STRATEGY MAP EFFECTS ON MANAGERS' STRATEGY REVISION JUDGMENTS

by

Brian Dean Knox

Bachelor of Science in Accountancy, Arizona State University, 2011

Master of Accountancy, Southern Utah University, 2013

Submitted to the Graduate Faculty of

Joseph M. Katz Graduate School of Business in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy in Business Administration

University of Pittsburgh

2017

UNIVERSITY OF PITTSBURGH

JOSEPH M. KATZ GRADUATE SCHOOL OF BUSINESS

This dissertation was presented

by

Brian Dean Knox

It was defended on

June 30, 2017

and approved by

Jongwoon (Willie) Choi
Assistant Professor of Business Administration and Fryrear Faculty Fellow in Accounting
University of Pittsburgh

Marc N. Coutanche
Assistant Professor
University of Pittsburgh
Department of Psychology

John H. Evans III
Katz Alumni Professor and Area Director of Accounting
University of Pittsburgh

Dhinu Srinivasan
Associate Professor of Accounting
University of Pittsburgh

Dissertation Advisor: Donald V. Moser
Professor of Business Administration and Dean's Excellence Faculty Fellow in Accounting
University of Pittsburgh

ABSTRACT

STRATEGY MAP EFFECTS ON MANAGERS' STRATEGY REVISION JUDGMENTS

Brian Dean Knox, PhD

University of Pittsburgh, 2017

Managers make *strategy revision judgments*, which are judgments that affect how well the firm can revise its strategy when new information comes to light. Using two studies, my dissertation examines how formatting the firm's strategy as a strategy map affects two types of strategy revision judgments. First, I study middle managers' judgments on passing along new information to upper management. Second, I study managers' judgments of the relevance of new information and the appropriateness of the firm's strategy.

In my first study, I find that middle managers are more likely to withhold new information from upper management when they feel that information would be less impressive to upper management. Middle managers also tend to punish their subordinates with less positive performance evaluations when the subordinates provide them with such less impressive information. However, middle managers with sufficient experience who receive a strategy map are more likely to pass along such less impressive information to upper management than those with comparable experience who do not receive a strategy map. In my second study, I find that receiving a strategy map improves managers' judgements of the relevance of new information. I also find that receiving a strategy map improves managers' judgments about the appropriateness

of the firm's strategy in light of this new information. However, I find this latter effect depends on whether it is easy to understand the cause-and-effect relationships depicted in a strategy map.

Finally, in an extensions chapter, I propose three neuroimaging studies that extend the above studies. One of these neuroimaging studies more fully describes the motivation, theory, and method of the study. This study approaches the relationship between strategy maps and strategy revision differently, extending prior research that suggests a strategy map leads workers to better allocate effort between short-term focused and long-term focused activities. I hope to provide evidence on the neural processing, and thus the thought processes, that underlie this prior finding. Such evidence would improve practitioners' predictions of how long the effect would persist over time, which informs practitioners about whether to revise the firm's strategy to include a strategy map.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I acknowledge and thank the many individuals who have contributed to my doctoral education and the completion of this dissertation. Without their help, I could not have completed it. Each committee member—Don Moser (chair), Willie Choi, Marc Coutanche, Harry Evans, and Dhinu Srinivasan—has made a direct positive impact on this dissertation, and I thank them for it. Don Moser, specifically, has patiently and diligently assisted me in a multitude of ways. Even the most glowing words of gratitude seem to do no justice to his service throughout these past years. For their helpful comments, I thank Kristy Towry, Paul Fischer, and participants at the 2016 AAA/Deloitte Foundation/J. Michael Cook Doctoral Consortium. I thank Mandy M. Cheng and Kerry A. Humphreys for providing me with their experimental instrument, which helped inspire Study 2. This dissertation has benefitted greatly from the hardworking Katz doctoral office, including Dennis Galletta, Chris Gursky, and Carrie Woods, who I heartily thank for their assistance at every step in the process. Lastly, I acknowledge and thank my wife, my children, and my faith for vital assistance throughout the completion of this dissertation.

1.0 INTRODUCTION

A firm's *strategy* is a plan of action that is expected to move the firm "from its present position to a desirable but uncertain future position" (Kaplan and Norton 2001a, 176). Managers often invest substantial resources to develop a strategy that is likely to move the firm toward a desired future state. Managers must then revise the firm's strategy as new information comes to light and as circumstances change. Otherwise, the firm will likely veer off the path toward its desired future state. This dissertation examines how the presentation format of the firm's strategy affects managers' ability to implement necessary strategy revisions. I examine this through two separate experiments, focusing on a widely-used format for presenting the firm's strategy: a strategy map. A *strategy map* is a chain of strategic goals that are joined together by arrows to communicate that one goal causes the next. These cause-and-effect relationships communicated by a strategy map are called *causal linkages*. Ultimately, my two studies provide evidence that a strategy map can improve managers' ability to revise the firm's strategy when necessary, if the strategy map meets certain criteria.

In my first study, I find that middle managers tend to withhold information from upper management when that information contradicts the firm's strategy. I then investigate whether receiving a strategy map helps alleviate this biased treatment of information that contradicts the

firm's strategy.[1] Upper management relies on middle managers to pass along information from lower levels of the firm. Without such information, upper management will have a more difficult time evaluating and revising the firm's strategy (Kaplan and Norton 2008; Dutton et al. 1997). In the absence of a strategy map, I find that middle managers are less likely to pass along information to upper management when that information contradicts the firm's strategy because they believe that such information is less impressive to upper management. I also find evidence that middle managers tend to evaluate their subordinates' performance less positively when the subordinates provide middle managers with information that contradicts the firm's strategy than when the subordinates provide information that supports the firm's strategy. These are troubling findings because upper management needs information that contradicts the firm's strategy—even more than information that supports the strategy. Top managers are likely already convinced that the firm's current strategy is the right strategy, thus, it is more difficult to convince them that a change is necessary than that the firm's current course is optimal (see Huelsbeck et al. 2011).

I find that experienced middle managers (i.e. ≥ 2 years of management experience) are more likely to pass along to upper management information that contradicts the firm's strategy when they receive a strategy map than when they do not receive a strategy map. Using mediation analysis, I find that this is because experienced middle managers who receive a strategy map believe that such contradictory information is more impressive to upper management than experienced middle managers who do not receive a strategy map. This result suggests that upper

[1] Throughout Section 3, I refer to middle managers withholding *knowledge*, rather than *information*. This is because I conduct Study 1 using the tradition of prior research on knowledge sharing (e.g. Hwang et al. 2009; Jensen and Meckling 1992). However, I refer to middle managers withholding *information*, rather than *knowledge*, in the remaining sections of this dissertation (e.g. Section 1, Section 2, and Section 6). This is done to align the findings of Study 1 to the more general questions of how managers use information for strategy revision. Thus, for the purposes of this dissertation, *knowledge* and *information* are used interchangeably.

management is better able to appropriately revise the firm's strategy when they provide middle managers with a strategy map, assuming the middle managers have sufficient experience. Interestingly, I do not find that receiving a strategy map affects middle managers' bias in their evaluation of subordinates' performance.

In my second study, I test whether a strategy map improves, first, managers' judgments of how relevant new information is for revising the firm's strategy and, second, their judgments of whether that new information suggests the firm's strategy is appropriate or inappropriate. Both judgments affect how accurately managers can use new information in strategy revisions, but judgments about the appropriateness of the firm's strategy can have a stronger effect on strategy revision. This is because poor judgments of information relevance can lead managers to discard relevant information and erroneously consider irrelevant information, but such poor judgments do not directly lead them down the wrong strategic path. In contrast, poor judgments of firm strategy appropriateness *directly* lead managers to pursue courses of action that exacerbate existing mismatches between the firm's strategy and the business environment, i.e. making it more likely that inappropriate strategies are kept and appropriate strategies are changed.

Cheng and Humphreys (2012) find that receiving a strategy map improves both of the above judgments, but other prior research suggests that the effects of a strategy map can depend on causal linkage strength (see Webb 2004). Causal linkages are strong when it is easy to understand why the cause leads to the effect (Luft 2004 outlines several reasons firms have difficulty creating strong causal linkages). As such, I manipulate two features of a strategy map in my second study: the presence of a strategy map's causal linkages and the strength of those causal linkages. I find that the causal linkages of a strategy map, regardless of their strength, improve managers' judgments regarding the relevance of new information (like Cheng and

Humphreys 2012). However, I then find that only strong causal linkages improve managers' judgments about the appropriateness of the firm's strategy. In one subsample, weak causal linkages even lead to *worse* judgments about the appropriateness of the firm's strategy. These results suggest that a strategy map effect discovered by prior research depends on strong causal linkages, specifically a strategy map effect that improves managers' judgments related to strategy revision.

My two studies, combined, support Kaplan and Norton's (2008) suggestion that receiving a strategy map benefits a firm by improving the type of information managers receive and the judgments they make once they receive this information. That is, a strategy map benefits a firm by improving managers' strategy revision judgments. The first study shows how a strategy map helps information flow in a less biased manner to decision makers and the second study shows how a strategy map helps decision makers recognize relevant new information and inappropriate strategies, given that new information. I note that neither study finds the effect of a strategy map to be a night-and-day reversal from bad judgments to good judgments. Instead, both effects are incremental and most likely occur at the margin. That said, I believe my findings are still economically significant. First, judgments about new information are pervasive and constant. New information is continuously arriving and managers must continuously respond to that new information by making judgments like those examined in this dissertation. Therefore, even an improvement at the margin of each judgment could substantially improve a firm's ability, over time, to stay on track toward its desired future state in the face of a changing business environment. Second, managers' strategy revision judgments can have enormous economic implications. Kaplan and Norton (2008) highlight new information's role in "Intel's switch from memory chips to microprocessors, Honda's shift in emphasis from motorcycles to automobiles,

and 3M Corporation's marketing of Post-it notes" (275). When new information suggests a *major* revision to the firm's strategy, managers' judgments related to that information have outsized financial effects, making even minor improvements economically significant.

This dissertation also includes a proposed extensions chapter in which I propose three studies that use neuroscience theory and neuroimaging techniques to extend the two studies described above. I develop one of these three proposals into a proposed study that lays out the motivation, hypotheses, and method with sufficient detail that the proposed study could feasibly be executed soon after completion of this dissertation. This proposed study seeks to shed light on how thought processes change when a strategy map is provided. Farrell et al. (2012) find that a strategy map improves how workers allocate their effort between long-term focused and short-term focused activities. My proposed neuroimaging study has implications for the persistence of this effect over time because some thought process changes are likely to be more persistent than others.

This proposed study differs from my first and second study, because it addresses a different aspect of the relationship between a strategy map and strategy revision. My two completed studies address how a strategy map affects strategy revision judgments by altering how managers consider information that is used to revise the firm's strategy. In contrast, my proposed study, once completed, would provide evidence of whether a strategy map's effect on workers' effort allocations is persistent over time, which evidence could lead managers to consider using or not using a strategy map to communicate the firm's strategy. Changing how the firm's strategy is communicated amounts to a revision of the format of the firm's strategy. Thus, any results from my proposed study, once completed, would be like the article summaries that managers in my second study judge for relevance to revising the firm's strategy.

2.0 BACKGROUND

2.1 OBSTACLES TO STRATEGY REVISION

Given that firms exist to serve their owners, managers entrusted with decision rights within firms are expected to create a plan of action that they expect to accomplish the owners' desires. A plan of action that is designed to accomplish long-term fundamental goals is called a *strategy* (Kaplan and Norton 2001a; 2001b). Most owners, as rational economic decision makers, invest in a firm with the object of increasing their long-term wealth. This motivation is assumed throughout this chapter. However, not all firms can maximize owner wealth using exactly the same strategy. Instead, the optimal strategy will vary from firm to firm based on contextual factors, such as market share, product position, cost function, supply chain, customer base, human capital, etc. (see Porter 1996). For example, Walmart's strategy is to use its superior supply chain and bargaining power to sell goods at the lowest price feasible. Apple, in contrast, does not shy away from high prices for its products. Instead, Apple's strategy is to ensure that its products are perceived as being *different* from competitors' and therefore worth a premium price. Managers' must discern the optimal strategy for the firm given its unique contextual factors, otherwise known as its business environment.

Once a strategy has been developed, managers then must revise the strategy as new information comes to light, continuously optimizing the firm's strategy to be the best possible plan of action for maximizing shareholder wealth, given the information available to managers. New information could reflect a change in the underlying environment or it could be the revelation of a previously unknown but longstanding aspect of the environment. Either way,

managers can successfully revise the firm's strategy only if they become aware of this new information and can understand its implications, updating their prior assumptions about how the firm should best navigate its environment. Beer et al. (2005) emphasize the importance of attending to information from a changing environment (see also Day and Schoemaker 2005).

> The environment is abundant with changes: changing customer demands and preferences, technological advances, global competitors, innovative strategies. This leads us to consider that organizations modify and adapt (and thus evolve) their designs in response to environmental and organizational changes. In a rapidly changing environment, such as that faced by contemporary organizations, organizational fitness – the capacity to learn and adapt – becomes especially important. This entails fusing existing organizational capabilities with new capabilities to fit new circumstances (Beer et al. 2005, 447).

My dissertation focuses on two obstacles to optimal strategy revision that managers face. The first obstacle to optimal strategy revision is that important information might not come to managers' attention because those responsible for passing information upward through the firm often have an interest in the outcome (see Dutton et al. 1997; Prendergast 1993). For example, Park et al. (2011) find evidence of middle managers' tendency to withhold bad news from CEOs—that is, news that contradicts the firm's strategy. This leads CEOs to persist in suboptimal strategies. Park et al. emphasize the detrimental effect this has on CEOs themselves, who later must answer to shareholders for poorly performing strategies. In contrast, I focus on the detrimental effect this has on the firm's strategy and on firm performance, providing direct experimental evidence of this phenomenon and its causes.

The first obstacle is best described in hierarchical terms: the obstacle of middle managers withholding information from upper management when they believe this information contradicts the firm's strategy. The second obstacle is not a hierarchical issue, but rather can be present at any level of the firm where the relevance of new information and the appropriateness of the firm's strategy must be judged. This obstacle arises because managers have finite cognitive resources for evaluating a seemingly endless stream of new information. Therefore, managers

must quickly decide the relevance of new information as it comes to them so as not to waste time and cognitive resources on irrelevant information and then must quickly determine what relevant information indicates about the firm's strategy. Managers' cognitive limitations can sometimes cause them not to recognize the relevance of new information or to misinterpret whether the new information suggests the firm's strategy is appropriate or inappropriate (see Campbell et al. 2015). Thus, the second obstacle is that managers' cognitive limitations inevitably lead to noisy judgments about new information's relevance and about the strategy's appropriateness.

2.2 STRATEGY MAPS AND STRATEGY REVISION

A strategy map is a format for presenting the firm's strategy that could help firms overcome the two obstacles to proper strategy revision detailed in Section 2.1. Strategy maps cannot be fully understood without considering their place within the balanced scorecard framework. More than two decades ago, Robert Kaplan and David Norton published their first book on the *balanced scorecard*, a framework for developing both strategic goals and accompanying performance measures that indicate progress toward those goals (Kaplan and Norton 1996). The key innovation of the balanced scorecard was to require these goals and performance measures to be spread across four categories: *financial*, *customer*, *internal processes*, and *learning and growth*. Three of these categories are non-financial, and Kaplan and Norton suggest that success in these categories precedes financial success. By balancing a firm's strategy across four categories that include both lagging financial performance and leading non-financial performance, Kaplan and Norton sought to encourage managers to balance their effort across short-term *and* long-term

outcomes, rather than solely focusing their energy on short-term financial outcomes to earn immediate performance bonuses.

Within five years, Kaplan and Norton (2001a; 2001b) added strategy maps to this framework. A strategy map is a visual representation of the suggested leading-lagging nature of strategic goals within the balance scorecard's four categories. A strategy map links goals from different categories together, with arrows to communicate that accomplishing one goal causes accomplishment of the next, i.e. causal linkages. Typically, learning and growth goals lead to internal processes goals, internal processes goals lead to customer goals, and customer goals lead to financial goals.

Accounting researchers have found strategy maps to have several judgment and decision making effects, some positive and some negative. Banker et al. (2004) find that managers' performance evaluations weight more heavily performance measures that are connected to the chain of goals on a strategy map over those that are freestanding apart from a strategy map (see also Banker et al. 2011). Choi et al. (2013, 2012) suggest that decision makers can treat performance measures as surrogates for the strategy itself (which a strategy map attempts to embody), potentially leading to suboptimal decisions. Relatedly, Mastilak et al. (2012) report that receiving a strategy map leads managers to evaluate worker's performance without regard to their own beliefs about outcome controllability, instead deferring to the outcome controllability suggested by the strategy map. Webb (2004) finds that, depending on the perceived strength of the cause-and-effect relationships between causally-linked goals, strategy maps can increase goal commitment.[2] Similarly, Farrell et al. (2012) present evidence that receiving information about

[2] Webb (2004) uses the more general term *strategic performance measurement system*, which is a term that includes balanced scorecards with strategy maps.

the cause-and-effect relationships between goals—like what is done in a strategy map—helps workers allocate effort more optimally between short-term and long-term activities.

Kaplan and Norton (2008) suggest that strategy maps also affect managers' judgments that relate to the two obstacles to proper strategy revision that I discuss in Section 2.1. Academic research related to this suggestion has been limited and mixed. On the organizational level, Gimbert et al. (2010) suggest that strategy reformulation improves when firms establish cause-and-effect relationships between strategic goals, which is a characteristic of a strategy map. This improvement could be due to strategy maps helping firms overcome obstacles to proper strategy revision, although there is no direct evidence of this presented by Gimbert et al. Likewise, Decoene and Bruggeman (2006) conduct a case study and find that balanced scorecards (including strategy maps) help align middle managers' motivations to that of upper management, potentially alleviating the obstacle of middle managers withholding information from upper management. However, this study might not generalize beyond the firm in question. In contrast, Bisbe and Malagueño (2012) find that Gimbert et al.'s findings are limited to firms implementing a strategy map in environments with low levels of change. This is not the result that would be expected if receiving a strategy map helps managers overcome obstacles to strategy revision, because these obstacles should become more and more pronounced in high-change environments.

On an individual level, there are two prior studies that suggest a strategy map can help overcome the identified obstacles to proper strategy revision. Accordingly, in this dissertation I conduct two new studies to answer key questions left unanswered by these prior studies. First, Taylor (2010) presents evidence that receiving a strategy map helps managers overcome his or her cognitive biases. This result indirectly suggests that receiving a strategy map would help

middle managers share important information with upper management, helping firms overcome the first obstacle I describe in Section 2.1. However, it is also clear that Tayler does not directly test this, and such interpretations are not fully supported by his results. The biases tested in his experiment are distinct from those I expect middle managers to face. Also, he only finds a strategy map to help overcome managers' biases when it is accompanied by a mechanism that allows managers to choose their own performance measures. As such, Study 1 (beginning in Section 3.0) effectively extends Tayler's results by testing whether a strategy map can overcome a different set of manager biases and whether this effect obtains when managers do not choose their own performance measures.

Second, Cheng and Humphreys (2012) provide direct evidence that receiving a strategy map helps managers recognize the relevance of new information and the appropriateness of the firm's strategy in light of that new information. This provides direct evidence that receiving a strategy map helps overcome the second obstacle identified in Section 2.1. However, this result might not apply when the causal linkages of a strategy map are weak—something Cheng and Humphreys do not test (see also Luft 2004; Webb 2004). To determine if Cheng and Humphreys' findings depend on causal linkage strength, I conduct Study 2 (beginning in Section 4.0).

3.0 STUDY 1: THE EFFECT OF A STRATEGY MAP ON DECISIONS TO SHARE KNOWLEDGE

3.1 INTRODUCTION

A strategy map is a popular format for presenting the firm's strategy. When using a strategy map, upper management explicitly and visually asserts causal linkages between the firm's strategic goals by drawing arrows between these goals. Kaplan and Norton (2008; 2001a; 2001b) explain that strategy maps' explicit causal linkages help managers revise the firm's strategy when such revisions are necessary. They suggest that this improvement arises for two main reasons. First, they expect strategy maps to improve upper management's awareness of when a strategy revision is necessary (Kaplan and Norton 2008; 2001a; see also Campbell et al. 2015). Second, they explain that subordinates are more aware of ground-level circumstances that necessitate strategy revisions than their supervisors, and they suggest that strategy maps improve these employees' ability to recognize and communicate these ground-level circumstances to upper management. (Kaplan and Norton 2008; 2001a).

Recent research provides evidence supporting the first point. Campbell et al. (2015) explain how Store 24 executives could have recognized their strategy was failing about one year earlier than they did if they had conducted strategy map-related statistical analysis. Tayler (2010) finds that strategy maps direct managers' attention to focus on the outcomes and timing to be expected if the strategy is successful. Tayler shows that this helps mitigate motivated reasoning that otherwise biases managers' decision making. Cheng and Humphreys (2012) find that a strategy map's causal linkages alter how managers later recall the firm's strategy such that

managers make better judgments about the relevance and strategic implications of new external information when they receive a strategy map.

However, this prior research does not examine Kaplan and Norton's second point that strategy maps positively affect how subordinates communicate about revising the firm's strategy. These communications involve *specific knowledge*, a term used by Jensen and Meckling (1992) to describe knowledge about the firm's ground-level realities that is held by subordinates and is both hard to aggregate and costly to communicate. Specific knowledge is typically conveyed to upper management by middle managers. My study examines the extent to which middle managers share different types of specific knowledge with upper management and examines whether receiving a strategy map positively affects this knowledge sharing. My first research question asks whether middle managers share specific knowledge differently based on whether the specific knowledge disconfirms the firm's strategy or confirms the firm's strategy. I hereafter refer to these two types of specific knowledge as *disconfirming specific knowledge* and *confirming specific knowledge*, respectively. Middle managers are often compensated using subjective performance measures. Prior research suggests that such measures would lead middle managers to share disconfirming specific knowledge less than they share confirming specific knowledge because disconfirming specific knowledge contradicts upper management (e.g. Park et al. 2011; Prendergast 1993). My second research question then asks whether receiving a strategy map improves middle managers' knowledge sharing, specifically their willingness to share disconfirming specific knowledge to upper management. I summarize these research questions as follows.

1) Do middle managers share objectively useful disconfirming specific knowledge less than they share objectively useful confirming specific knowledge?

2) Does receiving a strategy map improve how middle managers share disconfirming specific knowledge?

To answer these research questions, I conduct an experiment in which participants act as middle managers who, according to the scenario provided, receive objectively useful specific knowledge from a subordinate. Participants first indicate how likely they are to pass along this specific knowledge to upper management, and then they indicate how receiving this specific knowledge affects their overall performance evaluation of the subordinate who provided it. I manipulate whether the specific knowledge is disconfirming or confirming. I also manipulate whether the firm's strategy is presented as a strategy map or as a list of the firm's strategic goals without causal linkages between them. I hereafter refer to a list of goals without causal linkages as a *non-causal list* and treat it as the baseline condition for how upper management presents the firm's strategy.

My study has four major findings, two concerning my first research question and two concerning my second research question. First, I find that middle managers indicate that they are less likely to pass along disconfirming specific knowledge than confirming specific knowledge. I find evidence that this is because middle managers believe that upper management will have a worse impression of them if they pass along disconfirming specific knowledge. This belief also motivates middle managers to discount the credibility and relevance of disconfirming specific knowledge. Second, I find that middle managers give less positive performance evaluations to subordinates who communicate disconfirming specific knowledge than to subordinates who communicate confirming specific knowledge. Like my first finding, I find that this effect is also driven by the impression middle managers believe upper management will have of them and by middle managers' beliefs about the credibility and relevance of the specific knowledge.

Third, among middle managers with at least two years of management experience, I find that receiving a strategy map has a positive effect on the likelihood that they will pass along disconfirming specific knowledge. I find evidence that this is because those who receive a strategy map expect that sharing disconfirming specific knowledge will impress upper management more than those who do not receive a strategy map. Fourth, I do *not* find that receiving a strategy map mitigates middle managers' biased performance evaluations of subordinates who provide them with disconfirming specific knowledge.

