

Strategy Selection, Surrogation, and Strategic Performance Measurement Systems

JONGWOON (WILLIE) CHOI,* GARY W. HECHT,†
AND WILLIAM B. TAYLER‡

Received 24 August 2011; accepted 6 June 2012

ABSTRACT

Strategic performance measurement systems operationalize firm strategy with a set of performance measures. A consequence of such alignment is the tendency for managers to lose sight of the strategic construct(s) the measures are intended to represent, and subsequently act as though the measures *are* the constructs of interest, a phenomenon referred to as *surrogation*. We investigate how involvement in strategy selection affects managers' propensity to exhibit surrogation. We predict and find that strategy selection reduces surrogation. Surprisingly, we do not find that engaging in strategy deliberation, a key process underlying strategy selection, reduces surrogation. Thus, managers' involvement in the actual choice of strategy appears to be both a necessary and sufficient condition to mitigate surrogation. Our paper broadens understanding of factors that influence surrogation, such as the effects

*University of Pittsburgh; †University of Wisconsin–Madison; and ‡Brigham Young University.

Accepted by Merle Erickson. The authors appreciate the helpful comments and suggestions from an anonymous reviewer, as well as from Scott Emett, Jeff Hales, Kathryn Kadous, Michael Majerczyk, Axel Schulz, Chad Simon, Steve Smith, Tom Vance, David Wood, and Donnie Young, as well as participants at the American Accounting Association Annual Meeting, the Accounting, Behavior and Organization Midyear Meeting, Brigham Young University's Accounting Research Symposium, and Emory University's Experiment-Based-Research Brownbag and workshop participants at Georgia Tech, Monash University, and University of Melbourne. We also appreciate the research assistance of Song Wang and Ye Wang. This research benefited from a Dean's Research Grant from the Goizueta Business School at Emory University.

of different aspects of managers' strategic involvement and buy-in. Further, by documenting how managers behave *within* (as opposed to simply *with*) strategic performance measurement systems, we highlight the potential for managers to endogenously influence the effectiveness of such systems.

1. Introduction

Firms often rely on strategic performance measurement systems to facilitate managers' strategic decision-making. By aligning strategic constructs, initiatives, and objectives with related performance measures, these measurement systems function as a framework that organizes the firm's information environment around its strategy. Among many other benefits, this alignment enables managers to make strategic decisions supported by empirical data. However, another consequence of aligning strategy and performance measures is that managers may lose sight of the strategic construct(s) the measures are intended to represent, and subsequently act as though the (imperfect) measures *are* the constructs of interest (Ijiri [1967, 1975], Kaplan and Norton [1996, pp. 218–219], Choi, Hecht, and Tayler [2012]). This phenomenon, referred to as *surrogation*, can hinder management's strategic decisions.

Although surrogation is potentially detrimental in many scenarios, it is particularly so in settings in which strategic performance measurement systems are most beneficial: when strategic constructs are abstract or ill-defined. For instance, suppose a firm pursuing a customer-centric strategy gauges customer satisfaction using satisfaction survey scores. To the extent that those scores imperfectly measure customer satisfaction, managers' strategy-related judgments and decisions may be flawed if they surrogated customer satisfaction with the survey scores. The potential for surrogation to mitigate the effectiveness of strategic performance measurement systems suggests a need for a deeper understanding of the surrogation phenomenon.

We investigate how involvement in strategy selection affects managers' propensity to surrogate strategic constructs with performance measures. Although managers often engage in strategy formulation and selection (Kaplan and Norton [1992, 1996], Kauffman [2010]), research on strategic performance measurement systems generally investigates managerial decision making in settings involving an *assigned* strategy (e.g., Lipe and Salterio [2000], Banker, Chang, and Pizzini [2004]). Likewise, prior work on surrogation does not consider the effects of involving managers in strategy selection. Thus, whether managers' surrogation of strategic constructs generalizes to scenarios in which managers are engaged in the selection of strategy remains an empirical question.

