 4 appendix

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.

| | | |
|---|--|---|
| <p>Business name: Erectic.</p> <ul style="list-style-type: none"> - Investment: <u>high-rise construction</u>. - Predicted project features: <ul style="list-style-type: none"> - 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. <p>(In this particular industry, NPV is an unreliable predictor of project success.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Refit.</p> <ul style="list-style-type: none"> - Investment: <u>high-rise construction</u>. - Predicted project features: <ul style="list-style-type: none"> - 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. <p>(In this particular industry, NPV is an unreliable predictor of project success.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Pressbloom.</p> <ul style="list-style-type: none"> - Investment: <u>national newspaper</u>. - Predicted project features: <ul style="list-style-type: none"> - Newspapers printed: 30000 a day. - Number of weekly advertisers: 48. - Ink that is not discarded due to impurities: 3000L a day. <p>(In this particular industry, NPV is a reliable predictor of project success.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Grown Media.</p> <ul style="list-style-type: none"> - Investment: <u>national newspaper</u>. - Predicted project features: <ul style="list-style-type: none"> - Newspapers printed: 50000 a day. - Number of weekly advertisers: 80. - Ink that is not discarded due to impurities: 5000L a day. <p>(In this particular industry, NPV is a reliable predictor of project success.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: FreightCog.</p> <ul style="list-style-type: none"> - Investment: <u>railway</u>. - Predicted project features: <ul style="list-style-type: none"> - Railway lines built: 5 a decade. | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |

Figure B.38: Experiment 8 high alignment, high reliability display. Cropped for space (full display has eight projects).

B. Chapter 4 appendix

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.

| | | |
|---|---------------------------------------|--|
| <p>Business name: Enfuel. - Investment: <u>oil well</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Refinera. - Investment: <u>oil well</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Altchip. - Investment: <u>microchip</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Microxy. - Investment: <u>microchip</u>. - 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.)</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |
| <p>Business name: Solgistics. - Investment: <u>shipping logistics</u>. - 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%.</p> | Project ranking: <input type="text"/> | Budget allocation: \$ <input type="text"/> |

Figure B.39: Experiment 8 high alignment, numerical reliability display. Cropped for space (full display has eight projects).

B. Chapter 4 appendix

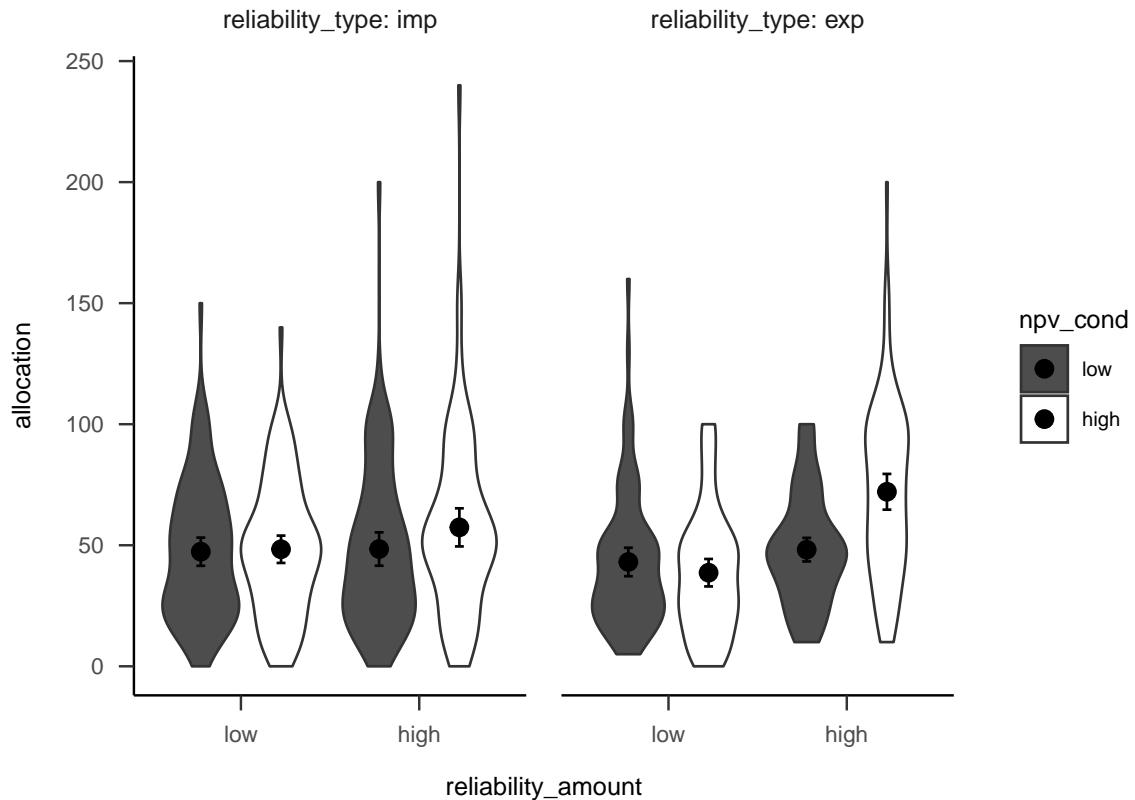


Figure B.40: Mean project allocation, for the low alignment condition. Error bars represent 95% confidence intervals.

This interaction was significant for both the high alignment condition, ; and the low alignment condition, . However, there was a significant effect of NPV in the low verbal reliability condition in the high alignment condition, $\Delta M = 18.69$, 95% CI [2.87, 34.52], $t(113.10) = 3.17$, $p = .012$; but not for the low alignment condition, $\Delta M = 3.75$, 95% CI [-12.07, 19.57], $t(113.10) = 0.64$, $p > .999$.