My results have significant implications for practice in today's fast-paced and highly-competitive economy. For a firm to be successful, upper management must create a strategy that fits well with the firm's business environment and must be able to change that strategy as environmental conditions change. Disconfirming specific knowledge can be *more* useful to revising the firm's strategy than confirming specific knowledge because upper management already believes the content communicated by confirming specific knowledge. My findings show that, given subjective performance measures, middle managers are more likely to pass along confirming specific knowledge than disconfirming specific knowledge, thus biasing the type of specific knowledge upper management is likely to hear. Also, because middle managers give more positive performance evaluations to subordinates who communicate confirming specific knowledge, subordinates are likely to skew the type of specific knowledge they provide to middle managers toward confirming specific knowledge. Over time, this bias in the type of specific knowledge middle managers prefer to receive and pass along could dramatically and negatively affect upper management's ability to make crucial adjustments to the firm's strategy. Although some components of these findings are in line with prior analytical and empirical research, I am unaware of any prior research that provides behavioral evidence of this double

effect of middle managers' bias against disconfirming specific knowledge. Specifically, prior research that examines agents' willingness to report disconfirming specific knowledge does not examine agents' biased evaluations of their subordinates who provide them with disconfirming specific knowledge (e.g. Park et al. 2011). Thus, my study is useful to practitioners because it provides a more complete understanding of the problem than prior research.

Furthermore, my finding that a strategy map can mitigate one downside of subjective performance measures contributes both to practice and to the subjective performance measurement literature. This finding contributes to practice by supporting Kaplan and Norton's point that receiving a strategy map can improve the communication of specific knowledge to upper management. In a certain way, I provide empirical evidence that helps explain the popularity of balanced scorecards (and their accompanying strategy maps). My study suggests that practitioners may want to use a strategy map in combination with subjective performance measures to reduce some of the downsides of subjective performance measures. My study also demonstrates how research on subjective performance measures can be expanded by considering how upper management's framing of the firm's strategic goals affects middle managers' behavior in response to such performance measures.

I note that my findings do not suggest a strategy map overcomes *all* the obstacles impeding disconfirming specific knowledge. That is, I do not find support for the notion that receiving a strategy map alleviates middle managers' bias in their performance evaluations of subordinates, suggesting that firms that use a strategy map are just as likely to experience that obstacle over time as firms that do not use a strategy map. Also, I only find that a strategy map benefits middle managers' knowledge sharing when those middle managers have achieved a minimum level of experience. I interpret the experience requirement as evidence that a strategy

map affects hypothesis testing norms that usually require a minimum threshold of time to internalize. However, the minimum level of management experience suggested in my results, i.e. two years of management experience, is hardly an unrealistic expectation for most middle managers in today's firms.

3.2 BACKGROUND

3.2.1 Knowledge Sharing and Subjective Performance Measures

One line of knowledge sharing research focuses on the behavioral effects of mandatory knowledge sharing among agents (e.g. Hannan et al. 2013; Towry 2003). For example, Towry (2003) considers a mutual monitoring system in which knowledge of one's effort level is shared with a peer agent. She examines how much effort this monitoring system elicits from the monitored agent when the peer agent must then report this information to a supervisor rather than speaking directly to the worker in question. She finds that the effectiveness of this policy depends on how well the team has forged a team identity. Similarly, Hannan et al. (2013) examine a mutual monitoring policy in which teammates are required to share knowledge. In this case, the authors find that sharing this knowledge amplifies workers' pre-existing tendencies toward using this information to compete or collude. Haesebrouck et al. (2015) examine how knowledge sharing affects voluntary helping behaviors when these helping behaviors mandatorily include knowledge sharing as well. Furthermore, the expansive literature on the effect of relative performance information on effort overlaps considerably with examinations of

the effect of mandatory knowledge sharing (e.g. Tafkov 2012; Hannan et al. 2008; Kerr et al. 2007).

However, these studies lack significant examination of how agents *voluntarily* share knowledge. A separate strand of knowledge sharing research examines how incentives can motivate voluntary knowledge sharing. I rely on this latter strand of research to establish the theoretical background for my study. Typically, these studies begin with the assumption that an agent possesses private knowledge, and then the studies examine how to optimally motivate this agent to share his or her private knowledge for the benefit of the firm (e.g. Hwang et al. 2009; Raith 2008; see also Baiman and Demski 1980). An interesting element of these studies is how motivation to invest high levels of effort interacts with motivation to report private knowledge truthfully.

Hwang et al. (2009) demonstrate that when agents have more private knowledge or when the sharing of their knowledge is more valuable to the firm, the firm relies more on output-based performance measures and less on input-based performance measures (see also Raith 2008). Output-based performance measures compensate agents based on firm output rather than how much work, i.e. input, the agent puts into achieving that output. This type of performance measure motivates agents to share their private knowledge because doing so is likely to help firm output, thus assisting agents in achieving their output-based goals. This reasoning is in line with the *revelation principle*, which holds that in any private knowledge situation there exists a contract that induces truthful reporting of this private knowledge (see Myerson 1982; 1979; Gibbard 1973). Output-based performance measures are often part of that revelatory, truth-inducing contract.

Although output-based performance measures can induce truthful reporting, they can lead risk-averse agents to invest low levels of effort. The output of the firm is influenced by things outside the agent's control. For risk-averse agents, lack of control over output makes output-based performance measures less attractive because the agent might invest high levels of effort but nonetheless achieve low output. This intuition is confirmed by Hwang et al. (2009), who find that when firm output is noisier, firms rely less on output-based performance measures. One way firms optimize agents' effort with output-based performance measures is by combining them with *subjective performance measures*. Subjective performance measures can be advantageous to induce high levels of effort from risk-averse agents because supervisors can adjust for uncontrollable events that lead to poor output (see Bol 2008; Gibbs et al. 2004). Thus, subjective performance measures can reduce agents' risk of investing high levels of effort but nonetheless achieving low output.

However, subjective performance measures can negatively affect their knowledge sharing, the opposite effect as output-based performance measures: positively affecting effort but negatively affecting truthful reporting. Prendergast (1993) develops a model demonstrating how subjective performance measures can negatively affect truthful reporting. His model shows that subjective performance evaluation can lead agents to be less truthful in their reports, because they tend to report what the supervisors want to hear to curry favor on subjective performance measures.[3] Prendergast demonstrates the necessary tradeoff between incentives to motivate high levels of effort, i.e. subjective performance evaluations, and incentives that motivate truthful reporting, i.e. objective performance evaluations such as output-based measures. My study is like

[3] An extensive "yes men" literature has developed additional models and gathered empirical evidence that refine and build upon Prendergast's study (for example, see Gneezy et al. 2017; Garciano and Posner 2005; Bentley MacLeod 2003; Khatri and Tsang 2003; Ewerhart and Schmitz 2000; Avery et al. 1999).

this model in that I assume a firm uses subjective performance measures for middle managers, in addition to any objective performance measures. I therefore expect these subjective performance measures to lead middle managers to be biased against reporting knowledge that contradicts the firm's strategy. As explained in Section 3.3.3, I expect that a strategy map is one way firms reduce subjective performance measures' detrimental effect on truthful reporting without sacrificing their positive effect on effort motivation.

3.2.2 Specific Knowledge and Strategy Revision

A firm's strategy is a set of goals that upper management believes will move the firm from its current state to a more desired state, which is typically a state of increased long-term shareholder wealth (Kaplan and Norton 2008, 2001a). However, upper management may not possess the knowledge necessary to craft the optimal strategy. Jensen and Meckling (1992) argue that knowledge of the firm's ground-level realities is asymmetrically distributed within the firm such that those with higher hierarchical rank have less of this knowledge. This knowledge of ground-level conditions is referred to by Jensen and Meckling as *specific knowledge*. Senior managers in the firm cannot possess as much specific knowledge as the collective knowledge of employees working at the ground level of the firm due to cognitive limitations (see also Hwang et al. 2009; Raith 2008; Beer et al. 2005; Brews and Hunt 1999; Dutton et al. 1997).[4] Given the

[4] Dutton et al. (1997) summarize these cognitive limitations as follows. "While organizational environments are becoming increasingly complex, dynamic and interdependent…the information-processing capacity of the top management group remains stable and is inadequate for detecting, interpreting and handling these environmental challenges…. It is often middle managers rather than the top managers who have their hands on the 'pulse of the organization' and are closer to customers and other stakeholders" (Dutton et al. 1997, 407).

limitations of upper management's knowledge of the firm's ground-level realities, the firm's strategy may be sub-optimal, i.e. it may not be a *good fit* for its environment.

An example of specific knowledge used by Jensen and Meckling (1999) is the understanding a machine operator has of the machine he or she works with. This specific knowledge is not easily aggregated, and it is costly for the operator to transfer this knowledge to supervisors, i.e. it is specific knowledge. If upper management issues a firm strategy that is built on the assumption that the firm can improve throughput by hiring more skilled machine operators, specific knowledge about the machine's operation could suggest a revision to upper management's strategy. For example, if the machine operator knew from experience that the constraint on the machine's throughput was not operator skill level but rather the accessibility of replacement parts, then the machine operator's specific knowledge could help upper management revise the firm's strategy.[5]

My research questions are based on the premise that upper management relies on middle managers as conduits for the flow of specific knowledge to upper management (see Dutton et al. 2001; 1997; Mintzberg and Waters 1985). I expect that middle managers "have their hands on the 'pulse of the organization'" (Dutton et al. 1997, 407) because they are recipients of specific knowledge from subordinates who are directly involved in the ground-level realities of the firm. This premise is in line with numerous prior studies, some of which are in the management literature, cited above, and some of which are in the auditing literature. Kadous et al. (2017) summarize this literature on auditor knowledge sharing by arguing that lower-level auditors are

[5] Per Kaplan and Norton (2008, 254), "many ideas for new strategic options can arise from employees closest to customers and processes." For example, employees' specific knowledge is argued to have "triggered Intel's switch from memory chips to microprocessors, Honda's shift in emphasis from motorcycles to automobiles, and 3M Corporation's marketing of Post-it notes" (Kaplan and Norton 2008, 275).

those at the front lines of the audit, gathering most of the evidence, and if they do not voluntarily report issues to their supervisors, those supervisors are unlikely to find the issues by themselves (see, as cited by Kadous et al. 2017, Bennett and Hatfield 2013; Vera-Muñoz et al. 2006; Rich et al. 1997).

3.3 HYPOTHESIS DEVELOPMENT

3.3.1 Impression Management and Disconfirmation Bias

Given upper management's cognitive limitations, middle managers should not automatically pass along *all* specific knowledge to upper management. Rather, middle managers must judge the likely usefulness of specific knowledge to upper management and only pass along specific knowledge that is likely to be useful. As stated above, specific knowledge tends to include qualitative elements that make it costly to aggregate. My study focuses on the likelihood that middle managers will pass along subordinates' specific knowledge to upper management. I describe this as *reporting likelihood* or the likelihood that middle managers, once they receive useful specific knowledge, will pass it along to upper management. (Reporting likelihood is one of the main dependent variables in my study.)

As argued above, subjective performance measures can lead middle managers to be less truthful in their reporting, favoring reports that are likely to impress upper management over reports that are unlikely to impress upper management (Prendergast 1993). This prediction can be understood as an effect of two psychological influences: (1) *impression management*, which leads to (2) *disconfirmation bias*. Impression management is one person's effort to affect others'

impression of him or her (Bozeman and Kacmar 1997). Upper management can influence future raises and promotions for middle managers, especially through subjective performance evaluations (Bol 2008). This influence could lead middle managers to engage in impression management to secure good subjective performance measures. Middle managers consider whether passing along different types of specific knowledge will impress upper management (Bozeman and Kacmar 1997; see also Jones 1964).

If a middle manager passes along specific knowledge suggesting that the firm's strategy is *not* appropriate for its environment, i.e. disconfirming specific knowledge, upper management could interpret this as an indication that middle managers are not doing enough to implement the firm's strategy. That is, disconfirming specific knowledge could be interpreted by upper management as an indication that middle managers are investing a low level of effort.[6] In contrast, upper management is unlikely to interpret confirming specific knowledge as an indication that middle managers are investing a low level of effort. Thus, impression management theory leads to the prediction that middle managers are likely to prefer passing along confirming specific knowledge more than disconfirming specific knowledge.

Disconfirmation bias is a cognitive bias characterized by individuals scrutinizing evidence that disconfirms their preferred outcome more than they scrutinize evidence that confirms their preferred outcome, because this additional scrutiny helps conjure up reasons for rejecting disconfirming evidence (Edwards and Smith 1996; Lord et al. 1979; see also Taber and

[6] Ittner and Larcker (2003) suggest that managers often do not empirically test the financial outcome effect of non-financial strategic goals because managers only use non-financial strategic goals they consider to have a "self-evident" (91) connection to financial outcomes. Huelsbeck et al. (2011) find that when senior managers are convinced of the firm's strategy they are unlikely to change that strategy based on a disconfirming signal (see also Campbell et al. 2015). Based on these studies, upper management is likely to be convinced of the firm's strategy. Thus, middle managers likely expect upper management to favor a *user error* interpretation of disconfirming specific knowledge, i.e. interpreting disconfirming knowledge as an indication that middle managers are not implementing the firm's strategy correctly.

Lodge 2006; Kunda 1990). As stated above, middle managers likely prefer passing along specific knowledge that makes a relatively positive impression on upper management, especially by showing that the firm's strategy is a good fit. Based on disconfirmation bias theory, then, middle managers who receive disconfirming specific knowledge are likely to subject that knowledge to more scrutiny than those who receive confirming specific knowledge. Thus, middle managers who receive disconfirming specific knowledge are likely to conjure up reasons that knowledge is irrelevant or not credible, which leads to lower reporting likelihood for disconfirming specific knowledge than for confirming specific knowledge.[7]

H1: Middle managers are less likely to report disconfirming specific knowledge to upper management than confirming specific knowledge.

This prediction is important because disconfirming specific knowledge can be *more* valuable to upper management than confirming specific knowledge (e.g. see Campbell et al. 2015; Huelsbeck et al. 2011). Upper management already agrees with the underlying facts conveyed by confirming specific knowledge and is already prepared to follow the course of action suggested by confirming specific knowledge. In contrast, disconfirming specific knowledge runs counter to upper management's current beliefs about the firm, suggesting a course of action that upper management likely, *a priori*, considers imprudent. It can take significant repetition of disconfirming signals for a firm to change course. For example, Store 24 received poor feedback from customers about its new strategy for two years before upper management was convinced the new strategy was failing (Campbell et al. 2015).

[7] Because my study examines when middle managers inappropriately withhold specific knowledge, I use specific knowledge that is objectively useful to upper management.

3.3.2 Middle Managers' Evaluations of Subordinates

Specific knowledge flows to middle managers from subordinates who are better informed about the ground-level conditions of the firm, and middle managers then decide whether to report this specific knowledge to upper management. Middle managers also often give subjective performance evaluations for subordinates, affecting subordinates' future raises and promotions. I expect both impression management and disconfirmation bias to influence how middle managers evaluate subordinates' performance (which is hereafter referred to as *subordinate evaluation*).

I expect this because middle managers are often compensated using both objective and subjective performance measures. While subjective performance measures motivate middle managers to withhold disconfirming specific knowledge from upper management, objective performance measures still typically motivate middle managers to seek objectively strong outcomes. Middle managers could interpret disconfirming specific knowledge as a sign that subordinates themselves are not doing enough to make the firm's strategy work, which would reduce middle managers' chances of achieving objective performance measures. This is the same interpretation of disconfirming specific knowledge that middle managers project onto upper management. That is, the same disconfirming specific knowledge that middle managers expect to give upper management the unfavorable impression that middle managers are investing a low level of effort also gives middle managers the unfavorable impression that subordinates are investing a low level of effort. As such, I expect middle managers to give less positive performance evaluations to subordinates who provide disconfirming specific knowledge than to subordinates who provide confirming specific knowledge.

Furthermore, disconfirmation bias leads middle managers to scrutinize disconfirming specific knowledge more than they scrutinize confirming specific knowledge, giving middle

managers more opportunity to internally criticize that disconfirming specific knowledge as less credible and less relevant. A subordinate who provides less credible and less relevant specific knowledge is likely to be viewed less favorably and is likely to receive a less positive subordinate evaluation than a subordinate who provides specific knowledge that is viewed as more credible and more relevant.

> **H2:** Middle managers evaluate subordinates who provide disconfirming specific knowledge less positively than they evaluate subordinates who provide confirming specific knowledge.

3.3.3 A Strategy Map and Impression Management

When upper management uses a strategy map, they arrange the firm's strategic goals in a causal sequence, with arrows drawn between goals to indicate how they are believed to drive and cause each other. The arrows that adjoin goals in a strategy map can be called *causal linkages* because these arrows suggest a relationship of cause-and-effect between strategic goals. A strategy map differs from alternative methods of communicating strategic goals that do not define a causal sequence among the goals, i.e. non-causal lists. Causally connecting strategic goals is a defining characteristic of *strategic performance measurement systems*, in general, which are defined as systems of performance measurement that attempt to align managers' incentives to the firm's strategy (Gimbert et al. 2010; see also Webb 2004). Strategy maps, together with a balanced scorecard, form a type of strategic performance measurement system (Kaplan and Norton 2004; 2001a). I focus on strategy maps because the balanced scorecard is one of the most widespread strategic performance measurement systems (Rigby and Bilodeau 2009). It is possible, although I do not directly examine it, that the results of this study could also apply to other types of strategic performance measurement systems.

My second research question asks whether receiving a strategy map improves middle managers' judgments about sharing disconfirming specific knowledge. I expect that receiving a strategy map *does* improve those judgments because it alters middle managers' incentives, changing how disconfirming specific knowledge affects impression management. Thus, I expect that receiving a strategy map moderates the relationship between the type of specific knowledge, i.e. confirming or disconfirming, and impression management.

This is because a strategy map presents middle managers with a set of hypotheses about how a firm's goals causally interact (see Kaplan and Norton 2001a) and the norm for hypothesis testing is that negative results can be commonplace (otherwise the hypothesis would not be worth testing). By explicitly stating these hypothesized relationships, upper management implicitly invites members of the firm to test these hypotheses. Indeed, this is one of the original purposes of strategy maps, per Kaplan and Norton. "The key, then, to implementing strategy is […] to test the hypotheses continually, and to use those results to adapt as required" (Kaplan and Norton 2001a, 176). Receiving a strategy map allows for an alternative reason for middle managers to pass along disconfirming specific knowledge apart from the middle manager investing a low level of effort. Instead, in the context of a strategy map, a high level of effort includes, or even requires, hypothesis testing and thus the reporting of negative results when they occur. This leads to the prediction, captured in my third hypothesis, that receiving a strategy map increases reporting likelihood for middle managers who receive disconfirming specific knowledge.

> **H3:** When middle managers receive a strategy map, they have higher reporting likelihood for disconfirming specific knowledge to upper management than when they do not receive a strategy map.

Using the same logic, I expect that receiving a strategy map mitigates at least part of H2, which predicts bias in middle managers' performance evaluations of subordinates who provide

them with disconfirming specific knowledge. I argue that H2 is due to middle managers interpreting disconfirming specific knowledge as a sign that subordinates are not doing enough to make the firm's strategy work. If receiving a strategy map alleviates middle managers' concerns about passing along disconfirming specific knowledge due to the norms of hypothesis testing, then middle managers who receive a strategy map could also be less concerned about subordinates participating in hypothesis testing.

H4: When middle managers receive a strategy map, they give higher evaluations to subordinates who provide disconfirming specific knowledge than when middle managers do not receive a strategy map.

3.4 METHOD

3.4.1 Experimental Design

I use an experiment to test H1 through H4. I manipulate two between-subject factors, each with two levels, resulting in a 2 × 2 design. The two between-subject factors are the presence of a strategy map and specific knowledge type. To manipulate the presence of a strategy map, I provide participants with either a strategy map or a non-causal list made up of the same strategic goals. The difference between the two levels of this factor is the presence or absence of the causal linkages used in a strategy map to indicate causal linkages between strategic goals. I manipulate specific knowledge by providing participants with either confirming specific knowledge or disconfirming specific knowledge. The confirming specific knowledge supports the firm's strategic goals and the disconfirming specific knowledge contradicts the firm's strategic goals. Participants are randomly assigned to these four between-participant conditions.

3.4.2 Participants

My participants are Amazon Mechanical Turk workers, who provide their responses online and are paid a flat fee of $1.50 for participation. Mechanical Turk workers have been shown to be effective proxies for employees in prior studies (see Farrell et al. 2017; also see Mason and Suri 2012 for a review). I follow Farrell et al.'s pre-screening procedure to make it more likely that participants' responses are meaningful. Specifically, at the time they participate in my study, all workers have a "master" qualification, more than 1,000 successfully completed tasks, and an approval rating of 95% or above for their work. I also only allow workers to participate if Amazon has verified they are United States residents. My initial sample includes 340 complete participant responses. I further screen this sample by excluding twelve responses that were from an IP address that was unacceptable (I exclude one IP address that is from outside of the United States, one IP address that is from the city where I was based at the time, and ten IP addresses that are duplicated among multiple participants).[8] Thus, my final sample consists of responses from 328 participants.

3.4.3 Experimental Task and Procedures

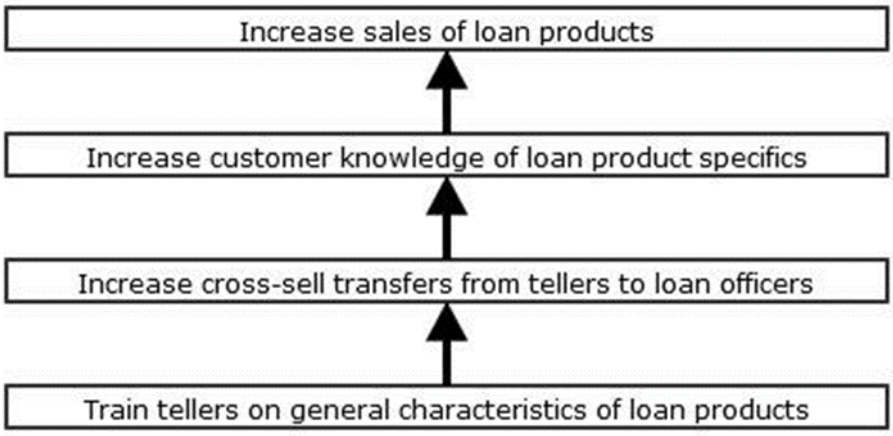
The experiment consists of six stages, is computer-based, and participants advance from screen to screen with no opportunity to return to prior screens. In stage one, participants learn of

[8] I exclude responses from IP addresses outside of the United States because I cannot independently confirm Amazon's verification procedures. I am aware of reports that some workers from outside the United States have circumvented those procedures in the past. I exclude IP addresses from the city where I was based at the time as an extra experimental control to exclude responses that may be from individuals at my institution or otherwise connected to me who might have known the study was being conducted and knew of its experimental design.

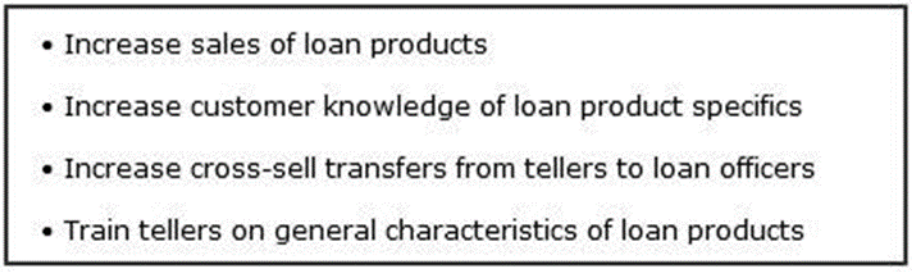
their role as a newly-hired regional manager at First Palisade Bank, a fictional retail bank in the United States. Participants learn that, as regional manager, they oversee three bank branches and report to the CEO. They also learn that First Palisade Bank employs two types of customer service representative: tellers (who perform banking services at the counter) and loan officers (who assist customers with purchasing and maintaining loan products at desks apart from the counter). The scenario indicates that upper management determines regional managers' future raises and promotions because I believe this is salient to middle managers in practice (Park et al. 2011).

Figure 1: Study 1 Presence of a Strategy Map Manipulation

Strategy Map Condition



Non-Causal List Condition



In stage two, participants read an email from the CEO describing the CEO's new strategy for First Palisade Bank. I manipulate whether this new strategy is a strategy map or a non-causal list of the same strategic goals (see Figure 1).[9] I maintain the order of the goals across conditions, while simply omitting, from the non-causal list condition, arrows that cue strategy map-like causal linkages.

In stage three, participants complete a distractor task to clear their working memory, because my research questions are most relevant to situations where middle managers retrieve the firm's strategy from long-term memory (see Cheng and Humphreys 2012). In the distractor task, participants have one minute to count the number of times the letter "c" appears in a short passage.[10] To maintain proper control over this distractor task, participants are not able to advance from this screen before one minute has passed and after a minute they are automatically advanced to the next screen where they input their count of how many times the letter "c" was used.