We leverage Kahneman and Frederick's [2002] notion of *attribute substitution* in developing our predictions. Attribute substitution occurs when an individual has a judgment to make regarding a complex "target attribute," and instead relies on a more easily accessible "heuristic attribute"

in his judgment process. A key condition necessary for attribute substitution (and, in turn, for surrogation) to occur is the relative inaccessibility of the target attribute (the strategic construct). Building on research on delegation and participative decision making (e.g., Milani [1975], Nouri and Parker [1998], Covalleski et al. [2003]), we predict that involving managers in strategy selection generates both *motivation* and *information* effects that increase the accessibility of strategic constructs, thereby reducing the likelihood that surrogation occurs. Notably, while the motivation effects relate primarily to involving managers in the selection of strategy, the information effects occur as a result of managers' deliberation of strategy, a key element of managers' strategy selection.

To examine the effects of strategy selection and strategy deliberation on surrogation, we conduct laboratory experiments in which business students use the computer game *Spore* (Electronic Arts [2008], Lawton [2008]) to design a virtual creature in line with a strategy. We employ a nested experimental design with four conditions. First, we vary whether participants receive incentive compensation tied to their creature design. Participants in the flat-wage condition implemented an assigned strategy absent incentive compensation, while participants in the incentive-compensation condition received piece-rate compensation tied to a single measure of their creature's ability. The flat-wage condition allows us to measure the baseline level of surrogation in our setting, providing a benchmark against which to compare the level of surrogation in other conditions. Nested within the incentive-compensation condition, we manipulate participants' strategy involvement at three levels: no involvement, choice, and deliberation. Participants in the no-involvement condition implemented an assigned strategy. Participants in the choice condition chose which strategy they would implement. And participants in the deliberation condition implemented an assigned strategy after providing a list of pros and cons of the assigned strategy.

After designing his or her creature, each participant viewed a series of 14 virtual creatures and made choices regarding the design of these creatures. These choices capture participants' tendency to use the incentivized measure as a surrogate for the implemented strategy. Importantly, these 14 choices are not tied to compensation in any way, allowing us to view behavior absent the influence of incentive compensation. Based on these choices, we compute participants' surrogation score, our key dependent variable. To further explore participants' propensity toward surrogation, we also collect data on participants' recall of strategy approximately one week after the experimental session.

We find evidence consistent with our hypotheses. As expected, we replicate the surrogation-inducing effect of incentive compensation reported by Choi, Hecht, and Tayler [2012]. Further, consistent with attribute substitution theory, involvement in strategy choice reduces the propensity to surrogate. However, we do not find evidence that deliberation of an assigned strategy mitigates surrogation, as participants in the deliberation condition

do *not* surrogate less than participants in the no-involvement condition. Finally, supplemental analyses suggest that the extent to which individuals “buy-in” to an implemented strategy is an important component of the effect of strategy involvement on surrogation. Collectively, these results suggest that involving managers in the actual choice of strategy is both a necessary and sufficient condition to mitigate surrogation.

Our study contributes to management accounting research and practice on multiple dimensions. First, our study advances literature on surrogation in that it provides evidence relating to the surrogation-mitigating effect of managers’ involvement in strategy selection. Because surrogation has the potential to negatively influence managers’ decisions, knowledge regarding factors that reduce surrogation is beneficial to academics and practitioners looking to harness the benefits (and reduce the costs) of aligning performance measures with firm strategy.

Second, our study is important to managers and management accountants who determine and facilitate managerial autonomy. Specifically, a frequently cited mantra of practitioner-oriented literature is the importance for firms to generate employee “buy-in” of organizational strategy and goals (e.g., Cokins [2005]). In a similar vein, extant academic literature focuses on the antecedents and (some) consequences of employees’ goal commitment (e.g., Webb [2004]). There exists less understanding, however, of the underlying process of *how* buy-in and goal commitment translate into enhanced performance. Our results suggest that one benefit is that involvement in strategy selection helps managers maintain a strategic perspective (i.e., surrogate less than if they had no involvement in strategy selection). Further, our results suggest that such a perspective is induced and maintained via selection of the strategy, but not necessarily via strategy deliberation. Thus, our study speaks to how the *type* of strategy-involvement influences managers’ behavior.