B.8.3 Discussion

I found that in the verbal reliability condition, participants allocated according to the reliability information, for both low and high alignment conditions. In the numerical reliability condition, 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 experiment shows that similar to the previous experiments, when controlling for presentation and domain, people still find it easier to allocate resources based

B. Chapter 4 appendix

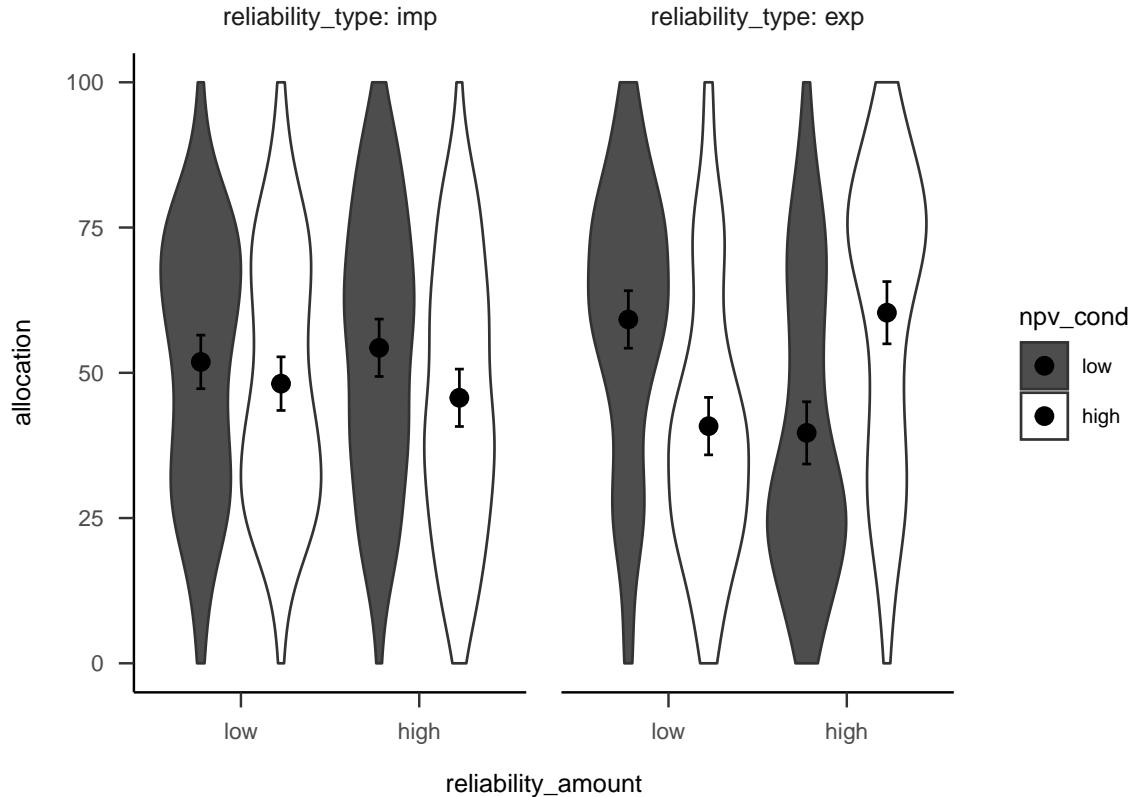


Figure B.41: Mean project allocation, for the high alignment condition. Error bars represent 95% confidence intervals.

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, I showed that people without much business experience also struggle to use range information in resource allocation to such an extreme extent that they do not seem to be using any coherent allocation strategy.

C

Chapter 6 appendix

Contents

| | |
|-----------------------------------|------------|
| C.1 Experiment 1 | 227 |
| C.1.1 Method | 228 |
| C.1.2 Results | 231 |
| C.2 Experiment 2 | 236 |
| C.2.1 Method | 239 |
| C.2.2 Results | 241 |

This appendix contains supplementary materials and analyses for the two experiments reported in Chapter 4.

C.1 Experiment 1

Hypothesis C.1—Similarity manipulation check for negative anecdote.

In the negative valence condition, 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 positive anecdote. In the negative valence condition, the correlation between allocation and similarity rating will be negative

C. Chapter 6 appendix

Hypothesis C.3—Relationship between allocation and specific-relevance

for positive 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, I asked participants to rate the relevance of the anecdote to the target project. I 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.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, I asked participants to rate the relevance of the anecdote to other projects in the same industry. I 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

Participants

Power analysis I determined the sample size for Experiment 1 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. I used data from Wainberg (2018), Jaramillo et al. (2019), and Hoeken and Hustinx (2009, Study 3) to determine realistic

C. Chapter 6 appendix

means and standard deviations for the evidence and similarity factors, and then calculated the sample size that would allow for an expected power of at least 80% according to the power functions.

I used data from Wainberg (2018) to determine the predicted means for the anecdote conditions. Specifically, I used the values from the anecdote + statistics, anecdote + enhanced statistics, and statistics only conditions as the values for my high similarity condition (because in Wainberg (2018) the anecdote was always of a similar case) for the corresponding anecdote conditions. 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, I used the same mean value for both these conditions.

I hypothesised that there will only be an effect of similarity for the anecdote only and anecdote + statistics conditions. As such, I used the data from Hoeken and Hustinx (2009, Study 3) to determine the corresponding mean values for the low similarity condition. Specifically, I multiplied each predicted mean by the Cohen's d_z of the similarity effect in Hoeken and Hustinx (2009, Study 3).

To determine the predicted standard deviation, I re-analysed the data from Jaramillo et al. (2019) Experiment 2 and Hoeken and Hustinx (2009, Study 3) to determine the coefficient of variation (CV) of each condition. I then converted them to a standard deviation value in the relevant scale by multiplying the mean of the CV values by the predicted means from above.

As seen in Figure C.1, the power analysis suggests that a minimum sample size of 294 ($42 \cdot 7$) is required for the interaction effect with an expected power of at least 80%.

Method

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.