Figure 2: Study 1 Specific Knowledge Type Manipulation

Disconfirming Specific Knowledge Condition

"Customers talk to loan officers more often now but don't get a better understanding of the specifics of loan products. In fact, a lot of customers are unhappy about taking the time to talk to a loan officer."

Confirming Specific Knowledge Condition

"Customers talk to loan officers more often now and get a better understanding of the specifics of loan products. In fact, a lot of customers are happy about taking the time to talk to a loan officer."

[9] As in the experiment by Cheng and Humphreys (2012), I have participants assume the role of a *newly hired* manager and provide them with a *new* strategy from the CEO. This helps avoid concerns that participants would assume prior beliefs, which could create confounding motivated reasoning-related behavior (see Tayler 2010).
[10] I use a passage of a public domain, English text acquired through Project Gutenberg: *www.gutenberg.org*.

In stage four, participants receive feedback from one of their subordinate branch managers. This feedback consists of either confirming or disconfirming specific knowledge, as shown in Figure 2. On the next screen is stage five, where participants use a scale of -3 (very unlikely) to +3 (very likely) to indicate the likelihood that they would report the specific knowledge presented in stage four to the CEO. Participant responses on this screen provide my main dependent measure of reporting likelihood.

In stage six, participants respond to various post-experiment questions. They respond to recognition questions that test how easily they can access the firm's strategic goals from long-term memory. They also answer questions to test whether they made the causal inferences I expect them to make when they are given a strategy map. Participants then indicate how they perceive reporting the branch managers' specific knowledge to the CEO would affect the CEO's impression of them. They also indicate, in separate questions, how credible and relevant they perceive the branch manager's specific knowledge to be. They answer a question on how the specific knowledge they received from the branch manager affects how they would evaluate the performance of that branch manager. Lastly, they answer manipulation check questions and provide demographics.

3.4.4 Measures and Constructs

Specific knowledge type is an indicator variable I use to distinguish between participants given disconfirming specific knowledge (coded as one) and participants given confirming specific knowledge (coded as zero). I use the indicator variable *strategy map*, which is coded as one when participants receive the firm's strategy in the form of a strategy map and coded as zero when participants receive the firm's strategy in the form of a non-causal list of the same strategic

goals. *Reporting likelihood* is the dependent measure for which participants indicate the likelihood that they would report the specific knowledge they received to upper management, measured on a 7-point scale from -3 (very unlikely) to +3 (very likely). *Subordinate evaluation* is also a dependent measure, drawn from participants' responses to the question, "Does the branch manager's feedback lead you to have a more positive or more negative evaluation of the branch manager's overall performance?" This variable is measured on a 7-point scale from -3 (much more negative) to +3 (much more positive).

CEO impression is made up of participants' responses to the question, "Do you believe reporting the branch manager's feedback to the CEO leads the CEO to have a more positive or more negative impression of you?" It is measured on a 7-point scale from -3 (much more negative impression of you) to +3 (much more positive impression of you). I ask two questions that I use to create a latent variable, *subjective usefulness*. I describe the factor analysis used to create this latent variable below in Section 3.5.1. The two questions I ask are called *perceived credibility* and *perceived relevance*.[11] *Perceived credibility* is a participant's response to the question, "How credible do you find the branch manager's feedback?" measured on a 7-point scale, from -3 (not at all credible) to +3 (very credible). *Perceived relevance* is a participant's response to the question, "How relevant do you believe the branch manager's feedback is to evaluating the CEO's new strategy?" measured on a 7-point scale from -3 (very irrelevant) to +3 (very relevant).

[11] I also asked *perceived importance* which is a participant's response to the question, "How important do you find the branch manager's feedback?" measured on a 7-point scale from -3 (very unimportant) to +3 (very important). However, because this question is overly broad and less closely connected to the concept of disconfirmation bias, I do not consider it as a question measuring disconfirmation bias. Confirming this intuition, if I include *perceived importance* as a predictor of my latent variable measuring disconfirmation bias, the goodness of fit statistics in my SEM models in Section 4 are substantially reduced.

3.5 RESULTS

3.5.1 Descriptive Statistics and Factor Analysis

In Table 1, I provide mean and standard deviations for several variables, broken out by experimental condition. Panel A shows *reporting likelihood*, and Panel B shows *subordinate evaluation*. Panel C shows the latent variable *subjective usefulness*, and Panel D shows *CEO impression.*

Table 1: Study 1 Descriptive Statistics

Panel A: Reporting Likelihood: means (standard deviations)

	Non-causal List	Strategy Map	*Row Averages*
Confirming Specific Knowledge	2.31 (1.05) n = 78	2.39 (0.81) n = 83	2.35 (0.93) n = 161
Disconfirming Specific Knowledge	1.48 (1.57) n = 82	1.69 (1.36) n = 85	1.59 (1.47) n = 167
Column Averages	1.88 (1.40) n = 160	2.04 (1.17) n = 168	1.96 (1.29) n = 328

Panel B: Subordinate Evaluation: means (standard deviations)

	Non-causal List	Strategy Map	*Row Averages*
Confirming Specific Knowledge	1.81 (1.02) n = 78	1.92 (0.98) n = 83	1.86 (1.00) n = 161
Disconfirming Specific Knowledge	0.89 (1.35) n = 82	0.75 (1.50) n = 85	0.82 (1.42) n = 167
Column Averages	1.34 (1.28) n = 160	1.33 (1.39) n = 168	1.33 (1.34) n = 328

Table 1: Study 1 Descriptive Statistics (continued)

Panel C: Subjective Usefulness: means (standard deviations)

	Non-causal List	Strategy Map	*Row Averages*
Confirming Specific Knowledge	0.05 (0.73) n = 78	0.13 (0.70) n = 83	0.09 (0.71) n = 161
Disconfirming Specific Knowledge	-0.22 (0.93) n = 82	0.04 (0.72) n = 85	-0.09 (0.84) n = 167
Column Averages	-0.09 (0.85) n = 160	0.08 (0.71) n = 168	0.00 (0.78) n = 328

Panel D: CEO Impression: means (standard deviations)

	Non-causal List	Strategy Map	*Row Averages*
Confirming Specific Knowledge	1.66 (1.15) n = 78	1.82 (1.09) n = 83	1.74 (1.12) n = 161
Disconfirming Specific Knowledge	0.23 (1.38) n = 82	0.46 (1.37) n = 85	0.35 (1.38) n = 167
Column Averages	0.93 (1.46) n = 160	1.13 (1.41) n = 168	1.03 (1.44) n = 328

Below I examine H1 through H4 using a latent variable, *subjective usefulness*, to proxy for the role of disconfirmation bias in my statistical tests. I name the variable *subjective usefulness* because when a person exhibits disconfirmation bias, he or she is likely to subjectively downgrade the perceived usefulness of the evidence that disconfirms his or her preferred outcome. I use a latent variable rather than individuals' raw responses because I want to extract the central underlying construct of disconfirmation bias, which is only partly captured by each raw response.[12] I create this latent variable using load factors derived from factor

[12] This decision is also supported empirically. When I use raw responses for disconfirmation bias, either alone or together, the model's goodness of fit statistics are substantially reduced.

analysis of my two observed variables related to disconfirmation bias, which I interpret as proxies for the overall concept of usefulness: *perceived credibility* and *perceived relevance.* Only one factor has an eigenvalue close to one (eigenvalue = 0.99). My two observed variables have the same load factor: *perceived credibility* (0.705) and *perceived relevance* (0.705). *Subjective usefulness* has a mean of 0.00, a median of 0.339, and a standard deviation of 0.784.

3.5.2 Tests of H1

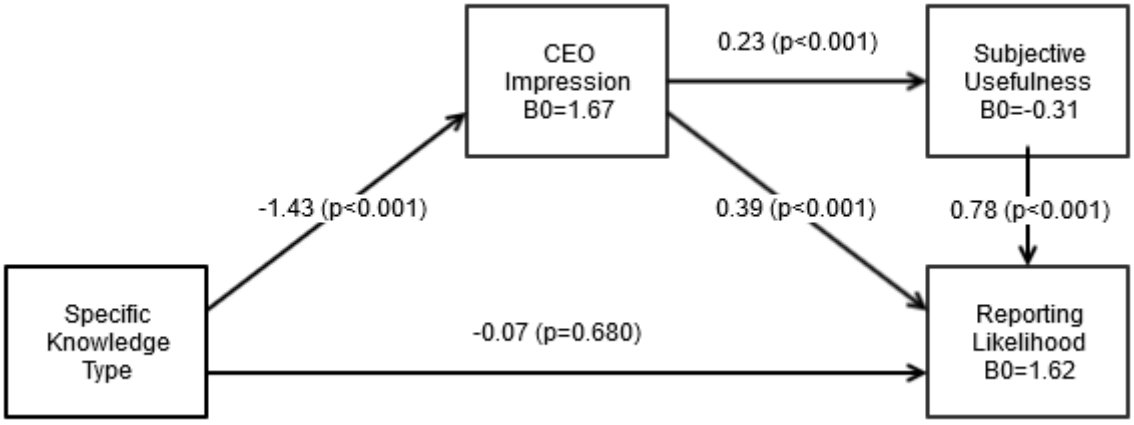
For all statistical tests of H1, I only use responses from participants who received a non-causal list. I do this because I hypothesize in H3 that *strategy map* affects *reporting likelihood*, and therefore the best assurance of internal validity for my H1 tests is to test H1 using only responses from participants who did not receive a strategy map. If I were to test H1 using responses from participants in both the strategy map and non-causal list condition, it would be difficult to disentangle H1 and H3 effects.

In H1, I predict middle managers are less likely to report disconfirming specific knowledge than confirming specific knowledge. I test this using a two-sample t-test that compares the variable *reporting likelihood* between the two groups indicated by the variable *specific knowledge type*. These two groups reflect participants who received disconfirming specific knowledge (*specific knowledge type* = 1) and confirming specific knowledge (*specific knowledge type* = 0). I find that *specific knowledge type* has a significant negative effect on *reporting likelihood* ($t = 5.58$, $p < 0.001$ one-tailed), meaning that *reporting likelihood* for participants given disconfirming specific knowledge (mean = 1.59) is significantly lower than it is for participants given confirming specific knowledge (mean = 2.35). This result is consistent with H1, suggesting that middle managers are biased against reporting disconfirming specific

knowledge to upper management. Results supporting H1 are important because H1 suggests that upper management may not receive the crucial disconfirming information they need in order to revise the firm's strategy.

In Section 3.3.1, I argue that the effects predicted in H1 are caused by impression management and disconfirmation bias, with the former leading to the latter. This argument suggests a series of mediating relationships. Accordingly, I use SEM analysis to examine whether my measure of impression management mediates the effect of *specific knowledge type* on *reporting likelihood* and whether *subjective usefulness* mediates some of the effect of impression management on *reporting likelihood*. SEM analysis uses a set of regression models to examine multiple mediation pathways simultaneously. This makes SEM analysis an effective statistical test for my H1 predictions about the mediating roles played by impression management and then disconfirmation bias.

Figure 3: Study 1 SEM Model for H1



Goodness of fit statistics: n = 160; χ^2 (1) = 0.447 (p = 0.504); CFI = 1.000; RMSEA < 0.001

As for my SEM model (shown in Figure 3, along with results and goodness of fit statistics), the model first tests a simple linear regression to show that *specific knowledge type*

has a significant negative effect on *CEO impression* ($p < 0.001$ one-tailed).[13] Then the model shows that *CEO impression* has a significant positive effect on *subjective usefulness* ($p < 0.001$ one-tailed). This shows that when specific knowledge is perceived to be less impressive to upper management, participants tend to view it as less credible and less relevant. Lastly, the model tests all three predictors (*specific knowledge type*, *CEO impression*, and *subjective usefulness*) in a multiple linear regression predicting *reporting likelihood*. For comparison, above I report that in a univariate t-test, which omits any mediators, *specific knowledge type* is a significant predictor ($t = 5.58$, $p < 0.001$ one-tailed) of *reporting likelihood*. In this final SEM step, however, *specific knowledge type* is non-significant ($p = 0.680$ two-tailed) as a predictor of *reporting likelihood*, while my mediators (*CEO impression* and *subjective usefulness*) are significant predictors of *reporting likelihood* (both with $p < 0.001$, one-tailed). This suggests that the effect of *specific knowledge type* on *reporting likelihood* observed in my initial tests of H1 are fully explained by my measures of impression management and disconfirmation bias (Baron and Kenny 1986). Furthermore, I conduct a Sobel test on the *CEO impression* mediation path (which mediates *specific knowledge type* and *reporting likelihood*), without including *subjective usefulness* for simplicity. I find a significant indirect effect ($p < 0.001$ one-tailed) and a non-significant direct effect ($p = 0.954$ two-tailed). This result further supports an interpretation of complete mediation.

The SEM model supports my reasoning behind H1. First, the results from the SEM model are consistent with impression management affecting disconfirmation bias, and with both disconfirmation bias and impression management playing a role in dissuading middle managers from reporting disconfirming specific knowledge. Importantly, the SEM model shows that my *subjective usefulness* measure does not explain the entirety of the effect of impression

[13] All SEM analysis in this study is completed using Stata/SE, version 14.2.

management, because *CEO Impression* is still a significant predictor of *reporting likelihood* even with *subjective usefulness* included in the model.

3.5.3 Tests of H2

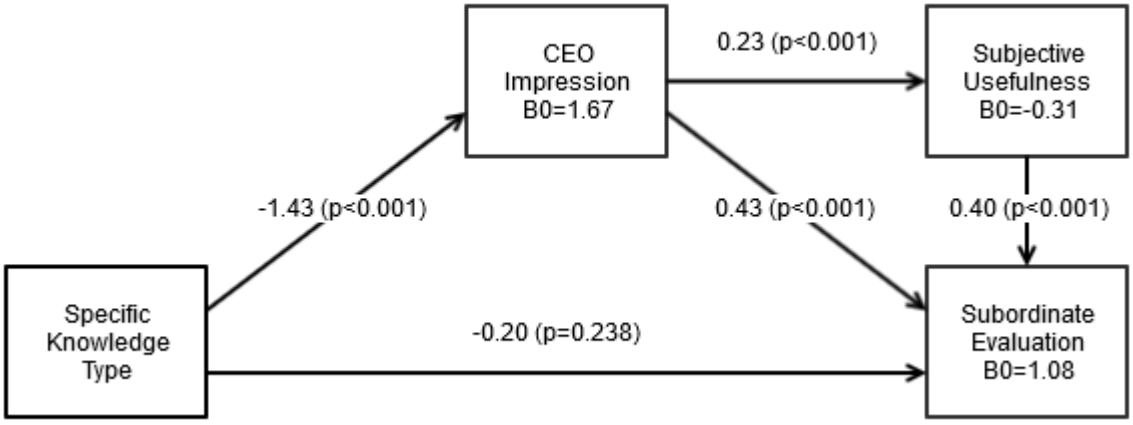
Similar to my tests for H1, all statistical tests of H2 only use responses from participants who received a non-causal list. Again, I hypothesize in H4 that *strategy map* affects *subordinate evaluation*. I can, therefore, more effectively test H2 by excluding responses from participants in the strategy map condition. In H2, I predict that middle managers will evaluate subordinates who provide disconfirming specific knowledge less positively than subordinates who provide confirming specific knowledge. I test this hypothesis using a two-sample t-test that compares the variable *subordinate evaluation* across the two groups indicated by the variable *specific knowledge type*. In a univariate t-test, I find that *specific knowledge type* has a significant negative effect on *subordinate evaluation* ($t = 7.66$, $p < 0.001$ one-tailed), meaning that *subordinate evaluation* for participants given disconfirming specific knowledge (mean = 0.82) is significantly lower than it is for participants given confirming specific knowledge (mean = 1.86).

This result is consistent with H2, suggesting that middle managers are biased against subordinates who provide them with disconfirming specific knowledge. Over time, this bias could lead subordinates to shift their efforts away from providing middle managers with disconfirming specific knowledge and toward providing middle managers with confirming specific knowledge (see Kassing 2009; 2002).[14] This would exacerbate upper management's

[14] The long-term implications of middle managers' bias hearken to Machiavelli's (1532/1988) famous quotation from *The Prince*. "And what physicians say about consumptive diseases is also true […], that at the beginning of the illness, it is easy to treat but difficult to diagnose, but […], as time passes it becomes easy to diagnose but difficult to treat" (11). That is, disconfirming specific knowledge is often an early signal of problems with the firm's strategy.

difficulty in acquiring crucial disconfirming specific knowledge to help revise the firm's strategy because middle managers, who are already more reluctant to pass along disconfirming specific knowledge to upper management, would have less disconfirming specific knowledge to consider in the first place.

Figure 4: Study 1 SEM Model for H2



Goodness of fit statistics: n = 160; χ^2 (1) = 0.447 (p = 0.504); CFI = 1.000; RMSEA < 0.001

In Section 3.3.2, I argue that the effects predicted in H2 are caused by disconfirmation bias and impression management, like H1. I test this reasoning using an SEM model, shown in Figure 4 (along with goodness of fit statistics). All but the final regression in this model are identical to the results detailed above in my tests of H1. The final regression in the SEM model shows a positive significant effect of the mediators *CEO impression* (p < 0.001 one-tailed) and *subjective usefulness* (p < 0.001 one-tailed) on *subordinate evaluation*. I find that *specific knowledge type* has a non-significant effect (p = 0.238 two-tailed) on *subordinate evaluation*.

By the time these problems manifest themselves in the form of poor lagging financial indicators it may be considerably more difficult to fix those problems. This is in line with Park et al. (2011) who suggest that biased reporting often leads to failed strategies and that shareholders often fire or otherwise penalize CEOs over firms with failed strategies.

Specific knowledge type has a significant effect on *subordinate evaluation* in univariate tests but has a non-significant effect on *subordinate evaluation* in the final model, suggesting complete mediation (Baron and Kenny 1986). I conduct a Sobel test on the *CEO impression* mediation path (which mediates *specific knowledge type* and *reporting likelihood*), again omitting the subjective usefulness pathway, and find a significant indirect effect ($p < 0.001$ one-tailed) as well as a significant direct effect ($p = 0.016$ one-tailed). This result contradicts my less formal suggestion of complete mediation based on *specific knowledge type*'s non-significance in the final model. Instead, this Sobel test suggests that even though my reasoning for H2 is generally supported, there are likely to be additional reasons that middle managers give biased subordinate evaluations apart from what is captured by my variables for impression management and disconfirmation bias.

3.5.4 Tests of H3

In H3, I predict that managers who receive a strategy map will be more likely to report disconfirming specific knowledge to upper management than middle managers who receive a non-causal list. As a test of this hypothesis, I use two-sample t-tests to directly examine the effect of receiving a strategy map on the dependent variable of *reporting likelihood*, restricting this test to those who receive disconfirming specific knowledge. Inconsistent with H3, I do not find a significant effect of *strategy map* on *reporting likelihood* ($p = 0.169$ one-tailed).

However, this may be because the norms associated with a strategy map are not active among inexperienced managers. Managers need to be exposed to norms for a sufficient period of time in order for those norms to enter into their thought process when making decisions (see Korte 2009 for a useful summary of prior research on how social norms are learned over time

through socialization). When I repeat this t-test, further restricting the sample to include only participants who receive disconfirming specific knowledge and who have management experience at or above the median number of years (management experience ≥ 2 years), *strategy map* has a significant, positive effect on *reporting likelihood* (n = 91, p = 0.010 one-tailed). This result supports H3 for this subsample. These results suggest that receiving a strategy map can improve middle managers' judgments about passing along disconfirming specific knowledge, with the caveat that this effect appears only when middle managers have sufficient experience to internalize the associated norms.

In Section 3.3.3, I predict that the strategy map effect would be due to an improvement in how middle managers think disconfirming specific knowledge would be received. This suggests a moderating relationship between receiving a strategy map and the H1 coefficient between *specific knowledge type* and *CEO impression*. To confirm this reasoning, I conduct multi-group path analysis using the SEM model from my tests of H1. I restrict this test to those who have at least two years of management experience, because these participants' responses support H3.[15] I also relax my restriction on *specific knowledge type* to include both conditions. The results of this analysis, along with goodness of fit statistics, are shown in Figure 5. In this multi-group path analysis, I estimate two separate SEM models based on two subsamples: the subsample of participants who receive a strategy map and the subsample of participants who receive a non-causal list. I hold constant between the two SEM estimation models all intercepts and all coefficients other than the coefficient of *specific knowledge type* predicting *impression management*.

[15] When I conduct the multi-group path analysis on the full sample of participants (regardless of management experience level), the chi-square test comparing the H1 coefficient between the two groups is non-significant (p = 0.241). This result is in line with my prior non-significant t-test of H3 using the full sample.

Figure 5: Study 1 SEM Model for H3

Goodness of fit statistics: n = 186; χ^2 (6) = 238.587 (p < 0.001); CFI = 1.000; RMSEA < 0.001

Figure 5 shows that the coefficient of *specific knowledge type* predicting *impression management* is less negative for those who receive a strategy map (B = -1.02, p < 0.001 one-tailed) than for those who do not receive a strategy map (B = -1.83, p < 0.001 one-tailed). I conduct a chi-squared comparison on these two coefficients and find that they differ at a significant level (χ^2 = 8.84, p = 0.003). This result supports my reasoning that receiving a strategy map changes middle managers' impression management calculus about disconfirming specific knowledge.

The effect observed in H3 demonstrates one way in which implementing a strategy map can be beneficial. My H3 results are also consistent with Kaplan and Norton's (2008; 2001a; 2001b) suggestion that receiving a strategy map improves overall communication of specific knowledge. Receiving a strategy map appears to increase truthful reporting among middle managers who are given subjective performance measures, and I have no evidence to suggest that this improvement would be accompanied by a decrease in the type of effort motivated by

those same subjective performance measures (on the contrary, Farrell et al. 2012 find that the communication of causal linkages *improves* effort allocation).

3.5.5 Tests of H4

I predict in H4 that middle managers who receive a strategy map, compared to middle managers who receive a non-causal list, are less biased in their performance evaluations of subordinates who provide disconfirming specific knowledge. As a test of this hypothesis, I use two-sample t-tests to directly examine the effect of receiving a strategy map on the dependent variable of *subordinate evaluation*, restricting this test to those who receive disconfirming specific knowledge. Inconsistent with H4, I do not find a significant effect of *strategy map* on *subordinate evaluation* ($p = 0.733$ one-tailed). Like in H3, I repeat this test among participants who have management experience at or above the median number of years (management experience ≥ 2 years). I find that, even among this subsample, *strategy map* has a non-significant effect on *subordinate evaluation* ($n = 91$, $p = 0.634$ one-tailed). These results do not support H4, suggesting that receiving a strategy map does not alleviate middle managers' bias against subordinates who provide disconfirming specific knowledge.

It is unclear why H4 is unsupported while H3 is supported even though both hypotheses are based upon the same theoretical grounds. From a statistical point of view, I note that my Sobel test from H2 suggest that impression management only partially explains the effect of *specific knowledge type* on *subordinate evaluation*. Although receiving a strategy map might affect the path indirectly leading from *specific knowledge type* to *subordinate evaluation* through *impression management*, the effect of a strategy map on that path might be overwhelmed by the direct path from *specific knowledge type* to *subordinate evaluation*. From a practical point of

view, this means that middle managers evaluate subordinates' performance based on complex reasoning, some of which is unmeasured in my experiment and apparently unaffected by receiving a strategy map.

As a supplemental test of H4, I repeat my multi-group path analysis from H3 (not tabulated). This time, however, I hold constant between the two strategy map conditions all coefficients and intercepts except for the coefficient of *specific knowledge type* predicting *subordinate evaluation*. I find a significant difference ($\chi^2 = 7.03$, $p = 0.008$) between the coefficient from participants who receive a strategy map ($B = -0.492$, $p = 0.01$ two-tailed) and the coefficient from participants who did not receive a strategy map ($B = 0.09$, $p = 0.68$ two-tailed).

Interpreting this supplemental test must be done conservatively because my experiment lacks the measures necessary for narrowing down the various possible reasons for this offsetting effect. Importantly, I do *not* have sufficient evidence from this test to conclude that receiving a strategy map *worsens* middle managers' evaluation of subordinate performance. Instead, I am limited to suggesting that this result likely captures some alternative rationale underlying subordinate evaluation judgments—beyond impression management—which suggests poorer evaluations for subordinates who provide disconfirming specific knowledge. Furthermore, I can suggest that receiving a strategy map does not mitigate the effect of this unmeasured rationale.[16]

[16] That is, without a strategy map, this alternative rationale covaries more fully with my impression management measures and thus the coefficient for the *specific knowledge type* indicator variable (which captures any such alternative rationales) is non-significant. With a strategy map, however, alternative rationales should covary less fully with my impression management measures because receiving a strategy map changes how *specific knowledge type* affects impression management. Thus, among those who receive a strategy map, the indicator variable that captures alternative negative effects of *specific knowledge type* on *subordinate evaluation* is significant and negative, compensating for the weakened impression management pathway.