Finally, we extend beyond extant literature on strategic performance measurement systems by examining what managers do *within* (as opposed to simply *with*) such systems. More specifically, we expose the potential for managers’ role within the strategic performance measurement system to endogenously influence the system’s effectiveness, and, ultimately, whether the underlying goals of the system are attained. Such considerations are crucial to understanding the evolutionary nature of firms’ strategic performance measurement systems.

Section 2 provides the background and hypotheses. Section 3 describes our experiment and related procedures. Section 4 presents results. Section 5 concludes.

2. *Background and Hypotheses*

2.1 STRATEGIC PERFORMANCE MEASUREMENT SYSTEMS

A central role of accounting measures is to represent constructs of interest (Ijiri [1967, 1975]). For example, GAAP earnings serve to reflect

a firm's income in a given reporting period. Such objectification is necessary because the underlying constructs are often abstract and complex, making them difficult for decision-makers to work with directly (Ijiri [1967, 1975]). Although accounting measures are often imperfect operationalizations of the constructs they are intended to represent (Baker [1992, 2000], Bushman and Indjejikian [1993], Bushman, Indjejikian, and Penno [2000], Prendergast [2002], Budde [2007]), they are still useful media through which decisions related to the constructs can be made.

This representational role of accounting measures is evident in a firm's strategic performance measurement system. In this system, firm strategy is explicitly linked to, and (imperfectly) represented by, a set of performance measures (Kaplan and Norton [1992], Malina, Nørreklit, and Selto [2007]). Representing firm strategy with performance measures fulfills several key functions, such as the communication, evaluation, and development of strategy (Langfield-Smith [1997], Ittner and Larcker [1998, 2001], Chenhall [2003]). Strategically linked performance measures signal the set of organizationally desirable behaviors, and inform employees of how their actions affect others within the firm and the firm as a whole (Malina and Selto [2001, 2004]). In doing so, strategic performance measurement systems facilitate a shared understanding of organizational objectives. Additionally, performance measures serve as tangible evidence useful not only for assessing how well the firm is executing its strategy, but also for empirically testing the strategy (Bromwich [1990], Campbell et al. [2008]). Finally, performance measures foster strategic learning by helping managers discover appropriate ways to refine strategy (Kaplan and Norton [1996]). Key to all of these functions of strategic performance measurement systems is that the measures be "transparent," such that managers can "see through" the measures back to the strategy (Kaplan and Norton [1996]).

2.2 SURROGATION

Representing strategy with performance measures is not without costs. A particular consequence is that employees can myopically focus on performance measures and subsequently act as though the measures *are* the constructs of interest (Ijiri [1967, 1975], Choi, Hecht, and Tayler [2012]). This phenomenon, referred to as *surrogation*, is potentially detrimental, especially in settings where strategic performance measurement systems are most useful. For example, consider a manager responsible for assessing whether a customer-centric strategy is appropriate. If that manager has surrogated the strategic construct of customer satisfaction with customer satisfaction survey scores, and those scores are not positively associated with financial performance (or some other organizational goal), then the manager may conclude that the strategy is a failing one. To the extent that the survey scores are an incomplete or imperfect proxy for customer satisfaction, the manager's inferences may be erroneous.

