INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Skill, luck, and company performance: how the outcome bias shapes our perception of a firm's competence

Wouter Anton Hamers

Master in Management

Supervisor:
Prof. Dr. Francisco Guilherme Serranito Nunes, Assistant Professor,
ISCTE Business School

June, 2021

BUSINESS
SCHOOL

Department of Marketing, Strategy and Operations

Skill, luck, and company performance: how the outcome bias shapes our perception of a firm's competence

Wouter Anton Hamers

Master in Management

Supervisor:
Prof. Dr. Francisco Guilherme Serranito Nunes, Assistant Professor,
ISCTE Business School

June, 2021

Resumo

Estudos recentes mostram que o fator sorte tem um papel muito mais significativo no sucesso das empresas do que o esperado. As pessoas tendem a subestimar o papel da sorte, tanto na sua vida pessoal como nos negócios, graças ao viés do resultado (*outcome bias),* o que as leva a tomar em consideração os resultados de uma decisão quando avaliam a qualidade da mesma. Tal é problemático pois nos negócios a sorte pode fazer com que boas decisões deem origem a maus resultados e vice-versa. O presente estudo teve como propósito investigar a presença de um viés do resultado numa experiência conduzida com estudantes universitários, considerando um caso real em que uma empresa toma decisões arriscadas em situação de incerteza. Os participantes foram distribuídos aleatoriamente por três cenários de resultados (positivo, negativo e sem resultados) e foi lhes pedido que avaliassem a qualidade das decisões e a competência da empresa com base no cenário atribuído. A análise evidenciou que associar um resultado positivo a decisões de gestão resulta numa avaliação significativamente mais positiva das mesmas decisões do que se os resultados associados forem negativos. Contrariamente, a análise do grupo sem resultados originou o mesmo padrão que o grupo com resultados positivos, indicando que o grupo em estudo vê as decisões de gestão como positivas no cenário apresentado. No entanto, este grupo pode também estar a sobrestimar o nível de competências da empresa devido à complexidade da situação descrita. Será necessária pesquisa futura para testar esta hipótese. Concluindo, o estudo indica que as pessoas são capazes de julgar austeramente boas decisões de gestão quando conhecem que os resultados das mesmas são negativos e vice-versa.

Palavras Chave: Sorte; viés do resultado; tomada de decisão; performance organizacional; estratégia; psicologia

JEL Classification: D81, C91

Abstract

Research shows that luck plays a much larger role in determining company success than we often think. People are prone to underestimate the role of luck, in life and in business because of the outcome bias, which causes people to take outcome information into account when evaluating the quality of a decision. This is problematic because in business, luck can cause good decisions to lead to bad outcomes and vice versa. This study aimed to provide evidence of an outcome bias in a case study-based experiment conducted among university business students, considering a real-world situation in which a company makes risky decisions under uncertainty. Participants were randomly divided over three conditions (positive, negative and no outcome) and asked to evaluate the quality of the decisions and competence of the company. Analysis reveals that attaching a positive outcome to management decisions results in significantly better evaluations of decision quality compared to attaching a negative outcome. Contrarily, evaluation ratings were equally high in the no outcome version and the positive outcome version, indicating the people in group no outcome see the decisions as reasonable in the context given, but may also be overestimating the competence-level of the company due to the complexity of the situation. Further research is needed to confirm this interpretation. The outcome bias demonstrated in the study indicates that people may blame companies too harshly for making good decisions with negative consequences. Likewise, people may praise companies too much for making bad decisions that fortunately turn out well.

Keywords: luck; outcome bias; management decisions; company performance; strategy; psychology

JEL Classification: D81, C91

Table of Contents

List of Tables and Figures

1. Introduction

1.1 Background and relevance

Explaining why some firms succeed where others fail is perhaps the central question in management sciences. In the strategic management field, scientists like Michael E. Porter (Porter, 1979, 1980, 1985) and Jay Barney (Barney, 1991) provide theories about the determinants of company success and ways to gain competitive advantage. In essence, these theories imply that good management practices lead to competitive advantage and therefore, to superior performance. However, more recent research shows that luck has a much greater influence on the performance of a company than we often attribute to it. Cambridge Dictionary (2020) describes luck as "the force that causes things to happen, not as a result of efforts or abilities, but due to chance". Therefore, it is uncontrollable, unpredictable, and independent of the actions of a firm's employees. Mauboussin (2012, 2014) concludes that "great business success combines skill with a lot of luck". Studies performing simulations of (fictional) companies, like Raynor et al. (2009), conclude that finding patterns the natural variation of a firm's performance can create the illusion of skill, while in reality the firm was just lucky. Further supporting evidence was found in research on high performing firms profiled in popular management literature. Studies show that firms for which was claimed to possess the "blueprint for success" showed a significant drop in performance after being profiled in the literature, indicating that claimed success formula was merely an illusion (e.g. Kahneman, 2011; Rosenzweig, 2007). The aforementioned research is joined by a steady line of sources referring to luck as an explanation for performance differences between firms (e.g. Denrell, 2005; Liu 2019, 2020; Liu & de Rond, 2016; Wiggins & Ruefli, 2002). Because luck plays a role, little can be learned from high performance alone. Luck breaks the link between skill and results. The more influence luck has, the less well skill-level is reflected by financial performance (Mauboussin, 2012).

Considering that previous research already provides substantial evidence on the role of luck in company performance, this dissertation focuses on the cognitive bias which leads people to underestimate the role of luck, in life and in business, known as the "outcome bias". Outcome bias is the phenomenon in which evaluators tend to take information about the outcome into account when evaluating the quality of a decision itself (Baron & Hershey, 1988). This is problematic because the evaluator has available a different information set than the decision maker, who typically faces uncertainty. On top of that, luck can cause bad decisions to lead to good results, which makes outcome-based evaluations irrational. Previous research has already

shown that a general impression of a company (e.g. high performance) can influence one's opinions on other aspects of that company. A high performing firm will automatically appear to have a strong culture, a smart CEO and motivated employees. A low performing firm will appear to have the opposite. In this way, performance colors our perception of a company's abilities (Rosenzweig, 2007). Humans have the strong tendency to build coherent and satisfying cause and effect stories out of outcome information, even if it is not rationally justified (Gottschall, 2013). Outcome-based evaluations in complex decision-making, like business decisions, is irrational and should therefore be avoided at all times.

Research on the outcome bias has shown its influence on decision evaluations under various circumstances. Baron & Hershey (1988) and Emerson et al. (2010) proved its presence in medical dilemmas and monetary gambles. Gino et al. (2009) found an outcome bias in ethically relevant contexts, while Damnjanović et al. (2019) provides evidence in the parental decision-making context considering childcare dilemma's. In the management field, König-Kersting et al. (2020) and Marshall & Mowen (1993) have proven the presence of an outcome bias in decision evaluations of an individual financial agent and salesperson. All studies showed that the outcome of a decision influenced the evaluation of that decision, which is not rationally justified. While a lot has been studied already, the outcome bias has not yet been measured directly in a real-world-based case study in which a company makes risky decisions under uncertainty (a situation in which luck played a role). One could ask why we still need to study outcome bias in this particular situation when there is already broader evidence on the bias. Crusius et al. (2012) states that, especially in regard to cognitive biases, transferability from one situation to another cannot be taken for granted, even if they appear to be highly similar. Studying the outcome bias on a company level is important since people are exposed to stories of company success (and failure) frequently and need to be informed about their irrational tendencies. Even more so, university business students are learning about determinants of company success on a daily basis. Hence, it is important they become aware of the cognitive biases that create irrational beliefs about company success.

1.2 Research aim, research question & research objective

This study aimed to identify an outcome bias in a real-world-based situation where a company makes risky decisions under uncertainty. The research sought to provide evidence that participants give higher ratings to management decisions that resulted in positive outcomes than decisions resulting in negative outcomes. In addition, this study was seeking to investigate the relationship between the outcome of management decisions and the type of words individuals

use to describe these decisions. The study considered the following research question: "*How does the outcome bias influence university business students' evaluation of the managerial decision-making competence of a company which made risky decisions under uncertainty?"*

In order to identify an outcome bias, a case-study based experiment was conducted among master's business students from ISCTE Business School in Lisbon, Portugal. The real-life company under consideration was the Swedish-Swiss engineering firm ABB. Rosenzweig (2004) dedicated an article to their fluctuating performance, showing that luck played a substantial role in their financial results. The methods section will provide further information on this matter. Based on Rosenzweig's findings, a case-study was created describing ABB's management practices with different outcomes attached (positive, negative and no outcome). Then, the students were asked to evaluate ABB's management practices, which provided insights into how outcome information influences these evaluations.

The remainder of this dissertation is laid out as follows. Section 2 provides an in-depth overview of the existing literature on luck in company performance and the outcome bias. Section 2 also presents the conceptual framework for the experiment and the hypotheses. Section 3 describes the methods used to conduct the experiment and the data analysis techniques. Subsequently, section 4 presents the results of the experiment and section 5 discusses the theoretical contributions and practical implications. The last section sums up the most important insights of the research, including its limitations and suggestions for future research.

2. Literature Review

2.1 Traditional strategic management

For decades, academics have been interested in the determinants of company success and sources of competitive advantage. The main goal of strategic management research is to explain differences in performance among firms. Why do some succeed where others fail? The reigning explanation for the uniqueness of a firm's performance is based on the concept of competitive advantage (Wiggins & Ruefli, 2002). "Successful firms are seen as having capabilities, processes, and routines that set them apart" (Denrell, 2005). The concept of competitive advantage first appeared in the early work of Ansoff (1968) and was later popularised by the work of Harvard Business School professor Michael E. Porter, who focused on a firm's unique positioning in its competitive environment as a source of competitive advantage. According to Porter, there are five competitive forces at play in a firm's environment. (1) competition in the industry, (2) threat of new entrants, (3) power of suppliers, (4) power of customers and (5) threat of substitutes (Porter, 1979, 1980, 1985). A firm should analyse these forces and come up with a fitting strategy to it to gain competitive advantage.

About a decade later, American professor Jay Barney developed the "resource-based view (RBV)" as an alternative view on strategic management. RBV focuses on a firm's unique resources and capabilities as a source of competitive advantage (Barney, 1991). The academic work of Porter and Barney created a solid foundation for strategic management and the sources of competitive advantage. Therefore, it is no surprise that the majority of scholars, business press, business books, consultants etc. believe that a firm's management decisions and practices matter most when creating competitive advantage and superior performance. According to this view, successful firms posess capabilities, processes, and routines that set them apart from their competitors. The main assumption made is that good practices bring good results. Therefore, people are interested in learning about them and want to uncover their secrets (Ma, 2002; Denrell, 2005; Kahneman, 2013; Liu, 2020).

2.2 The role of luck in high performance

From the start of the 21st century onwards, more academic work started appearing on the influence of luck on the performance of firms. Jay Barney himself was one of the first to mention it in a paper written in 1986, describing a difference between the expected and the actual value of a firm's strategy as a "manifestation of a firm's good fortune and luck" (Barney, 1986). However, this study saw luck as an "error-term" and "by-product" of strategy. Luck certainly did not take center stage.

First, we need to define what "luck" actually is. The Cambridge Dictionary defines luck as "the force that causes things, especially good things, to happen to you by chance and not as a result of your own efforts or abilities" (Cambridge Dictionary, 2020). Skill is defined as "an ability to do an activity or job well, especially because you have practiced it" (Cambridge Dictionary, 2020) Skill can be identified on an individual level as well as on a company-level. Luck can have many different explanations and depends on many factors (Hafer & Gresham, 2008). However, for this research a general and more simple definition will be applied which is in line with Cambridge Dictionary's definition: luck happens, not as a result of efforts or abilities, but due to chance. Luck should be seen not as some sort of indefinable force, but as an event which is uncontrollable, unpredictable, and independent of the actions of the company's actors. Think of competition, government regulation, exchange rates, economy, but also more simple events like happening to be "in the right place at the right time". Skill, however, is in a firm's control and can be practiced.

The role of luck in company performance has been studied extensively over the last decades. Perhaps the most influential book on the role of luck in business in recent years is "The Success Equation: Untangling Skill and Luck in Business, Sports, and Investing" from Michael J. Mauboussin (Mauboussin, 2012). Mauboussin explains that "great business success and gaining competitive advantage combines skill with a lot of luck. You can not get there by relying on either skill or luck alone. You need both" (Mauboussin, 2012, 2014). Being sufficiently skilled is important as a basis for success. For example, Gompers et al. (2006) show that serial entrepreneurs who have experience in starting a business have higher chances of succeeding than first-time entrepreneurs due to their network, knowlegde, and experience in the market. Nevertheless, still only 30% of the serial entrepreneurs will succeed in their next venture (Gompers et al. 2006), indicating the presence of luck. Mauboussin states that luck breaks the link between skill and results. When luck has little influence, a good process will always have a good outcome (e.g. chess or olympic sprinting). When luck has large influence, good processes will only occasionally bring good outcomes. In that case it is very reasonable to expect that a different outcome could have occurred, either positive or negative. Thus, when luck plays a role, outcome does not reflect the decision quality or skill very well. One should focus more on the process followed towards that outcome instead (Mauboussin, 2012). Furthermore, he points out that when several companies within a given industry all become more skilled (e.g. by copying best practices of others), luck starts to play a bigger role. Mauboussin calls this the "paradox of skill". Nowadays, "absolute skill" is at the highest level its ever been, meaning that there are more companies with a high level of skill. However,

"relative skill" has never been narrower, meaning that the companies with high skill levels are much closer/ similar to each other than they have been in previous decades. When skill levels are so similar, more is left for luck to make a difference (Mauboussin, 2012). Similarly, Liu & de Rond (2016) describe severe competition between companies as "an important but often neglected source of randomness." Competition leads to randomness because it removes obvious opportunities and equalizes expected returns.