3.6 CONCLUSION

In this study, I have examined how middle managers react to different types of specific knowledge: disconfirming and confirming. I find evidence that both impression management and disconfirmation bias lead middle managers to report disconfirming specific knowledge to upper management less than confirming specific knowledge. I also find evidence that impression management and disconfirmation bias lead middle managers to give more positive performance evaluations to subordinates who provide confirming specific knowledge than to subordinates who provide disconfirming specific knowledge. These results extend prior knowledge sharing research by providing experimental evidence that middle managers given subjective performance measures are likely to be biased against sharing knowledge that might indicate they are investing a low level of effort and are likely to encourage subordinates to be similarly biased. Based on these results, the specific knowledge that upper management receives may be far more confirming of the firm's strategy than the firm's ground-level reality actually indicates. Although upper management might be able to adjust for this, information economics theory would suggest this departure from fully truthful reporting is a real loss of overall value and leads to a second-best knowledge sharing solution.

However, I find that receiving a strategy map alters middle managers' incentives, making it more likely that middle managers will pass along disconfirming specific knowledge. This extends the subjective performance measurement literature by highlighting how management framing of such measures can positively affect agents' reaction to them. This study suggests that upper management can overcome biased reporting due to subjective performance measurement by convincingly communicating that disconfirming specific knowledge is welcomed rather than

frowned upon. This communication might be effective whether it takes place through a strategy map or otherwise.[17]

It is also important to note that I only find a strategy map effect among middle managers with at least two years of management experience.[18] But it is not unreasonable to expect middle managers generally to have at least two years of management experience before being promoted to such a position. One limitation of my study is that it assumes firms use subjective performance measures to determine raises and promotions. I do not test a control condition in which middle managers are evaluated using less prominent subjective performance measures or none at all. Instead, I rely on prior research to suggest how subjective and objective performance measures would differ, and I expect subjective performance measures to be a driving cause of the impression management mediation in my study (see Prendergast 1993). While this leads to a more parsimonious experimental design, omitting explicit objective performance measures in this study suggests a future area for research: examining the treatment of specific knowledge among middle managers in settings where subjective performance measures play a relatively less prominent role. More emphasis on objective performance measures might alter the impression management calculus in a way that could eliminate or amplify the strategy map effect I observe in this study.

Another possible limitation of my study is that I do not directly examine the effect a strategy map has on upper management's willingness to revise the firm's strategy once senior

[17] The difficulty being that such communications could be disregarded as cheap talk by agents if upper management does not incur a meaningful cost to make them, as with a strategy map.

[18] One could also plausibly interpret the fact that I only found this effect among those with high experience as evidence that those with high experience likely also had more security in their current jobs, making them more willing to pass along disconfirming specific knowledge because they had less fear of supervisor retribution. I do not expect this alternative interpretation to significantly change the predictions from this study.

managers receive disconfirming specific knowledge. Ittner and Larcker (2003) suggest that upper management often does not empirically test the causal relationships of a strategy map because senior managers only select strategic goals they consider to have "self-evident" (91) causal relationships. Similarly, Huelsbeck et al. (2011) find that when senior managers are convinced of a strategy map it is difficult to change their mind with a disconfirming signal (see also Campbell et al. 2015). In this study, I interpret these findings from prior research to be emblematic of firm strategies generally, suggesting that upper management generally has significant inertia in favor of pursuing the firm's chosen strategy, whatever the strategy's format may be. As reviewed in Section 3.1, some prior research suggests that using a strategy map significantly helps managers recognize when a strategy revision is necessary (Cheng and Humphreys 2012; Tayler 2010). However, I cannot completely rule out the possibility that Ittner and Larcker, as well as Huelsbeck et al., have documented upper management convictions that are unique to strategy maps and that are not captured by the prior literature I cite in Section 3.1. Thus, it is possible that senior managers are more reluctant to revise the firm's strategy when they use a strategy map. If this is the case, it represents a cost of using a strategy map that my study does not consider, and which may offset the knowledge sharing benefit for middle managers that I observe. I leave to future research the task of disentangling a strategy map's effect on upper management's willingness to revise the firm's strategy in the face of disconfirming specific knowledge.

4.0 STUDY 2: DOES THE EFFECT OF A STRATEGY MAP ON APPROPRIATENESS JUDGMENTS DEPEND ON STRONG CAUSAL LINKAGES?

4.1 INTRODUCTION

A strategy map is a popular format for presenting the firm's strategy in which upper management visually and explicitly links the firm's strategic goals in a chain of cause-and-effect relationships, typically by drawing arrows between these goals. These arrows are called *causal linkages* because they communicate that accomplishing one strategic goal is at least partly responsible for achieving the next strategic goal in the chain. Framing the firm's strategic goals as a strategy map can lead to several benefits, including improvements in managers' judgments and decisions (Cheng and Humphreys 2012; Tayler 2010; Kaplan and Norton 2008; 2004; 2001b; Webb 2004; see also Gimbert et al. 2010; Beer et al. 2005). Researching these benefits helps managers decide whether to implement a strategy map and how to implement it.

Prior research has also identified informative limitations to the benefits of a strategy map. For example, Webb (2004) finds evidence suggesting that a strategy map can improve goal commitment among managers, but that this effect relies on the strength of the cause-and-effect linkages between strategic goals. Specifically, he hypothesizes and finds that a causally-linked strategy that has "strong cause–effect links will result in greater commitment to financial performance goals" (931). A strong causal linkage, per Webb, is one that is "plausible and informative" (929). He manipulates the strength of the causal linkages, in part, by varying how easy it is for managers to understand exactly how the cause leads to the effect. He describes this aspect of the weak causal linkage condition as follows.

> The cause–effect linkages among the nonfinancial objectives and measures are also designed to be weak. The percentage of […] employees that report to work on time is presented as being causally related to customer ratings of staff friendliness, but this linkage is questionable (939).

In the current study, I examine another judgment effect of a strategy map and test whether this effect depends on strong causal linkages like those used by Webb (2004). Specifically, Cheng and Humphreys (2012) find that a strategy map affects vital manager judgments related to strategy revision, and I test whether these effects remain when a strategy map has weak causal linkages. Cheng and Humphreys (2012) find that providing managers with a strategy map improves, first, judgments about whether new information is relevant to evaluating the firm's strategy (hereafter referred to as *relevance judgments*) and, second, judgments about whether that new information suggests the strategy is appropriate or inappropriate (hereafter referred to as *appropriateness judgments*).[19]

Both judgments are essential for optimal strategy revision. Although I test both judgments, I primarily focus my analysis on appropriateness judgments because these judgments require managers to have a deeper understanding of the firm's strategy and are thus more likely to depend on causal linkage strength. Also, appropriateness judgments are arguably more important than relevance judgments because one's assessment of the appropriateness of the strategy directly suggests a course of action related to strategy revision, which can have a damaging effect on the firm if this course of action is not the correct action. Relevance judgments can suggest that information is relevant to revising the firm's strategy, but do not suggest what this information means to revising the firm's strategy.

[19] Cheng and Humphreys require participants to make these relevance and appropriateness judgments based on what they recall from memory about the firm's strategy, as might be expected in practice.

In my experiment, participants make relevance and appropriateness judgments like those described by Cheng and Humphreys (2012). I first find results comparable to theirs in that I find that a strategy map improves participants' relevance judgments. However, my study's primary contribution to this line of research is providing evidence that improved appropriateness judgments due to a strategy map's causal linkages depends on strong causal linkages. In supplemental analysis, I find evidence consistent with this interaction being caused by two separate mechanisms. In one subsample of participants, I find that strong causal linkages have a positive effect on appropriateness judgments. In another subsample of participants, I find that weak causal linkages have a negative effect on appropriateness judgments.

The results of this study provide additional guidance to firms considering a strategy map by demonstrating whether a firm should expect the same beneficial judgment effects from weak causal linkages as from strong causal linkages. It is especially important to know what features of a strategy map—such as the strength of its causal linkages—are necessary conditions for the benefits associated with a strategy map because strategy maps differ widely in terms of the strength of their causal linkages. Luft (2004) makes this point as well (utilizing, like Webb 2004, the term *strategic performance measurement systems* to describe performance measurement systems that use causally-linked strategic goals, such as a strategy map).

> A casual perusal of cases and practitioner articles describing strategic performance measurement systems (SPMS) yields examples of proposed causal links that range from the highly probable (for example, future revenues will be higher if more customers intend to purchase again) to the dubious (for example, customer satisfaction will increase if employees make more effort to cross-sell firm products) (Luft 2004, 959).

Luft (2004) goes on to identify three reasons that weak causal linkages might exist. First, the costs for strengthening the linkages may be too high, especially when processes are new or otherwise "not well understood by anyone in the organization" (963). Second, those who create the firm's strategy may have conflicting understandings of the cause-and-effect relationships that

drive firm performance. A strategy map would therefore be "the result of political compromise between these conflicting understandings rather than evidence-based determination of the actual drivers of profit" (963). Third, individuals who contribute to the firm's strategy might "lobby for using nonfinancial measures that make them look best rather than the nonfinancial measures that contribute most to firm value" (963).

Overcoming any of these can be costly, but my findings suggest only strong causal linkages are beneficial to managers' appropriateness judgments and weak causal linkages may even be detrimental to them. Prior research suggests that firms often do not quantify a strategy map's causal linkages (Ittner and Larcker 2003), and other prior studies cast doubt on the benefits associated with such quantification of causal linkages (Farrell et al. 2012; Huelsbeck et al. 2011). My findings suggest that apart from quantifying a strategy map's causal linkages, it is beneficial for firms to ensure the causal linkages of a strategy map are strong. By ensuring causal linkages are strong, managers can maximize the strategy revision-related benefits of a strategy map.

4.2 HYPOTHESIS DEVELOPMENT

4.2.1 Memory and Managers' Judgments

Firm strategies are management's best guess of the optimal path for increased shareholder wealth. Meanwhile, managers are constantly receiving new information (see Day and Schoemaker 2005). Each piece of new information might lead to an important revision to

firm management's best guess of the optimal firm strategy.[20] Thus, some of the most important judgments managers make are relevance judgments and appropriateness judgments because these judgments affect revising the firm's strategy. In practice, managers often must make relevance and appropriateness judgments without immediate access to the firm's strategy. Instead they must make these judgments based on their understanding of the strategy encoded in memory.

As such, Cheng and Humphreys (2012) expect a strategy map to lead to differences in relevance and appropriateness judgments due to differences in how the firm's strategic goals are encoded in memory when managers are given a strategy map. *Encoding* refers to how the brain converts current cognitions into constructs that are stored as memories, categorized, and connected to other memories, allowing that current cognition to be recalled from memory later (e.g. Nickerson and Adams 1979; Anderson 1974; Shiffrin and Atkinson 1969; see also Otmakhova et al. 2013; Bliss and Collingridge 1993). There is enough overlap between how well managers understand the firm's strategy and how well they encode the firm's strategy in memory that I treat these concepts interchangeably. Understanding the firm's strategy well is effectively the same as having encoded in memory the strategic goal details, interconnections, and categories of the firm's strategy.

To examine the effect of a strategy map on memory, it is important to distinguish between working memory and long-term memory. For example, Cheng and Humphreys (2012) give participants a distractor task to clear their working memory, meaning participants must rely on their long-term memory to make relevance and appropriateness judgments. *Working memory* is a temporary cache of recent cognitions that allows one to easily recall these most recent thoughts. *Long-term memory*, in contrast, is the collective body of cognitions encoded to

[20] Or, also important, this new information could improve management's confidence in the firm's current strategy.

memory through categorization and through connection with other memories, with the object being to allow these cognitions to be recalled in the future (see Nickerson and Adams 1979; Anderson 1974; Shiffrin and Atkinson 1969; see also Otmakhova et al. 2013; Bliss and Collingridge 1993). Encoding a cognition to memory effectively means encoding it in *long-term* memory.

Because Cheng and Humphreys (2012) use a distractor task, cognitions that participants must make to complete the distractor task occupy participants' working memory, forcing out previous cognitions about the firm's strategy. After the distractor task clears participants' working memory, participants can only complete relevance judgments and appropriateness judgments by recalling the cognitions about the firm's strategy they have encoded in long-term memory.[21] Participants make these judgments in response to new external information, i.e. summaries of six articles from a business journal. That is, they determine if these article summaries are relevant to the task of evaluating the firm's strategy and then determine whether the article summaries suggest the firm's strategy is appropriate or inappropriate for the current business environment. Participants in the strategy map condition make better relevance and appropriateness judgments than those who receive a random list of goals or those who receive

[21] By way of analogy, working memory is like an office desk. Items on the desk are easily retrieved because they are in essentially the same state as they were when placed on the desk and because they are easily identifiable and within line of sight. Working memory contains the cognitions that are most recent, and these cognitions are typically easy to recall in the same state as when they originated. Clearing the working memory with a distractor task is like clearing items off the desk to view a large map. Items on the desk are removed and must be filed away (like encoding them in long-term memory) or discarded (like forgetting them). In this analogy, long-term memory is like a filing cabinet. Once one has finished with the map, i.e. once the distractor task is complete, one removes the map and has a clear desk, i.e. a cleared working memory. If one wants to use an item that had been on the desk before viewing the map, the item must not be discarded (i.e. it must not be forgotten) and one must be able to find where it is in the filing cabinet (i.e. one must be able to recall it from long-term memory). Thus, by clearing working memory, Cheng and Humphreys place an emphasis on managers' ability to retrieve strategy-related cognitions previously stored in long-term memory.

goals organized by balanced scorecard category (judgments from the latter two conditions are statistically equivalent).

Thus, Cheng and Humphreys (2012) find that the communication of causal linkages improves both types of managers' judgments because causal linkages help participants recall the firm's strategy. In my first hypothesis, I predict this same effect for relevance and appropriateness judgments. Namely, I express my expectation that a strategy map's causal linkages improve managers' relevance and appropriateness judgments.

> **H1:** Managers who receive a strategy map make better relevance judgments and better appropriateness judgments than managers who receive a firm's strategy without causal linkages.

I generally expect the effect of causal linkages, as hypothesized in H1, to be stronger among those with less experience. This is because greater background knowledge can improve judgments such as relevance and appropriateness judgments, potentially substituting for the judgment-improving effects of a strategy map's causal linkages. The more experience a manager has, the more likely it is that he or she has background knowledge that already equips him or her with stable framework of understanding for firm strategies and strategic goals (these frameworks could also be called *schema*; see Arbib 2002 for a review). These mental frameworks should help highly experienced managers recall the firm's strategy for later judgments, leaving less room for detectable differences due to causal linkages.

4.2.2 Causal Linkage Strength and Causal Relatedness

Although causal linkages communicate to managers that there should be a cause-and-effect relationship between strategic goals, this does not automatically mean the manager can make sense of that relationship. The strength of the causal linkage also plays a role, as shown by

Webb (2004). One important factor determining whether causal linkages are strong or weak is how easy it is to understand how the cause leads to the effect. In a study that can also be applied to managers, Pennington and Hastie (1992) report that jurors' decisions are driven by whether the evidence coherently fits into a causal story. That is, jurors make their decisions based on whether they can understand why the perpetrator's actions (the cause) lead to the harm or the infraction in question (the effect). If the relationship between the two is unclear—such as when jurors cannot decide between multiple pathways from the cause to the effect—then jurors are less likely to convict. Pennington and Hastie refer to this as the *Story Model*, and I expect it to describe managers' decisions in the face of causal linkages as well.

I use the term *causal relatedness* to describe how easy it is to understand why the cause leads to the effect, and therefore the term also describes causal linkage strength.[22] The key feature of causal linkages with low causal relatedness is that they have multiple uncertain pathways through which the cause could lead to the effect. Because of these multiple uncertain pathways, it is difficult to fully understand why the cause leads to the effect. This is a characteristic that is shared by the weak causal linkage from Webb (2004), the two weak causally-linked measures reported by Luft (2004), and the weak causal linkages that I highlight below in this section (see Niven 2006). Pennington and Hastie's (1992) Story Model is congruent with a larger literature on causal relatedness, which suggests that the easier it is to narrow down the reason causally-linked information is causally related, i.e. the greater its causal relatedness, then the more easily the information is read and recalled, affecting individuals'

[22] In this study, I use the term *causal relatedness* interchangeably with causal linkage strength. One could also describe causal relatedness as the ease with which one forms a causal story from a causal linkage, with *forming a causal story* being tantamount to understanding how the cause leads to the effect in a causal linkage. In this case, the terminology used has no effect on the hypothesized effects.

judgments (Mason and Just 2004; Singer et al. 1992; Duffy et al. 1990; Keenan et al. 1984; Myers et al. 1987; see also Zwaan and Radvansky 1998; Stock et al. 1991).[23]

Luft (2004) confirms that weak causal linkages are prevalent in practice. Niven (2006) provides several more examples of strategy maps from practice. These strategy maps' causal linkages differ in causal relatedness. In one example, he reprints a strategy map used by a government agency that makes a causal linkage between the goal *optimize use of interns* and the goal *ensure regulatory compliance*. There are several realistic cause-and-effect relationships that might exist, making it difficult to decide on one such relationship, limiting one's understanding of this causal linkage and ultimately meaning low causal relatedness. First, interns may be recent graduates whose knowledge of regulatory details may be fresh. Second, there might be regulations that affect the usage of interns and optimizing their use allows for compliance with these regulations. Third, optimizing intern usage might free up experienced workers to spend more time on regulatory compliance. Because it is difficult to interpret this causal linkage, this causal linkage is like the "questionable" causal linkage in Webb's (2004, 939) experiment, and would likely be included in what Luft (2004) refers to as "dubious" (959) causal linkages.

In another example, Niven (2006) reprints Brother International's strategy map, which has an outer circular ring of arrows and four goal categories inside this ring of arrows. None of these strategic goals in these goal categories is explicitly connected to any other strategic goal.

---

[23] For example, consider the following causally-linked pair of sentences (used by Mason and Just 2004). *Joey's crazy mother became furiously angry with him. The next day his body was covered with bruises.* These two sentences present a story that is coherent and has only one plausible relationship between Joey's mother getting angry with him (the cause) and Joey being bruised the next day (the effect). Thus, a causal linkage between these two would be strong because it has high causal relatedness and should be easy to encode in long-term memory (Myers et al. 1987). In contrast, another causally-linked sentence pair demonstrates low causal relatedness (also used by Mason and Just 2004). *Joey went to a neighbor's house to play. The next day his body was covered with bruises.* There are multiple realistic reasons that going to a neighbor's house to play (the cause) leads to Joey being bruised the next day (the effect). Therefore, it is a weaker causal linkage and would be harder to encode in long-term memory.

However, the ring of arrows clearly communicates causal linkages between the categories, with one goal category somehow leading to the next goal category. This is another way in which causal linkages can be weak: because they join multiple causes to multiple effects, leaving several possible causal relationships among several possible combinations of goals. In their examples explaining strategy maps, Kaplan and Norton (2008; 2004; 2001b) also often use unspecific causal linkages connecting groups of goals.

4.2.3 Causal Relatedness and Appropriateness Judgments

The first judgments managers make about new information are relevance judgments, which require only a surface-level understanding of the firm's strategy. Relevance judgments are akin to remembering the title of a chapter or the bottom line of an income statement. Cheng and Humphreys (2012) argue that relevance judgments are a "categorization process whereby managers try to fit the external information item to their cognitive representation" of the firm's strategy (903-904). The critical test of a relevance judgment is whether the new information "belongs" to the same category as the firm's strategic goals "and therefore is relevant to" the firm's strategy (904; see Markman and Wisniewski 1997 for a review of categorization; see also Corneille et al. 2006). Relevance judgments are an economical way to quickly flag relevant information, perhaps for a subsequent appropriateness judgment, without engaging in the time-consuming task of analyzing the meaning of irrelevant new information. As such, I do not expect causal linkage strength to affect relevance judgments significantly because I expect the

incremental benefit of high causal relatedness to be minimal for this relatively simple recall task.[24]

For appropriateness judgements, however, a more profound level of detail about the firm's strategy must be recalled. Determining whether new information suggests the firm's strategy is appropriate or inappropriate requires managers to recall sufficient details about the firm's strategic goals to understand the predictions and assumptions associated with those strategic goals. Then managers must compare the new information with those predictions and assumptions. Failure to properly recall enough of the details about the firm's strategic goals to understand the predictions and assumptions of these goals can lead to an incorrect understanding of the appropriateness of the firm's strategy, making it more difficult to correctly revise the firm's strategy. Prior research suggests that greater causal relatedness improves one's recall of details about causally-linked material (e.g. Mason and Just 2004; Keenan et al. 1984). This leads to the prediction that causal relatedness moderates a greater causal linkage benefit for appropriateness judgments. That is, the predicted H1 benefit from causal linkages is greater for appropriateness judgments when causal linkages have greater causal relatedness.

H2: The positive effect that causal linkages have on appropriateness judgments is greater when strategic goals have high causal relatedness than when the strategic goals have low causal relatedness.

[24] An example of a relevance judgment from an academic setting is when a new edition of an accounting journal is released. A researcher might browse the articles in this new edition to determine which, if any, are relevant to his or her research projects. To judge the relevance of these new articles, the researcher will likely first read the title and abstract to determine the topical and/or methodological category of these new articles and then compare these categories to the categories he or she recalls from his or her current research projects. This is more economical than reading each new article to determine if it is relevant, and it is not necessary to recall every detail of one's current research projects to compare their categories to the new article categories. Once a researcher determines an article is relevant, the researcher likely recalls the details of his or her corresponding research project(s) to compare them with details he or she reads from the body of the article text to determine the implications of this new article on his or her project(s)—like an appropriateness judgment.

As with H1, I generally expect this effect more among those with low experience. Low experience individuals are less likely than high experience individuals to have background knowledge that will otherwise help these individuals recall details of the firm's strategy. The background knowledge associated with experience could improve one's recall of firm details independent of the strategic goals' causal relatedness and effectively make low causal relatedness causal linkages stronger. Like my reasoning with H1, I expect those with low experience to have less background knowledge and thus greater differences between the effect of strong causal linkages and weak causal linkages.

4.3 METHOD

4.3.1 Experimental Design

In my 2 × 2 experiment, I manipulate one between-subject factor (the strategy map manipulation) and one within-subject factor (the causal relatedness manipulation). For my between-subject strategy map manipulation, I randomly assign each participant to one of two levels: strategy map and non-causal list. I make these manipulations by altering how strategic goals are presented to participants. For both conditions, I present participants with two strategic goals at a time. I refer to each goal-pair as a *mini-strategy*, and I present a total of sixteen mini-strategies to each participant (I explain this further below). Those in the strategy map condition receive, for each mini-strategy I present to them, an arrow that indicates a causal linkage between the two strategic goals of the mini-strategy. Those in the non-causal list condition receive the same mini-strategies, but I do not present them with any arrows to indicate causal

linkages between the strategic goals of these mini-strategies.[25] I make this manipulation between-subject rather than within-subject to maintain experimental control. Specifically, if a participant sees an arrow to indicate a causal linkage in one mini-strategy but sees *no* arrow to indicate a causal linkage in the next mini-strategy, that participant could be primed by the first mini-strategy to mentally infer a causal linkage between the two goals of the second mini-strategy. Such an outcome could confound my strategy map manipulation.

Before describing my within-subject manipulation, I explain how I create the mini-strategies that I present to participants. Again, mini-strategies are pairs of strategic goals, and I present participants with one mini-strategy at a time. I present each participant with sixteen mini-strategies, for a total of thirty-two strategic goals. I create these mini-strategies from eight four-goal *strategic themes*. For my purposes, a *strategic theme* is a chain of four goals—one from each balanced scorecard category—with a unifying business topic. No mini-strategy contains strategic goals from different strategic themes, and therefore each strategic theme evenly yields two mini-strategies. Each participant sees the same thirty-two goals from the same eight strategic themes.