While not investigating surrogation directly, recent studies offer a psychology-based account for why surrogation occurs (Schkade and

Kahneman [1998], Kahneman and Frederick [2002], Kahneman et al. [2006]). This explanation centers around the notion of *attribute substitution*, in which individuals rely on an easily accessible “heuristic attribute” when making a judgment about a less accessible, complex “target attribute.” Strack, Martin, and Schwarz’s [1988] experimental study provides an apt example of attribute substitution. In these experiments, college students answer two questions: 1) “How many dates did you have last month?” and 2) “How happy are you with your life in general?” The order in which students answered these questions was manipulated between subjects. When students answered the happiness question first, the correlation between their answers to the two questions was negligible. However, when students answered the dating question first, the correlation jumped to 0.66. In this example, the target attribute is overall happiness, a rather ambiguous construct. Consistent with attribute substitution, students appeared to rely on the highly accessible (and quantifiable) number of dates they had last month to a greater extent when they answered the happiness question second than when they answered the happiness question first.

For attribute substitution to occur, three conditions must be met (Kahneman and Frederick [2002, p. 54]). First, the target attribute must be relatively inaccessible. Second, the heuristic attribute must be highly accessible. Third, the substitution of the target attribute with the heuristic attribute must not be consciously rejected.

The extent to which these conditions are met likely depends on contextual factors related to the nature and use of the strategic performance measurement system itself. Choi, Hecht, and Tayler [2012] provide evidence related to one such factor—the nature of measure-based incentive compensation. Specifically, they find that managers’ propensity to exhibit surrogation is greater when incentive pay is tied to a single measure of a strategic construct than when it is tied to multiple measures of a strategic construct. Their theory suggests that, relative to a scenario characterized by flat-wage compensation, the attention-directing effects of incentive compensation heighten the accessibility of the heuristic attribute (i.e., the compensated measures), thereby increasing the likelihood of satisfying the second condition necessary for attribute substitution, and hence surrogation, to occur. However, in conjunction with this attention-directing effect, compensation on *multiple* measures also activates cognitive operations that *decrease* the likelihood that the third necessary condition is met. Ultimately, Choi, Hecht, and Tayler’s [2012] results speak to the effect of the form of incentive compensation on individuals’ propensity to exhibit surrogation, and how this effect plays out via the *second* and *third* attribute-substitution conditions. In the next subsection, we develop predictions regarding managers’ involvement in strategy selection and how this factor influences the *first* attribute-substitution condition, the accessibility of the target attribute (strategic construct).

2.3 STRATEGY SELECTION AND SURROGATION

Prior work on surrogation, as well as much of the prior research on strategic performance measurement systems (e.g., see Lipe and Salterio [2000] and Banker, Chang, and Pizzini [2004]), focus on settings in which managers make decisions regarding an assigned strategy. However, managers' interaction with strategy often extends beyond mere implementation to include strategy formulation and selection (Giles [1991], Kaplan and Norton [1992, 1996], Sterling [2003], Drake, Wong, and Salter [2007], Malshe and Sohi [2009], Kauffman [2010]). This raises the natural question of whether managers' surrogation of strategic constructs generalizes to settings in which managers are involved in strategy selection.

We investigate how involvement in strategy selection affects surrogation within a setting where incentive compensation is tied to a single measure of the strategic construct. We use this setting for two reasons. First, research shows that the adoption of participative management practices such as involving managers in strategy selection often goes hand-in-hand with incentive compensation (Athey and Roberts [2001], Nagar [2002], Prendergast [2002], Bester and Kräbmer [2008], Devaro and Kurtulus [2010]). Thus, our investigation of the effects of managers' involvement in strategy selection incorporates features common to our setting of interest. Second, a prerequisite of an investigation of how involvement in strategy selection affects surrogation is that managers exhibit some degree of surrogation. The results of Choi, Hecht, and Tayler [2012] suggest that this requirement is likely to be satisfied when managers' compensation is tied to a single measure of a strategic construct. Our first hypothesis predicts a replication of this result. Specifically, we predict that incentive compensation tied to a single measure of a strategic construct increases surrogation relative to flat-wage compensation.

H1: Managers surrogate the strategic construct more when compensated on a single measure of a strategic construct than when compensated with a flat wage.