The best performing companies are often the ones that are the most admired for their way of doing business. Wiggins & Ruefli (2002) showed that (1) a very small minority of firms enjoy superior economic performance, and (2) it very rarely persists for long time frames. Over the years, there have been several studies trying to explain why these superior companies are so successful. The most popular examples include "In Search of Excellence" (Waterman & Peters, 1982), "Built to Last" (Collins & Porras, 1996) and "Good to Great" (Collins, 2001). The approach of their research was generally the same: they started with a population of firms and identified the most successful among them. Then they examined their management practices and looked for patterns (critical success factors) associated with that success. Finally, they distilled those patterns into a framework and claimed that if the reader of the book used that framework to steer their behaviors they can achieve those same results (Liu, 2009; Mauboussin, 2012; Raynor et al., 2009).

As it turns out, the majority of successful companies featured in popular management literature show a significant drop in performance in the period after being profiled in the literature. Due to the role of randomness and situational factors, their perceived "greatness" could be explained by the fact that companies naturally experience variations in performance over time. (Raynor et al., 2009; Kahnemann, 2011; Mauboussin, 2012; Pluchino et al., 2018; Liu, 2020). Research on the high-performing companies mentioned in the three management books show that only one in four were remarkable, the rest were mediocre firms catching lucky breaks. Rosenzweig (2007) points out that for "In Search of Excellence", only one-third of the companies grew faster than the overall stock market over the next five years, the rest failed to keep up. Of the 35 companies for which data was available, only five improved their profitability, while 30 declined. The authors of "Built to Last" claimed to provide "a master blueprint for building organizations that will prosper long into the future." This turned out to be a delusion. The performance of the 18 "visionary" companies regressed sharply in the five years after the study ended. Also the results presented in "Good to Great" turned out to be seriously flawed (Rosenzweig, 2007). Finding patterns in randomness can create the illusion of skill. The majority of the firms just got lucky (Raynor et al., 2009; Kim et al, 2016). The books

all claimed to provide a "blueprint for success". The results however show this was merely an illusion. What really happened to these failing firms was a strong "regression to the mean". "Regression to the mean is a statistical phenomenon that can make natural variation in repeated data look like real change. It happens when unusually large or small measurements tend to be followed by measurements that are closer to the average" (Barnett, Pols, & Dobson, 2005). High performing firms are likely to become predictably worse because their performance deviates from the average. The more a firm's performance deviates from the average, the more likely it is performance will start moving towards the average (Liu, 2019).

To visualize how easily we can be led astray when considering a complex system, like the performance of a company in an industry, Raynor et al. (2009) conducted a simulation. They considered 20 years of performance for 100 firms, in which each year's performance equals the previous year's performance plus a random "bump" (luck), which can be positive or negative. The 100 firms all start with performance level 0, and the "bump" is normally distributed ($M = 0$ and $SD = 1$). The simulation is known as a "random walk". Figure 2.1 shows the results for six of the 100 firms. Our intuition suggests there are three "big winners" and three "big losers". However, all firms in the simulation were identical, and any differences are exclusively consequences of randomness. Therefore, there is nothing to learn about individual competences of these firms.

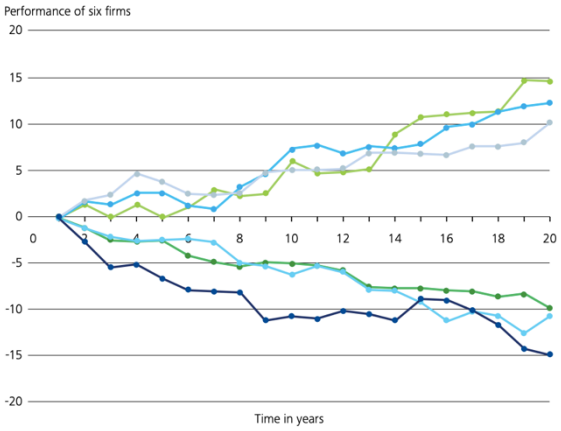

Figure 2.1 The illusion of greatness (Raynor et al., 2009)

Exceptionally high firm performance can even be an indicator of managerial incompetence and risk-taking. High performance is a rare event, and therefore deviates from the mean of performance, which implies it has high variance. "For example, if we list the firms with the highest return on equity in any given year, which are the firms that are most often on the cover of business magazines, the list will probably be dominated by firms from industries with a high variance in returns" (Denrell, 2005). Taking risks brings high variance, either negative or positive, but the firms with poor performance will be more likely to fail and to be removed through bankruptcy or by exiting the industry. After all, 50% of businesses fail in their first five years (VentureBeat, 2019). Thus, any sample of active firms will be biased toward successful firms. As a result, the observed association between the risky practice and performance will be positive. Therefore, if managers or consultants want to learn about the determinants of high performance by studying existing firms, their conclusions may be highly

misleading (Denrell, 2005; Denrell & Liu, 2012). As Mauboussin (2012) points out: "The trouble is that the performance of a company always depends on both skill and luck, which means that a given strategy will succeed only part of the time. So attributing success to any strategy may be wrong simply because you are sampling only the winners. The more important question is: how many of the companies that tried that strategy actually succeeded?" On top of that, Denrell (2005) points out that luck is cumulative, meaning that if the strategy succeeds, it will bring advantages over a longer time period in the form of resources, which will result in even more resources the next years, etc. This creates a snowball effect of positive financial results, all caused by one lucky break, not necessarily by superior capabilities.

All aforementioned research provides evidence that luck plays a substantial role in the success of a company. Because luck plays a role, financial performance does not necessarily reflect the skill-level of a company. As a counterargument to luck, people often say that "chance favors the prepared mind" and, "the harder I work, the luckier I become" (Liu & de Rond, 2016). While this is true to some extent, there is no way to improve your luck, because anything you do to improve a result can reasonably be considered skill (Mauboussin, 2012). Indeed, a company can adopt traits like an open-minded view and an agile organizational structure to maximize their return when luck occurs, but that still does not guarantee that good luck will befall on you.

2.3 Psychology behind our interest in high performing companies

This section discusses why people have the tendency not to recognize luck, and why we are so extremely interested in high performance. Coming back to "In Search of Excellence", "Built to Last" and "Good to Great". What exactly caused the conclusions of these management books to be so incorrect? After all, years of extensive research was done trying to explain why these companies were so successful. Rosenzweig (2007), author of the book "the Halo effect" provides an explanation. Essentially, the three books made the same mistake. They selected the companies *based* on their financial performance, so based on the dependent variable, which is outcome. Therefore, their findings regarding management practices (the "blueprint" for success) were all biased by this performance. They relied on data that was not independent of performance. Apart from that, the studies gathered enormous amounts of data, which created the illusion of rigourous research. They did not catch the causes of high performance at all, but rather attributions based on performance. "Pick any group of highly successful companies and look backwards, relying on articles in the business press and on retrospective interviews, and you may well find that they are said to have strong cultures, solid values, and a commitment to

excellence. Pick a group of comparison companies that are good but not outstanding, and they are likely to be described in somewhat lesser terms." When a company is doing well, with rising sales and high profits, observers infer that is has a smart strategy, a visionary leader and motivated employees. When sales fall and profits shrink, many are quick to conclude employees became complatent, customers were neglected etc. (Rosenzweig, 2007a, 2007b). This phenomenom is called the Halo effect, originally identified by Edward Thorndike in 1920. A positive impression in one area (high performance) influences one's opinion in other areas (management practices).

The Halo effect occurs frequently. Rosenzweig concludes that stories of success and failure consistently exaggerate the impact of leadership styles and management practices on firm outcomes, and thus their message is rarely useful. One of the examples can be found in studies of CEO turnovers. One study found that CEOs were frequently fired for factors well outside their control, e.g. during recessions or when the industry is suffering (Jenter & Kanaan, 2015). A negative impression (bad financial performance) influences one's opinion on the skill-level of the CEO. Additionaly, there are multiple examples of companies which were praised as long as performance was high, but when performance fell, everything was suddenly seen in a different light. Examples include Dell, Cisco Systems, IBM and ABB (Rosenzweig, 2007). Rosenzweig even dedicated a full article to ABB's specific situation (Rosenzweig, 2004). This article will be elaborated later in this dissertation. Similarly, Mintzberg et al. (1996) conducted a study on Honda's success in the US market for motorcycles. They found that the consultancy firm BCG wrote a report trying to explain Honda's success. BCG's main belief was that Honda had a pre-determined and well-executed strategy, while Honda's executives admitted that at first, they had no clue what they were doing, which resulted in taking huge risks that happened to turn out well. The Halo effect is proven in experimental studies as well. Staw (1975) proved that people attribute different sets of characteristics to groups they believe are high or low performers. The groups had done just as well, but telling participants about performance was enough to make them believe that the high performers were more cohesive and better in communicating, and the low performers were the opposite. Likewise, Meindl & Ehrlich (1987) found that the words used to describe leaders were highly dependent on the company's performance. In successful companies the leader was seen to be visionary, charismatic and a good communicator, in bad times the same leader appears to be confused, indecisive or arrogant. Lastly, in large-scale surveys like the Fortune magazine's survey of corporate reputations the Halo effect is also present. Here financial performance significantly influenced perceptions of firm quality (McGuire et al., 1990).

The Halo effect is very closely related to, although not the same as, the psychological phenomenon known as *the outcome bias*. The outcome bias is an error that occurs in evaluating a decision when the outcome of that decision is already known (e.g. Baron & Hershey, 1988; Damnjanović et al., 2019; König-Kersting et al., 2020). People tend to take outcome information into account in a way that is not rationally justified. Outcome-based evaluations are considered irrational, because information that is available only after a decision is made is irrelevant to the quality of the decision (Neumann & Morgenstern, 1945). After all, the decision maker does not know the outcome at the time of the decision and typically faces uncertainty. To a certain degree however, judging decisions by their outcomes can be a useful heuristic, but only when the situation allows it. It is intuitive and it saves time and energy. At the end of the day, it is easier to draw conclusions from outcome than to conduct a thorough research (Kahneman, 2011).

However, Baron & Hershey (1988) point out that we have to be careful not to overgeneralize evaluating decisions based on their outcome to situations where it is inappropriate, like complex decision-making for example, in which randomness and situational factors play a big role. To illustrate a complex decision, consider a doctor who has to decide if he is going to perform surgery on a patient. There is a 50% chance the patient will die, and a 50% chance the patient will live. The doctor decides to go ahead with the surgery. Baron & Hershey (1988) show that in case the patient dies, evaluations of the decision are much more negative than in case the patient lives, while the circumstances in which the decision is made are identical. Even when the patient made the decision on his own, the outcome bias was still present. Consider another example: a person under the unfluence of alcohol decides to drive home in his car. Obviously a bad decision. However, he gets home safely and nobody gets hurt. In this case a bad decision is followed up by a good outcome. The outcome bias will create the belief that drunk driving is not that bad, and he will most likely do it again. In this manner, near-disasters can be interpretet as successes (Dillon & Tinsley, 2008). Outcome-based evaluations are irrational because it is impossible for the decision maker to know the outcome at the time of the decision. Hence, “reasonable decisions are often criticized by Monday-morning quarterbacks who think they might have decided otherwise, and decision makers end up being punished for their bad luck” (Baron & Hershey, 1988). Even though we should not judge based on outcome, as soon as our brain processes the outcome information, it tries to build a coherent and satisfying story around it. Humans have a constant tendency to look for cause and effect relationships. As Gotschall (2013) states: “the storytelling mind is allergic to uncertainty, randomness, and coincidence. It is addicted to meaning. If the storytelling mind

cannot find meaningful patterns in the world, it will try to impose them. The storytelling mind is a factory that churns out true stories when it can, but will manufacture lies when it can't". Additionaly, Parnell & Dent (2009) and Taleb (2007) identified that people are more likely to assign luck for bad outcomes rather than good. Success is attributed to character, hard work, preparation etc., while failure is attributed to factors outside of one's control. Decisions with positive outcomes are seen as causal, when they were actually lucky. However irrational and wrong, this cognitive bias also has a positive (evolutional) side. It positively influences people's motivation, esteem and willingness to keep taking risks (Liu & de Rond, 2016).

Because people take outcome information into account, they are likely to underestimate the role of luck (randomness and situational factors) in company performance. For that reason, "the halo effect and outcome bias combine to explain the extraordinary appeal of books that seek to draw operational morals from systematic examination of successful businesses" (Kahneman, 2011). Considering outcome information in complex decision-making situations where errors are costly (like a company's management decisions) is irrational and should be avoided. Nevertheless, research shows that people tend to rely more on outcome information when the complexity of the decision-making process rises (Damnjanović et al., 2019; Kahneman, 2011).