I make my within-subject causal relatedness manipulation by simply altering how goal-pairs from the four-goal strategic themes are grouped into mini-strategies. I assign four of the strategic themes from which I create mini-strategies to each of my two causal relatedness levels: high causal relatedness and low causal relatedness. Strategic goals in a strategy map are generally expected to flow through the four balanced scorecard categories in the following order: *learning and growth* goals to *internal processes* goals, *internal processes* goals to *customer*

[25] Also, as detailed in Section 4.3.3, I explicitly describe the strategy map as being a "strategy map" that "arranges the firm's goals into cause-and-effect relationships." Adding this description to the strategy map condition is in line with the procedures followed by Cheng and Humphreys (2012).

goals, and *customer* goals to *financial* goals (Kaplan and Norton 2004; 2001a; 2001b). For themes in the high causal relatedness condition, I make mini-strategies by drawing strategic goal-pairs from adjacent balanced scorecard categories in that generally expected order. That is, if a strategic theme is in the high causal relatedness condition, I create one mini-strategy from the theme's learning and internal processes goals and another mini-strategy from the theme's customer and financial goals.[26] Because these goals are from adjacent categories, they are more easily bridged, making it easier to understand the cause-and-effect relationships between them.

I create mini-strategies in the low causal relatedness condition by pairing strategic goals from non-adjacent balanced scorecard categories. Specifically, I skip a category in the chain to find goal-pairs to make into mini-strategies. If a strategic theme is in the low causal relatedness condition, I make one mini-strategy from the theme's learning and customer goals and another from the theme's internal process and financial goal.[27] The fact that mini-strategies in the low causal relatedness condition use strategic goals from non-adjacent categories makes it likely that participants can plausibly infer multiple ways that one goal causes the next. Thus, in this condition, it is relatively more difficult to narrow down possible causal relationships to just one.

---

[26] By random assignment, I counterbalance which four themes are assigned to the two causal relatedness conditions. That is, for each theme, half of the participants learn of its four strategic goals through two high causal relatedness mini-strategies and the other half of participants learn of its four strategic goals through two low causal relatedness mini-strategies. Counterbalancing helps control for idiosyncratic differences between how the themes are written. When I add an indicator variable to account for this random assignment in Section 4.4.3, it does not change either of my chief hypothesis tests for H1 and H2. This indicator variable is not significant (coefficient = 0.04, p = 0.49 two-tailed) when added to my test of relevance judgments (see regression in Table 3 Panel A). However, the indicator variable *is* significant (coefficient = 0.13, p = 0.03 two-tailed) when added to my test of appropriateness judgments (see regression in Table 4 Panel A). This significant result does not change the interpretation of my results.

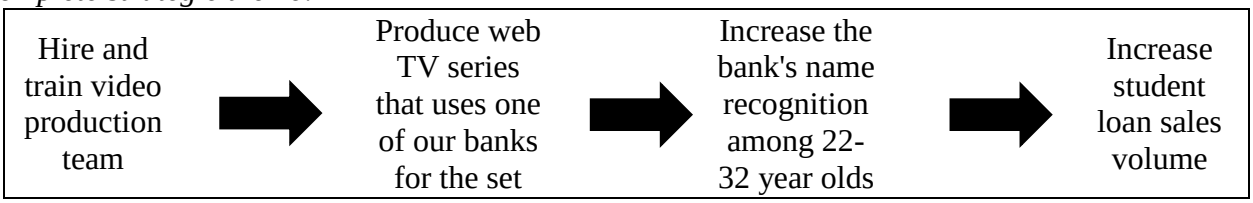
[27] Some prior research suggests that indirect causal relationships can strengthen over time. That is, if a person learns that A causes B and that B causes C, then over time—especially if one has a chance to sleep—the indirect relationship between the A and C items grows stronger in long-term memory (Coutanche et al. 2013; Ellenbogen et al. 2007; see also Schlichting and Preston 2016; 2015). Although I use weak mini-strategies that technically conform to a similar A-to-C pattern of indirect causal relationships, I do not expect my results to be affected by this memory phenomenon observed by prior research. First, the timeframe of my study is too short. Second, I randomize the order in which mini-strategies are presented, meaning the B item is likely difficult for participants to connect to the A and C items.

To avoid order effects, I randomize the order in which mini-strategies are presented in both causal relatedness conditions.

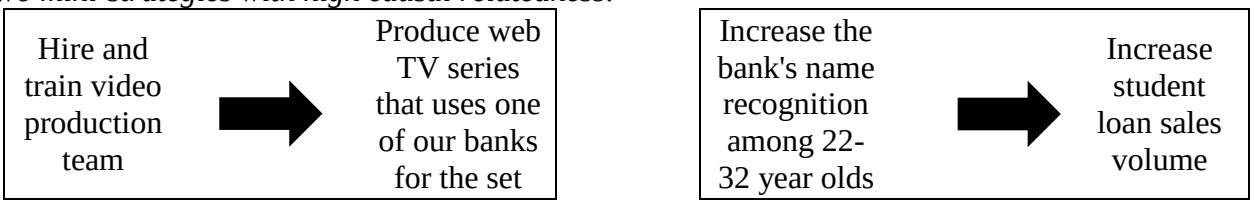
I provide an example of both causal relatedness conditions in Figure 6. This figure shows a complete strategic theme and how I group pairs of goals from that strategic theme into high and low causal relatedness mini-strategies. The mini-strategies in this figure are in the strategy map condition, i.e. with arrows to indicate causal linkages. If these mini-strategies were in the non-causal list condition, the causal relatedness manipulation would operate in the same way, but without any arrows between strategic goals.

Figure 6: Study 2 Causal Relatedness Manipulation

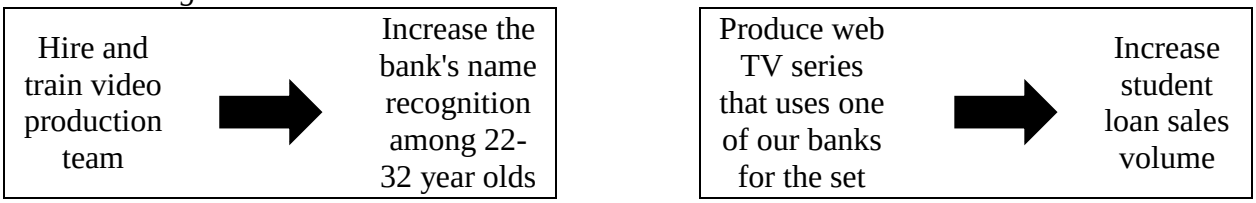
Complete strategic theme:



Two mini-strategies with high causal relatedness:



Two mini-strategies with low causal relatedness:



One critical advantage of manipulating this factor within-subject rather than between-subject is that it adds realism. Actual firm strategies often include a multitude of strategic goals and some of these goals will have high causal relatedness and some of these goals will have low

causal relatedness. Testing how participants respond to both causal relatedness conditions is more useful for informing practitioners because it better simulates what managers face in practice.

4.3.2 Participants

Participants are Amazon Mechanical Turk workers. I screen participants using requirements similar to those tested and validated by prior studies (e.g. Farrell et al. 2017; also see Mason and Suri 2012 for a review). Namely, I screen for participants who have the "master" designation, have a favorable task completion rating of at least 95%, and have successfully completed 1,000 or more tasks. Using these screening methods, I gather 375 complete participant responses. I further screen these responses for those from an IP address that was unacceptable (I exclude three IP addresses from outside of the United States, one IP address from the city where I was based at the time, and twenty-two IP addresses that are duplicated among multiple responses).[28] To ensure participant responses are meaningful, I give participants three separate attention check questions, in which I instruct them which answer choice to select.[29] I exclude results from the 19 participants who failed any of these attention checks. The final sample consists of 330 complete participant responses. Participants earn a flat fee for completing the

[28] I exclude responses from IP addresses outside of the United States because I cannot independently confirm Amazon's verification procedures. I am aware of reports that some workers from outside the United States have circumvented those procedures in the past. I exclude IP addresses from the city where I was based at the time as an extra experimental control to exclude responses that may be from individuals at my institution or otherwise connected to me who might have known the study was being conducted and knew of its experimental design.
[29] These attention checks are straightforward about which answer is the correct answer. If a participant misses one of these attention checks, I must conclude that this participant is not paying sufficient attention to the questions or instructions for their responses to be meaningful.

study ($3) and have an opportunity for a $1 bonus if they pay close attention throughout the study (I describe this bonus opportunity below).

4.3.3 Experimental Task and Procedures

The experiment consists of three stages. Stage one is the encoding stage, in which I present sixteen mini-strategies and participants encode them in long-term memory. First, participants learn that they have recently been hired as a regional manager in a fictional bank in the United States (i.e. *the firm*), and that upper management has recently developed a new strategy for the firm. Participants in the strategy map condition learn that the firm's strategy "is in a format called a strategy map," which "arranges the firm's goals into cause-and-effect relationships." This is like the instructions provided by Cheng and Humphreys (2012) to participants in the strategy map condition. Explicitly stating that the intent of a strategy map is to communicate causal linkages helps ensure a clear manipulation between my strategy map and non-causal list conditions.

Participants in all conditions then learn of the firm's strategic goals, mini-strategy by mini-strategy. These mini-strategies are provided one at a time with no opportunity to review prior mini-strategies. No mini-strategy is repeated. I also randomize the order in which the sixteen mini-strategies are presented. For each mini-strategy, the first strategic goal is shown for ten seconds and then both strategic goals are shown for ten seconds. An arrow is shown between the two goals if the participant has been assigned to the strategy map condition. I advance participants from mini-strategy to mini-strategy automatically to ensure consistent timing across all mini-strategies. Otherwise, participants might develop an affinity toward the high causal

relatedness mini-strategies, affecting how long they spend reading those mini-strategies and thus altering the overall salience of those mini-strategies in later judgments.[30]

In stage two, participants first perform a distractor task to clear their working memory. In this distractor task, participants have one minute to count the number of times the letter "c" appears in a short passage on an unrelated topic. To maintain proper control over this distractor task, participants are not permitted to advance from this screen until one minute has elapsed. After one minute, participants are automatically advanced to the next screen where they input how many times they counted the letter "c." Then participants are instructed that regional managers are responsible for assisting in the ongoing evaluation of the appropriateness of the firm's strategy and that they will see several article summaries that may or may not be relevant to that evaluation. In a random order, participants read article summaries and for each article summary respond to (1) a question about the relevance of the article summary to evaluating the firm strategy and (2) a question about whether the article summary suggests the firm's strategy is appropriate or inappropriate. These questions test the primary dependent variables of interest to this study, i.e. relevance and appropriateness judgments, and are comparable to those questions used by Cheng and Humphreys (2012).

In the final stage, stage three, I ask participants several demographic and control questions. I explain these control questions in Section 4.3.5, in which I describe the control variables that I use for my analysis. As a tool to ensure participants pay close attention to the goals presented in stage one, even though participants are located remotely, I give participants a short quiz at the end of stage three. In this quiz, they are shown a set of goals and indicate whether they saw these goals in stage one. Half of the goals shown in the quiz are foil goals,

[30] Also, this pattern is like prior studies on causal relatedness (e.g. Mason and Just 2004).

designed to be plausibly mistaken for the actual goals seen. Participants receive a $1 bonus if they complete the quiz with a score of 75% or more correct. The quiz appears to be an effective attention-holding tool for stage one (the average score on the quiz is 77%) and for the other stages as well (only 19 participant responses fail attention check questions—two of which are in stage two and one of which is in stage three).

4.3.4 Dependent Variables

In stage two, I generally follow Cheng and Humphreys' (2012) procedures and present participants with twelve article summaries, which I describe as originating from recent articles in a reputable business press. Four article summaries are irrelevant to the firm's strategy, four article summaries are relevant and suggest a strategic goal is appropriate, and the remaining four article summaries are relevant and suggest a strategic goal is inappropriate.[31] In other words, the twelve article summaries can be grouped evenly into three subcategories: irrelevant, appropriate, and inappropriate. In each of these three subcategories, I use two article summaries related to a high causal relatedness mini-strategy and two article summaries related to a low causal relatedness mini-strategy. This last design feature ensures I can perform statistical tests comparing scores across the two causal relatedness conditions.

For each article summary, I first ask participants, on a five-point scale, the degree to which they believe the article summary is relevant to the evaluation of the firm's strategy, where a higher score indicates more relevance and a lower score indicates less relevance. A score of

[31] Cheng and Humphreys (2012) use six article summaries, and divide them so two of them are in each of the categories described in this section. I am presenting a strategy with more goals than they do, making it reasonable to present participants with more article summaries without dramatically changing the balance of goals and article summaries.

zero, the lowest possible value, indicates that the item is irrelevant. This is the relevance judgment. Then, for each article summary participants rate as relevant (i.e. rated more than zero), I ask participants the degree to which the article summary suggests the firm's strategy is appropriate or inappropriate. I use a six-point force-choice scale for the appropriateness judgment, where a higher number indicates more appropriate and a lower number indicates more inappropriate. When forming the dependent variables for *relevance judgment* and *appropriateness judgment*, I reverse-code participant responses to irrelevant article summaries (for relevance judgments) and to inappropriate article summaries (for appropriateness judgments).[32] For the article summaries that I present, the end points of the scale are the most appropriate answer.[33] Therefore, *relevance judgment* and *appropriateness judgment* stand for participants' degree of correctness in the respective judgments, with higher always meaning more correct.

4.3.5 Independent Variables and Controls

I create indicator variables to represent my manipulations as follows. The indicator variable *strategy map* is coded as one when the participant receives a strategy map, i.e. causal linkages between strategic goals, and it is coded as zero when the participant receives a non-

[32] For example, assume there's an irrelevant item (which should be zero, meaning irrelevant) and a participant responds to this item with a four, meaning highly relevant. I code this response as a zero for my analysis. This can be interpreted to mean the response is as distant from the correct answer as possible. If another participant responds to the same irrelevant question with a one, meaning slightly relevant, that participant's response is analyzed as a three. This can be interpreted to mean that that response is *almost* the correct answer.

[33] Furthermore, even if an article summary were written in a way that suggests the correct answer is not at the end point of the scale, I use fixed effects control variables (described below) to adjust the means of responses to individual article summaries. This control is designed specifically to adjust for such qualitative or idiomatic differences between article summaries.

causal list, i.e. no causal linkages between strategic goals. The indicator variable *causal relatedness* is coded as one for relevance and appropriateness judgments that pertain to the six article summaries for high causal relatedness mini-strategies and coded as zero for judgments that pertain to the six article summaries for low causal relatedness mini-strategies. *Strategy map* necessarily varies by participant, and *causal relatedness* varies from article summary to article summary. In sum, each participant has six of each dependent variable observations from article summaries coded as zero for *causal relatedness* and six from article summaries coded as one for *causal relatedness*. I also use a set of eleven indicator variables that I refer to collectively as *article summary fixed effects*. As fixed effects indicator variables, no more than one variable is ever coded as one at a time (although for one of the article summaries, all fixed effects indicator variables are coded as zero). The *article summary fixed effects* variables reduce noise that otherwise may arise due to subjective or qualitative differences in the way individual article summaries are written.[34]

In stage three, I ask participants manipulation check questions, demographic questions, and control questions. For my check on the strategy map manipulation, I ask participants to think about those times two goals were presented on the screen at the same time, and then ask them to indicate how much they agree with the notion that the firm expects those two goals to have a cause-and-effect relationship. Participants respond on a scale of 1 (strongly disagree) to 6 (strongly agree). This reflects the construct I am attempting to manipulate with the strategy map manipulation: the explicit expectation of a cause-and-effect relationship.

[34] When I remove this set of variables it does not affect the significance or direction of my main test of H1, but my main test of H2 does move outside of the standard range of significance (the coefficient term testing H2 is $p = 0.08$ one-tailed). I expect this change to be due to statistical noise from recording appropriateness judgments about twelve different article summaries with varying content.

For my check on the causal relatedness manipulation, I present two of the mini-strategies that participants see in stage one: a high causal relatedness mini-strategy and a low causal relatedness mini-strategy. These mini-strategies are presented with or without causal linkages in accordance with the participants' strategy map condition. I instruct participants to assume there is indeed a cause-and-effect relationship between the two goals shown on the screen. Then I ask participants how much they felt it was "clear how the first goal causes the second goal" for each of the two mini-strategies, on scale of 1 ("very unclear") to 6 ("very clear"). By comparing answers across mini-strategies from different causal relatedness conditions, I can check to see if participants can understand the causal relationship for a mini-strategy with high causal relatedness more clearly than a mini-strategy with low causal relatedness.

My stage three demographic questions include age, gender, educational attainment, student status, and business education (i.e. whether participants have completed a degree in business). I also ask participants how many years of overall work experience they have and how many years of management experience they have. For use as control variables, I also ask participants their familiarity with the strategy map and with banking in general. Familiarity with banking is an especially important control variable given that the firm in the scenario for this experiment is a bank and the firm's strategic goals relate to banking.

4.4 RESULTS

4.4.1 Descriptive Statistics and Manipulation Checks

In Table 2, I present descriptive statistics for several variables of interest. *Relevance judgment* (Panel A) and *appropriateness judgment* (Panel B) values are broken out by experimental condition. Panel C shows the mean and, where applicable, standard deviation and quartile measures for participant age, gender, educational background, work experience, management experience, and familiarity with the balanced scorecard and the strategy map.[35]

Table 2: Study 2 Descriptive Statistics

Panel A: Relevance Judgment; Means (standard deviations)

	Non-causal list	Strategy Map	*Row Averages*
Low Causal Relatedness	2.96 (1.48) n = 1,014	3.03 (1.45) n = 966	2.99 (1.47) n = 1,980
High Causal Relatedness	2.89 (1.51) n = 1,014	3.01 (1.52) n = 966	2.95 (1.51) n = 1,980
Column Averages	2.92 (1.49) n = 2,028	3.02 (1.49) n = 1,932	2.97 (1.49) n = 3,960

[35] None of these variables differs significantly across strategy map conditions ($p > 0.05$ one-tailed—in either direction), except for familiarity with strategy maps. As would be expected, familiarity with strategy maps is significantly higher ($p = 0.03$, one-tailed) when participants are in the strategy map condition (mean = 1.2, standard deviation = 1.18) than when they are in the non-causal list condition (mean = 0.97, standard deviation = 1.11).

Table 2: Study 2 Descriptive Statistics (continued)

Panel B: Appropriateness Judgment; Means (standard deviations)

	Non-causal list	Strategy Map	*Row Averages*
Low Causal Relatedness	4.17 (1.59) n = 801	4.07 (1.61) n = 746	4.12 (1.60) n = 1,547
High Causal Relatedness	4.22 (1.65) n = 799	4.28 (1.58) n = 727	4.25 (1.61) n = 1,526
Column Averages	4.20 (1.62) n = 1,600	4.17 (1.60) n = 1,473	4.18 (1.61) n = 3,073

Panel C: Descriptive Statistics

	N	Mean	SD	Lower Quartile	Median	Upper Quartile
Age	330	40.32	19.42	31	37	46
Gender (Female)	330	0.48 (156)	—	—	—	—
Number of Business classes	330	2.46	4.59	0	1	2
Business degree	330	0.10 (32)	—	—	—	—
Current Student	330	0.06 (19)	—	—	—	—
Work Experience	330	17.88	11.04	10	15	25
Management Experience	330	3.70	5.32	0	2	5
Educational Attainment:						
High School/GED	330	0.15 (50)	—	—	—	—
Some college	330	0.35 (114)	—	—	—	—
Bachelors	330	0.42 (138)	—	—	—	—
Graduate	330	0.08 (28)	—	—	—	—
Familiarity: (0 to 4 scale)						
Balanced Scorecard	330	0.57	0.97	0	0	1
Strategy Map	330	1.05	1.13	0	1	2
Banking	330	2.32	1.04	2	2	3

I performed two manipulation checks, described in Section 4.3.5, to ensure that my two experimental manipulations match the constructs they are intended to reflect. First, in the strategy map manipulation check, I ask participants how much they agree that the firm expects a

cause-and-effect relationship between mini-strategies presented in stage one. This manipulation, as expected, varies significantly ($p < 0.01$ one tailed) between participants who received a strategy map (mean = 5.50; standard deviation = 0.87) and participants who received a non-causal list (mean = 3.98; standard deviation = 1.61). The strategy map manipulation passes its manipulation check because there is more of an expectation of a cause-and-effect relationship among those in the strategy map condition than among those in the non-causal list condition. Participants provided with the causal linkages of a strategy map have a more overt rationale for attempting to understand the causal relationships between strategic goals.

I check the causal relatedness manipulation by presenting two mini-strategies previously presented to participants and by asking whether it is clear *how* the first goal causes the second, under the assumption that the two goals of the mini-strategy do indeed have a cause-and-effect relationship. One of these mini-strategies has high causal relatedness and the other has low causal relatedness. I ask this check question on a six-point scale from 1 to 6. Participants report that it is significantly clearer ($p < 0.01$ one-tailed) how the goals cause each other when the mini-strategy has high causal relatedness (mean = 5.46; standard deviation = 0.86) than when the mini-strategy has low causal relatedness (mean = 2.88; standard deviation = 1.55). This results supports my expectation that the causal relatedness manipulation affects how easily one can understand the cause-and-effect relationship of a causal linkage.

4.4.2 Tests of H1

As hypothesized in H1, I expect that those who receive a strategy map with causal linkages between the goals will make better relevance and appropriateness judgments than those who do not. To test this, I use a mixed multiple linear regression model that appropriately

accounts for repeated measures. That is, this form of regression accounts the fact that I collect twelve relevance and twelve appropriateness judgment observations from each participant rather than erroneously assuming each dependent variable observation is from an independent, randomly-assigned participant.

My first regression model uses *relevance judgment*—which measures the correctness of participants' relevance judgments—as the dependent variable. The predictors in this model are *strategy map* and *article summary fixed effects* (a set of indicator variables that identify the article summary a dependent variable observation comes from). As shown in Table 3 Panel A, I find support for H1, in the form of a significant positive coefficient for *strategy map* ($p = 0.046$ one-tailed) predicting *relevance judgments*. This result replicates Cheng and Humphreys' (2012) result that the causal linkages of a strategy map improve relevance judgments. My confirmation of this result is important because I use a different set of participants, a unique scenario, and I increase the number of strategic goals the firm has. Still, I find that a strategy map's causal linkages help managers judge how relevant new information is to strategy revision.[36] Practically speaking, managers who efficiently and accurately judge the relevance of new information are less likely to waste time on irrelevant information and are less likely to discard truly relevant information that the firm needs to revise its strategy.

My second regression to test H1 is the same as the first, except that it predicts *appropriateness judgment* as the dependent variable. These results are shown in Table 3 Panel B. Inconsistent with H1, I find that *strategy map* is not a significant predictor of *appropriateness judgment* ($p = 0.47$ one-tailed). Although this is technically a failed test of H1, it is still

[36] In line with my reasoning for H1, *causal relatedness* is not a significant predictor ($p = 0.35$ two-tailed) if I add it as a predictor to the models in Table 3 Panel A and Table 3 Panel B.

consistent with the overall appropriateness judgment reasoning from Section 4.2. Namely, appropriateness judgments require managers to recall substantial details about the firm's strategy. It is hardly inconceivable that a group of causal linkages bridging a mixture of low causal relatedness and high causal relatedness strategic goals would be insufficient to improve managers' recall of such details. This mixture of causal relatedness is most likely why my results for H1—as they pertain to appropriateness judgments—differ from Cheng and Humphreys' results.