Our remaining two hypotheses pertain to managers' involvement in strategy selection. Research on participative decision-making identifies two benefits that can arise when managers are involved in strategy selection. The first is an information effect (Becker and Green [1962], Miller and Monge [1986], Kren [1992], Covaleski et al. [2003]). That is, when managers are involved in the selection of strategy, they acquire information that helps them understand the differential costs and benefits of pursuing the possible strategies. In determining which strategy to implement, managers use this information to compare and contrast the various strategies. In contrast, managers not involved in the selection of strategy (e.g., those for whom the strategy is assigned) do not acquire (or benefit from) strategy-oriented information to the same degree.

The second benefit of manager involvement in strategy selection is a motivation effect. Involvement in strategy selection imbues managers with a sense of ownership (Shields and Shields [1998], Becker and Huselid [1999], Chen and Jones [2009]). In turn, involved managers identify with and internalize organizational objectives (Milani [1975], Levine and Tyson [1990], Coleman [1996], Menon [2001]), and their commitment to organizational goals is greater than when there is no strategy selection involvement (Kearney and Hays [1994], Nouri and Parker [1998], Liden, Wayne, and Sparrowe [2000]). By adopting the firm's perspective to a greater degree, managers involved in the selection of strategy may maintain a greater focus on the strategic construct while implementing the chosen strategy than those not involved in the strategy-selection process.

In terms of attribute substitution, the information and motivation effects decrease the likelihood that the first attribute-substitution condition is met by heightening the accessibility of the strategic construct (the target attribute). Thus, building on our first hypothesis, we predict that allowing managers to choose strategy will decrease the surrogation induced by incentive compensation. Formally, we express Hypothesis 2 as follows:

H2: Managers surrogate the strategic construct less when they are allowed to choose strategy than when they are assigned a strategy.

A key distinction between the information and motivation effects described above is that the latter relate primarily to managers' involvement in strategy choice, while the former could be driven by choice or by mere strategy deliberation, a key process underlying strategy selection. Disentangling these two effects is difficult, as strategy choice typically implies some level of deliberation, while strategy deliberation often (but not always) involves choice, even if that choice is not formally implemented (e.g., an individual may deliberate and make a private choice without actually outwardly executing a formal choice). We provide a preliminary investigation of this partially separable construct. Specifically, we test the hypothesis that deliberating on a strategy's pros and cons reduces surrogation, even absent the ability to choose the strategy. Formally, we express Hypothesis 3 as follows:

H3: Managers surrogate the strategic construct less when they deliberate on strategy than when they do not deliberate on strategy.

3. Method

We test our hypotheses via three experiments. For all three experiments, we adapt Choi, Hecht, and Tayler's [2012] experimental instrument. Participants used computers to work through self-guided instructions and to perform the tasks described below. We recruited participants from courses at a *BusinessWeek* top five business school. One hundred fifty-seven students participated in Experiment 1 (average age = 20 years, 64% male); 78 students

participated in Experiment 2 (average age = 20 years, 60% male); 92 students participated in Experiment 3 (average age = 19, 55% male).

3.1 STRATEGY IMPLEMENTATION VIA CREATURE DESIGN

We use the computer game *Spore* (Electronic Arts [2008], Lawton [2008]) as the platform for the experimental tasks. Instructions indicated that:

The purpose of the game of *Spore* is to create a creature that interacts with other life forms in a way that enables it to become the dominant species. You can become the dominant species by making friends with other species, or by overpowering them.

After reading the instructions, participants designed a virtual creature using the *Creature Creator* module of *Spore*, in accordance with one of the two available domination strategies:

Attack Strategy:	<i>Design a creature that can attack other life forms such that it will become the dominant species on the planet.</i>
Socialize Strategy:	<i>Design a creature that can socialize with other life forms such that it will become the dominant species on the planet.</i>

Figure 1 presents a screenshot of this creature-design phase of the experiment. In designing their creature, participants could purchase various body parts using an endowment of 500 “DNA points.” The 500-DNA-point budget provided participants with the resources necessary to design a creature in line with the strategy, but forced participants to select from among multiple ways of implementing their strategy. Each body part available for purchase had a posted DNA-point cost, as well as a list of abilities the body part would add to the creature. Participants could add parts to their creatures as long as their DNA-point budget was not exhausted. Parts could also be removed for a full DNA-point refund at any time.