Rosenzweig (2007) emphasizes that the business world is not a place of clear causal relationships, where a given set of actions leads to predictable results, but one that is more tenuous and uncertain. "The task facing executives is to gather appropriate information, evaluate it thoughtfully, and make choices that provide the best chances for the company to succeed, all the while recognizing the fundamental nature of uncertainty in the business world" (Rosenzweig, 2007). While success studies can certainly help in creating a competent business and can act as a source of inspiration and motivation, they should not be treated as how-to manuals for becoming exceptionally successful (Raynor et al., 2009; Kim et al, 2016). "We should stop showing our students how a limited number of stars have risen to levels that others are unlikely to achieve" (Liu & de Rond, 2016). As mentioned before, when luck plays a role, financial performance does not reflect skill-level very well. Therefore, little can be learned from high or low performance alone. Being aware of this fact can help people and companies to make better and more rational decisions.

2.4 Research on the outcome bias across disciplines

Decision making is an essential human cognitive process, occurring on a daily basis. Decisions can be relatively simple, like choosing what kind of clothes to wear or what to have for dinner.

They can also be more complex, like medical dilemmas, risky business decisions or choosing what kind of career to pursue. Every decision we make has its corresponding outcome(s), which can either be positive or negative. Earlier in this section we have already seen that people's evaluation of a decision as "good" or "bad" is highly dependent on the outcome of the decision. Early research on this matter was conducted by Baruch Fischhoff in 1975 on the phenomenon "hindsight bias". Fischhoff discovered that people consistently exaggerate what could have happened in foresight. Meaning when someone finds out that an outcome has occurred, it increases the perceived likelihood of that outcome, making the judge think this this was the only outcome that could have occurred. In addition, the judges of the decisions are unaware of the effect that outcome knowledge has on their perceptions, creating the tendency to believe that others should have been able to anticipate outcomes much better than was actually the case. People even misremember their own predictions in this way, exaggerating in hindsight what they knew in foresight. The individuals are simply unable to recover an uninformed state of mind (Camerer et al., 1989; Fischhoff, 1975). This is what Tversky and Kahneman refer to as "creeping determinism" (Tversky & Kahneman, 1974).

Baron & Hershey (1988) were the first to come up with the label "outcome bias". They proved there is a *direct effect* of outcome on evaluation of decisions by conducting five experiments on medical dilemmas and monetary gambles (Baron & Hershey, 1988). Consistent outcome bias was found in all five experiments, even in cases where the probability of outcomes was known to the participants. One of their experiments concerned the medical dilemma described in the previous section, clearly showing an outcome bias. They also found that, even in experiments where chance outcomes are determined by a roulette wheel, making it completely random, an outcome bias was present (Baron & Hershey, 1988)..

Further research on the outcome bias has been conducted in a variety of scientific fields. König-Kersting, et al. (2020) conducted experiments in which an agent makes risky financial decisions for a principal. The study showed that the principal's evaluations and financial rewards for the agent were strongly affected by the random outcome of the investment, showing strong outcome bias. A random success in which the agent did not follow the principal's request resulted into a better evaluation than a failed investment that followed their explicit request (König-Kersting, et al., 2020). Three decades earlier, Marshall & Mowen (1993) came to a similar conclusion after conducting an experiment in which a salesperson had to decide to pursue one of two possible companies concerning a large sales order. They found that solely varying the outcome of the decision in the scenarios substantially influenced the evaluation of the salesperson, displaying the presence of an outcome bias (Marshall & Mowen, 1993). In the

medical field, Emerson et al. (2010) have shown that medical experts judge two identical manuscripts differently when only the outcome differs. Both versions had five puposefully placed errors in them. However, the experts believed to have found more errors in the negative version than in the positive version (Emerson, et al., 2010). Gino et al. (2009) studied the existence of outcome effects in ethically-relevant contexts, providing strong evidence that people judge the ethicality of others' behavior differently depending on the outcome. They also state that the outcome bias in ethical domains is driven by an intuitive impulse, not by rationality (Gino et al., 2009). As a last example, Damnjanović et al. (2019) tested the presence of the outcome bias in the parental decision-making context using childcare dilemmas. Along with other studies, they concluded that when the complexity of a decision rises, the outcome bias becomes stronger. On the contrary, when an individual is more involved in a decision, it yields a weaker outcome bias than in low-involvement decisions (Damnjanović et al., 2019).

While a lot of research on the outcome bias has been done already, it has not yet been measured directly in a real-world-based case study in which a company makes risky decisions under uncertainty (a situation where luck played a role). One could ask why we still need to study outcome bias in this particular situation when there is already broader evidence on the bias. Crusius et al. (2012) states that, especially in regard to cognitive biases, transferability from one situation to another cannot be taken for granted, even if they appear to be highly similar. Studying the outcome bias on a company level is important because people are exposed to stories of company success (and failure) on a daily basis. It is particularly important for business master's students, since their main focus is to understand what kind of practices makes a company successful. Therefore, it is crucial to become aware of the cognitive biases that create irrational beliefs about the determinants of company success.

2.5 Conceptual model and hypotheses

The above-mentioned literature all proved an outcome bias of some sort, while also creating a solid conceptual model for future research. The studies all used a highly similar model to prove the presence of an outcome bias. Usually participants are presented with an identical decision in an identical situation, once followed by a positive and once by a negative outcome and asked to evaluate the decision, most often on a scale (e.g. ranging from 1 - very bad to 7 - very good). An outcome bias has occurred when the calculated difference between the mean evaluations of decisions followed by a positive and a negative outcome is statistically significant (Baron & Hershey, 1988; Damnjanović et al., 2019).

Oucome bias studies aim to show that outcomes of decisions lead people to see the decisions themselves in a different light (Gino et al., 2009). Most outcome bias studies however, are missing a "control situation". This situation is necessary to determine the "base value" of a certain decision. Because we use a scale from 1 to 7, it is assumed that the number 4, that is- the middle position (decision is neither good nor bad) is the base value which we compare the results to. However, if we ask a group of participants to evaluate the decisions without knowing the outcome, we can actually determine the real base evaluation value. Since most decisions are considered bad or good to start with, the base evaluation value of a decision does not necessarily have to be 4. Comparing results to the control group also allows us to see if evaluations are skewed more by positive or by negative outcomes, which can provide insight into the relative contribution of different outcomes on the evaluation of the decision (Damnjanović et al., 2019).

The aim of this research was to identify an outcome bias in a real-world-based situation where a company makes risky decisions under uncertainty. Based on previous research, a case study was created in which the outcome of the decisions (financial performance) was proven to be influenced by randomness and situational factors (luck). Hence, no solid claims can be made about the skill-level of the firm. By identifying an outcome bias, this study showed how outcome information influences the perceived skill-level of the company. The following hypotheses were tested:

Hypothesis 1a: Evaluation ratings of management decisions are higher when the outcome of a decision is positive, compared to when the outcome is negative or when no outcome is given.

Hypothesis 1b: Evaluation ratings of CEO skill are higher when the outcomes of his behavior are positive, compared to when the outcomes are negative or when no outcome is given.

Hypothesis 1c: Evaluation ratings of the future decision-making capability of the company are higher when the current outcomes of the decisions are positive, compared to when the current outcomes are negative or when no outcome is given.

Hypothesis 2: A negative outcome creates a stronger relative outcome bias effect than a positive outcome.

Hypothesis 3: The type of words attributed to the company's management decisions are more positive when the outcomes of the decisions are positive, compared to when the outcomes are negative or when no outcome is given.

3. Methods

3.1 Participants

A total of 98 master's students from ISCTE Business School in Lisbon participated in the experiment. The experiment was conducted during three online video classes at ISCTE and served as an introductory exercise for the students. On average participants were 22,5 years old (*SD* = 1.21). 57,1% were female and 82,7% of Portuguese nationality. The remaining 17,3% of participants were Chinese, French, German, Slovenian, Indian, Polish, Indonesian and Croatian. 92,9% of the participants followed a MSc in Management at ISCTE. Annex E shows the summary statistics of participants in greater detail.

3.2 Research design and procedure

Participants were asked to read a short case study and answer a set of questions. The case study used in the experiment was based on Philip Rosenzweig's article about the company ABB (Rosenzweig, 2004). Between 1988 and 2000, ABB's performance was exceptional compared to its competitors and it consistently ranked among the top companies in the world. But then, between 2000 and 2004, performance faltered and ABB struggled. However, Rosenzweig discovered that ABB's management practices and way of approaching business remained exactly the same, implying that external, uncontrollable influences (luck) significantly affected ABB's financial performance over this 16-year period. There were also no signs of complacency within ABB, considering that great efforts were made to adapt to the (ever-changing) market. Management practices stayed the same, financial performance changed. As a result, business press, newspapers and scholars started to see ABB's practices in a bad light. The same practices they used to explain ABB's rise between 1988 and 2000, were also used to explain their fall between 2000 and 2004. Because luck influenced ABB's financial performance, in reality no solid claims can be made about the quality of their management practices.

Firstly, the case study described the management practices of company "WH" by identifying several strategic decisions made in times of uncertainty. Leadership characteristics were also described. All names were changed to prevent respondents recognizing the real-life situation of ABB. As Baron & Hershey (1988) point out, it is essential to provide a detailed description of the situation in which the decisions were made, as participants may feel that describing a complex situation with too much simplicity is implausible. Therefore, descriptions were made in a detailed and objective manner, only stating factual information.

An online questionnaire in the Google Forms platform was used to gather the data. The experiment was conducted in March 2021. Inspired by previous research on the outcome bias in different scientific contexts (Baron & Hershey, 1988; Damnjanović et al., 2019; Emerson, et al., 2010; Gino et al., 2009; König-Kersting et al., 2020; Marshall & Mowen, 1993), an experimental design was used, randomly assigning participants to three different groups. Each group read a different version of the case study, using a between-subject design to avoid the participants noticing there were different versions. All versions were *identical* except for the last section, which varied in outcome. The outcome was either positive ($N = 33$), negative ($N = 33$) or no outcome was given ($N = 32$). The group which received no outcome served as the control group. In the positive version, the last section stated that the company had excellent financial results. Revenues increased, net profit tripled, and the company was worth tens of billions of dollars. Oppositely, in the negative version, participants read that sales growth had faltered and profits decreased, a financial crisis had hit and debts reached troubling heights. All versions of the case study can be found in annex A. Then, individuals were asked to evaluate individual management decisions, CEO skill, and future decision making competence of the company. The dependent variable was the evaluation of each decision, made by the participants on a scale from 1 (terrible decision) to 7 (excellent decision). Additionaly, in the last question participants were asked to "select the words you feel are most suitable to describe WH's management practices", having to choose between the words "visionary", "inspiring", "excellent", "courageous", "arrogant", "unfocused", "incompetent", and "chaotic". Individuals were asked to form their answers purely based on the information available to them. No further instructions were given, aiming to capture the participants' natural opinion. Decisions were made anonymously with no communication between the respondents. The research question was not revealed to participants at any time before or during the experiment. Annex B presents the questions in the experiment in greater detail.

3.3 Data analysis techniques

The power analysis was used to calculate the statistical power of the experiment. The probability of detecting a statistically significant outcome bias effect by omnibus F-test ($\alpha = 0.05$) of the effect size of 0.7, as reported by Baron & Hershey (1988) and Damnjanović et al. (2019), for a sample of 98 subjects amounts to 99.9%. The quantatative data was analyzed through SPSS version 26 and Excel 2020. For all statistical tests a significance level of 0.05 was applied.

Hypotheses 1a, 1b, 1c and 2 concern measuring the susceptibility to the outcome bias. The outcome bias was measured by analyzing the mean differences between evaluations of the (1) positive, (2) negative and (3) no outcome version by one-way between-subjects ANOVA and corresponding Post Hoc Multiple Comparisons tests. The use of ANOVA for ordinal data (such as a Likert scale) is slightly controversial and no consensus on this subject has been reached yet in the research community. Sullivan & Artino (2013) conclude that ANOVA can be used to analyze Likert scale responses if the situation is suitable. Norman (2010) comes to the same conclusion, stating that ANOVA can be used with Likert data with no fear of "coming to the wrong conclusion". Moreover, ANOVA requires the distribution to be normal. This requirement is fulfilled by the Central Limit Theorem (CLT). CLT states that when sample size (n) for each group > 30, normality of distribution and means can be assumed. On top of that, ANOVA is the standard procedure used in the aforementioned research on the outcome bias. In order to obtain similar and comparable results, this research follows the same framework.

Hypothesis 3 concerns a measurement of association, studying whether the type of outcome is associated with the words participants use to describe the company's management decisions. Because this concerned testing two nominal variables, chi-square tests (non-parametric) were conducted for every word individually. Chi-square tests whether the variables are independent or related. If there is a relation, Cramer's V shows the strength of that relationship (McHugh, 2013). It is a widely used descriptive statistical tool to test variables measured at the nominal level and therefore the most suitable analysis to test these associations.

4. Results

Introductory note: from a rational point of view, information that is available only after a decision is made is irrelevant to the quality of the decision. Outcome information does not provide meaningful additional information about a decision process. Therefore, outcome should not have an effect on the evaluation of a decision.