Table 3: Study 2 Tests of H1

Panel A: Test of H1 for Relevance Judgments

$Relevance\ Judgment_{it} = \beta_0 + \beta_1\ Strategy\ Map_{it} + Article\ Summary\ Fixed\ Effects_i + u_i + \varepsilon_{it}$

	Coefficient	z-value	p-value (***one-tailed***)
Strategy Map	0.10	1.69	***0.046***
Article Summary Fixed Effects	—	(868.08 χ^2)	<0.01
Constant	4.35	54.63	<0.01

R^2 *(Relevance Judgment): 0.172; Wald* χ^2 *(12) = 870.93 (p < 0.01)*

Panel B: Test of H1 for Appropriateness Judgments

$Appropriateness\ Judgment_{it} = \beta_0 + \beta_1\ Strategy\ Map_{it} + Article\ Summary\ Fixed\ Effects_i + u_i + \varepsilon_{it}$

	Coefficient	z-value	p-value (***one-tailed***)
Strategy Map	<0.01	0.08	***0.47***
Article Summary Fixed Effects	—	(1239.52 χ^2)	<0.01
Constant	3.93	28.94	<0.01

R^2 *(Appropriateness Judgment): 0.277; Wald* χ^2 *(12) = 1239.69 (p < 0.01)*

Panel C: Supplemental Boundary Tests of H1

	***β₁ Strategy Map* (p-value, one-tailed)**	
	Below-median subsample	Above-median subsample
Work Experience	0.18 (0.02)	0.02 (0.37)
Management Experience	0.12 (0.08)	0.09 (0.24)
Age	0.13 (0.04)	0.05 (0.28)
Business degree	0.09 (0.07)	0.11 (0.27)
Familiarity with Banking	0.20 (0.04)	0.06 (0.16)

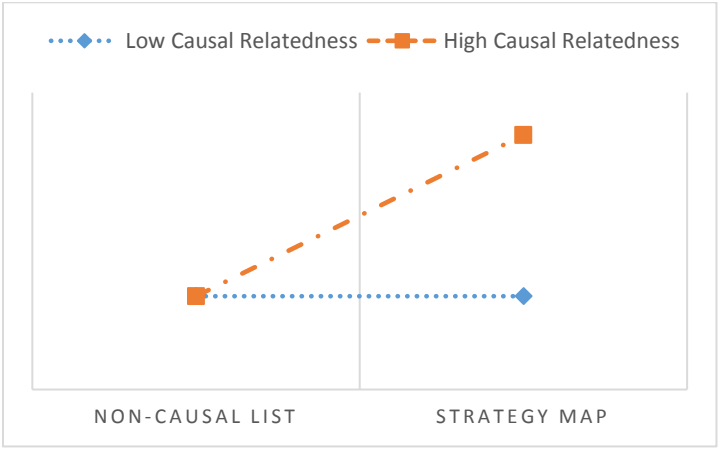
Furthermore, I expect that the effect observed in H1 should be weaker among certain subsamples because these subsamples should already possess more background knowledge and thus more alternative means of recalling surface-level information about strategic goals, which is necessary for relevance judgments, without the help of a strategy map. In a supplemental test of H1, I create subsamples based on above-median and below-median work experience (median = 15), management experience (median = 2), age (median = 37), as well as two subsamples based on whether the participants have a business degree (business degree n = 32). Then, I re-run my H1 test in each of the four subsamples. As summarized in Table 2 Panel B, for both median splits, the *strategy map* coefficient is significant and positive, at either the 95% or 90% confidence level, for those in the below-median subsamples (or those without a business degree) and non-significant at the 90% confidence level, for those in the above-median subsamples (or those with a business degree).[37] Familiarity with banking also follows the pattern shown above, with high familiarity leading causal linkages to have a weaker effect than low familiarity. This supplemental test suggests that strategy maps have more of an effect on relevance judgments for managers with less experience. As might be expected from the non-significant effect of my H1 test on appropriateness judgments, I test and do not find support (not tabulated) for an H1 effect on appropriateness judgment among the subsamples shown in Panel A of Table 3 ($p > 0.05$ one-tailed in all cases).

[37] Work experience ($p < 0.01$ two-tailed), management experience ($p = 0.03$ two-tailed), and possession of a business degree ($p = 0.04$ two-tailed) are significant positive predictors of *relevance judgment* when I add each of them, one at a time, as a predictor in the regression shown in Table 4 Panel A. Participant age is not a significant predictor when tested as a continuous variable ($p = 0.47$ two-tailed), but is significant at the 90% confidence level ($p = 0.07$ two-tailed) if I use an indicator variable coded as one for participant's at or above the median age. Familiarity with banking is not a significant predictor whether I test it as a continuous variable ($p = 0.56$ two-tailed) or a dichotomous variable ($p = 0.38$ two-tailed).

4.4.3 Tests of H2

I hypothesize in H2 that managers' appropriateness judgments are improved by causal linkages more when those causal linkages connect goals with high causal relatedness than when those goals connect goals with low causal relatedness. I predict this because causal linkages explicitly indicate to managers that they should try to interpret a cause-and-effect relationship between strategic goals, whereas high causal relatedness allows managers to succeed in this interpretation, which is what I expect to ultimately affect appropriateness judgments. This prediction describes an interaction between my two manipulated variables. Therefore, if this hypothesis is correct, the interaction term between *causal relatedness* and *strategy map* will have a significant and positive coefficient when predicting participants' appropriateness judgments. This prediction for mean *appropriateness judgment* is shown in Figure 7, broken out by experimental condition. Effectively, H2 predicts that the slope of the high causal relatedness line will be steeper than the slope of the low causal relatedness line.

Figure 7: Study 2 Predicted H2 Interaction Effect for Mean Appropriateness Judgment



Like my first hypothesis, I test H2 with a mixed multiple linear regression model. This regression model uses *appropriateness judgment* as the dependent variable. The predictors in this

model are *causal relatedness*, *strategy map*, the interaction between *strategy map* and *causal relatedness*, and *article summary fixed effects*. I show these results in Table 4 Panel A.

Consistent with H2, I find a significant positive coefficient for the interaction term (p = 0.02 one-tailed), suggesting that *strategy map* has a greater improving effect on *appropriateness judgment* when *causal relatedness* is coded as one.[38] This finding confirms that greater causal relatedness enhances the effect of causal linkages. This interaction is also evident from Figure 8, which shows actual appropriateness judgment means.

To further explore this effect, I perform simple effects regressions broken out by causal relatedness condition (not tabulated). Simple effects tests such as this help interpret interactions by testing one term at different levels of the other term, which is held constant. However, when I do this, I find that *strategy map* has a non-significant negative effect on appropriateness judgments in the low causal relatedness condition (p = 0.13 one-tailed) and in the high causal relatedness condition a positive effect on appropriateness judgments that is only significant at the 90% confidence level (p = 0.09 one-tailed). My simple effects tests could fail because there are more than one effect contributing to the interaction reported in Table 4 Panel A. I address this below, after discussing my subsample analysis, to provide additional context for my failed simple effects test.

[38] In Section 4.2.3, I do not expect an H2 effect for relevance judgments because these judgments require a relatively superficial degree of recall of the firm's strategy. Consistent with this, when I re-run my regression for H2 (not tabulated) with *relevance judgment* as the predicted variable, I find a non-significant interaction term (p = 0.19 one-tailed).

Table 4: Study 2 Tests of H2

Panel A: Test of H2

$Appropriateness\ Judgment_{it} = \beta_0 + \beta_1 Strategy\ Map_{it} + \beta_2 Causal\ Relatedness_{it} + \beta_3 Strategy\ Map \times Causal\ Relatedness_{it} + Article\ Summary\ Fixed\ Effects_i + u_i + \varepsilon_{it}$

	Coefficient	z-value	p-value (***one-tailed***)
Strategy Map	-0.09	-1.20	0.23
Causal Relatedness	0.02	0.28	0.78
Strategy Map × Causal Relatedness (H2)	0.20	2.03	***0.02***
Article Summary Fixed Effects	—	(1,242.23 χ^2)	<0.01
Constant	3.92	28.14	<0.01

R^2 *(Appropriateness Judgment): 0.279; Wald* χ^2 *(14) = 1252.66 (p < 0.01)*

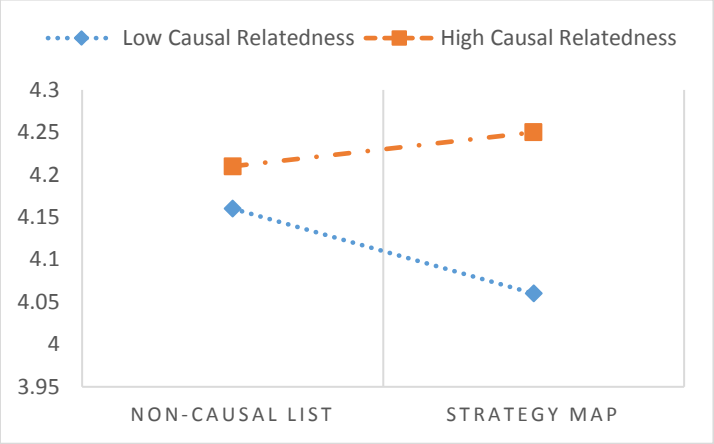
Panel B: Supplemental Boundary Tests of H2

	***β₃Strategy Map × Causal Relatedness* (p-value, one-tailed)**	
	Below-median subsample (or *no* business degree)	Above-median subsample (or *has* business degree)
Work Experience	0.34 (<0.01)	0.07 (0.30)
Management Experience	0.12 (0.19)	0.25 (0.03)
Age	0.27 (0.02)	0.11 (0.20)
Business degree	0.17 (0.04)	0.31 (0.15)
Familiarity with Banking	0.48 (<0.01)	0.12 (0.14)

Panel C: Supplemental Simple Effects for H2

	***β₁Strategy Map* (p-value, one-tailed)**	
	Low Causal Relatedness	High Causal Relatedness
Low Work Experience (>15 years; n = 161)	-0.05 (0.33)	0.30 (<0.01)
High Management Experience (≥ 2 years n = 186)	-0.18 (0.04)	0.06 (0.27)

Figure 8: Study 2 Appropriateness Judgment Means



As detailed in Section 4.2.3, I generally expect the H2 effect to be stronger among those with low experience. The subsample analysis shown in Table 3 Panel B largely supports this, with stronger H2 effects in the below-median subsamples for work experience, age, familiarity with banking, and in the subsample for those without a business degree. However, management experience has the opposite pattern, in that those with above-median management experience appear to have a stronger H2 effect than those with below-median management experience. Furthermore, in this regression (not tabulated), *strategy map* has a significant negative effect (β = -0.18, p = 0.04 one-tailed) for those with above-median management experience. This result suggests that the H2 effect among the above-median management experience subsample is capturing unique downsides to causal linkages that are weak, rather than unique upsides to causal linkages that are strong.[39]

[39] Work experience and management experience are correlated but not perfectly: r = 0.50. There are 67 participants with above-median management experience *and* below-median work experience. Under the assumption that the effects of work experience and management experience are separable and distinct, I include these participants' responses in my main tests of both subsamples. When I exclude these participants, the negative simple effect for the subsample with above-median management experience (and now above-median work experience) remains significant (β = -0.24, p = 0.03 one-tailed) while the positive simple effect for the subsample with below-median work experience (and now below-median management experience) is only significant at the 90% confidence level (β = 0.21, p = 0.06 one-tailed).

One possible reason for this result, post hoc, is that management experience provides a unique type of background knowledge that increases individuals' ability to create innovative pathways through which the cause could lead to the effect. Those with high experience may experience encoding failure because they can think of more pathways for low causal relatedness causal linkages. That is, if managers with high experience generate an abundance of possible relationships between the cause and the effect for low causal relatedness mini-strategies, this could exhaust their working memory capacity and interfere with the process of encoding these strategic goals from working memory in long-term memory.

This tentative explanation fits well with some key prior research. Miller (1956), for example, suggests a maximum capacity for items help in working memory at around seven items (with some variations depending on item complexity and individual differences, see Prat et al. 2011). Furthermore, Chang and Birkett (2004) suggest that higher levels of management experience are accompanied by increased skills in balancing creativity. If that result is applied to this setting, it would be consistent with high management experience accompanying an increased ability to think of creative routes that a cause leads to an effect—potentially *too many* routes to hold in working memory when the causal linkage has low causal relatedness. If this is the correct explanation for my results among those with above-median management experience, then my results suggest it is a phenomenon unique to management experience. My subsample analysis of age, work experience, familiarity with banking, and the possession of a business degree do not show this pattern.

This also might explain my failed simple effects analysis. If my H2 effect is a composite of two effects—high causal relatedness amplifying causal linkages' effect on appropriateness judgments and low causal relatedness weakening this effect—then splitting these two effects out

by causal relatedness condition could weaken the H2 effect such that each condition yields a non-significant simple effect (at the 95% confidence level). Panel C of Table 3 is a supplemental test that supports this interpretation of my failed simple effects test. In this panel, I re-test for H2 simple effects among the two subsamples that best represent the two possible effects: those with below-median work experience (who best exemplify high causal relatedness amplifying the effect of causal linkages) and those with above-median management experience (who best exemplify low causal relatedness weakening the effect of causal linkages). I find significant simple effects in opposite direction among these two subsamples, suggesting that high causal relatedness helps some process causal linkages while low causal relatedness inhibits others from processing causal linkages.[40]

4.5 CONCLUSION

A strategy map is a critical component of a balanced scorecard (Kaplan and Norton 2001b) and is expected to benefit the firm in several ways (e.g. Gimbert et al. 2010; Tayler 2010; Kaplan and Norton 2008). Firms expect these benefits to offset the costs of implementing a strategy map. One of these benefits is an improvement in managers' ability to discern the relevance of new external information and managers' ability to determine whether the firm's

[40] An examination of adjusted means for *appropriateness judgment* in the four experimental conditions of these subsamples is directionally consistent with my amplifying and weakening effect explanation (means are adjusted for *article summary fixed effects*). If participants have below-median work experience, appropriateness judgment means for those who receive a non-causal list are 3.53 and 3.54 for low and high causal relatedness conditions, respectively. But those who receive a strategy map have appropriateness judgment means of 3.41 in the low causal relatedness condition and 3.71 in the high causal relatedness condition. If participants have above-median management experience, appropriateness judgment means of those who receive a non-causal list are 3.63 and 3.59 for low and high causal relatedness conditions, respectively. But those who receive a strategy map have a mean of 3.38 in the low causal relatedness condition and a mean of 3.43 in the high causal relatedness condition.

strategy is appropriate given this new information (Cheng and Humphreys 2012). However, it is unclear from prior research if this benefit of a strategy map depends on strong causal linkages.

In this study, I conduct an experiment that tests how causal linkage strength, i.e. *causal relatedness*, impacts the effect of a strategy map on relevance and appropriateness judgements. First, I find evidence that receiving a strategy map improves relevance judgments, a finding in line with prior literature. This allows managers to better determine the relevance of new information by testing if that information belongs to the same category as the firm's strategic goals. This finding is still worth replicating because it uses a different sample, a different scenario, and a different set of goals that make up a strategy map.

I do *not* find evidence that receiving a strategy map's causal linkages alone improve appropriateness judgments. However, I do find evidence that when causal linkages connect high causal relatedness goals, the result is better appropriateness judgments. Together, these results enrich the knowledge of researchers who examine the judgment and decision making effects of a strategy map as well as the knowledge of practitioners who hope to successfully implement a strategy map. Specifically, these findings suggest practitioners can maximize the beneficial appropriateness judgment of a strategy map by ensuring its causal-linkages are strong.

One limitation of this study is the time constraint I place on how participants encode the firm's strategic goals in long-term memory. It is possible that if participants were given more time they could commit low causal relatedness goals to memory, thereby diminishing any differences in appropriateness judgments between this condition and the high causal relatedness condition. However, even if this were the case, the practical implications of this study would largely remain the same. My results still suggest that firms are best off with strong causal linkages, i.e. causal relationships with high causal relatedness, thus maximizing the benefit

obtained from a strategy map's causal linkages. Managers' time is scarce, and if causal linkages with low causal relatedness strategic goals require extra time to commit to memory, that in and of itself would make a strategy map less beneficial to the firm.

This study is also limited to examining the effect of receiving a strategy map in isolation. Strategy maps are components of a balanced scorecard framework that includes performance measures in each of the balanced scorecard categories. Also, balanced scorecard frameworks are usually distributed within the firm by way of regular meetings and substantial training during the development and rollout stages, as well as regular follow-up meetings (Kaplan and Norton 2008). While my study lacks these other contextual features that typically accompany a strategy map, I do this purposefully. Prior research has suggested that receiving a strategy map has a differential effect apart from its contextual trappings within the larger balanced scorecard (Cheng and Humphreys 2012). This study follows in that vein by stripping away the context of a strategy map and looking at the incremental effect of this particularly visible component of a balanced scorecard. Also, by limiting these contextual features, I can better control the laboratory environment and better ensure my statistical inferences actually do relate to the constructs associated with strategy maps, rather than being driven by something else in the balanced scorecard framework. I acknowledge, though, that this experimental design choice differs, in large part, from the way decision makers encounter a strategy map in practice. It is entirely possible that some contextual features of a balanced scorecard suppress, overwhelm, or interact with the effects found in this study.

Given this latter limitation, one potential extension of this study would be to examine the effect of receiving causal linkages with high causal relatedness and low causal relatedness within a more robust and complete balanced scorecard context. Prior research shows that when

performance measures are strategically-linked (meaning that they line up with the strategic goals in a strategy map), managers pay more attention to them (Banker et al. 2004). Similarly, it is possible that the causal relatedness effect observed in this study extends to the performance measures that are linked to the strategic goals in a strategy map. Managers may differentially attend to performance measures linked to strategic goals with high causal relatedness over performance measures linked to goals with low causal relatedness.

5.0 PROPOSED NEUROIMAGING EXTENSIONS

5.1 HOW DO SUBORDINATES REACT TO MIDDLE MANAGERS' IMPRESSION MANAGEMENT?

Impression management describes one person's attempts to influence another person's impression of him or her. In Study 1, I find that middle managers make decisions about what information to pass along to upper management based on whether they think upper management will be impressed by that information—that is, based on impression management. I then find that middle managers' impression management also affects their evaluation of the subordinates who provide information to middle managers. Middle managers more positively evaluate the performance of subordinates who provide information that middle managers think will impress upper management than subordinates who provide information that middle managers think will *not* impress upper management.

In Study 1, I predict from these findings that long-term subordinate preferences for information will change because of this bias in middle managers' evaluations. Subordinates will learn the evaluative consequences of providing different types of information to middle managers and will preferentially seek out and pass along information that is perceived to impress upper management (see also Bol 2008). Effectively, my Study 1 prediction could arise for two reasons. First, subordinates may become biased toward passing along impressive information simply because they learn that middle managers are most likely to reward this type of feedback. If this holds true, subordinates with biased middle managers still likely view unimpressive information as valuable, but know not to bring it up with middle managers.

The second reason for subordinates to differentially pass along impressive information is that subordinates *internalize* middle managers' impression management and come to believe that unimpressive information is by itself less valuable. In this case, subordinates effectively engaging in impression management of upper management by proxy, even though upper management does not directly affect these subordinates' compensation. The second reason for this phenomenon would have far wider implications than the first, suggesting subordinates might undervalue unimpressive information in a vast array of different choices they make.

My proposed extension of Study 1 is a neuroimaging study that tests how much subordinates have internalized middle managers' impression management-based biases. Rather than behaviorally test a long list of contextual decisions subordinates need to make that might be affected by a devaluation of unimpressive information, imaging allows me to examine the degree to which subordinates' immediate brain activation in response to unimpressive information matches with them engaging in impression management themselves.

To test this, I assign all participants to act as a subordinate who decides whether to pass along information that they discover. However, these participants are divided into two conditions: one in which they receive middle manager evaluations that are biased against unimpressive information and one in which middle managers are not biased against unimpressive information. By using a prearranged set of middle manager responses, rather than actual participants playing the role of middle managers, this design avoids the costly imprecision that real participants often present. Real participants acting as middle managers are likely to be inconsistent in their decisions about rewarding or punishing subordinates and subordinate learning of these preferences might take a long time (using up costly minutes in an MRI scanner).

To analyze these results, I compare neural activation across the three types of stimuli: impressive information, unimpressive information, and neutral information that is neither impressive nor unimpressive.[41] I expect to find that, over time, subordinates' brain activation will become biased towards impressive information over unimpressive information when the middle manager is biased. I expect two patterns of brain activation to emerge among subordinates with biased middle managers. First, I expect the subordinates' neural activation pattern in those areas associated with impression management (see Farrow et al. 2015; Schilbach et al. 2006) are stronger when they receive impressive information than when they receive unimpressive information.[42] Second, I expect that brain regions more generally associated with anticipating reward (see Knutson et al. 2001) will activate more strongly when they decide to pass along impressive information than when they pass along unimpressive information.

Importantly, I compare activation in response to the three types of stimuli with the brain regions associated with anticipating future rewards and impression management for each individual. To locate these brain regions, I conduct two localizer tasks in which participants perform tasks identifying the brain regions that activate when participants anticipate future rewards and consider impression management implications. The first localizer task is patterned after the task used by Knutson et al. (2001), in which participants respond to different levels of anticipated reward. The regions identified in this task indicate an activation pattern for anticipating reward. The second localizer task is patterned after the experiment performed by

[41] The neutral information condition is useful for developing appropriate contrasts in neural activation between conditions. Neural activation needs to be compared to a baseline, and using impressive or unimpressive information as a baseline would be inadequate for discerning if the hypothesized effect is from subordinates reacting positively to impressive information or negatively to unimpressive information.

[42] Alternatively, the neural activation pattern could be U-shaped, with impressive and unimpressive information requiring impression management extensive processing with neutral information requiring little. Given limited literature on the predicted shape of the neural response to this kind of stimuli, I plan to test both plausible patterns.

Farrow et al. (2015). In this task, participants are shown questions about themselves and alternate between instructions to answer the questions so as to give the best impression of themselves (the impression management condition) and instructions to count how many letters are in the questions (the counting condition, i.e., a control condition). By comparing activation in this localizer task, I can locate the brain regions that activate when participants are engaging in impression management to make a good impression on others.

In this study, I provide important evidence about how subordinates react to biased middle managers. This is an important extension of Study 1. This study provides evidence of whether subordinates react to these middle managers by biasing one judgment or by mirroring middle managers' bias in a more pervasive and widespread manner to affect other judgments as well.

5.2 DIFFERENTIAL BRAIN ACTIVATION FROM A STRATEGY MAP

A *relevance judgment* is a manager's judgment of how relevant a piece of new information is to the task of evaluating the firm's strategy. Managers make this critical judgment as they work to continuously revise the firm's strategy to address new developments and previously unknown trends. In Study 2, I find that the causal linkages of a strategy map improve relevance judgments. That is, when managers learn the firm's strategic goals and causal linkages have been placed between these goals, they more accurately determine whether new information is relevant to the firm's strategy. I theorize that this is because causal linkages improve managers' recall of the firm's strategic goals from long-term memory, thus making it easier to recall those strategic goals and compare them with the new information.

In Study 2, however, I do not directly measure participants' recall of the firm's strategic goals in long-term memory. In this proposed study, I use neuroimaging techniques to extend Study 2 to test *whether* these goals are better recalled and *why* they are better recalled. In short, I test whether receiving a strategy map leads to sharper contrasts between the brain activation patterns for each goal, which is my proxy for improved recall of the firm's strategic goals.[43] To test this, I partially replicate the work of Just et al. (2010). They analyze the brain activation patterns produced by 60 concrete nouns and train a machine learning *classifier* algorithm on these activation patterns. Then the classifier is tested by whether it can predict which noun is represented in brain activation patterns that the algorithm has not previously seen. This classifier uses an approach called *multi-voxel pattern analysis*, which tests patterns of brain activation across a population of voxels rather than just comparing activation in one voxel at a time (see Coutanche 2013 for a description and review of this method).[44] Just et al. find that the algorithm can predict, with a relatively high level of accuracy, which word is being thought of by using the brain activation pattern. In effect, the classifier can *read the mind* of the participant. Importantly, they find that the more distinct the words being tested, the higher the accuracy of the classifier.

I expect strategic goals to function in the same way. Strategic goals and concrete nouns should both have recognizable brain activation patterns specific to each goal or noun. Concrete nouns have the advantage of being associated with some physical experience, which leads to

---

[43] Although it is possible to measure recall of the firm's strategic goals using a behavioral experiment without neuroimaging techniques, using neuroimaging allows me to provide a more complete answer of how a strategy map affects recall. That is, the paragraphs below explain how a classifier is used to test whether brain regions associated with the four categories are apt for discriminating between strategic goals. Thus, this proposed study not only tests whether the firm's strategic goals are better recalled, but also whether this improvement occurs *in the way* theorized by strategy map theory.

[44] A voxel is a unit of measurement of three-dimensional space in the brain used in fMRI studies to pinpoint brain activation. A voxel is typically the smallest unit of measurement available and varies in absolute size based on the resolution of the scan performed. Voxels are the three-dimensional analog to two-dimensional pixels.

unique and distinguishable locations for activation. For example, Just et al. (2010) find that concrete nouns that are interpreted as tools tend to be associated with brain areas that are involved with motor control, almost as if the meaning is tied up in simulating hefting the tool. However, strategy maps may still produce a recognizable brain activation pattern because, by design, they span four distinct categories of strategic goals: learning and growth, internal processes, customer, and financial goals. Managers with adequate experience will likely have unique experiences with these categories of strategic goals, leading to differentiable brain regions that activate when exposed to goals in these respective categories.

The brain activation patterns for better recalled strategic goals should be more distinct, making the classifier more accurate. By replicating the procedures of Just et al. (2010)—using the firm's strategic goals in place of concrete nouns—I train two separate classifiers on two separate groups of participants: participants who receive the firm's strategic goals *with* causal linkages and participants who receive the same strategic goals *without* causal linkages. The task participants complete, which produces the brain activation for the classifier algorithm, is a two-step learning and recognition task. Outside the scanner, participants learn the firm's strategic goals. Then, inside the scanner, participants complete a recognition task. The recognition task is a series of multiple choice questions. Participants are shown several foil goals (designed to look like the actual firm strategic goals) and one actual firm strategic goal. They attempt to select the actual strategic goal presented to them. This task is repeated in the scanner several times.