C. Chapter 6 appendix

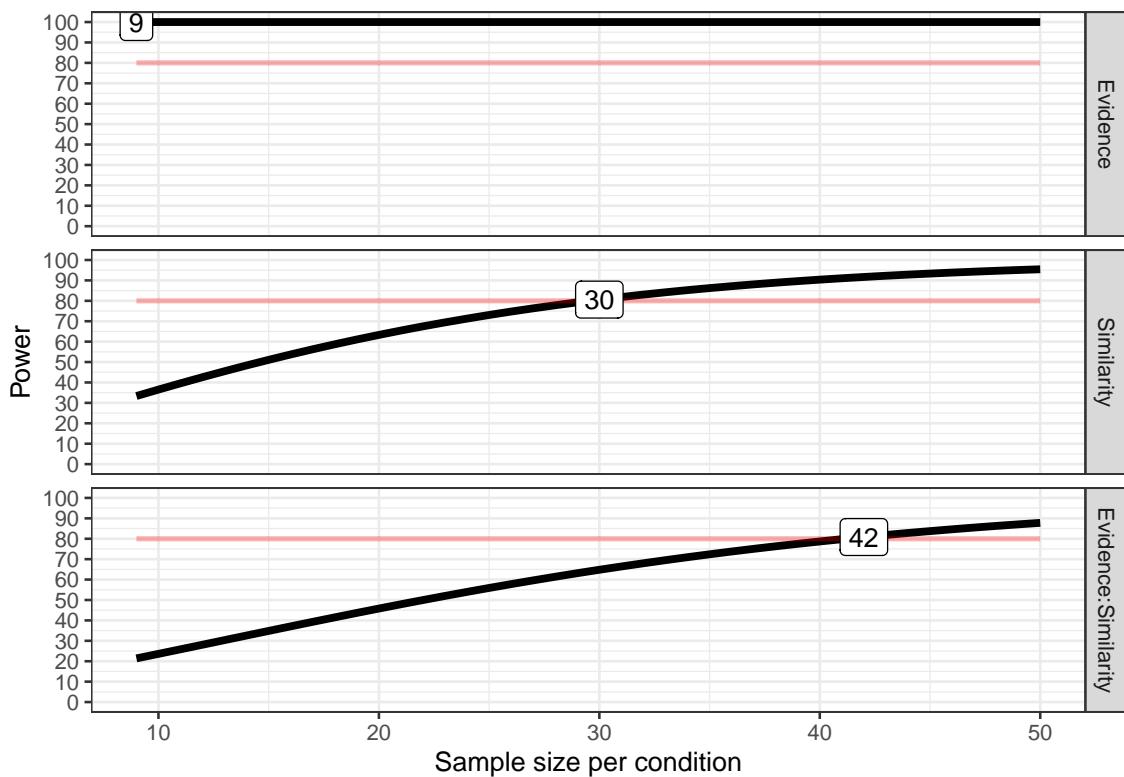


Figure C.1: Power curves for the similarity and anecdote effects.

Imagine you are a executive in a multi-business company and that you are presented with two projects to potentially invest in. Your job is to decide how to allocate the capital available in your budget between these two projects.

In a moment you will see a table that details the two target projects and relevant information about them.

Figure C.2: Experiment 1 general instructions. Note, the two boxes were split between two separate web-pages.

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.

C. Chapter 6 appendix

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.

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 one in which decisions can be made by separate units or people independently (Kenton (2021)).

Follow-up Figure C.7 shows the follow-up questions.

C.1.2 Results

Allocation

A two-way ANOVA was conducted to investigate the interaction of similarity (low and high) and anecdote conditions (anecdote only, statistics + anecdote, anecdote

C. Chapter 6 appendix

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.

C. Chapter 6 appendix

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.

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. Chapter 6 appendix

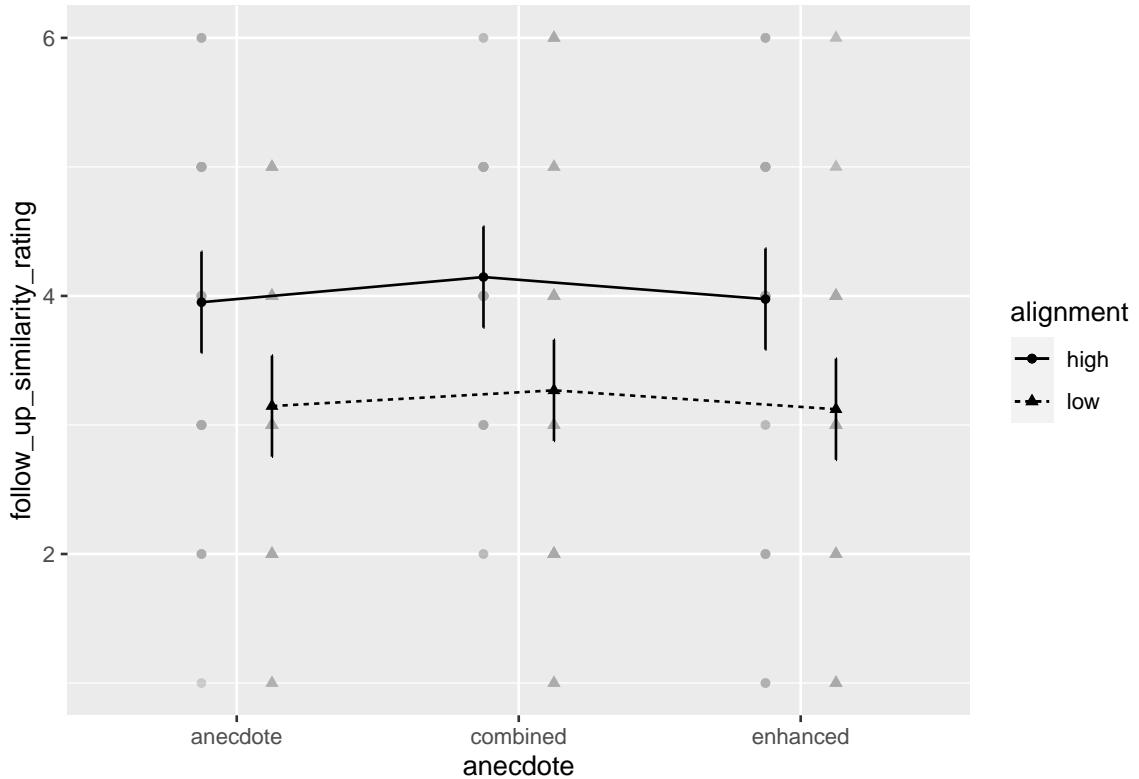


Figure C.8: Mean similarity rating of Project A (the target project) to the anecdote. Error bars represent 95% confidence intervals.

+ enhanced statistics). I found a main effect of anecdote type, $F(2, 240) = 14.14$, $p < .001$, $\hat{\eta}_p^2 = .105$; and a main effect of similarity, $F(1, 240) = 39.68$, $p < .001$, $\hat{\eta}_p^2 = .142$. interaction was not significant, $F(2, 240) = 2.39$, $p = .094$, $\hat{\eta}_p^2 = .020$. The difference between the anecdote only condition and the anecdote + enhanced statistics condition was not significant when excluding the anecdote + statistics condition, $M = -9.59$, 95% CI $[-22.27, 3.09]$, $t(240) = -1.49$, $p = .138$.

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, $F(1, 240) = 27.09$, $p < .001$, $\hat{\eta}_p^2 = .101$. Table shows participants' justifications for these ratings.