4.1 Outcome bias in evaluations of management decisions and -competence

Hypothesis 1a predicted that evaluations of management decisions would differ when different outcomes are attached to it. In order to test if there is a significant difference between evaluations of decisions with a positive, negative and no outcome, a one-way between-subjects ANOVA was conducted. The outcome (three versions: positive, negative, no outcome) being the independent variable and the evaluation being the dependent variable. When evaluating the *general quality of the management decisions*, the means of the three groups were found to be significantly different; $F(2, 95) = 23.636$, $p < .001$. Post hoc multiple comparisons using the Scheffe test revealed that the evaluations of management decisions in the negative version ($M = 4.06$, $SD = 1.12$) are significantly worse than both the positive ($M = 5.61$, $SD = 0.97$, $p < .001$), and the no outcome version ($M = 5.19$, $SD = 0.69$, $p < .001$). Even though there is a difference between the results of the positive version and the no outcome version, the difference was not found to be significant ($p = .208$).

In two other questions, respondents were asked to evaluate a specific management decision. Firstly, opinions were asked about the company's *"rapid growth through acquisition"* strategy. ANOVA revealed that the means of the three groups are significantly different; $F(2, 95) = 25.271$, $p < .001$. Because Levene's test of equality of variances showed a significant result (meaning variances are significantly different), a Welch test was conducted to ensure equality of means ($p < .001$). Then, a post hoc Games-Howell test showed that the evaluations of the acquisition strategy in the negative version ($M = 3.67$, $SD = 1.41$) are significantly worse than both the positive ($M = 5.55$, $SD = 0.94$, $p < .001$), and the no outcome version ($M = 5.16$, $SD = 0.99$, $p < .001$). Again, a difference can be observed between the positive and the no outcome version, but it is not significant ($p = .242$). Next, participants were asked to rate the company's *complex organizational structure*. ANOVA is once more showing significant differences between the means of the groups; $F(2, 95) = 14.407$, $p < .001$. Again, a Welch test was necessary to ensure the equality of means ($p < .001$). The Games-Howell test results are very similar to the previous analyses. Evaluation ratings of the complex organizational structure in the negative version ($M = 4.06$, $SD = 1.35$) are significantly worse than both the positive (*M*

= 5.45, *SD* = 0.67, $p < .001$) and the no outcome version (*M* = 5.03, *SD* = 1.12, $p = .007$). The differences between positive and no outcome were not found to be significant ($p = .166$).

To summarize the results of questions about management decisions: for all questions the negative outcome yields the worst decision evaluation ratings and positive outcomes yield the best evaluation ratings. As expected, no outcome produces the middle results. In support of hypothesis 1a, it can be seen that evaluation ratings of management decisions are indeed significantly higher for a positive outcome, compared to a negative outcome, showing a strong outcome bias. However, hypothesis 1a can only partly be accepted since no significant outcome bias was found between the positive and no outcome group. The results are highly similar. See annex C for the detailed analyses.

Hypothesis 1b predicted that evaluations of *CEO skill* would be more positive when the outcomes of his behavior are positive as well, compared to when the outcomes are negative or when no outcome is given. Like in the previous analyses, this hypothesis was tested by conducting a one-way between-subjects ANOVA. Results show significant differences between the means of the three groups; $F(2, 95) = 11.064$, $p < .001$. Post hoc multiple comparisons Scheffe test reveals that evaluations of CEO skill in the negative version (*M* = 4.39, *SD* = 1.22) are significantly lower than in the positive version (*M* = 5.58, *SD* = 0.97, $p < .001$) and in the no outcome version (*M* = 5.56, *SD* = 1.29, $p = .001$). Again, hypothesis 1b can only partly be accepted. The results show that evaluation ratings of CEO skill are indeed significantly higher when the outcomes of the CEO's behavior are positive, compared to when they are negative, demonstrating the presence of an outcome bias. However, again, no significant difference was found between the positive and the no outcome version ($p = .999$). In fact, the means of the two groups are practically identical (5.58 vs. 5.56). For more detailed information on this analysis, see annex C.

Hypothesis 1c addresses the *future decision-making capability* of the company. It forecasts that participants would give a higher rating to this capability when current outcomes are positive, compared to when current outcomes are negative or when no outcome is given. The one-way between-subjects ANOVA shows that there are significant differences between the means of the groups; $F(2, 95) = 20.582$, $p < .001$. A Welch test ensures equality of means. The post hoc Games-Howell test reports that the rating of the future decision-making capability of the company in the negative version (*M* = 3.82, *SD* = 1.51) is significantly worse than in the positive version (*M* = 5.42, *SD* = 0.97, $p < .001$) and in the no outcome version (*M* = 5.34, *SD* = 0.83, $p < .001$). Once again, hypothesis 1c can only partly be accepted because the means of the positive and the no outcome version are highly similar (5.42 vs. 5.34), hence no significant

difference was found (p = .931). Nonetheless, participants gave much higher ratings to the company's future decision-making competence in the positive version, compared to the negative version, strongly showing the presence of an outcome bias. Table 4.1 shows an overview of all evaluations of the decisions with a positive, negative and no outcome, annex C presents a detailed overview of all the analyses done.

Hypothesis 2 predicted that a negative outcome creates a stronger relative outcome bias effect than a positive outcome. Assuming that the no outcome version is the "base value" of the decision, the values of the negative and the positive version can be compared to this base value to see which one differs the most. All previous analyses show that the participant's answers in the negative outcome version differ significantly from those in the no outcome (base value) version. The average mean difference between negative and no outcome over all five questions is -1.257. For the positive version and the no outcome (base value) version, not one analysis showed a significant difference. The average mean difference between positive and no outcome over all five questions is .265. This value shows that participants give slightly higher ratings in the positive version than in the no outcome version, but it is not enough to detect a significant difference. The data also reveals that the average mean difference between a positive outcome and a negative outcome over all questions is 1.522. The results support hypothesis 2: attaching a negative outcome to the case study creates a stronger relative outcome bias effect than attaching a positive outcome.

Table 4.1 Overview of decision evaluations with positive, negative and no outcome

	Group	Mean	Std. Deviation	N
ManagementDecisions	Positive	5.61	.966	33
	Negative	4.06	1.116	33
	None	5.19	.693	32
AcquisitionStrategy	Positive	5.55	.938	33
	Negative	3.67	1.407	33
	None	5.16	.987	32
CEOskill	Positive	5.58	.969	33
	Negative	4.39	1.223	33
	None	5.56	1.294	32
MatrixStructure	Positive	5.45	.666	33
	Negative	4.06	1.345	33
	None	5.03	1.121	32
FutureDecisions	Positive	5.42	.969	33
	Negative	3.82	1.509	33
	None	5.34	.827	32

Note: Total N = 98

4.2 Word associations

Hypothesis 3 predicted that the participants attribute more positive words to the company's management decisions when the outcomes are positive as well, compared to when the outcomes are negative or when no outcome is given. Because it concerns nominal values, chi-square tests were conducted to see if there is a significant association between the outcome of the case study (positive, negative or no outcome) and the type of words participants attribute to the company. This was tested for eight different words: "visionary", "inspiring", "excellent", "courageous", "arrogant", "unfocused", "incompetent", and "chaotic". Visionary, inspiring, excellent and courageous were classified as being positive. Arrogant, unfocused, incompetent and chaotic were classified as being negative. To create an overview, figure 4.1 shows what percentage of participants selected a particular word. The graph uses percentages as the groups did not have equal sample size.

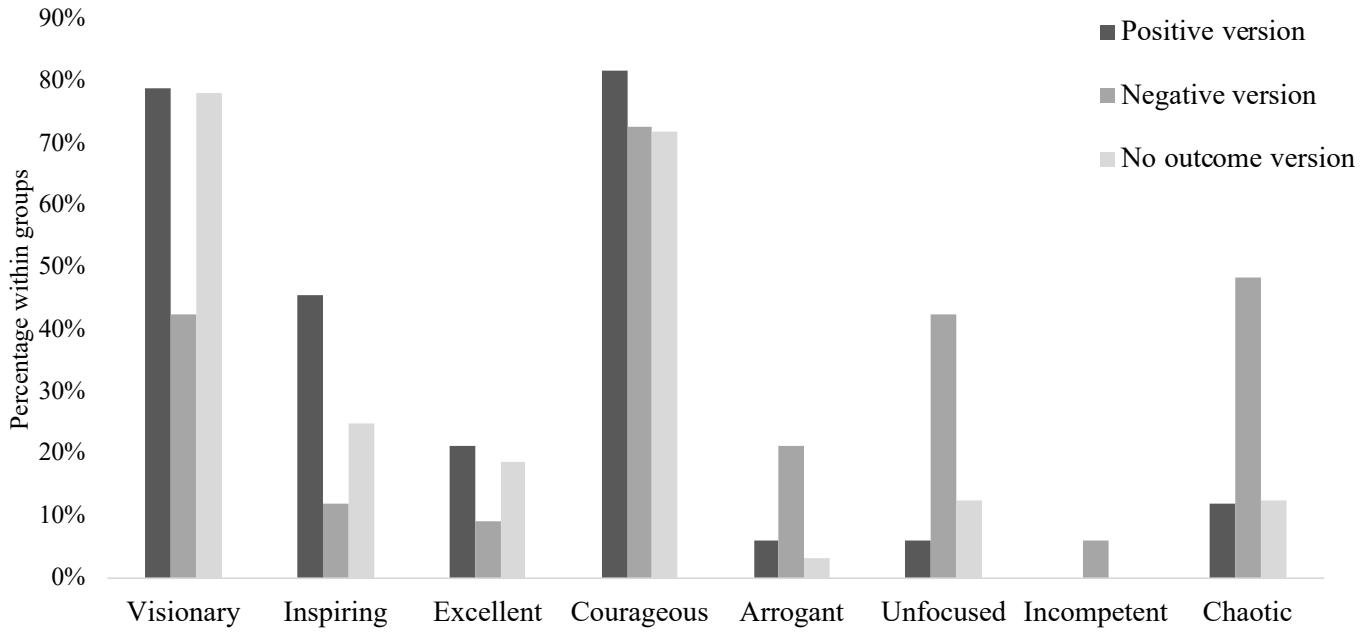

Figure 4.1 Percentual distribution of words associated with company WH

Firstly, the chi-square test shows a significant association between the outcome of the case study and the use of the word "*visionary*", X^2 (2, N = 98) = 12.73, p = .002. Chi-square tests comparing only two groups reveal that group positive (X^2 (1, N = 66) = 9.14, p = .003), as well as group no outcome (X^2 (1, N = 65) = 8.63, p = .003), were both more likely to describe the company as "visionary" than group negative, showing a moderate correlation, r(64) = .37, p = .003 and, r(63) = .36, p = .003 respectively. However, group positive and group no outcome were equally likely to use the word "visionary", X^2 (1, N = 65) = .004, p = .948.

Secondly, a chi-square test displays a significant association between the outcome of the case study and the use of the word "*inspiring*", X^2 (2, N = 98) = 9.34, p = .009. Group positive (X^2 (1, N = 66) = 8.94, p = .003) is more likely to describe the company as "inspiring" than group negative, showing a moderate correlation, r(64) = .37, p = .003. However, comparing group no outcome with both group negative (X^2 (1, N = 65) = 1.79, p = .181), and group positive (X^2 (1, N = 65) = 2.97, p = .085), showed these groups are equally likely to use the word "inspiring". No significant association was found.

For "*excellent*", no significant association with outcome was found, X^2 (2, N = 98) = 1.98, p = .372, meaning that the outcome of the case study did not significantly influence the use of the word "excellent". The variables are independent.

Similarly, for the word "*courageous*", the chi-square test shows no significant association with outcome, X^2 (2, N = 98) = 1.08, p = .584. Participants were equally likely to select the word "courageous", regardless of a positive, negative, or no outcome.

Moving on to the words classified as negative. When analyzing the word "*arrogant*", the assumption of a chi-square test, saying that only ≤ 20% of cells can have expected count less than 5 was violated. In this case the "likelihood ratio" provides more accurate information (see table 4.2 for clarification). A significant association was found between outcome and the use of the word "arrogant", X^2 (2, N = 98) = 6.50, p = .039. However, showing a weak correlation r(96) = .26, p = .035.

Table 4.2 Chi-square test for word "arrogant"

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	6.733[a]	2	.035
Likelihood Ratio	6.495	2	.039
N of Valid Cases	98		

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is 3.27.

When comparing the individual outcomes, the chi-square assumption was violated as well. Because this concerns 2x2 tables, the "Fisher's exact test" provides the most accurate information. As a result of using Fisher's exact test, suddenly there were no significant associations to be found anymore between any of the outcomes and the use of the word "arrogant" (see table 4.3 as an example). As mentioned before, the initial correlation was weak and therefore, a more conservative test does not show a significant association. Statistically, it cannot be claimed with sufficient confidence that the use of the word "arrogant" is dependent on the type of outcome in the case study.