Participants also complete a localizer task in the scanner. This task helps identify brain regions that activate among individual participants. The task involves evaluating several chains of strategic goals passing through each of the four balanced scorecard categories. This task provides evidence of the brain regions that activate when thinking about these four categories.

Then when I train the classifiers on activation patterns during the recognition task I can focus on brain regions specified in the localizer task. Based on my theory from Study 2, I expect the classifier that is trained on participants that receive the firm's strategic goals with causal linkages will have greater accuracy than the classifier trained on participants who receive the firm's strategic goals without causal linkages. This would provide direct evidence that causal linkages do indeed lead to better recall of strategic goals in long-term memory, confirming a key assertion of Study 2.

5.3 A NEUROIMAGING INVESTIGATION OF WHY COMMUNICATING CAUSAL LINKAGES IMPROVES WORKERS' EFFORT ALLOCATIONS

5.3.1 Introduction

Agents in a firm have limited effort to allocate among a nearly-infinite array of possible activities. Unfortunately, agents often allocate their effort between activities in a suboptimal way, favoring activities that more directly impact the agent's financial outcome and activities with more immediate rewards (e.g. Chen et al. 2015; Hannan et al. 2012; Farrell et al. 2012; Laverty 1996; Holmstrom and Milgrom 1991). Farrell et al. (2012) find that communicating the cause-and-effect relationship between non-financial quality activities in the short-term and long-term financial rewards leads to agents better allocating their time between activities. This finding has significant implications for practice, including the implication that firms could find it beneficial to implement a strategy map. A strategy map is a formal visual depiction of hypothesized cause-and-effect relationships between leading non-financial goals and lagging

financial goals. For the remainder of this proposal, I use the term *causal linkage* to mean cause-and-effect relationships, whether visually communicated in a strategy map or otherwise.

However, practitioners may not have as much confidence from Farrell et al.'s (2012) findings as they need in order to decide if implementing a strategy map is beneficial to the firm. There are two possible explanations for why agents who learn of causal linkages better allocate effort than those who do not. These two explanations, detailed below, differ in terms of how long the effect is likely to last, which impacts decision makers significantly. If this effect dissipates quickly, then it is unlikely that a strategy map substantially benefits the firm through improved effort allocations. Thus, it is less likely to be worth implementing a strategy map. If, instead, the effect persists over the long-term, then it is more likely that a strategy map substantially benefits the firm through improved effort allocations. Thus, it is more likely to be worth implementing a strategy map.

The first explanation for why communicating causal linkages improves workers' effort allocations is that poor effort allocations from workers who do not learn of causal linkages are driven by an irrational and distorted view of utility over time that favors short-term rewards. This is also called *hyperbolic discounting* because it involves inappropriate discounting of the present value of future rewards, often approximated by a hyperbolic function (e.g. Epper et al. 2011; Laibson 1997). Hyperbolic discounting theory suggests that many agents suffer from a deep and systematic preference of short-term rewards over long-term rewards. Based on this theory, receiving information on a causal linkage might only temporarily increase the salience of long-term rewards, leading to better effort allocation only in the short-term. Once that salience has dissipated, agents should return to their prior suboptimal effort allocation. This explanation suggests that a strategy map's effect on agents' effort allocations could be short-lived. If this

explanation is accurate, a strategy map is unlikely to improve workers' effort allocations for long, thus suggesting a strategy map is less beneficial to the firm.[45]

In contrast, the second explanation for causal linkages affecting workers' effort allocations is that poor effort allocations in the absence of causal linkages are due to the overuse of model-free reinforcement learning on the part of agents. Model-free reinforcement learning is one of two complementary systems in the brain—the other system being model-*based* reinforcement learning—that together help one learn and optimize future actions. Model-free reinforcement learning is merely trial-and-error learning in which one does not develop any theory for *why* a certain outcome is obtains from a certain alternative, but merely notes that it did obtain from a certain alternative. Model-based reinforcement learning, in contrast, is guided learning based on a theory for events—or a "model"—that attempts to predict the future beyond what has been observed in the past. It is also cognitively costly. Most learning is a combination of these two, but an overuse of model-free reinforcement learning can lead to an agent ignoring even obvious interdependencies and interconnections between activities, instead focusing on the average outcome previously observed from each alternative (Herrnstein and Prelec 1991). This explanation of Farrell et al.'s (2012) results would suggest that the causal linkage communicated to agents serves as a model that allows agents otherwise prone to overusing model-free reinforcement learning to more appropriately balance that form of reinforcement learning with model-based reinforcement learning. Furthermore, this explanation suggests that a strategy map's effect on agents' effort allocations might be long-lasting. If this explanation is accurate, a

[45] Or, alternatively, this result could simply mean that to maintain the strategy map benefit of improvements in effort allocation over time, managers need to regularly continuously remind workers of causal linkages to keep these connections salient to workers. Regardless, this finding would suggest an additional cost to implementing a strategy map: continuous reminders.

strategy map is likely to improve workers' effort allocations for an extended period, thus suggesting a strategy map is more beneficial to the firm.

My study seeks to disentangle these two explanations for why communicating causal linkages improves agents' effort allocations. By using functional magnetic resonance imaging (hereafter referred to as *fMRI*), I record agents' brain activation while they perform a task like that used by Farrell et al. (2012). Then I compare this pattern of brain activation to agent's brain activation on tasks previously used to study hyperbolic discounting (Ballard and Knutson 2009; see also Kable and Glimcher 2007) and model-free/model-based reinforcement learning (Gläscher et al. 2010; see also Schönberg et al. 2007). In so doing, I provide evidence of agents' neural processing (and thus, thought patterns) during task completion, providing practitioners with important evidence about the persistence of a strategy map's effect on effort allocation and helping them decide whether implementing a strategy map is beneficial to the firm.

In the abstract, I define *strategy revision judgments* as manager judgments that affect how well the firm can revise its strategy when new information comes to light. This study does not directly test how a strategy map affects how managers make these judgments. Rather, the results of this study provide managers with information about strategy maps that could prompt a manager to revise the firm's strategy to include a strategy map or to eliminate a strategy map. This differs from the tack taken in my first two studies. My first study, in Section 3, examines how a strategy map affects information sharing within a firm, thus affecting the information available for use to revise the firm's strategy. My second study, in Section 4, examines how a strategy map affects managers' judgments about the relevance of new information and the appropriateness of the firm's strategy, judgments that precede appropriately using new information to revise the firm's strategy.

Together these two studies fit well with the idea of examining a strategy map's effect on the how managers *consider* information when revising the firm's strategy, either by altering how middle managers' share information to upper management when that information is relevant to strategy revision or by affecting how managers process new information that is relevant to strategy revision when they receive it. The results of this study, in contrast, examine one way a strategy map could *be* the information managers consider when revising the firm's strategy. The results of this study could easily be an article summary presented to managers in Study 2. Thus, this study joins a stream of prior studies that examine the broader benefits or costs of a strategy map that could affect its attractiveness to firms (e.g. Farrell et al. 2012; Mastilak et al. 2012; Banker et al. 2011; Banker et al. 2004).[46]

5.3.2 Hypothesis Development

5.3.2.1 Causal Linkages' Effect on Workers' Effort Allocations

Farrell et al. (2012) conduct an important experiment, in which participants perform the same task over several rounds. The quality of the participant's output in one round affects the price per unit of his or her output in the next round. This is the causal linkage of interest to their study. Thus, in each round participants choose between short-term focused activities, which yield short-term rewards (i.e. creating as many units of output as possible), and long-term focused activities, which yield long-term rewards (i.e. maintaining high quality output so as to raise the price in the next round). Farrell et al. manipulate how much communication participants receive

[46] Contrast these studies against those examining the principal studies prompting my first and second studies (Cheng and Humphreys 2012; Tayler 2010), which focus on the effect of a strategy map on how managers make strategy revision judgments.

about this causal linkage. In one condition, a control condition, the participants are merely told that prices can change from round to round. In a second condition, the *qualitative* condition, participants learn that higher quality output in one round will raise the price per unit of output in the next round, but participants do not know the numerical value of this relationship. In the last condition, the *quantitative* condition, participants are given a detailed description of the exact numerical relationship between quality in one round and prices in the next round.

In the control condition, the researchers find that some participants learn about the causal linkage on their own, but about thirty percent fail to produce high quality output that raises the next period's price—suggesting this group ignores the causal linkage between current-period quality and future-period price. In the second condition, where the causal linkage is communicated but not quantified, the researchers find that participants' allocation of effort between quality and quantity of output significantly improves. In the final condition, in which numerical information about the causal linkage is communicated, they find little improvement upon the results from the second condition. Together these findings suggest that communicating causal linkage between long-term focused activities in the current period and long-term rewards improves workers' effort allocation.

This finding has significant implications for firms considering a strategy map. Strategy maps visually depict hypothesized cause-and-effect relationships between leading non-financial goals and lagging financial goals. A strategy map can be interpreted as a large-scale implementation of the qualitative causal linkage Farrell et al. (2012) communicate to participants in the second condition of their experiment. However, it is still unclear from their findings how long this effect will persist, because the researchers provide limited evidence as to *why* participants better allocate their effort when causal linkages are communicated.

5.3.2.2 Hyperbolic discounting

A broad spectrum of literature suggests agents in a firm often allocate effort between long-term focused and short-term focused activities in a way that is suboptimal to the firm (e.g. Chen et al. 2015; Hannan et al. 2012; Farrell et al. 2012; Laverty 1996; Holmstrom and Milgrom 1991). This suboptimal allocation is consistent with two separate theories about *intertemporal choice*, that is, choices that agents make over time. In this section, I examine an economics theory for deficiencies in intertemporal choice, while in section 5.3.2.2, I examine a psychology and neuroscience theory for deficiencies in intertemporal choice. Prior research in economics shows that individuals often prefer immediate rewards over financially equivalent future rewards. This is sometimes called *hyperbolic discounting* because of the shape of the preference function often observed in these irrational intertemporal preferences.[47]

In order to compare future rewards and immediate rewards, individuals must discount future rewards based on their desired rate of return over time. That is, if alternative #1 is to take $1.00 today and alternative #2 is to take $1.25 next year, one cannot choose between the two without having an expectation of how much one expects to earn from $1.00 over the course of a year, i.e. his or her desired rate of return. If one's desired rate of return were more than 25%, taking the dollar immediately would be the better choice because in a year's time it would be worth more than $1.25.[48] Better put, the present value of $1.25 in a year, which is the amount of money in the present that is expected to grow into $1.25 in a year's time, is less than the present

[47] Technically, the following paragraphs lay out the mathematical approximation known as *quasi*-hyperbolic discounting (Laibson 1997; originally developed by Phelps and Pollak 1968). This approximation of hyperbolic discounting provides largely the same predictions and is easier to manage mathematically.
[48] This assumes that risks are factored into one's desired rate of return.

value of $1.00. Discounting future rewards to their present value is an obligatory component of almost all business calculations.

In contrast, some research has found that human behavior is characterized by an irrational preference for immediate rewards over future rewards. Laibson (1997) approximates hyperbolic discounting in a utility function similar to the following.

$$U_t = u_t + \beta \times \delta(u_{t+1})$$

The $\delta(\cdot)$ term represents a discounting function that appropriately discounts the utility of future rewards (i.e. u_{t+1}). The β term is between 0 and 1 and represents the irrational discounting of future rewards. Thus, returning to my example above, I assume that one's rate of return is less than 25%, meaning the $\delta(\cdot)$ function would render utility of alternative #2 to be greater than the utility of alternative #1. A rational individual in this case would choose alternative #2. But the β term can discount this future reward, making the future reward less attractive and making one opt for alternative #1 at time $t = 1$. The smaller the β term the greater the preference for immediate rewards. Interestingly, at time $t = 0$, the β term has no effect on one's preference between the two alternatives, because both $1 (in the future) and $1.25 (further in the future) are discounted by the same amount: β. Thus, at $t = 0$ one could plan on choosing alternative #2, the more long-term oriented alternative with higher overall utility. Then, at $t = 1$, one would suddenly find it less attractive to wait for the $1.25 (which has been discounted by the β term, while the $1.00 is no longer discounted by the β term). Then one changes course and takes the immediate $1.00, instead of following through with the original plan to select the higher utility, long-term oriented alternative.

One everyday example of this very pattern is the tradition of new year's resolutions, which often begin well-intentioned and spirited, but typically fail over time. Wiseman (2008)

reports that new year's resolutions have a success rate of about 12%, a poor rate of success that appears to be consistent with hyperbolic discounting. That is, on New Year's Day, i.e. $t = 0$, the resolution appears to be achievable, with the future costs of the resolution being at least equal to the benefits over the long-term. As the year progresses, i.e. at $t > 0$, the utility of future rewards decrease by a factor of β relative to the disutility of now-present costs, often leading to the abandonment of the resolution.

The fact that hyperbolic discounting is observed so widely suggests that the β term is a stubborn feature of one's utility function and is difficult to change (see O'Donoghue and Rabin 1999). However, I expect that it is possible for an emphasis on causal linkages between short-term costs and long-term rewards can offset β for a short time. Returning to the new year's resolution example, there is often a short time during which temporary excitement for the new commitment increases the salience of the long-term rewards. The increased salience of future period rewards effectively increases β. In the below equation, I represent this by replacing β with a function, $\sigma(\beta, t^*)$, with t^* representing a temporary increase in β in period t^*.

$$U_t = u_t + \sigma(\beta, t^*) \times \delta(u_{t+1})$$

Using this explanation of Farrell et al.'s (2012) findings, the effect of communicating causal linkages to workers is relatively temporary. Once period t^* passes, β returns to standard rates and irrational hyperbolic discounting continues unabated. Using this interpretation, when Farrell et al. communicate causal linkages to participants, it positively affects participants only in time t^*, which lasts throughout their experiment. Based on this interpretation, sometime after the length of time it takes for participants to complete their experiment, the salience of the long-term rewards should fade and short-term rewards would look more and more attractive in comparison. Thus, suboptimal discounting would return. This interpretation suggests that a strategy map,

which is costly to implement and maintain, does not have a sustainable beneficial effect on workers' effort allocations. I capture this interpretation in my first hypothesis.

> **H1:** When a firm communicates to workers that causal linkages exist between long-term focused activities in the current period and long-term reward, this improves their effort allocation between long-term focused activities and short-term focused activities in the current period by temporarily disrupting workers' hyperbolic discounting.

5.3.2.3 Model-free and Model-based Reinforcement Learning

Intertemporal choice is affected by what individuals predict to arise from different alternatives, and cognitive inefficiencies in these predictions could explain Farrell et al.'s (2012) findings. Psychology and neuroscience theory suggest that individuals adjust their predictions of the future over time using two complementary systems in the brain: *model-free* reinforcement learning and *model-based* reinforcement learning. Model-free reinforcement learning simply tracks average prior outcomes from different alternatives, producing a likely outcome number for each of the available alternatives. This is a useful tool that helps humans learn most skills. For example, it is useful in learning how to shoot a basketball. The practice involved in developing expertise in basketball provides the brain with millions of data points for determining which muscles to use and when to use them to make a basket. The analysis of these many data points would be the role of model-free reinforcement learning.

Model-based reinforcement learning, in contrast, requires one to develop a declarative theory for cause and effect to better predict the expected outcome. Model-based reinforcement learning serves to economize the trial-and-error approach of model-free reinforcement learning; one can rule out certain activities and focus on other activities based on the model, without the need for extensive practice at these activities first. Continuing with my basketball example, the benefits of practice are amplified when a coach provides examples of good form for shooting a basketball. The coach's example is a model that provides a theoretical basis for predicting what

muscle movements and timing are likely or unlikely to make a basket, above and beyond one's own observations from his or her own practice. This form of reinforcement learning is more cognitively costly because one must create and interpret a model, then extrapolate the implications of this model for one's choice of alternatives.

Typically, model-free and model-based systems work together. However, sometimes one system is overused (see Herrnstein and Prelec 1991). If a model is unavailable or difficult to form, model-free reinforcement learning can be overused and lead to suboptimal decisions. For example, it can be difficult to discern interrelationships between different alternatives over time, if one only judges from past experience without creating a model from that experience. For example, assume there are two alternative basketball shots: a jump shot and a layup.[49] A layup, on average, is effective 50% of the time and a jump shot, on average, is effective 25% of the time. An offensive player using model-free reinforcement learning can detect these averages over time, leading such a player to favor layups over jump shots.

However, what is missing from this example is that once defenders have been bested by a layup, they begin to anticipate the next layup and change their defensive tactics to block it more effectively. Thus, a layup that follows another layup is somewhat less effective. A layup that follows two consecutive layups is even less effective and so forth. As defenders alter tactics to better block layups, they make it relatively easier for the offensive player to successfully complete a jump shot. This cause-and-effect relationship between alternatives is a model because it predicts the rates of success in the next round without necessarily relying on past experience with these alternatives. Model-free reinforcement learning is ill-equipped for altering the success rate of different alternatives based on unexperienced context. Based on this cause-and-effect

[49] This example is inspired by an example relayed by Farrell et al. (2012).

relationship between alternatives, i.e. based on this model, the optimal strategy is to pursue a mixture of layups and jump shots that includes far more jump shots than suggested by the original success rates (the exact mixture depending on how quickly defenders adapt to layups). However, without a model for how a layup attempted in this round affects the probable outcomes for next round, the offensive player will forego this strategy in favor of a far more layup-heavy strategy based on a 50% layup success rate versus 25% jump shot success rate. Thus, an overuse of model-free reinforcement learning can lead to suboptimal outcomes because it ignores interrelationships between alternatives over time.

Effectively, Farrell et al.'s (2012) findings can be explained as an overuse of model-free reinforcement learning. Based on this explanation, communicating causal linkages provides a model so workers can use a model-based approach to effort allocations. Alternatively, relying on model-free learning leads some workers to overuse that system and suboptimally allocate effort. I have no theoretical basis for expecting the positive effect of having a cause-and-effect model to diminish over time. This leads to my second hypothesis.

> **H2:** When a firm communicates to workers that causal linkages exist between long-term focused activities in the current period and long-term reward, this improves their effort allocation between long-term focused activities and short-term focused activities in the current period by communicating a model to facilitate model-based reinforcement learning.

5.3.2.4 Measuring Neural Activation

Prior research shows brain activation in differentiable brain regions for both hyperbolic discounting (Ballard and Knutson 2009; see also Kable and Glimcher 2007) and model-free/model-based reinforcement learning (Gläscher et al. 2010; see also Schönberg et al. 2007). This prior research is important because it suggests that one should be able to identify neural activation, and thus thought processes, that are consistent with these explanations. By examining

the neural activation that accompanies workers' effort allocation decisions, I can provide evidence as to which of my two above hypotheses is correct. Measuring neural activation also avoids difficulties associated with participants lacking self-knowledge of their thought processes.[50]

Activation of a neuron, i.e. neuronal activation, refers to the signal sent by one neuron to another, which is the fundamental phenomenon underlying cognitive activity in the brain. Typically, the activation of individual neurons, or even the activation of small groups of neurons, is not directly measured in human brain research. Directly measuring this activation would involve the surgical application of probes, which is extremely invasive and results in some degree of permanent damage, raising many technical, medical, and ethical concerns. Instead, brain activity is often measured by close proxies for neuronal activation. These proxies typically capture activation on a much larger scale than the individual neuron, instead representing whole regions of the brain, which are often made up of tens of thousands, hundreds of thousands, or millions of neurons.

In the case of fMRI, the close proxy for neuronal activation is the blood-oxygen levels in different regions of the brain, which correspond to intensive neuronal activation in those regions. Blood-oxygen levels correspond to neuronal activation because oxygenated blood provides trophic support for activated neurons, given that neuronal activation depletes important blood-borne resources in the neuron. When a brain region undergoes intensive activation, oxygenated blood re-routes toward this activated brain region to resupply that region and this change within the brain can be detected by measuring the magnetic resonance of brain regions in response to a

[50] I do not review the brain regions specified by these studies in this section because my method involves localizer tasks that capture each participant's unique pattern of brain region activation during the tasks.

magnetic pulse that originates outside the brain (see Ogawa et al. 1990). Magnetic resonance responses generate a contrast between oxygenated blood levels in different regions of the brain.[51] Thus, the signal being tracked is often called *blood-oxygen level dependent* contrast, or *BOLD contrast*. In the case of this proposed study, I expect this BOLD contrast to support either H1 or H2.

5.3.3 Method

5.3.3.1 Experimental Design

To replicate the results from Farrell et al.'s (2012) study that are most relevant to my research question, I employ a 1 × 2 experimental design with one between-subject manipulation. This between-subject manipulation is the qualitative communication of a causal linkage, and it is manipulated at two-levels: causal linkage and no causal linkage. As described below, the experimental task is also derived from Farrell et al. In the causal linkage condition, I communicate to participants that quality in one period affects output price in future periods. In the no causal linkage condition, I simply inform participants that output price may change in future periods.

For simplicity, I omit Farrell et al.'s (2012) condition in which they provide a quantitative substantiation of the causal linkage because their results suggest the quantitative communication of causal linkages provides little improvement in effort allocations above communication of qualitative cause-and effect relationships. Both of my experimental conditions complete the same

[51] Magnetic resonance imaging generates an image from different regions within the scanning field, i.e. different regions of brain tissue, based on their response to a magnetic pulse. By repeating this magnetic pulse over and over in quick succession, fMRI can provide real-time depictions of changes in blood-oxygen level from different stimuli.

experimental procedures described below. Also, I expect that endogenous groupings of participants into learning and non-learning types, as observed by Farrell et al. This is because some participants likely tend toward completing tasks through an overuse of model-free reinforcement learning. These groupings are endogenous however, and do not reflect an experimental manipulation.

One key problem with BOLD contrasts is that they tend to be statistically noisy. There are several reasons for this and intensive statistical pre-processing techniques have been developed to remove as much of the noise as possible and extract the underlying signal. Other pre-processing techniques are useful for communicating studies to other researchers, but these can remove some of the richness of the data and can actually make BOLD signals harder to find. One such pre-processing technique is normalization of the brain to a standard map of the brain across participants. Everyone's brain is wired differently and brain regions vary slightly as to where they are, what they look like, and how much they activate given a stimulus. Normalization removes some of the idiosyncrasies of individual brains and overlays each participants' brain activation onto a standardized map of the brain that can be readily communicated to other researchers. When comparisons across conditions are completed after normalization, after the unique patterns of each brain are stripped away, tests for significant differences rely on whether brain activation patterns align across individuals.

Instead, I use *localizer tasks* to enhance an otherwise noisy BOLD contrast, which can utilize idiosyncratic information prior to normalization. Localizer tasks serve as baseline comparisons that identify the brain regions that activate uniquely for each individual participant's brain as they complete a task known to engage certain types of thought processes. Rather than compare groups of participants in my two experimental conditions against each

other—a comparison that would take place after the normalization process where idiosyncratic differences are lost—I compare participants' activation pattern in Farrell et al.'s (2012) task against their own activation patterns during localizer tasks. This increases the statistical power of my analysis. These localizer tasks, described below, are patterned after prior tasks used to research hyperbolic discounting and model-free/model-based reinforcement learning.

5.3.3.2 Participants

Participants are drawn from a pool like that used by Farrell et al. (2012): undergraduates enrolled in upper-division accounting courses. This participant pool has several advantages. First, this pool matches Farrell et al.'s, a fact that limits potential confounds that otherwise could interfere with interpreting my results in light of theirs. Second, this participant pool is readily available and relatively easy to schedule in the MRI scanner. This is important because only one participant can be in the scanner at a time and if participants are hard to schedule, it may take an excessive period of time to collect the MRI data. Third, I expect this group to have significant background knowledge due to their advanced position in their respective degree programs. This background knowledge means they can complete the task using thought processes close to those used in practice. I pay participants a competitive flat rate as a general incentive for participating in an experiment inside the MRI scanner, which can be noisy, cramped, and otherwise uncomfortable. I also pay participants appropriate piece-rate compensation based on their performance, like the performance pay contracts described in Farrell et al. (2012), Gläscher et al. (2010), and Ballard and Knutson (2009).

5.3.3.3 Experimental Procedures

Before entering the scanner, participants are screened for right-handedness (to ensure similar brain lateralization) and other prerequisite conditions for their participation (such as not being pregnant and not having pieces of metal in their bodies).[52] In a separate room, participants are trained on using two five-key keyboards to manipulate the sandwich assembly task used by Farrell et al. (2012).[53] The causal linkage between quality and price is not included in this practice session. Once participants are proficient at the mechanics of the task, they enter the scanner. The first localizer task is the task given to participants by Ballard and Knutson (2009).[54] This task measures neural activation during discounting stimuli as participants are offered rewards of various levels of magnitude and delay.[55] I discuss how these results are used to test H1 in the next section.