C. Chapter 6 appendix

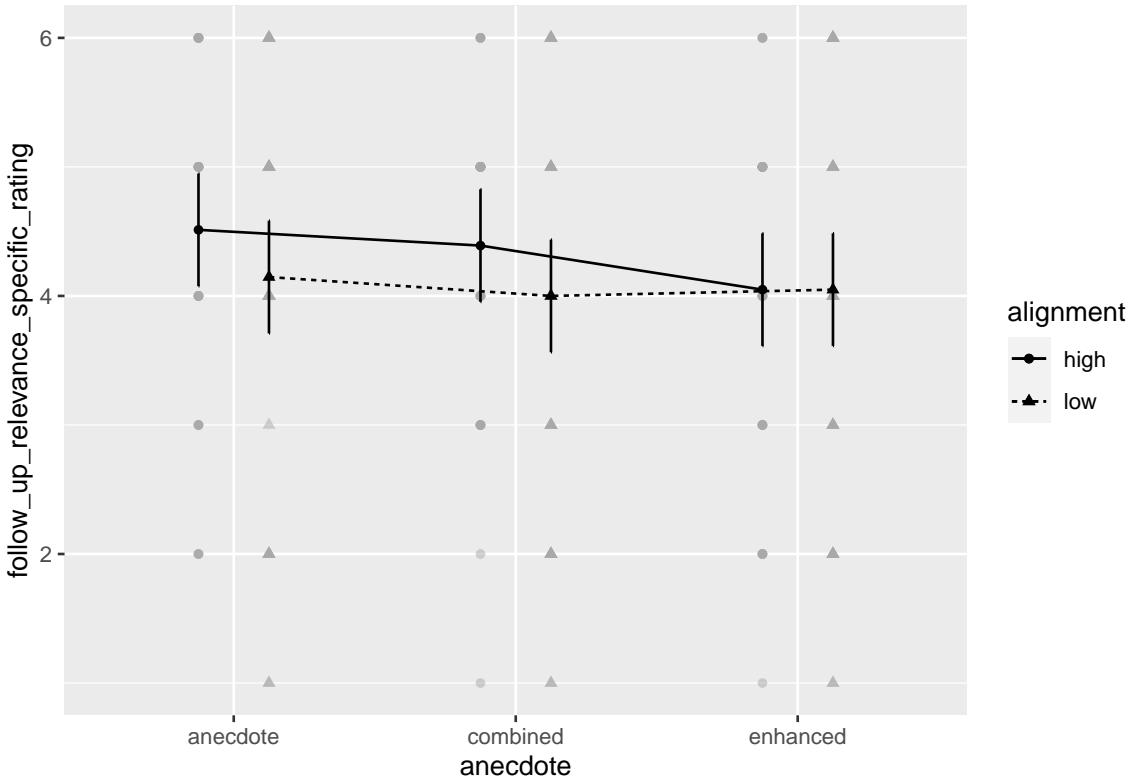


Figure C.9: Mean rating of how relevant participants thought the anecdote was to Project A (the target project). Error bars represent 95% confidence intervals.

Follow-up

Figure C.9 shows participants' ratings of the specific relevance question. I did not find a significant effect of evidence type $F(2, 240) = 0.80, p = .450, \hat{\eta}_p^2 = .007$; or similarity, $F(1, 240) = 1.94, p = .165, \hat{\eta}_p^2 = .008$. The interaction was also not significant, $F(2, 240) = 0.49, p = .615, \hat{\eta}_p^2 = .004$.

Figure C.10 shows participants' ratings of the general relevance question. There was no main effect of similarity, $F(1, 240) = 3.18, p = .076, \hat{\eta}_p^2 = .013$, or interaction of similarity and evidence type, $F(2, 240) = 0.24, p = .787, \hat{\eta}_p^2 = .002$. However, there was an unexpected main effect of evidence type, $F(2, 240) = 3.80, p = .024, \hat{\eta}_p^2 = .031$. A contrast analysis with Bonferroni correction revealed that there was a the anecdote only condition had significantly higher rating than the anecdote + statistics condition, $\Delta M = 0.57, 95\% \text{ CI } [0.06, 1.09], t(240) = 2.70, p = .022$. However, the difference between the two anecdote + statistics conditions was not significant, $\Delta M = -0.39, 95\% \text{ CI } [-0.90, 0.12], t(240) = -1.84, p = .202$.

C. Chapter 6 appendix

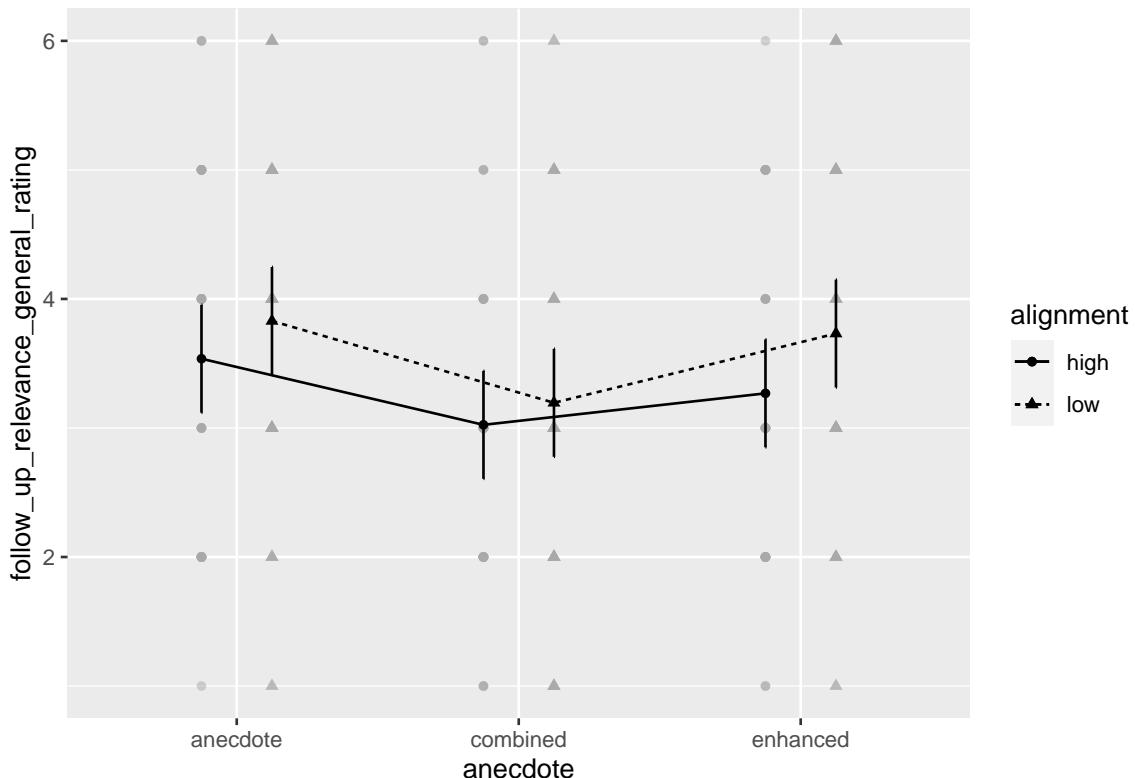


Figure C.10: Mean rating of how relevant participants thought the anecdote was to other oil projects. Error bars represent 95% confidence intervals.