Table 4.3 Chi-square test for word "arrogant" comparing negative and no outcome

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.924[a]	1	.026		
Continuity Correction[b]	3.391	1	.066		
Likelihood Ratio	5.486	1	.019		
Fisher's Exact Test				.054	.030
N of Valid Cases	65				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 3.94.

b. Computed only for a 2x2 table

For "*unfocused*", the chi-square test shows a significant association with outcome, X^2 (2, N = 98) = 15.26, p < .001. 2x2 chi-square tests reveal that group negative was more likely to describe the company as "unfocused" than both group positive (X^2 (1, N = 66) = 11.88, p = .001), and group no outcome (X^2 (1, N = 65) = 7.27, p = .007), showing a moderate correlation, r(64) = -.42, p = .001 and, r(63) = -.33, p = .007 respectively. However, Fischer's exact test (2-sided) shows us that group positive and group no outcome were equally likely to use the word "unfocused", p = .427.

For "*incompetent*", the likelihood ratio shows us that no significant association with outcome was found, X^2 (2, N = 98) = 4.44, p = .109. In fact, only two out of 98 respondents selected this answer, both from group negative. Nonetheless, results show that the outcome of the case study did not significantly influence the use of the word "incompetent".

Lastly, the word "*chaotic*". Chi-square reveals a significant association between outcome and using the word "chaotic", X^2 (2, N = 98) = 15.49, p < .001. Group negative was more likely to describe the company as "chaotic" than both group positive X^2 (1, N = 66) = 10.33, p = .001), and group no outcome (X^2 (1, N = 65) = 9.88, p = .002), showing a moderate correlation, r(64) = -.40, p = .001 and, r(63) = -.39, p = .002 respectively. Group positive and group no outcome were equally likely to use the word "chaotic", p = 1.

Hypothesis 3 predicted that the type of words attributed to the company's management decisions are more positive when the outcomes of the decisions are positive as well, compared to when they are negative or when no outcome is given. To summarize the results, four words were classified as being positive (visionary, inspiring, excellent and courageous) and four words as being negative (arrogant, unfocused, incompetent and chaotic). For the positive words, results show that participant's decisions to select the words "visionary" and "inspiring" is significantly associated with the type of outcomes in the case study. Generally, group positive

was most likely to select these words. Contrarily, outcome did not play a part in whether participants selected the words “excellent” or “courageous”. Regarding the negative words, only “unfocused” and “chaotic” generated significant associations, meaning that outcome did influence the participants’ decisions to select these particular words. Group negative was more likely to select “unfocused” and “chaotic” than the other two groups. For “arrogant” and “incompetent”, the association with outcome was not found. In addition, equal to the outcome bias analysis, not one association was found between group positive and group no outcome, meaning that whether participants read a positive version or a no outcome version did not influence the attribution of certain words. Hypothesis 3 forecasted that the participants’ attribution of words is dependent on whether they read a positive, negative, or no outcome. It predicted that a positive outcome would generate more positive attributions, and a negative outcome would generate more negative attributions. This was found to be true for four (two positive and two negative) out of the eight words specified. For the other four words the hypothesis does not prove to be true. Participants’ selection of these particular words are independent of the outcome of the case study.

5. Discussion

The present research provides evidence of the presence of an outcome bias in a real-world-based situation where a company makes risky decisions under uncertainty. For all evaluations of management decisions, outcome affected the ratings of decision quality. The quality of the management decisions are seen as much better when a positive outcome is attached to the decisions compared to when a negative outcome is attached. The same is true for evaluations of CEO skill-level and future decision-making competence, a positive outcome resulted into much better evaluations than a negative outcome. Over all five questions, the mean difference between a positive and negative outcome is a substantial 1.522. These results are in line with the hypotheses and the findings of previous outcome bias studies used as input for this research (Baron & Hershey, 1988; Damnjanović et al., 2019; Emerson, et al., 2010; Gino et al., 2009; König-Kersting et al., 2020; Marshall & Mowen, 1993). The outcomes of decisions lead people to see the decisions themselves in a different light. Judging from the results, participants see a bad outcome as a sign of bad decision making, and a good outcome as a sign of good decision making.

However, not all results support the hypotheses. From the perspective of the "base value", that is – the results from the no outcome version, only the negative outcome shows a statistically significant difference. Participants evaluate a decision with a negative outcome as worse than a decision without outcome. However, it does not matter if a positive outcome or no outcome is attached to a management decision because the evaluations ratings from these two groups are equally high (5.52 vs. 5.26 respectively). The participants were not biased by a positive outcome. The results suggest that, in the no outcome version, participants see the decisions as reasonable in the context given and assume that the company is competent and qualified to make the right decisions, often without realizing the risk and uncertainty associated with choosing a strategy. Business decisions are rarely as clear-cut as presented in the case study. As a result, participants may feel that presenting a complex situation (in which the odds are not entirely clear) with too much simplicity is implausible. Hence, people can develop the belief that the decision maker had more information available to them than actually was the case, which also leads to making a biased judgement (Baron & Hershey, 1988). Earlier, we have seen an example of this in the Honda case (Mintzberg et al., 1996). Of course, there was no possibility to give participants any information about the odds of specific outcomes because they were simply not available in this particular context. In addition, the participants were not aware of the fact that there were any versions which included an outcome, strengthening the assumption that company WH knew what it was doing. Although not yet supported by

quantative data, this could provide a possible explanation for the lack of difference between the positive and no outcome version. In this particular context, attaching no outcome has the same effect as attaching a positive outcome, creating the assumption that the company is competent and we can learn something from their management practices. Considering that the effects for a positive outcome and no outcome are identical, it can also be concluded that attaching a negative outcome to management decisions has much greater influence on the evaluation ratings than attaching a positive outcome.

Regarding the type of words participants use to describe the company's management decisions. In line with the hypothesis, the use of words "visionary", "inspiring", "unfocused" and "chaotic" are dependent on the type of outcome participants read. As expected, participants who read the positive outcome are more likely to select "visionary" and "inspiring" and participants who read the negative outcome are much more likely to select "unfocused" and "chaotic". Like in the outcome bias results, attaching a positive outcome or no outcome did not make a difference. Contrary to the hypothesis, the use of the words "excellent", "courageous", "arrogant" and "incompetent" are independent of the outcome the participants read. All groups were equally likely to select these words. A possible explanation could lie in the more extreme nature of the words, compared to the words for which an association with outcome was found. Considering that the mean average for group positive in the outcome bias analysis is 5.52 out of 7, participants evaluated the management decision as good, but not "excellent" (number 7 represented an "excellent decision", see annex B for clarification). The same works in the other direction. Considering that the mean average for group negative is 4.0 out of 7, evaluations prove to be significantly worse, but not bad enough to select "arrogant" or "incompetent". In contrast, "Courageous" was selected very often (by 76% of the participants) and was almost equally distributed among the different outcomes. Seemingly, the word leaves room for interpretation. A bad but brave decision could be labelled as "courageous", while a good and brave decision could also be considered "courageous".

To conclude, for four out of eight words under consideration a dependency was found between whether the participants selected the words and the type of outcome they were assigned to. Participants in the positive outcome selected significantly more positive words while the negative outcome had significantly more negative words attached to it. For the remaining four words this dependency was not found. Considering that a significant outcome bias is found in the previous analysis, it is reasonable to assume the nature of the words was either too extreme ("excellent", "arrogant" and "incompetent") or too generic ("courageous"). For the extreme words the outcome bias effect proves to be not powerful enough.

5.1 Contributions to theory

These findings extend prior research on the outcome bias from a variety of scientific fields. In the management field, König-Kersting et al. (2020) and Marshall & Mowen (1993) have already proven that the outcome bias is present when evaluating the decisions of an individual financial agent or salesperson. The present research shows that the outcome bias also extends to more complex decision-making situations on company-level, which complies to the psychological research done by Kahneman (2011). Similarly to Damnjanović et al. (2019) and Marshall & Mowen (1993), a third "no outcome" situation was added to the experiment which served as the control group. However, contrary to their findings, no evaluation difference was found between attaching a positive outcome or no outcome to the decisions. The main difference between these studies and the present study lies in the complexity of the situation described in the experiment. This can present valuable insights into the effects of the outcome bias when decision-complexity increases. Damnjanović et al. (2019) already reported that when decision complexity rises, the outcome bias becomes stronger. The present research may provide evidence that, when decision complexity rises and odds become less clear, people make more assumptions about the information they *think* was available to the decision maker. This reasoning is compliant with the findings of Baron & Hershey (1988) and Mintzberg et al. (1996). Although not yet supported by quantative research, it may explain why in the no outcome version the assumption is strengthened that the company is competent and qualified to make the right decisions, often not taking into account the risk and uncertainty associated with choosing a strategy. Hence, in this particular context, applying a positive outcome or no outcome does not make a difference for evaluations of the decisions. Moreover, this confirms the findings of Damnjanović et al. (2019) that the relative effect of attaching a negative outcome is larger than when a positive outcome is attached.

Lastly, this research confirms experimental research on "the Halo effect" (McGuire et al., 1990; Meindl & Ehrlich, 1987; Staw, 1975), showing that people attribute a different set of words to groups they believe are high performers according to the outcome. However, the present research shows that the outcome effect was not strong enough to make participants choose attributions which are too extreme (e.g. "arrogant" or "excellent").

5.2 Practical implications

This research also has practical implications. Previous studies show that luck plays a much greater role in the performance of a company than we often attribute to it. Considering that luck breaks the link between skill and result, judging a company by its financial results is irrational

(Mauboussin, 2012). In case of ABB, luck played a big role. Therefore we cannot conclude their management practices were good or bad to start with (Rosenzweig, 2004). The outcome bias demonstrated in the study indicates that people may blame companies too harshly for making sensible decisions that have unlucky outcomes. Likewise, people may praise companies too much for making average or poor decisions that fortunately turn out well. The results suggest that we may confuse the evaluations of the decisions with the evaluations of the outcomes themselves. That whenever an outcome of a complex decision is good or bad, we automatically make assumptions about the skill-level of the decision maker. Because the sample for the experiment was exclusively made up of university business students, it indicates that the majority of students allow their impression of a company to be colored too much by financial results. Instead, in order to make a good assessment of a company's abilities, people should focus more on the process towards financial results rather than the financial results themselves. Exceptional success stories of companies should not be seen as a "blueprint for success", but more as a source of inspiration. The results can also help managers in recognizing they are subject to cognitive biases as well. It can teach them to stay humble when experiencing successes. To not become too sure of your own abilities as a company, and to think more rationally about the nature of firm performance.

6. Conclusions and Recommendations

This research aimed to identify an outcome bias in university business students' evaluation of the managerial decision-making competence of a company which made risky decisions under uncertainty. Based on the analysis of the evaluation ratings participants gave to management decisions with a positive, negative and no outcome, it can be concluded that attaching a positive outcome to a management decision results in a significantly better evaluation of decision quality and company competence than when a negative outcome is attached to the same decision. This indicates that university business students have the tendency to see good financial performance as a sign of good decision-making, and bad financial performance as a sign of bad decision-making. The outcome automatically creates an assumption about the quality of the decision and the skill-level of the company. These results confirmed the hypotheses. Contrary to the hypotheses, no difference in evaluation was found between attaching a positive outcome or no outcome to the management decisions, indicating that people see the decisions as reasonable in the context given, and assume that the company is competent and qualified to make the right decisions when no outcome is attached. A possible explanation for the high ratings in the no outcome version could be that, in complex decision-making situations, people see the simplicity of the case study as implausible. As a result, people intuitively make assumptions about information they *think* was available to the decision maker, even if in reality it was not. This however, can also lead to a biased judgement (Baron & Hershey, 1988; Mintzberg et al., 1996). Moreover, it can be concluded that the outcome bias influences the type of words people attribute to a company, generally attributing more positive words (e.g. "visionary") to a positive outcome, and more negative words (e.g. "chaotic") to a negative outcome. However, the outcome bias effect was not strong enough to make participants choose more extreme words like "excellent" or "incompetent". Decisions in the positive version were rated as good, but not exceptional. Likewise, decisions in the negative version were rated as bad, but not terrible.

6.1 Strengths and limitations

Before measuring the presence of an outcome bias, it had to be clarified why outcome-based evaluations are considered irrational in a business context. Hence, the literature review started with providing evidence on the role of luck in the performance of a company, after which it dove more into the cognitive biases that prevent us from recognizing luck. It also described the framework to measure susceptibility to the outcome bias. In line with existing literature, the outcome bias was measured by an experimental design. A between-subject design was adopted to avoid giving the participants any foreknowledge on outcome information. Participants

completed only one of the three versions. In contrast to traditional outcome bias studies, this research used a case study based on a real-life situation of a company. This is beneficial because it actually proves the presence of an outcome bias in reality. Therefore, it is more plausible the results from the experiment apply to the real world as well, compared to research using fictitious data.

Naturally, basing the case study on a real-world situation brings its corresponding limitations, mainly because the information available cannot be controlled by the researcher. Therefore, information cannot be formulated as specific as in conventional experimental studies. Even though the case study describes the decision context in as much detail as possible, business situations are known to be complex. Therefore, it is highly probable the participants believe that the decision makers had more information available to them than described in the case study, providing a possible explanation for the fact that no evaluation differences were found between a positive and no outcome. Although it is a limitation and it blurs the measurement of the outcome bias, it is consistent with reality. Business decisions are simply not as clear-cut as, for example, the flip of a coin. The odds are often unclear and influenced by many external factors. As a result, assumptions are easily made. Lastly, as an additional limitation, the array of words participants could select to characterize the company should have been more diverse and more carefully selected to be able to draw better conclusions. For example, asking participants to choose between two antonyms for a clearer comparison.