The second localizer task is the task used by Gläscher et al. (2010) in which participants demonstrate their tendency toward model-free or model-based reinforcement learning of a task. By performing this task, I gain a baseline observation of the brain region networks that activate during model-free and model-based reinforcement learning. This task entails navigating through a tree of images, and one must construct a mental model of this decision tree to maximize the outcome from the interrelated alternatives on this tree. I expect some participants to favor model-free reinforcement learning and thus to be "non-learners." In the below section, I discuss how data from this localizer task is used to test H2.

---

[52] Upon completion of this proposed study, I will document the model number(s) for the MRI scanner(s) used.
[53] Each finger on one hand controls one of the five ingredient categories used by Farrell et al. (2012). Only two fingers are used on the other hand: first to indicate that the participant wishes to remove an ingredient and second to complete the order.
[54] Although I refer to these as the first and second localizer tasks, they are presented in a counterbalanced order.
[55] Because I also collect behavioral results from this localizer task, I could feasibly calculate participants' irrational discounting factors, which I could then use to further test whether hyperbolic discounting behavior predicts behavior in Farrell et al.'s (2012) task.

After the two localizer tasks have been completed, participants complete the task used by Farrell et al. (2012). That is, participants simulate the assembly of sandwiches to meet customer orders that appear on the computer screen. The number of mistakes each participant makes in the current round affects the price participants receive per sandwich in the next round. I inform participants in the causal linkage condition of this causal linkage before they begin the task. After participants complete this task over several rounds, participants exit the scanner and complete a post-experiment questionnaire, including demographic questions. Participants are paid flat wages as well as any performance pay they earned from the second localizer task or the sandwich assembly task. Performance pay from the first localizer task is sent to participants after the delay specified in that task (see Ballard and Knutson 2009).

5.3.3.4 Planned Analysis of Localizer Tasks

I use the two localizer tasks to test H1 and H2, respectively. As explained above, these localizer tasks are chosen because they engage thought processes that correlate to H1 and H2. Based on results from the first localizer task, I can locate brain regions that activate when participants irrationally discount future rewards. From there, I compare activation during the Farrell et al. (2012) task to these hyperbolic discounting brain regions. If H1 is correct, these two sets of activation patterns should be more highly correlated in the no causal linkage condition than in the causal linkage condition.

In the second localizer task, I expect participants' behavioral results to divide them into two endogenous groups: learners and non-learners (see Gläscher et al. 2010). Based on H2, the non-learner group has a tendency to overuse model-free reinforcement learning rather than develop a model. I expect these non-learners to also overuse model-free reinforcement learning in the no causal linkage condition, but I expect them to incorporate more model-based

reinforcement learning in the causal linkage condition. To test H2, I compare non-learner's neural activation between the localizer task and the task used by Farrell et al. (2012). Support for H2 would come in the form of non-learners' neural activation between the two tasks being more highly correlated in the no causal linkage condition than in the causal linkage condition.

6.0 CONCLUSION

Prior research suggests strategy maps can have broad effects on managers' judgments and firm outcomes (e.g. Cheng and Humphreys 2012; Mastilak et al. 2012; Gimbert et al. 2010; Tayler 2010). This line of prior research provides some limited evidence that a strategy map improves strategy revision judgments, with two studies providing especially relevant results. The first study, by Tayler (2010), provides evidence that receiving a strategy map can reduce managers' biased reactions to new information that contradicts a strategy the managers themselves chose. However, his experiment does not directly test middle managers' reaction to information that contradicts upper management's chosen strategy.

In Study 1, I use impression management theory and find a different motivation at play among middle managers, who generally do not create the firm's strategy: middle managers more often withhold information the more unimpressive it is to upper management and information that contradicts the firm's strategy is seen as relatively unimpressive. Among managers with two or more years of management experience, I find that receiving a strategy map helps mitigate this bias in what information they are willing to pass along. Furthermore, I find *no* evidence that receiving a strategy map ameliorates another negative behavior of middle managers: providing performance evaluations to subordinates that are biased against those providing information that contradicts the firm's strategy and is thus unimpressive to upper management. My results also contribute to those of Tayler's because his results suggest that a strategy map only mitigates managers' biases when those managers choose their own performance measures. In contrast, I find that receiving a strategy map affects middle managers' willingness to pass along unimpressive information even when they do not choose their own performance measures.

In the second highly relevant prior study, Cheng and Humphreys (2012) provide evidence that receiving a strategy map improves two judgments that are essential to revising the firm's strategy based on new information. These two strategy revision judgments are, first, managers' judgment of how relevant the new information is to revising the firm's strategy (i.e. relevance judgments), and second, managers' judgments of whether the firm's strategy is appropriate in light of the new information (i.e. appropriateness judgments). However, Cheng and Humphreys only test strong causal linkages in their experiment. A separate study (see Webb 2004), suggests that weak causal linkages in a strategy map are less beneficial than strong causal linkages. Weak causal linkages are abundant in practice due to ignorance of the cause-and-effect relationships, philosophical differences among upper management agents creating the strategy map, or incentives to find strategic goals and accompanying performance measures that are easier to meet regardless of how well they fit together in causal linkages (Luft 2004). I find that a strategy map affects relevance judgments regardless of causal linkage strength, but only positively impacts appropriateness judgments when the causal linkages are strong. Among one subsample, weak causal linkages even worsen appropriateness judgments.

At the end of each study, I detail limitations and potential extensions (see Sections 3.6 and 4.5). In Section 5.0, I propose three additional studies that utilize neuroscience theory and neuroimaging techniques. These three proposals are especially important because neuroimaging tools are uniquely poised to expand and clarify existing theories about the effect of receiving a strategy map and about the mental processes that affect strategy revision judgments. This is because my two studies, like many prior studies on this topic, are broadly characterized as framing effects studies. That is, the primary research questions of my two studies tend to revolve around the effect of changing the way information is framed: in this case, by framing information

about the firm's strategy as a strategy map. It can be difficult to extrapolate results from framing effects studies to the business world because seemingly minor differences in the way information is framed can dramatically alter the findings presented. The number of dimensions on which the framing of information can be changed is often so numerous that one cannot possibly control the context by clever experimental design alone. Neuroimaging studies, in contrast, promise to observe underlying neural processes, which provide especially useful insight into why a certain framing effect is found. Thus, with a greater understanding of the thought processes underlying an effect, one can better predict when it would be seen in practice.

Indeed, the neuroimaging proposals in Section 5.0 include one proposed study that effectively examines thought processes underlying a previously found framing effect in the interest of making stronger predictions about its persistence over time. Farrell et al. (2012) find that receiving a strategy map improves how workers allocate effort between activities that benefit the firm (and the worker) in the short-term and activities that benefit the firm (and the worker) in the long-term. However, it is unclear from this study whether this effect would be expected to last long into the implementation of a strategy map. I propose a neuroimaging study that will provide evidence of neural processing, and thus thought processes, which helps provide a better prediction of how long a strategy map is likely to improve these effort allocations.

In sum, my dissertation provides knowledge that is important for practice because managers face significant obstacles to properly revising the firm's strategies when new information comes to light. Prior literature provides limited evidence suggesting a strategy map might help firms overcome these challenges. My findings contribute to this prior literature by providing more robust and detailed evidence of how receiving a strategy map helps managers' strategy revision judgments. As such, these two studies provide guidance to practitioners on how

best to design a strategy map if the intended outcome is proper strategy revision. And lastly, my dissertation includes proposals that serve to point the way forward for research on strategy map judgment effects by including neuroscience theory and neuroimaging techniques. This, I believe, will be a particularly impactful tool in answering future questions about how a strategy map affects managers' judgments.

REFERENCES

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*(4), 451-474.

Arbib, M. A. (2002). Schema theory. In *The Handbook of Brain Theory and Neural Networks* (pp. 830-834). MIT Press.

Avery, C., Resnick, P., & Zeckhauser, R. (1999). The market for evaluations. *American Economic Review*, 564-584.

Baiman, S., & Demski, J. S. (1980). Economically optimal performance evaluation and control systems. *Journal of Accounting Research*, 184-220.

Banker, R. D., Chang, H., & Pizzini, M. J. (2004). The balanced scorecard: Judgmental effects of performance measures linked to strategy. *The Accounting Review*, *79*(1), 1-23.

Banker, R. D., Chang, H., & Pizzini, M. (2011). The judgmental effects of strategy maps in balanced scorecard performance evaluations. *International Journal of Accounting Information Systems*, *12*(4), 259-279.

Ballard, K., & Knutson, B. (2009). Dissociable neural representations of future reward magnitude and delay during temporal discounting. *Neuroimage*, 45(1), 143-150.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of personality and social psychology*, *51*(6), 1173.

Beer, M., Voelpel, S. C., Leibold, M., & Tekie, E. B. (2005). Strategic management as organizational learning: developing fit and alignment through a disciplined process. *Long Range Planning*, *38*(5), 445-465.

Bennett, G. B., and R. C. Hatfield. 2013. The effect of the social mismatch between staff auditors and client management on the collection of audit evidence. *The Accounting Review* 88 (1): 31-50.

Bentley MacLeod, W. (2003). Optimal contracting with subjective evaluation. *The American Economic Review*, *93*(1), 216-240.

Bisbe, J., & Malagueño, R. (2012). Using strategic performance measurement systems for strategy formulation: Does it work in dynamic environments?. *Management Accounting Research*, *23*(4), 296-311.

Bliss, T. V., & Collingridge, G. L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*, *361*(6407), 31.

Bol, J. C. (2008). Subjectivity in Compensation Contracting. *Journal of Accounting Literature, 27*, 1-24

Bozeman, D. P., & Kacmar, K. M. (1997). A cybernetic model of impression management processes in organizations. *Organizational Behavior and Human Decision Processes*, *69*(1), 9-30.

Brews, P. J., & Hunt, M. R. (1999). Learning to plan and planning to learn: resolving the planning school/learning school debate. *Strategic Management Journal*, *20*(10), 889-913.

Campbell, D., Datar, S. M., Kulp, S. L., & Narayanan, V. G. (2015) Testing Strategy with Multiple Performance Measures: Evidence from a Balanced Scorecard at Store24. *Journal of Management Accounting Research*. In-Press.

Chang, L., & Birkett, B. (2004). Managing intellectual capital in a professional service firm: exploring the creativity–productivity paradox. *Management Accounting Research*, *15*(1), 7-31.

Chen, X., Cheng, Q., Lo, A. K., & Wang, X. (2015). CEO contractual protection and managerial short-termism. *The Accounting Review*, 90(5), 1871-1906.

Cheng, M. M., & Humphreys, K. A. (2012). The Differential Improvement Effects of the Strategy Map and Scorecard Perspectives on Managers' Strategic Judgments. *The Accounting Review*, 87(3), 899-924.

Choi, J., Hecht, G. W., & Tayler, W. B. (2012). Lost in translation: The effects of incentive compensation on strategy surrogation. *The Accounting Review*, *87*(4), 1135-1163.

Choi, J. W., Hecht, G. W., & Tayler, W. B. (2013). Strategy selection, surrogation, and strategic performance measurement systems. *Journal of Accounting Research*, *51*(1), 105-133.

Corneille, O., Goldstone, R. L., Queller, S., & Potter, T. (2006). Asymmetries in categorization, cerceptual discrimination, and visual search for reference and nonreference exemplars. *Memory & Cognition*, *34*(3), 556-567.

Coutanche, M. N. (2013). Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us?. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(3), 667-673.

Coutanche, M. N., Gianessi, C. A., Chanales, A. J., Willison, K. W., & Thompson-Schill, S. L. (2013). The role of sleep in forming a memory representation of a two-dimensional space. *Hippocampus*, *23*(12), 1189-1197.

Day, G. S., & Schoemaker, P. J. (2005). Scanning the periphery. *Harvard Business Review*, *83*(11), 135.

Decoene, V., & Bruggeman, W. (2006). Strategic alignment and middle-level managers' motivation in a balanced scorecard setting. *International Journal of Operations & Production Management*, *26*(4), 429-448.

Dimoka, A. (2012). How to conduct a functional magnetic resonance (fMRI) study in social science research. *MIS Quarterly*, *36*(3), 811-840.

Duffy, S. A., Shinjo, M., and Myers, J. L. (1990). The effect of encoding task on memory for sentence pairs varying in causal relatedness. *Journal of Memory and Language*, 29(1), 27-42.

Dutton, J. E., Ashford, S. J., O'Neill, R. M., Hayes, E., & Wierba, E. E. (1997). Reading the wind: How middle managers assess the context for selling issues to top managers. *Strategic Management Journal*, *18*(5), 407-423.

Dutton, J. E., Ashford, S. J., O'Neill, R. M., & Lawrence, K. A. (2001). Moves that matter: Issue selling and organizational change. *Academy of Management Journal*, *44*(4), 716-736.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, *71*(1), 5.

Ellenbogen, J. M., Hu, P. T., Payne, J. D., Titone, D., & Walker, M. P. (2007). Human relational memory requires time and sleep. *Proceedings of the National Academy of Sciences*, *104*(18), 7723-7728.

Epper, T., Fehr-Duda, H., & Bruhin, A. (2011). Viewing the future through a warped lens: Why uncertainty generates hyperbolic discounting. *Journal of Risk and Uncertainty*, 43(3), 169-203.

Ewerhart, C., & Schmitz, P. W. (2000). "Yes men", integrity, and the optimal design of incentive contracts. *Journal of Economic Behavior & Organization*, *43*(1), 115-125.

Farrell, A. M., Grenier, J. H., and Leiby J. (2017). Scoundrels or Stars? Theory and Evidence on the Quality of Workers in Online Labor Markets. *The Accounting Review*, 92(1), 93-114

Farrell, A. M., Kadous, K., & Towry, K. L. (2012). Does the communication of causal linkages improve employee effort allocations and firm performance? An experimental investigation. *Journal of Management Accounting Research*, *24*(1), 77-102.

Farrow, T. F., Burgess, J., Wilkinson, I. D., & Hunter, M. D. (2015). Neural correlates of self-deception and impression-management. *Neuropsychologia*, *67*, 159-174.

Garicano, L., & Posner, R. A. (2005). Intelligence failures: An organizational economics perspective. *The Journal of Economic Perspectives*, *19*(4), 151-170.

Gibbard, A. (1973) Manipulation of Voting Schemes. *Econometrica*, 41: 587-601

Gibbs, M., Merchant, K. A., Stede, W. A. V. D., & Vargus, M. E. (2004). Determinants and effects of subjectivity in incentives. *The Accounting Review*, *79*(2), 409-436.

Gimbert, X., Bisbe, J., & Mendoza, X. (2010). The role of performance measurement systems in strategy formulation processes. *Long Range Planning*, *43*(4), 477-497.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4), 585-595.

Gneezy, U., Gravert, C., Saccardo, S., & Tausch, F. (2017). A must lie situation–avoiding giving negative feedback. *Games and Economic Behavior*, *102*, 445-454.

Haesebrouck, K., Van den Abbeele, A. G., & Williamson, M. G. (2015). The effects of knowledge sharing on helping behavior: Implications for reward system design. Working paper: KU Leuven.

Hannan, R., R. Krishnan, and A. Newman. (2008). The effects of disseminating relative performance feedback in tournament versus individual performance compensation plans. The Accounting Review 83: 893–913.

Hannan, R. L., McPhee, G. P., Newman, A. H., & Tafkov, I. D. (2012). The effect of relative performance information on performance and effort allocation in a multi-task environment. *The Accounting Review*, *88*(2), 553-575.

Hannan, R. L., Towry, K. L., & Zhang, Y. M. (2013). Turning up the volume: An experimental investigation of the role of mutual monitoring in tournaments. *Contemporary Accounting Research*, *30*(4), 1401-1426.

Herrnstein, R., and D. Prelec. (1991). Melioration: A theory of distributed choice. *The Journal of Economic Perspectives*, *5*(3): 137–156.

Holmstrom, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7, 24-52.

Huelsbeck, D. P., K. A. Merchant, & T. Sandino (2011) On Testing Business Models. *The Accounting Review 86*(5), 1631-1654

Hwang, Y., Erkens, D. H., & Evans III, J. H. (2009). Knowledge sharing and incentive design in production environments: Theory and evidence. *The Accounting Review*, *84*(4), 1145-1170.

Ittner, C. D., & Larcker, D. F. (2003). Coming up short on nonfinancial performance measurement. *Harvard Business Review*, *81*(11), 88-95.

Jensen, M. C., & Meckling, W. H. (1992). Specific and general knowledge and organizational structure. *Knowledge Management & Organizational Design*, 17-18.

Jensen, M. C., & Meckling, W. H. (1999). Specific knowledge and divisional performance measurement. *Journal of Applied Corporate Finance*, *12*(2), 8-17.

Jones, E. E. (1964). Ingratiation. *Appleton-Century-Crofts.*

Just, M. A., Cherkassky, V. L., Aryal, S., & Mitchell, T. M. (2010). A neurosemantic theory of concrete noun representation based on the underlying brain codes. *PloS One*, *5*(1), e8622.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*(12), 1625-1633.

Kadous, K., Proell, C. A., Rich, J., & Zhou, Y. D. (2017). Upward Communication of Audit Issues: The Effects of Issue Ambiguity and Intrinsic Motivation. Working Paper.

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard: translating strategy into action*. Harvard Business Press.

Kaplan, R. S., & Norton, D. P. (2001a). Having trouble with your strategy?: Then map it. *Harvard Business Review*, 78(5).

Kaplan, R. S., & Norton, D. P. (2001b). The strategy-focused organization: How balanced scorecard companies thrive in the new business environment. *Harvard Business Press.*

Kaplan, R. S., & Norton, D. P. (2004). The strategy map: guide to aligning intangible assets. *Strategy & Leadership*, *32*(5), 10-17.

Kaplan, R. S., & Norton, D. P. (2008). The execution premium: linking strategy to operations for competitive advantage. *Harvard Business Press*.

Kassing, J. W. (2002). Speaking Up Identifying Employees' Upward Dissent Strategies. *Management Communication Quarterly*, *16*(2), 187-209.

Kassing, J. W. (2007). Going Around the Boss Exploring the Consequences of Circumvention. *Management Communication Quarterly*, 21(1), 55-74.

Kassing, J. W. (2009). "In Case You Didn't Hear Me the First Time" An Examination of Repetitious Upward Dissent. *Management Communication Quarterly*, *22*(3), 416-436.

Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior*, *23*(2), 115-126.

Kerr, N., L. Messe, D. Seok, E. Sambolec, R. Lount, and E. Park. (2007). Psychological mechanisms underlying the Kohler motivation gain. *Personality and Social Psychology Bulletin* 33: 828–841

Khatri, N., & Tsang, E. W. (2003). Antecedents and consequences of cronyism in organizations. *Journal of Business Ethics*, *43*(4), 289-303.

Knutson, B., Adams, C. M., Fong, G. W., & Hommer, D. (2001). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, *21*(16), RC159.

Korte, R. F. (2009). How newcomers learn the social norms of an organization: A case study of the socialization of newly hired engineers. *Human Resource Development Quarterly*, *20*(3), 285-306.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, *112*(2), 443-478.

Laverty, K. J. (1996). Economic "short-termism": The debate, the unresolved issues, and the implications for management practice and research. *Academy of Management Review*, *21*(3), 825-860.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*(11), 2098.

Luft, J. (2004). Discussion of: Managers' commitment to the goals contained in a strategic performance measurement system. *Contemporary Accounting Research* 21 (4): 959–964.

Machiavelli, N. (1988). The Prince (Q. Skinner & R. Price, Trans.). Cambridge, UK: *Cambridge University Press*. (Original work published in 1532)

Markman, A. B., & Wisniewski, E. J. (1997). Similar and different: The differentiation of basic-level categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(1), 54.

Mason, R. A., & Just, M. A. (2004). How the Brain Processes Causal Inferences in Text A Theoretical Account of Generation and Integration Component Processes Utilizing Both Cerebral Hemispheres. *Psychological Science*, *15*(1), 1-7.

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44*(1), 1-23.

Mastilak, C., Matuszewski, L., Miller, F., & Woods, A. (2012). Evaluating conflicting performance on driver and outcome measures: the effect of strategy maps. *Journal of Management Control*, *23*(2), 97-114.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Mintzberg, H., & Waters, J. A. (1985). Of strategies, deliberate and emergent. *Strategic Management Journal*, 6(3), 257-272.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language*, 26(4), 453-465.

Myerson, R., (1979). Incentive Compatibility and the Bargaining Problem. *Econometrica*, 47:61-73.

Myerson, R., (1982). Optimal Coordination Mechanisms in Generalized Principal-Agent Problems. *Journal of Mathematical Economics*, 10: 67-81.

Nickerson, R. S., & Adams, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology*, *11*(3), 287-307.

Niven, P. R. (2006). *Balanced scorecard step-by-step: Maximizing performance and maintaining results*. 2nd Edition. John Wiley & Sons.

O'Donoghue, T., & Rabin, M. (1999). Doing it now or later. *American Economic Review*, 103-124.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences*, *87*(24), 9868-9872.

Otmakhova, N., Duzel, E., Deutch, A. Y., & Lisman, J. (2013). The hippocampal-VTA loop: the role of novelty and motivation in controlling the entry of information into long-term memory. In *Intrinsically motivated learning in natural and artificial systems* (235-254). Springer Berlin Heidelberg.

Park, S. H., Westphal, J. D., & Stern, I. (2011). Set up for a fall: the insidious effects of flattery and opinion conformity toward corporate leaders. *Administrative Science Quarterly*, *56*(2), 257-302.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of Personality and Social Psychology*, *62*(2), 189.

Phelps, E. S., & Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, *35*(2), 185-199.

Porter, Michael E. "What is strategy?." *Harvard Business Review* (1996).

Prat, C. S., Mason, R. A., & Just, M. A. (2011). Individual differences in the neural basis of causal inferencing. *Brain and Language*, *116*(1), 1-13.

Prendergast, C. (1993). A theory of "yes men". *The American Economic Review*, 757-770.

Raith, M. (2008). Specific knowledge and performance measurement. *The Rand Journal of Economics*, *39*(4), 1059-1079.

Rich, J. S., I. Solomon, and K. T. Trotman. 1997. The audit review process: A characterization from the persuasion perspective. *Accounting, Organizations and Society* 22: 481–505.

Rigby, D., & Bilodeau, B. (2009). Management tools and trends 2009. *Bain & Company Inc*, 1-14.

Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., & Vogeley, K. (2006). Being with virtual others: Neural correlates of social interaction. *Neuropsychologia*, *44*(5), 718-730.

Schlichting, M. L., & Preston, A. R. (2015). Memory integration: neural mechanisms and implications for behavior. *Current Opinion in Behavioral Sciences*, *1*, 1-8.

Schlichting, M. L., & Preston, A. R. (2016). Hippocampal–medial prefrontal circuit supports memory updating during learning and post-encoding rest. *Neurobiology of Learning and Memory*, *134*, 91-106.

Schönberg, T., Daw, N. D., Joel, D., & O'Doherty, J. P. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience*, *27*(47), 12860-12867.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review*, *76*(2), 179.

Singer, M., Halldorson, M., Lear, J. C., & Andrusiak, P. (1992). Validation of causal bridging inferences in discourse understanding. *Journal of Memory and Language*, *31*(4), 507-524.

Stock, W. A., Robinson, D. H., & Pridemore, D. R. (1991). The effects of evidence statements in a passage on the acceptability of bridging inferences. *Contemporary Educational Psychology*, *16*(3), 254-264.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755-769.

Tafkov, I. D. (2012). Private and public relative performance information under different compensation contracts. *The Accounting Review*, *88*(1), 327-350.

Tayler, W. B. (2010). The balanced scorecard as a strategy-evaluation tool: The effects of implementation involvement and a causal-chain focus. *The Accounting Review*, 85(3), 1095-1117.

Towry, K. L. (2003). Control in a teamwork environment—The impact of social ties on the effectiveness of mutual monitoring contracts. *The Accounting Review*, *78*(4), 1069-1095.

Vera-Muñoz, S., J. L. Ho, and C. W. Chow. 2006. Enhancing knowledge-sharing in public accounting firms. *Accounting Horizons* 20 (2): 133-155.

Webb, R. (2004). Managers' commitment to the goals contained in a strategic performance measurement system. *Contemporary Accounting Research*, *21*(4), 925-958.

Wiseman, R. (2008). Quirkology: the curious science of everyday lives. *Pan Macmillan*.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162.