I conducted regression analyses to determine the relationship between allocations and the follow-up ratings of similarity and relevance. As seen in Figure C.11, similarity ratings are negatively correlated to allocations, $b = -3.51$, 95% CI $[-5.67, -1.36]$, $t(244) = -3.21$, $p = .002$. Finally, as seen in Figure C.12 similarity ratings are positively correlated to specific relevance ratings, $b = 0.30$, 95% CI $[0.17, 0.43]$, $t(244) = 4.56$, $p < .001$.

C.2 Experiment 2

Hypothesis C.8—Similarity manipulation check for positive anecdote. In the positive valence condition, allocations for the anecdote-only high similarity condition will be higher than those in the anecdote-only low similarity condition.

I expected to replicate the rating effects found in Experiment 1 in the Experiment 2 negative valence condition, and to find the reverse effects in the

C. Chapter 6 appendix

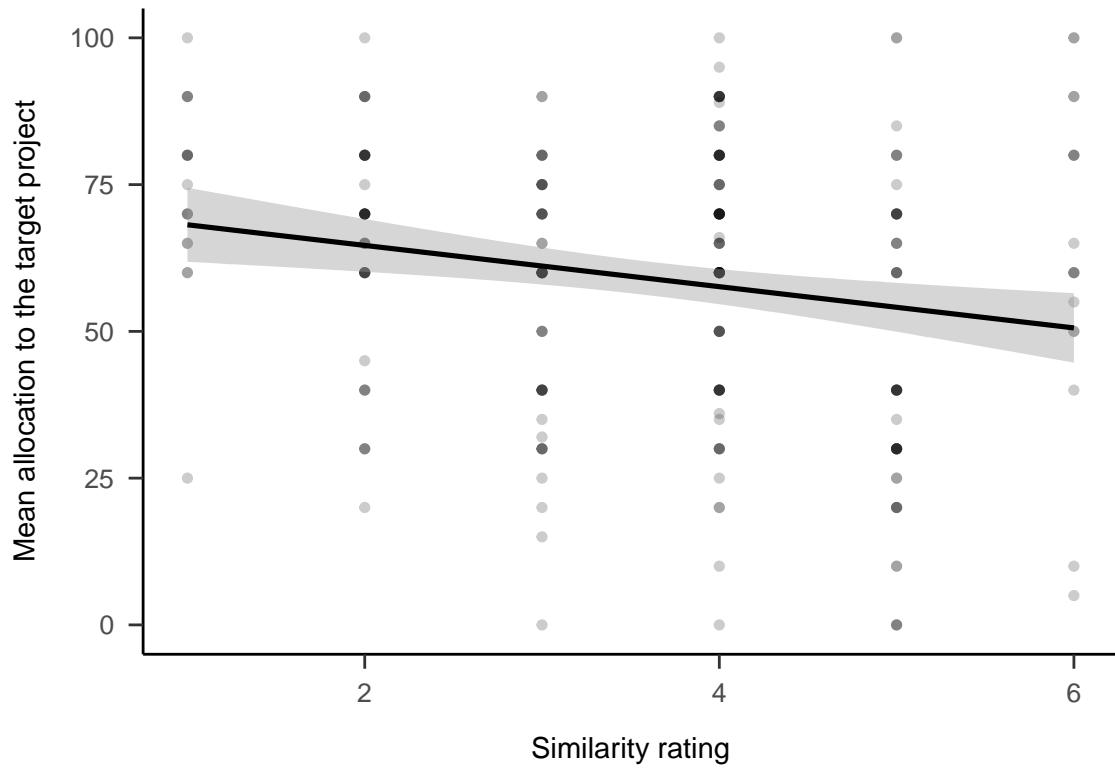


Figure C.11: Mean allocations to the target project by similarity rating

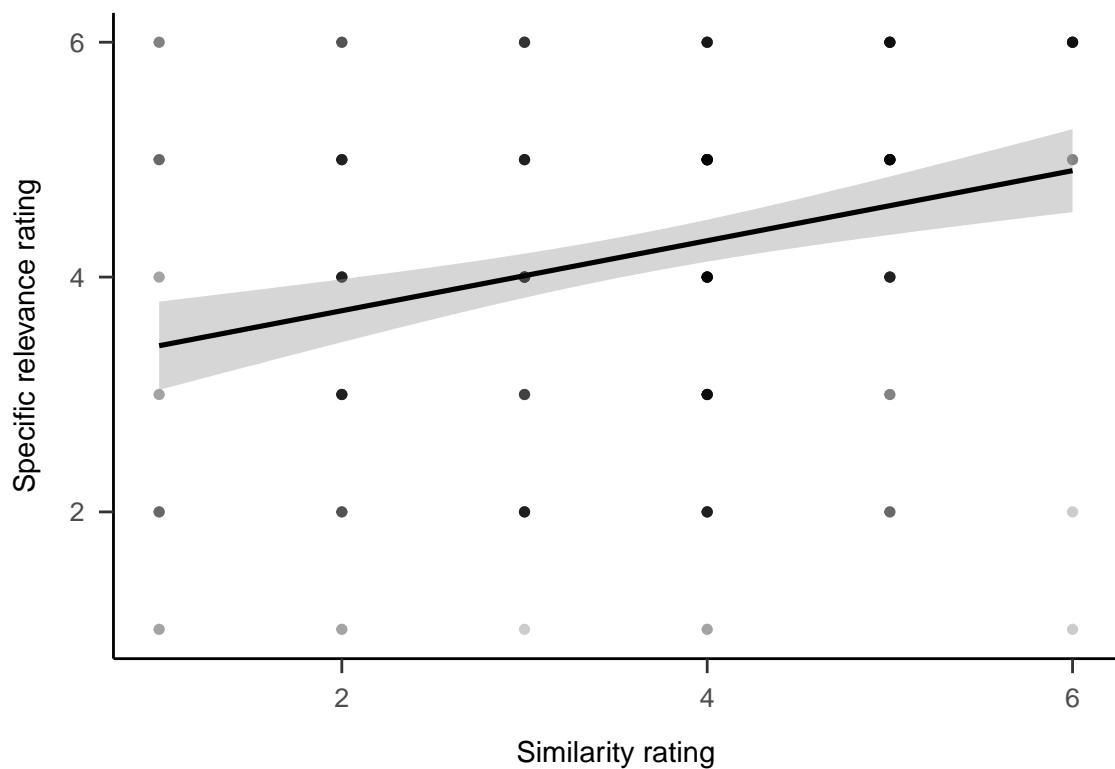


Figure C.12: Rating of how relevant participants considered the anecdote to the target project, by similarity rating