6.2 Suggestions for future research

A substantial amount of research on the outcome bias has already been done. The present research extends this knowledge by providing evidence of an outcome bias effect when evaluating the managerial decision-making competence of a real-life company. Considering the experiment showed unexpected evaluations of the no outcome version, future research could focus on investigating the difference between attaching a positive outcome and no outcome to different decisions. This would help in uncovering the situations where people are (not) biased by a positive outcome. Because the present research aimed to be as close to real-life as possible, it did not adopt different scenarios with alterations in outcome severity, decision-involvement, or decision appropriateness. Future research is needed to determine how these factors influence the effect of the outcome bias in this specific context. Besides, instead of using a between-subject design, a within-subject design could determine if people see it as appropriate to allow a decision's outcome to determine their assessment of the decision's quality, considering they

would be subject to all the levels of the experiment. Finally, perhaps the most important direction of future research may focus on developing ways to reduce the susceptibility to the outcome bias in business contexts. Most universities still neglect the role of luck when teaching their students about strategy. Perhaps a framework could be developed, specifying situations in which one should be aware not to focus on outcome (financial performance), but on intentions (did the company make the right decision under the circumstances?). In addition to traditional strategy classes, such a framework would provide students with a more accurate view on the determinants of company success. Simultaneously, it would help current business managers in becoming more aware of their irrational tendencies, allowing them to act accordingly.

References

Ansoff, I. H. (1968). *Corporate Strategy: An Analytic Approach to Business Policy for Growth and Expansion*. Penguin.

Barnett, A. G., Pols, J. C., & Dobson, A. J. (2004). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology*, *34*(1), 215–220. https://doi.org/10.1093/ije/dyh299

Barney, J. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Barney, J. B. (1986). Strategic Factor Markets: Expectations, Luck, and Business Strategy. *Management Science*, *32*(10), 1231–1241. https://doi.org/10.1287/mnsc.32.10.1231

Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, *54*(4), 569–579. https://doi.org/10.1037/0022-3514.54.4.569

Cambridge Dictionary. (2020, November 18). *Luck*. https://dictionary.cambridge.org/dictionary/english/luck

Cambridge Dictionary. (2020b, November 18). *Skill*. https://dictionary.cambridge.org/dictionary/english/skill

Camerer, C., Loewenstein, G., & Weber, M. (1989). The Curse of Knowledge in Economic Settings: An Experimental Analysis. *Journal of Political Economy*, *97*(5), 1232–1254. https://doi.org/10.1086/261651

Collins, J. (2001). *Good to Great: Why Some Companies Make the Leap and Others Don't* (1st ed.). HarperBusiness.

Collins, J., & Porras, J. (1994). *Built to Last: Successful Habits of Visionary Companies (Good to Great)* (3rd ed.). Harper Business.

Crusius, J., van Horen, F., & Mussweiler, T. (2012). Why process matters: A social cognition perspective on economic behavior. *Journal of Economic Psychology*, *33*(3), 677–685. https://doi.org/10.1016/j.joep.2011.09.004

Damnjanović, K., Ilić, S., Pavlović, I., & Novković, V. (2019). Refinement of outcome bias measurement in the parental decision-making context. *Europe's Journal of Psychology*, *15*(1), 41–58. https://doi.org/10.5964/ejop.v15i1.1698

Denrell, J. (2005). Should We Be Impressed With High Performance? *Journal of Management Inquiry*, *14*(3), 292–298. https://doi.org/10.1177/1056492605279100

Denrell, J., & Liu, C. (2012). Top performers are not the most impressive when extreme performance indicates unreliability. *Proceedings of the National Academy of Sciences*, *109*(24), 9331–9336. https://doi.org/10.1073/pnas.1116048109

Dillon, R. L., & Tinsley, C. H. (2008). How Near-Misses Influence Decision Making Under Risk: A Missed Opportunity for Learning. *Management Science*, *54*(8), 1425–1440. https://doi.org/10.1287/mnsc.1080.0869

Emerson, G. B., Warme, W. J., Wolf, F. M., Heckman, J. D., Brand, R. A., & Leopold, S. S. (2010). Testing for the Presence of Positive-Outcome Bias in Peer Review. *Archives of Internal Medicine*, *170*(21), 1934–1939. https://doi.org/10.1001/archinternmed.2010.406

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*(3), 288–299. https://doi.org/10.1037/0096-1523.1.3.288

Gino, F., Moore, D. A., & Bazerman, M. H. (2009). No Harm, No Foul: The Outcome Bias in Ethical Judgments. *SSRN Electronic Journal*, 1–51. https://doi.org/10.2139/ssrn.1099464

Gompers, P. A., Kovner, A., Lerner, J., & Scharfstein, D. S. (2006). Skill vs. Luck in Entrepreneurship and Venture Capital: Evidence from Serial Entrepreneurs. *SSRN Electronic Journal*, 1–29. https://doi.org/10.2139/ssrn.933932

Gottschall, J. (2013). *The Storytelling Animal: How Stories Make Us Human* (1st ed.). Mariner Books.

Hafer, J., & Gresham, G. G. (2008). Luck's Role in Business Success: Why it's too Important to Leave to Chance. *Journal of Behavioral and Applied Management*, *9*(3), 295–315. https://doi.org/10.21818/001c.17152

Jenter, D., & Kanaan, F. (2015). CEO Turnover and Relative Performance Evaluation. *The Journal of Finance*, *70*(5), 2155–2184. https://doi.org/10.1111/jofi.12282

Kahneman, D. (2011). *Thinking, Fast and Slow* (1st ed.). Farrar, Straus and Giroux.

Kim, S., Eberhart, R., & Armanios, D. (in press). What Makes a Winner? Quantifying Luck in Sustained Performance. *Department of Engineering and Public Policy, Carnegie Mellon University*. 10.13140/RG.2.2.19222.14401

König-Kersting, C., Pollmann, M., Potters, J., & Trautmann, S. T. (2020). Good decisions vs. good results: Outcome bias in the evaluation of financial agents. *Theory and Decision*, *90*, 1–31. https://doi.org/10.1007/s11238-020-09773-1

Liu, C. (2019). *Luck: A Key Idea for Business and Society (Key Ideas in Business and Management)* (1st ed.). Routledge.

Liu, C. (2020, March 1). MBA masterclass: when strategy is a matter of luck. *Financial Times.* https://www.ft.com/content/5f0afa98-4ce3-11ea-95a0-43d18ec715f5

Liu, C., & de Rond, M. (2016). Good Night, and Good Luck: Perspectives on Luck in Management Scholarship. *The Academy of Management Annals*, *10*(1), 409–451. https://doi.org/10.1080/19416520.2016.1120971

Ma, H. (2002). Competitive advantage: what's luck got to do with it? *Management Decision*, *40*(6), 525–536. https://doi.org/10.1108/00251740210433927

Marshall, G. W., & Mowen, J. C. (1993). An experimental investigation of the outcome bias in salesperson performance evaluations. *Journal of Personal Selling & Sales Management, 13*(3), 31–47. 10.1080/08853134.1993.10753956

Mauboussin, M. J. (2012). *The Success Equation: Untangling Skill and Luck in Business, Sports, and Investing* (10.7.2012 ed.). Harvard Business Review Press.

Mauboussin, M. J. (2014, July 22). *The Success Equation: Untangling Skill and Luck | Michael Mauboussin | Talks at Google*. YouTube. https://www.youtube.com/watch?v=1JLfqBsX5Lc&feature=youtube_gdata

McGuire, J. B., Schneeweis, T., & Branch, B. (1990). Perceptions of Firm Quality: A Cause or Result of Firm Performance. *Journal of Management*, *16*(1), 167–180. https://doi.org/10.1177/014920639001600112

McHugh, M. L. (2013). The Chi-square test of independence. *Biochemia Medica*, *23*(2), 143–149. https://doi.org/10.11613/bm.2013.018

Meindl, J. R., & Ehrlich, S. B. (1987). The Romance of Leadership and The Evaluation of Organizational Performance. *Academy of Management Journal*, *30*(1), 91–109. https://doi.org/10.5465/255897

Mintzberg, H., Pascale, R. T., Goold, M., & Rumelt, R. P. (1996). CMR Forum: The "Honda Effect" Revisited. *California Management Review*, *38*(4), 77–117. https://doi.org/10.2307/41165855

N., E., von Neumann, J., & Morgenstern, O. (1945). Theory of Games and Economic Behavior. *The Journal of Philosophy*, *42*(20), 550–554. https://doi.org/10.2307/2019327

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Parnell, J. A., & Dent, E. B. (2009). The role of luck in the strategy-performance relationship. *Management Decision*, 47(6), 1000–1021. https://doi.org/10.1108/00251740910966703

Pluchino, A., Biondo, A.E., & Rapisarda, A. (2018). Talent versus luck: the role of randomness in success and failure. *Advances in Complex Systems*, *21*(03n04), 1850014. https://doi.org/10.1142/s0219525918500145

Porter, M. E. (1989). How Competitive Forces Shape Strategy. *Readings in Strategic Management*, 133–143. https://doi.org/10.1007/978-1-349-20317-8_10

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

Porter, M. (1985). *Competitive Advantage: Creating and Sustaining Superior Performance (Hardcover)* (1st ed.). Free Press.

Raynor, M. E., Ahmed, M., & Henderson, A. (2009, April). *A Random Search for Excellence Why " great " company research delivers fables and not facts roi.* https://www2.deloitte.com/content/dam/insights/us/articles/a-random-search-for-excellence-why-great-company-research-delivers-fables-not-facts/DUP106_Random_Search_For_Excellence.pdf

Rosenzweig, P. M. (2007a). Misunderstanding the Nature of Company Performance: The Halo Effect and other Business Delusions. *California Management Review*, *49*(4), 6–20. https://doi.org/10.2307/41166403

Rosenzweig, P. M. (2007b*). The Halo Effect: . . . and the Eight Other Business Delusions That Deceive Managers* (Original ed.). Free Press.

Rosenzweig, P. M. (2004). What do we think happened at ABB? Pitfalls in research about firm performance. *International Journal of Management and Decision Making*, *5*(4), 267–281. https://doi.org/10.1504/ijmdm.2004.005642

Staw, B. M. (1975). Attribution of the "causes" of performance: A general alternative interpretation of cross-sectional research on organizations. *Organizational Behavior and Human Performance*, *13*(3), 414–432. https://doi.org/10.1016/0030-5073(75)90060-4

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. https://doi.org/10.4300/jgme-5-4-18

Taleb, N. N. (2007). *The Black Swan: The Impact of the Highly Improbable by Taleb, Nassim Nicholas* (1st (first) Edition) [Hardcover(2007)] (1st ed.). Random House.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

VentureBeat. (2019, May 14). *50% of businesses fail in their first 5 years. What's the secret for those that survive?* https://venturebeat.com/2019/05/13/50-of-businesses-fail-in-their-first-5-years-whats-the-secret-for-those-that-survive/

Waterman, R. H., & Peters, T. J. (1982). *In Search of Excellence: Lessons from America's Best-Run Companies* (Presumed First Edition). HarperTrade, New York, NY, U.S.A.

Wiggins, R. R., & Ruefli, T. W. (2002). Sustained Competitive Advantage: Temporal Dynamics and the Incidence and Persistence of Superior Economic Performance. *Organization Science*, *13*(1), 81–105. https://doi.org/10.1287/orsc.13.1.81.542

Annex A

Experiment case study

In 1988, two of Europe's leading engineering companies merged to form the new company "WH". The plan to merge the two companies came from one of the CEO's: Patrick Banik. Banik believed that a borderless Europe would weaken national power and could offer them huge global scale benefits. In 1988 WH immediately became world leader in three areas, (1) energy plants (coal, gas, oil etc.), (2) power transmission, and (3) power distribution, while being active in a few other technical industries as well. In the risky years after the merger, WH's management was defined by the following themes:

1. Acquisition: WH followed a "rapid growth through acquisition" strategy, in order to become a leading global player, acquiring some fifteen companies in their first year. These acquisitions were rapidly integrated, while also cutting headquarters staff to 10% of their previous size, aiming to save costs and work more efficiently. Next, WH started expanding into emerging markets in Europe and Asia. Due to regulations, Asia became more open to outside investment and industrial growth. These changes completely transformed WH's profile in just a few years, creating a very broad range of business.
2. CEO Banik's leadership style: originally Scandinavian, Banik owned a Stanford MBA, spoke English with American slang, was bold and sure of himself, and combined old world manners with American pragmatism and orientation for action. In addition, Banik was known to be a workaholic and rapid decision maker.
3. The matrix structure: Large multinational companies face big organizational challenges when trying to manage a complex set of departments. The challenge is to be successful in local markets and at the same time take advantage of global scale and scope. WH identified itself as a "multidomestic" company, not a global one. Meaning it was essentially a combination of many national companies. WH adopted a complex matrix structure: seven sectors divided into two axes: Business Areas and countries. WH had 51 Business Areas and 41 Country Managers, crossing in 1,300 separate companies. These separate companies further subdivided into 5,000 profit centers. This was all held together by the "hard glue": the management reporting system Abacus. The intention of the matrix structure was to bring customers in direct contact with the employees and "overcome big-company bureaucracy and still get the advantage of size".
4. Corporate Culture: Speed, autonomy and empowerment were key elements of WH's corporate culture. WH's founding itself was bold and certainly not without risk. As mentioned before, CEO Banik had an obsession for speed, encouraging people to take initiative and stick their neck out. Employees were encouraged to innovate in local markets and free to speak their minds.