Participants saw real-time updates of measures of their creature’s abilities as they added and removed parts to and from their creature. Among these were measures tied to their creature’s ability to “attack” (the strategic construct of interest for those implementing the attack strategy) and to “socialize” (the strategic construct of interest for those implementing the socialize strategy). For the strategic construct of attack, the measures reflected the creature’s ability to *bite*, *charge*, *spit poison*, and *strike*. For the strategic construct of socialize, the measures captured the creature’s ability to *sing*, *dance*, *charm*, and *pose*. All of these measures were quantified on a scale of 0 to 5, where 0 (5) reflects the lowest (highest) possible ability level. We informed participants that, in *Spore*, each specific attack (socialize) ability contributes equally to a creature’s *overall* ability to attack (socialize).

To hold constant the incentive to design a creature in line with the strategy across all conditions, we based a portion of participants’ compensation on a postexperiment assessment of their creature’s ability to achieve global



FIG. 1.—Screenshot of Creature Design. Above is a sample screenshot of the *Creature Creator* module of the computer game *Spore*. The left side of the screen displays various body parts that can be added to the creature. The upper-right corner of the screen displays various measures of the creature’s abilities. The second row of boxes on the upper right displays the four attack-ability measures: *bite* (level = 5), *charge* (level = 5), *spit* (level = 3), and *strike* (level = 1). The third row of boxes displays the four social-ability measures: *sing* (level = 1), *dance* (level = 2), *charm* (level = 4), and *pose* (level = 2). During the experiment, these measures were constantly available and adjusted real-time as participants added and removed parts from their creatures. “DNA points,” used as currency to purchase parts, are displayed in the bottom-left corner of the screen (currently DNA = 45). Buttons at the top center of screen allow participants to paint their creature and see their creature move in a simple environment.

domination. To calculate this compensation, a research assistant placed each participant’s creature in a *Spore* environment after the experiment had ended, and gave it commands to either attack or socialize with neighboring creatures (depending on the implemented strategy). The research assistant was not aware of the experimental conditions or the purpose of the study. The time to dominate (either by overpowering or by befriending) two computer-based creatures was recorded. On average, the higher a creature’s total attack (socialize) ability level (i.e., the sum of the four attack (socialize) ability levels), the faster the creature dominates other creatures.

To assure participants that they did not have to be experts at the game of *Spore* to receive the maximum compensation, we based this portion of their compensation on *relative* ability, rather than an absolute measure. Specifically, within each experimental condition, we matched each participant

with nine other randomly selected participants who implemented the same global-domination strategy. The top performer in each group (i.e., the participant whose creature required the least amount of time to dominate two creatures) earned \$20, the second best performer received \$18, and so on, down to \$2 for the worst performer.

3.2 INDEPENDENT VARIABLES

We employ a nested experimental design with four conditions. First, we manipulate whether participants receive incentive compensation tied to their creature design. In the flat-wage condition, participants received \$10 regardless of their creature's attack or social abilities. In contrast, participants in the incentive-compensation condition received \$2 per level of their creature's *bite* ability (if they implemented the attack strategy), or \$2 per level of their creature's *sing* ability (if they implemented the socialize strategy). For example, if a participant implementing the attack (socialize) strategy designed a creature with *bite* (*sing*) level of 4, then she received \$8.

Within the incentive-compensation condition, we manipulate participants' strategy involvement at three levels. Participants in the no-involvement condition implemented an assigned strategy, similar to participants in the flat-wage condition. Participants in the choice condition chose which strategy they would implement, and provided a list of pros and cons of the strategy. Participants in the deliberation condition were assigned a strategy to implement, and provided a list of the pros and cons of the strategy.