C. Chapter 6 appendix

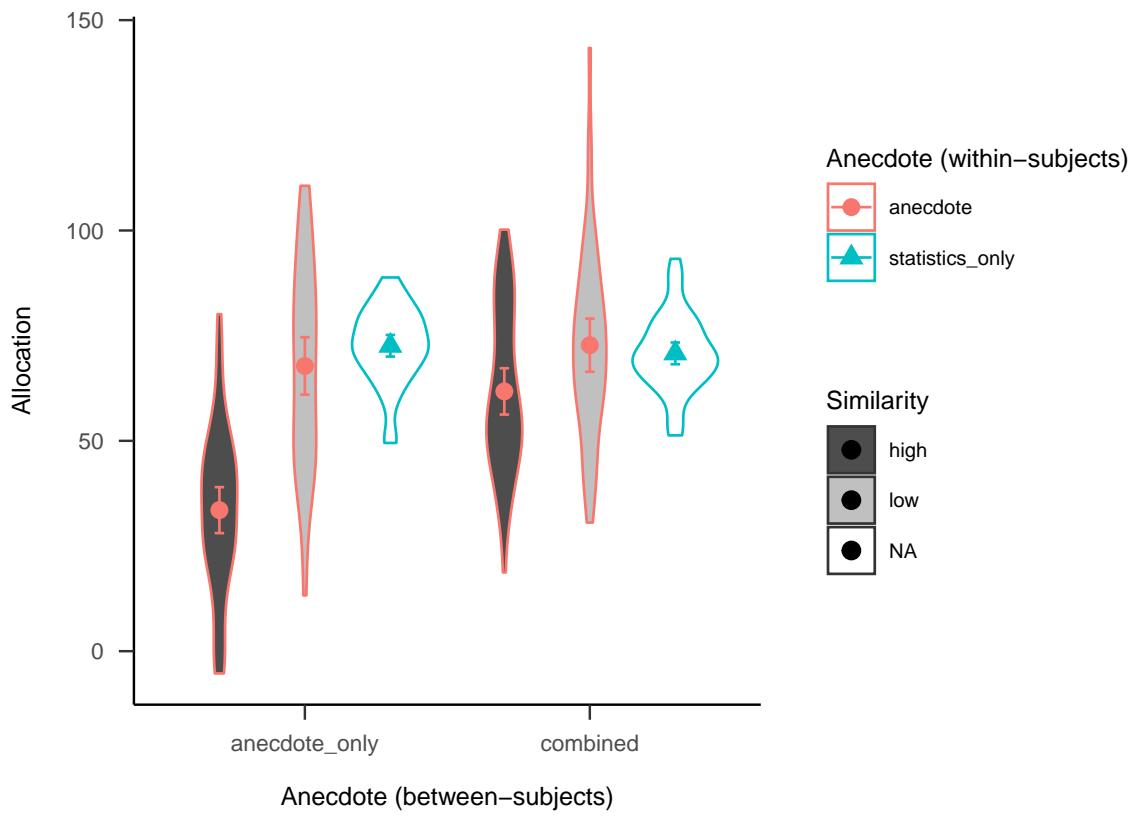


Figure C.13: Anecdotes Experiment 2 predicted data for the negative valence condition

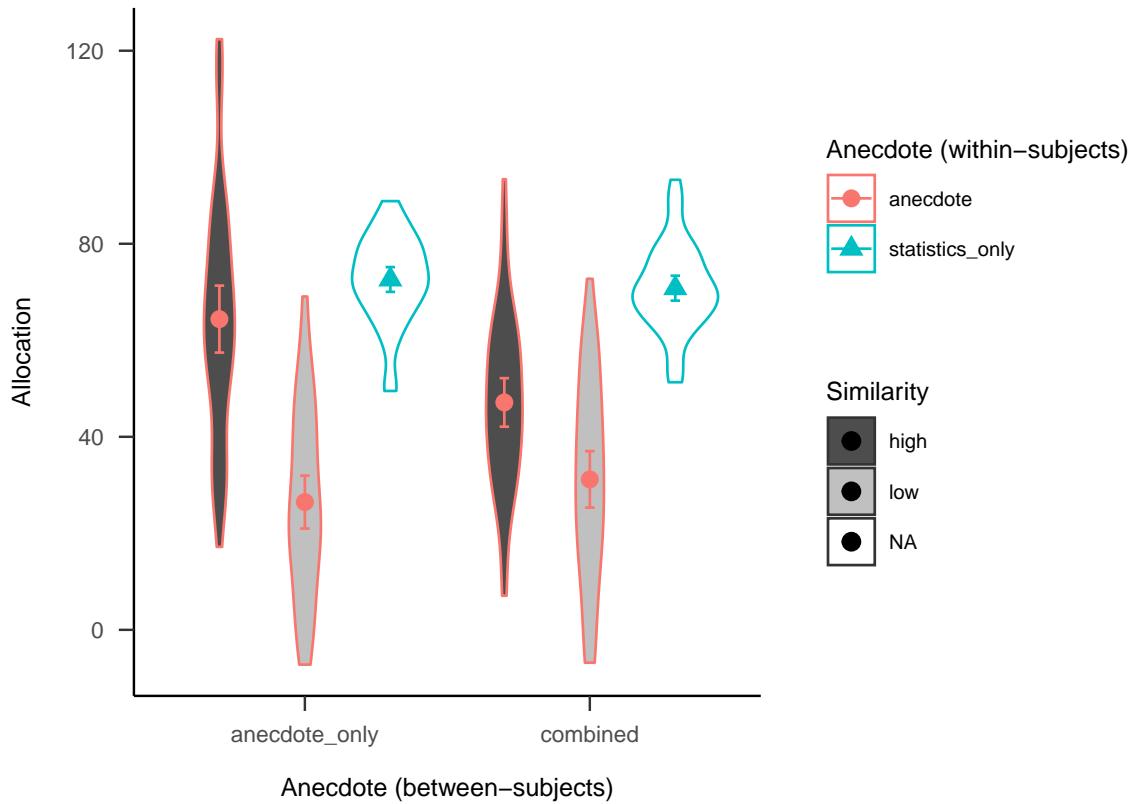


Figure C.14: Anecdotes Experiment 2 predicted data for the positive valence condition

C. Chapter 6 appendix

positive valence condition.

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

C.2.1 Method

Participants

Power analysis I conducted a power analysis through simulation of the effects implied by the hypotheses in [Experiment 2](#). I simulated data with the same mean values as Experiment 1 for the effects that were previously significant (i.e., similarity, statistics, and moderation effects), and no effect for the differences that were non-significant (as shown in Figures [C.13](#) and [C.14](#)). The null effect is 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. The analysis suggests a total sample size of 96 (48 × 2).