A few years later, WH started to move away from their reliance on heavy manufacturing and engineering, steering more towards services. WH was concerned about the future of power generation and transmission because of the upcoming maturity of this market and new potentially troubling regulations by the European Union. WH became a more "knowledge-based" company, focusing on intellectual property and software. WH resumed making acquisitions to facilitate the shift to services. It started buying companies operating in process control, industrial services and financial services, while selling some of the old businesses (e.g. trains & trams business), reducing their dependency on the power sector and shifting their core competences.

Positive outcome condition	*Negative outcome condition*
So how did the company do financially? After the merger in 1988, WH immediately became a world leader in three main segments, having US$18 billion in revenues. It established rapid international presence, employing 215.000 people by 1990. By 1991, it had acquired or taken minority positions in 60 companies representing investment worth US$3.6 billion, including two major acquisitions in North America. Between 1988 and 1996, revenues almost doubled, to US$ 34.7 billion, and net profit tripled, to US$ 1.2 billion. By the end of 1991, WH employed 10,000 people in central and Eastern Europe. The following year, that number doubled. By 1994, WH had 30,000 employees and 100 centers across Asia. After all the acquisitions and developments, WH's profile had changed completely. However, all the while, performance rose consistently. By 1996, WH's value as a company was put at more than US$ 40 billion.	So how did the company do financially? By 2000, sales growth had faltered and profits decreased rapidly. Their new ventures showed very little profits, and the acquisitions financed by loans had driven debt to troubling heights. A new CEO was appointed, yet performance continued to go down. On top of that, one of their past acquisitions from the 90s was the subject of asbestos claims, costing WH US$ 470 million for claims in 2001. In addition, a financial crisis hit in Asia late 90s. Banik resigned as chairman in 2001 and was replaced. His successor reviewed WH's broad range of business and decided to sell certain divisions of the company, but performance continued to slide. A US$ 1.5 billion loan was necessary so WH could avoid a liquidity crisis. In 2002, WH reported a loss of US$ 600 million. Debt raised to US$4.1 billion. Once valued at more than US$ 40 billion, WH was now worth less than US$4 billion. It had to restructure its activities, cut jobs, sell assets, and manage large liabilities.

Annex B

Experiment questions

"The task: This case study is aiming to capture your opinion on the management practices of the company WH under the circumstances described in the text. Please answer the following questions purely based on the information given in the text.

1. Based on the information given in the case, how would you evaluate the quality of the management decisions that WH made?

Mark only one oval.

	1	2	3	4	5	6	7	
Terrible decisions	O	O	O	O	O	O	O	Excellent decisions

2. Based on the information given in the case, how would you evaluate WH's decision to adopt a "rapid growth through acquisition" strategy?

Mark only one oval.

	1	2	3	4	5	6	7	
Terrible decision	O	O	O	O	O	O	O	Excellent decision

3. Based on the information given in the case, how skilled do you consider the CEO Patrick Banik to be?

Mark only one oval.

	1	2	3	4	5	6	7	
Not skilled at all	O	O	O	O	O	O	O	Extremely skilled

4. Based on the information given in the case, how would you evaluate WH's decision to adopt the complex organizational matrix structure?

Mark only one oval.

	1	2	3	4	5	6	7	
Terrible decision	O	O	O	O	O	O	O	Excellent decision

5. Based on the information given in the case, how would you judge WH's capability to make good decisions in the future?

Mark only one oval.

	1	2	3	4	5	6	7	
Not capable	O	O	O	O	O	O	O	Extremely capable

6. Please select the words you feel are most suitable to describe WH's management practices

Tick all that apply.

- O Visionary
- O Arrogant
- O Unfocused
- O Inspiring
- O Excellent
- O Courageous
- O Incompetent
- O Chaotic

Personal information

7. What is your gender?

Mark only one oval.

- O Male
- O Female

8. How old are you?

Mark only one oval.

- O 18 or under
- O 19
- O 20
- O 21
- O 22
- O 23
- O 24
- O 25
- O 26 or older

9. What is your nationality?

Mark only one oval.

- O Portuguese
- O Chinese
- O French
- O Italian
- O German
- O Spanish
- O Other…

10. Which program are you following at ISCTE? (e.g. MSc in Management, MSc in Marketing etc.)

Open answer.

Annex C

Outcome bias data analysis

General management decisions

Table C1: Descriptive statistics general management decisions

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	33	5.61	.966	.168	5.26	5.95	3	7
2	33	4.06	1.116	.194	3.66	4.46	2	6
3	32	5.19	.693	.122	4.94	5.44	4	6
Total	98	4.95	1.143	.115	4.72	5.18	2	7

Table C2: Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
ManagementDecisions	Based on Mean	2.776	2	95	.067
	Based on Median	2.493	2	95	.088
	Based on Median and with adjusted df	2.493	2	88.426	.088
	Based on trimmed mean	2.667	2	95	.075

Table C3: ANOVA general management decisions

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	42.112	2	21.056	23.636	.000
Within Groups	84.633	95	.891		
Total	126.745	97			

Table C4: Multiple Comparisons general management decisions

Scheffe

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	1.545*	.232	.000	.97	2.12
	3	.419	.234	.208	-.16	1.00
2	1	-1.545*	.232	.000	-2.12	-.97
	3	-1.127*	.234	.000	-1.71	-.54
3	1	-.419	.234	.208	-1.00	.16
	2	1.127*	.234	.000	.54	1.71

Rapid growth through acquisition strategy

Table C5: Descriptive statistics rapid growth through acquisition strategy

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	33	5.55	.938	.163	5.21	5.88	3	7
2	33	3.67	1.407	.245	3.17	4.17	1	6
3	32	5.16	.987	.175	4.80	5.51	4	7
Total	98	4.79	1.387	.140	4.51	5.06	1	7

Table C6: Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
AcquisitionStrategy	Based on Mean	6.653	2	95	.002
	Based on Median	4.483	2	95	.014
	Based on Median and with adjusted df	4.483	2	85.069	.014
	Based on trimmed mean	6.711	2	95	.002

Table C7: ANOVA rapid growth through acquisition strategy

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	64.766	2	32.383	25.271	.000
Within Groups	121.734	95	1.281		
Total	186.500	97			

Table C8: Robust Tests of Equality of Means rapid growth through acquisition strategy

	Statistic[a]	df1	df2	Sig.
Welch	20.511	2	61.955	.000

a. Asymptotically F distributed.

Table C9: Multiple Comparisons rapid growth through acquisition strategy

Games-Howell

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	1.879*	.294	.000	1.17	2.59
	3	.389	.239	.242	-.18	.96
2	1	-1.879*	.294	.000	-2.59	-1.17
	3	-1.490*	.301	.000	-2.21	-.77
3	1	-.389	.239	.242	-.96	.18
	2	1.490*	.301	.000	.77	2.21

*. The mean difference is significant at the 0.05 level.

CEO skill level

Table C10: Descriptive statistics CEO skill level

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	33	5.58	.969	.169	5.23	5.92	4	7
2	33	4.39	1.223	.213	3.96	4.83	3	7
3	32	5.56	1.294	.229	5.10	6.03	1	7
Total	98	5.17	1.285	.130	4.92	5.43	1	7

Table C11: Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
CEOskill	Based on Mean	1.299	2	95	.278
	Based on Median	1.187	2	95	.309
	Based on Median and with adjusted df	1.187	2	88.997	.310
	Based on trimmed mean	1.134	2	95	.326

Table C12: ANOVA CEO skill level

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	30.237	2	15.118	11.064	.000
Within Groups	129.814	95	1.366		
Total	160.051	97			

Table C13: Multiple Comparisons CEO skill level

Scheffe

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	1.182*	.288	.000	.47	1.90
	3	.013	.290	.999	-.71	.73
2	1	-1.182*	.288	.000	-1.90	-.47
	3	-1.169*	.290	.001	-1.89	-.45
3	1	-.013	.290	.999	-.73	.71
	2	1.169*	.290	.001	.45	1.89

*. The mean difference is significant at the 0.05 level.

Complex organizational matrix structure

Table C14: Descriptive statistics complex organizational matrix structure

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
1	33	5.45	.666	.116	5.22	5.69	4	7
2	33	4.06	1.345	.234	3.58	4.54	1	6
3	32	5.03	1.121	.198	4.63	5.44	2	7
Total	98	4.85	1.221	.123	4.60	5.09	1	7

Table C15: Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
MatrixStructure	Based on Mean	5.491	2	95	.006
	Based on Median	4.624	2	95	.012
	Based on Median and with adjusted df	4.624	2	89.596	.012
	Based on trimmed mean	5.728	2	95	.004

Table C16: ANOVA complex organizational matrix structure

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	33.675	2	16.837	14.407	.000
Within Groups	111.029	95	1.169		
Total	144.704	97			

Table C17: Robust Tests of Equality of Means complex organizational matrix structure

	Statistic[a]	df1	df2	Sig.
Welch	14.293	2	57.378	.000

a. Asymptotically F distributed.

Table C18: Multiple Comparisons complex organizational matrix structure

Games-Howell

					95% Confidence Interval	
(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
1	2	1.394*	.261	.000	.76	2.03
	3	.423	.230	.166	-.13	.98
2	1	-1.394*	.261	.000	-2.03	-.76
	3	-.971*	.307	.007	-1.71	-.23
3	1	-.423	.230	.166	-.98	.13
	2	.971*	.307	.007	.23	1.71

*. The mean difference is significant at the 0.05 level.

Future decision-making competence

Table C19: Descriptive statistics future decision-making competence

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
1	33	5.42	.969	.169	5.08	5.77	3	7
2	33	3.82	1.509	.263	3.28	4.35	2	7
3	32	5.34	.827	.146	5.05	5.64	4	7
Total	98	4.86	1.355	.137	4.59	5.13	2	7

Table C20: Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
FutureDecisions	Based on Mean	8.151	2	95	.001
	Based on Median	3.995	2	95	.022
	Based on Median and with adjusted df	3.995	2	68.422	.023
	Based on trimmed mean	7.971	2	95	.001

Table C21: ANOVA future decision-making competence

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	53.812	2	26.906	20.582	.000
Within Groups	124.188	95	1.307		
Total	178.000	97			

Table C22: Robust Tests of Equality of Means future decision-making competence

	Statistic[a]	df1	df2	Sig.
Welch	14.879	2	60.978	.000

a. Asymptotically F distributed.

Table C23: Multiple Comparisons future decision-making competence

Games-Howell

		Mean Difference (I-J)			95% Confidence Interval	
(I) Group	(J) Group		Std. Error	Sig.	Lower Bound	Upper Bound
1	2	1.606*	.312	.000	.85	2.36
	3	.080	.223	.931	-.46	.62
2	1	-1.606*	.312	.000	-2.36	-.85
	3	-1.526*	.301	.000	-2.25	-.80
3	1	-.080	.223	.931	-.62	.46
	2	1.526*	.301	.000	.80	2.25

*. The mean difference is significant at the 0.05 level.

Annex D

Word associations data analysis

Visionary

Table D1: Crosstab Visionary

			Visionary		
			No	Yes	Total
Group	Negative	Count	19	14	33
		Expected Count	11.1	21.9	33.0
		% within Group	57.6%	42.4%	100.0%
		Adjusted Residual	3.6	-3.6	
	None	Count	7	25	32
		Expected Count	10.8	21.2	32.0
		% within Group	21.9%	78.1%	100.0%
		Adjusted Residual	-1.7	1.7	
	Positive	Count	7	26	33
		Expected Count	11.1	21.9	33.0
		% within Group	21.2%	78.8%	100.0%
		Adjusted Residual	-1.9	1.9	
Total		Count	33	65	98
		Expected Count	33.0	65.0	98.0
		% within Group	33.7%	66.3%	100.0%

Table D2: General Chi-Square Test Visionary

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	12.730[a]	2	.002
Likelihood Ratio	12.500	2	.002
N of Valid Cases	98		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.78.

Table D3: Specific Chi-Square Test Visionary – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	9.138[a]	1	.003		
Continuity Correction[b]	7.679	1	.006		
Likelihood Ratio	9.410	1	.002		
Fisher's Exact Test				.005	.003
N of Valid Cases	66				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.00.

b. Computed only for a 2x2 table

Table D4: Specific Chi-Square Test Visionary – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.628[a]	1	.003		
Continuity Correction[b]	7.204	1	.007		
Likelihood Ratio	8.884	1	.003		
Fisher's Exact Test				.005	.003
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 12.80.

b. Computed only for a 2x2 table

Table D5: Specific Chi-Square Test Visionary – POSITIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.004[a]	1	.948		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.004	1	.948		
Fisher's Exact Test				1.000	.593
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.89.

b. Computed only for a 2x2 table

Inspiring

Table D6: Crosstab Inspiring

			Inspiring		
			No	Yes	Total
Group	Negative	Count	29	4	33
		Expected Count	23.9	9.1	33.0
		% within Group	87.9%	12.1%	100.0%
		Adjusted Residual	2.4	-2.4	
	None	Count	24	8	32
		Expected Count	23.2	8.8	32.0
		% within Group	75.0%	25.0%	100.0%
		Adjusted Residual	.4	-.4	
	Positive	Count	18	15	33
		Expected Count	23.9	9.1	33.0
		% within Group	54.5%	45.5%	100.0%
		Adjusted Residual	-2.8	2.8	
Total		Count	71	27	98
		Expected Count	71.0	27.0	98.0
		% within Group	72.4%	27.6%	100.0%

Table D7: General Chi-Square Test Inspiring

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	9.340[a]	2	.009
Likelihood Ratio	9.538	2	.008
N of Valid Cases	98		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.82.

Table D8: Specific Chi-Square Test Inspiring – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	8.943[a]	1	.003		
Continuity Correction[b]	7.391	1	.007		
Likelihood Ratio	9.381	1	.002		
Fisher's Exact Test				.006	.003
N of Valid Cases	66				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.50.

b. Computed only for a 2x2 table

Table D9: Specific Chi-Square Test Inspiring – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.790[a]	1	.181		
Continuity Correction[b]	1.037	1	.309		
Likelihood Ratio	1.816	1	.178		
Fisher's Exact Test				.215	.154
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.91.

b. Computed only for a 2x2 table

Table D10: Specific Chi-Square Test Inspiring – POSITIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.973[a]	1	.085		
Continuity Correction[b]	2.146	1	.143		
Likelihood Ratio	3.009	1	.083		
Fisher's Exact Test				.120	.071
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 11.32.

b. Computed only for a 2x2 table

Excellent

Table D11: Crosstab Excellent

			Excellent		
			No	Yes	Total
Group	Negative	Count	30	3	33
		Expected Count	27.6	5.4	33.0
		% within Group	90.9%	9.1%	100.0%
		Adjusted Residual	1.4	-1.4	
	None	Count	26	6	32
		Expected Count	26.8	5.2	32.0
		% within Group	81.3%	18.8%	100.0%
		Adjusted Residual	-.5	.5	
	Positive	Count	26	7	33
		Expected Count	27.6	5.4	33.0
		% within Group	78.8%	21.2%	100.0%
		Adjusted Residual	-.9	.9	
Total		Count	82	16	98
		Expected Count	82.0	16.0	98.0
		% within Group	83.7%	16.3%	100.0%

Table D12: General Chi-Square Test Excellent

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.979[a]	2	.372
Likelihood Ratio	2.132	2	.344
N of Valid Cases	98		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.22.

Table D13: Specific Chi-Square Test Excellent – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.886[a]	1	.170		
Continuity Correction[b]	1.061	1	.303		
Likelihood Ratio	1.932	1	.165		
Fisher's Exact Test				.303	.152
N of Valid Cases	66				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 5.00.

b. Computed only for a 2x2 table

Table D14: Specific Chi-Square Test Excellent – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1.271[a]	1	.260		
Continuity Correction[b]	.590	1	.442		
Likelihood Ratio	1.290	1	.256		
Fisher's Exact Test				.303	.222
N of Valid Cases	65				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 4.43.

b. Computed only for a 2x2 table

Table D15: Specific Chi-Square Test Excellent – POSITIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.062[a]	1	.804		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.062	1	.804		
Fisher's Exact Test				1.000	.525
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.40.

b. Computed only for a 2x2 table

Courageous

Table D16: Crosstab Courageous

			Courageous		
			No	Yes	Total
Group	Negative	Count	9	24	33
		Expected Count	8.1	24.9	33.0
		% within Group	27.3%	72.7%	100.0%
		Adjusted Residual	.5	-.5	
	None	Count	9	23	32
		Expected Count	7.8	24.2	32.0
		% within Group	28.1%	71.9%	100.0%
		Adjusted Residual	.6	-.6	
	Positive	Count	6	27	33
		Expected Count	8.1	24.9	33.0
		% within Group	18.2%	81.8%	100.0%
		Adjusted Residual	-1.0	1.0	
Total		Count	24	74	98
		Expected Count	24.0	74.0	98.0
		% within Group	24.5%	75.5%	100.0%

Table D17: General Chi-Square Test Courageous

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.077[a]	2	.584
Likelihood Ratio	1.115	2	.573
N of Valid Cases	98		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.84.

Table D18: Specific Chi-Square Test Courageous – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.776[a]	1	.378		
Continuity Correction[b]	.345	1	.557		
Likelihood Ratio	.781	1	.377		
Fisher's Exact Test				.558	.279
N of Valid Cases	66				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.50.

b. Computed only for a 2x2 table

Table D19: Specific Chi-Square Test Courageous – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.006[a]	1	.939		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.006	1	.939		
Fisher's Exact Test				1.000	.579
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.86.

b. Computed only for a 2x2 table

Table D20: Specific Chi-Square Test Courageous – POSITIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.905[a]	1	.341		
Continuity Correction[b]	.431	1	.511		
Likelihood Ratio	.909	1	.340		
Fisher's Exact Test				.389	.256
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.38.

b. Computed only for a 2x2 table

Arrogant

Table D21: Crosstab Arrogant

			Arrogant		
			No	Yes	Total
Group	Negative	Count	26	7	33
		Expected Count	29.6	3.4	33.0
		% within Group	78.8%	21.2%	100.0%
		Adjusted Residual	-2.6	2.6	
	None	Count	31	1	32
		Expected Count	28.7	3.3	32.0
		% within Group	96.9%	3.1%	100.0%
		Adjusted Residual	1.6	-1.6	
	Positive	Count	31	2	33
		Expected Count	29.6	3.4	33.0
		% within Group	93.9%	6.1%	100.0%
		Adjusted Residual	1.0	-1.0	
Total		Count	88	10	98
		Expected Count	88.0	10.0	98.0
		% within Group	89.8%	10.2%	100.0%

Table D22: General Chi-Square Test Arrogant

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	6.733[a]	2	.035
Likelihood Ratio	6.495	2	.039
N of Valid Cases	98		

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is 3.27.

Table D23: Specific Chi-Square Test Arrogant – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.216[a]	1	.073		
Continuity Correction[b]	2.058	1	.151		
Likelihood Ratio	3.381	1	.066		
Fisher's Exact Test				.149	.074
N of Valid Cases	66				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 4.50.

b. Computed only for a 2x2 table

Table D24: Specific Chi-Square Test Arrogant – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	4.924[a]	1	.026		
Continuity Correction[b]	3.391	1	.066		
Likelihood Ratio	5.486	1	.019		
Fisher's Exact Test				.054	.030
N of Valid Cases	65				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 3.94.

b. Computed only for a 2x2 table

Table D25: Specific Chi-Square Test Arrogant – POSITIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.318[a]	1	.573		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.324	1	.569		
Fisher's Exact Test				1.000	.512
N of Valid Cases	65				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 1.48.

b. Computed only for a 2x2 table

Unfocused

Table D26: Crosstab Unfocused

			Unfocused		
			No	Yes	Total
Group	Negative	Count	19	14	33
		Expected Count	26.3	6.7	33.0
		% within Group	57.6%	42.4%	100.0%
		Adjusted Residual	-3.9	3.9	
	None	Count	28	4	32
		Expected Count	25.5	6.5	32.0
		% within Group	87.5%	12.5%	100.0%
		Adjusted Residual	1.4	-1.4	
	Positive	Count	31	2	33
		Expected Count	26.3	6.7	33.0
		% within Group	93.9%	6.1%	100.0%
		Adjusted Residual	2.5	-2.5	
Total		Count	78	20	98
		Expected Count	78.0	20.0	98.0
		% within Group	79.6%	20.4%	100.0%

Table D27: General Chi-Square Test Unfocused

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	15.262[a]	2	.000
Likelihood Ratio	14.988	2	.001
N of Valid Cases	98		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.53.

Table D28: Specific Chi-Square Test Unfocused – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.880[a]	1	.001		
Continuity Correction[b]	9.983	1	.002		
Likelihood Ratio	13.032	1	.000		
Fisher's Exact Test				.001	.001
N of Valid Cases	66				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.00.

b. Computed only for a 2x2 table

Table D29: Specific Chi-Square Test Unfocused – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	7.265[a]	1	.007		
Continuity Correction[b]	5.848	1	.016		
Likelihood Ratio	7.603	1	.006		
Fisher's Exact Test				.012	.007
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.86.

b. Computed only for a 2x2 table

Table D30: Specific Chi-Square Test Unfocused – POSITIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.804[a]	1	.370		
Continuity Correction[b]	.219	1	.640		
Likelihood Ratio	.817	1	.366		
Fisher's Exact Test				.427	.321
N of Valid Cases	65				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 2.95.

b. Computed only for a 2x2 table

Incompetent

Table D31: Crosstab Incompetent

			Incompetent		
			No	Yes	Total
Group	Negative	Count	31	2	33
		Expected Count	32.3	.7	33.0
		% within Group	93.9%	6.1%	100.0%
		Adjusted Residual	-2.0	2.0	
	None	Count	32	0	32
		Expected Count	31.3	.7	32.0
		% within Group	100.0%	0.0%	100.0%
		Adjusted Residual	1.0	-1.0	
	Positive	Count	33	0	33
		Expected Count	32.3	.7	33.0
		% within Group	100.0%	0.0%	100.0%
		Adjusted Residual	1.0	-1.0	
Total		Count	96	2	98
		Expected Count	96.0	2.0	98.0
		% within Group	98.0%	2.0%	100.0%

Table D32: General Chi-Square Test Incompetent

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	4.021[a]	2	.134
Likelihood Ratio	4.436	2	.109
N of Valid Cases	98		

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is .65.

Table D33: Specific Chi-Square Test Incompetent – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.063[a]	1	.151		
Continuity Correction[b]	.516	1	.473		
Likelihood Ratio	2.835	1	.092		
Fisher's Exact Test				.492	.246
N of Valid Cases	66				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 1.00.

b. Computed only for a 2x2 table

Table D34: Specific Chi-Square Test Incompetent – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	2.001[a]	1	.157		
Continuity Correction[b]	.485	1	.486		
Likelihood Ratio	2.773	1	.096		
Fisher's Exact Test				.492	.254
N of Valid Cases	65				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is .98.

b. Computed only for a 2x2 table

Chaotic

Table D35: Crosstab Chaotic

			Chaotic		
			No	Yes	Total
Group	Negative	Count	17	16	33
		Expected Count	24.9	8.1	33.0
		% within Group	51.5%	48.5%	100.0%
		Adjusted Residual	-3.9	3.9	
	None	Count	28	4	32
		Expected Count	24.2	7.8	32.0
		% within Group	87.5%	12.5%	100.0%
		Adjusted Residual	1.9	-1.9	
	Positive	Count	29	4	33
		Expected Count	24.9	8.1	33.0
		% within Group	87.9%	12.1%	100.0%
		Adjusted Residual	2.0	-2.0	
Total		Count	74	24	98
		Expected Count	74.0	24.0	98.0
		% within Group	75.5%	24.5%	100.0%

Table D36: General Chi-Square Test Chaotic

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	15.492[a]	2	.000
Likelihood Ratio	14.899	2	.001
N of Valid Cases	98		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 7.84.

Table D37: Specific Chi-Square Test Chaotic – POSITIVE vs. NEGATIVE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	10.330[a]	1	.001		
Continuity Correction[b]	8.680	1	.003		
Likelihood Ratio	10.877	1	.001		
Fisher's Exact Test				.003	.001
N of Valid Cases	66				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 10.00.

b. Computed only for a 2x2 table

Table D38: Specific Chi-Square Test Chaotic – NEGATIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	9.876[a]	1	.002		
Continuity Correction[b]	8.259	1	.004		
Likelihood Ratio	10.411	1	.001		
Fisher's Exact Test				.003	.002
N of Valid Cases	65				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 9.85.

b. Computed only for a 2x2 table

Table D39: Specific Chi-Square Test Chaotic – POSITIVE vs. NONE

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.002[a]	1	.963		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.002	1	.963		
Fisher's Exact Test				1.000	.628
N of Valid Cases	65				

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 3.94.

b. Computed only for a 2x2 table

Annex E

Summary statistics participants

All respondents	Total	98	100%
Gender	Female	56	57,1%
	Male	42	42,9%
Age	Under 26	93	94,9%
	26 or older	5	5,1%
Nationality	Portuguese	81	82,7%
	Chinese	6	6,1%
	French	3	3,1%
	German	2	2,0%
	Slovenian	2	2,0%
	Indian	1	1,0%
	Polish	1	1,0%
	Indonesian	1	1,0%
	Croatian	1	1,0%
Program	MSc in Management	91	92,9%
	Erasmus	2	2,0%
	MSc in Marketing	2	2,0%
	Entrepreneurship	1	1,0%
	MSc in Business Admin.	1	1,0%
	Msc in Int. Management	1	1,0%