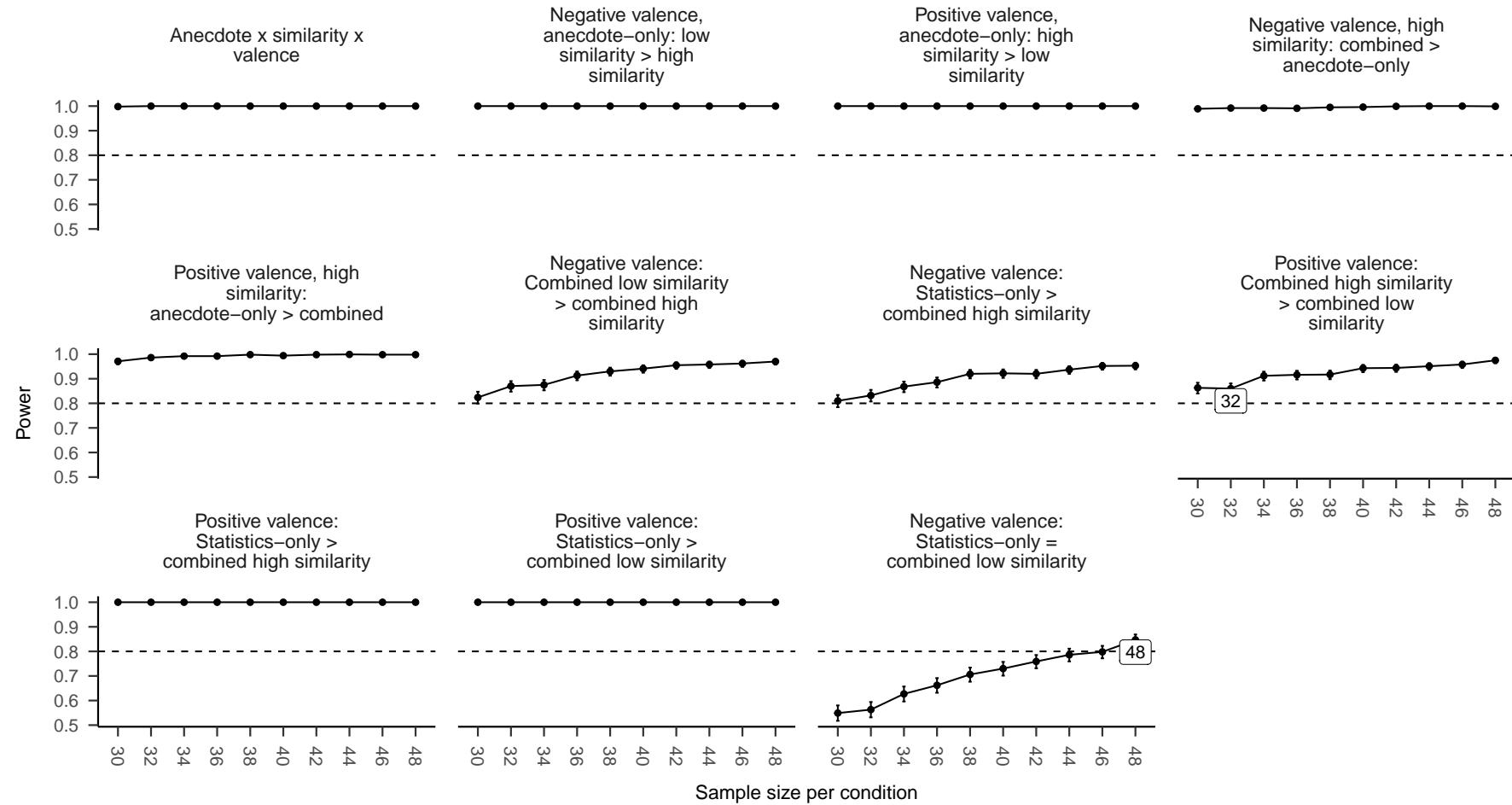


Figure C.15: Anecdotes Experiment 2 power curve. Labels indicate lowest sample size above 80% power.

C. Chapter 6 appendix

Instructions

Imagine you are an executive in a multi-business company and that you are presented with two projects to potentially invest in. Your job is to decide how to allocate the capital available in your budget between these two projects.

In total, you will see five of these project pairs (across five separate web pages). Each page will also contain relevant information about the projects.

Test yourself on the above instructions: How many pairs of projects will you see?

project pairs

Continue

Figure C.16: General instructions for Experiment 1. Border added for clarity.

Materials

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.

Allocation task The following were counterbalanced: 1. project variation (five latin square variations; the association of each display content with each within-subject condition); 2. anecdote variation (two variations), which is the association of each project display and being either the target or comparison project. The following were randomised: 1. table column order, and 2. project display order.

Follow-up questions Figure C.20 shows the follow-up questions.

Interstitial display

C.2.2 Results

Allocation

Similarity manipulation check The similarity manipulation worked as expected, with the negative valence anecdote-only low similarity condition being allocated significantly more than those in the high similarity condition, $\Delta M = -26.98$,

C. Chapter 6 appendix

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.

95% CI [−35.84, −18.12], $t(186.55) = -6.01$, $p < .001$. In the positive valence condition, those in the high similarity condition allocation significantly more than those in the low similarity condition, $\Delta M = 22.62$, 95% CI [13.77, 31.48], $t(186.55) = 5.04$, $p < .001$

Ratings

Similarity manipulation check Our similarity manipulation seems to have worked, with participants rating anecdotes in the high similarity conditions as more similar to the target than those in the low similarity condition, $F(1, 94) = 48.36$, $MSE = 2.26$, $p < .001$, $\hat{\eta}_G^2 = .082$.

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 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. Chapter 6 appendix

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.

C. Chapter 6 appendix

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.

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":

Figure C.21: An example of an interstitial display in Experiment 2.

C. Chapter 6 appendix

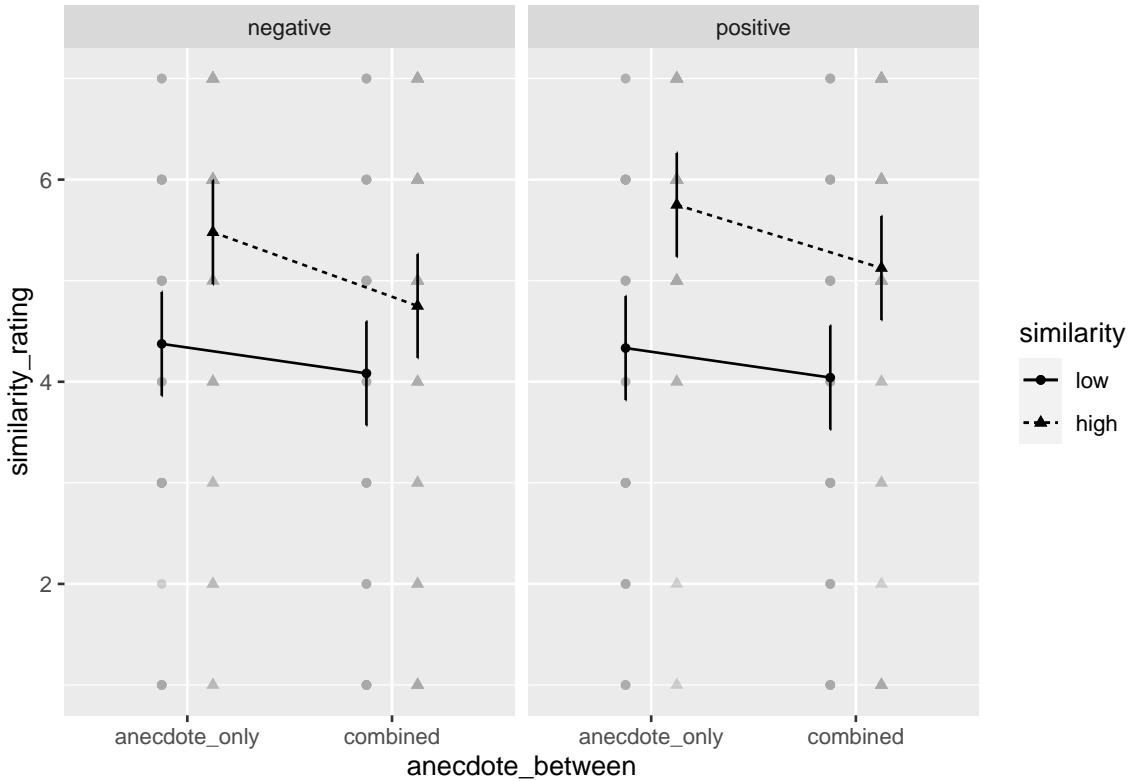


Figure C.22: Mean similarity rating of Project A (the target project) to the anecdote. Error bars represent 95% confidence intervals.

The relationship between allocation and specific-relevance is moderated by similarity In the negative valence condition, there was a 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 $[-1.53, 3.54]$, $t(376) = 0.90$, $p = .371$, whereas, in the low similarity condition, they were correlated negatively $\Delta M = -1.01$, 95% CI $[-3.54, 1.53]$, $t(376) = -0.90$, $p = .371$.

In the positive valence condition, there was a 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 $[-4.67, 0.42]$, $t(376) = -1.88$, $p = .061$, whereas, in the low similarity condition, they were correlated positively $\Delta M = 2.12$, 95% CI $[-0.42, 4.67]$, $t(376) = 1.88$, $p = .061$.

C. Chapter 6 appendix

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>
- 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>
- 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*. Retrieved February 24, 2021, from <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>
- 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>

References

- 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>
- Birshan, M., Engel, M., & Sibony, O. (2013). Avoiding the quicksand: Ten techniques for more agile corporate resource allocation [magazine]. *McKinsey Quarterly*, 6. <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>
- 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.
- 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.n97>
- 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.
- 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
- 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>
- 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

References

- 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>
- Cavagnac, M. (2005). Strategic managerial incentives under adverse selection. *Managerial and Decision Economics*, 26(8), 499–512. <https://doi.org/10/cwdbt>
- 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>
- 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. Retrieved July 24, 2020, from <http://search.ebscohost.com/login.aspx?direct=true&db=bsu&AN=91571432&zsite=bsi-live>
- 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/bx3wp>
- 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](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.
- 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>

References

- 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>
- 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.
- 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). A brief critical history of NPV. *BAA Conference*, 16.
- 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>
- 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>

References

- 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>
- 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>
- 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 [magazine]. *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., Svendsater, 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>
- 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>

References

- Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis*, 59, 41–51.
<https://doi.org/10/gjscr>
- 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.
- 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/gjscr>
- 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>
- 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/czwknn>
- 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/b5gfr>
- 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. Retrieved April 5, 2021, from
<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>

References

- 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.
- 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/bbrvh>
- 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 [magazine]. *McKinsey Quarterly*, (4), 15–17. Retrieved April 20, 2021, from <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>
- Kox, E. (2018). *Evaluating the effectiveness of uncertainty visualizations: A user-centered approach* (Masters thesis). University of Utrecht.
- 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.
- 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>
- 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 RESEARCH AND DEVELOPMENT ATLANTIC DARTMOUTH (CANADA). Retrieved May 4, 2021, from <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>
- 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>

References

- 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>
- 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., Uhlener, 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? [magazine]. *McKinsey Quarterly*, 3. <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/ftmkjt>
- 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.
- MacEachren, A. M. (1992). Visualizing Uncertain Information. *Cartographic Perspectives*, (13), 10–19. <https://doi.org/10/gjscq9>
- Mahoney, J. T., & Pandian, J. R. (1992). The resource-based view within the conversation of strategic management. *Strategic Management Journal*, 13(5), 363–380. <https://doi.org/10/cvv4cp>
- 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., & 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>

References

- 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>
- 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>
- 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.
- 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>
- 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>
- 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)
- 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. (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>
- Porter, M. E. (1998a). *Competitive Advantage: Creating and Sustaining Superior Performance*. Free Press. (Original work published 1985)
- Porter, M. E. (1998b, June 1). *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press. (Original work published 1980)

References

- 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>
- Rabin, M., & Weizsäcker, G. (2009). Narrow Bracketing and Dominated Choices. *American Economic Review*, 99(4), 1508–1543. <https://doi.org/10/fk4rmz>
- 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/gjscrn>
- 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. United States – New York. Retrieved April 6, 2021, from <http://search.proquest.com/docview/304937594/abstract/4B8166E9517494FPQ/1>
- 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>
- 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>
- 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.
- 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>

References

- Schwenk, C. R. (1984). Cognitive simplification processes in strategic decision-making. *Strategic Management Journal*, 5(2), 111–128. <https://doi.org/10/b2hsgw>
- 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>
- 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>
- Stutzer, M. (2013). Misperceptions of long-term investment performance Insights from an experiment. *The Journal of Behavioral Finance & Economics*, 3(1), 1–20.
- 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.
- 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>

References

- 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., Sedlmaier, M., & Möller, T. (2015). Decision making in uncertainty visualization. Retrieved May 4, 2021, from <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>
- 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>
- Vivaldi, E., & Coville, A. (2018). *How Do Policymakers Update?* http://barrett.dyson.cornell.edu/NEUDC/paper_573.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 (60th Anniversary Commemorative Edition)*. 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 Piercy, 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.
- 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>

References

- 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.
- 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>
- 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>