In the three conditions in which participants implement an assigned strategy (i.e., the flat-wage, deliberation, and no-involvement conditions), we randomly assigned half of the participants to the attack strategy and the other half to the socialize strategy. Doing so helps balance for the strategy implemented across conditions. Further, a 50–50 split corresponded to strategy choices made in pilot testing, and is not far from the actual distribution of choices in the strategy choice condition (approximately 63% chose attack and 37% chose socialize). However, distribution matching does not guarantee that those participants who *ex ante* would have preferred a given strategy were assigned that strategy. This relates to the question of participants' buy-in of each strategy (Cokins [2005], Tayler [2010]), which we discuss further in our supplemental analyses.

3.3 DEPENDENT VARIABLES

We elicited our primary dependent measure after participants finished designing their virtual creature. Participants viewed 14 successive pictures of other virtual creatures. With each of these creatures, participants saw two sets of potential realizations of the four ability measures corresponding to the strategy they implemented (*bite*, *charge*, *spit*, and *strike* for participants implementing the attack strategy, and *sing*, *dance*, *charm*, and *pose* for the participants implementing the socialize strategy). We labeled these potential realizations "modification packages." For each creature, we asked

participants to select the modification package that they believed would best implement the strategy. In addition, we told participants to assume that the DNA cost was equal across each pair of modification packages. Participants made their modification-package selections as they viewed each creature, and could not go back and change earlier selections. To control for an order effect, we created four randomly determined orders for the 14 pairings, and fully crossed them with our experimental conditions. Since results are unaffected by the order in which choices are made, we ignore this variable in subsequent analyses and discussion.

We code each individual modification-package selection as either consistent (1) or inconsistent (0) with surrogation with the measure on which the participant received incentive compensation during the initial creature design task (*bite* or *sing*). To illustrate, consider the sample pairing presented in figure 2, panel A, which includes the following modification-package pairing:

- Modification #1: *bite* = 2, *charge* = 4, *spit* = 1, and *strike* = 5
- Modification #2: *bite* = 3, *charge* = 2, *spit* = 4, and *strike* = 2

In this pairing, modification package #2 maximizes the measure *bite*. Thus, a selection of modification package #2 (modification package #1) would be coded as 1 (0).

Using this coding scheme, we calculate each participant’s “surrogation score,” which captures the proportion of times a participant chooses the modification package that maximizes the initially incentivized measure. Two of the 14 pairings serve as distracter tasks, and are not included when calculating participants’ surrogation score. In these two pairings, the modification package with the higher level of the compensated measures also has a higher total across all four ability measures, confounding surrogation and optimal modification-package selection. Table 1 presents the 14 modification-package pairings.

We create three variations of participants’ surrogation score. The *comprehensive* surrogation score is based on the full set of 12 modification-package selections. We use a subset of six pairings to construct an *opportunity cost* surrogation score. In these six pairings, selecting the surrogation-consistent modification package reflects an opportunity cost because the alternative modification package offers a higher total ability, which is a key determinant of a creature’s ability to attack or socialize. To illustrate, consider again the pairing shown in figure 2, panel A. While modification package #2 maximizes the *bite* level (*bite* = 2 for modification package #1; *bite* = 3 for modification package #2), it also has a lower total attack-ability level (attack measures sum to 12 for modification package #1, but only sum to 11 for modification package #2). Therefore, selecting the surrogation-consistent modification package in this pairing forgoes higher total attack-ability. This pairing, and the five pairings like it, comprise the subset of modification-package selections used to construct the *opportunity cost* surrogation score.

Panel A:

TASK 2

Please click on the check-box under the 'modification package' that you prefer for this creature.

Note that, regardless of the picture shown, the modification package you select represents the final set of abilities the creature will have. Also, assume that both modification packages cost the same, and that you have sufficient DNA points for one package, but not both.

<u>Modification #1</u>		<u>Modification #2</u>
Bite: 2		Bite: 3
Charge: 4		Charge: 2
Spit: 1		Spit: 4
Strike: 5		Strike: 2
Choose <input type="checkbox"/> # 1		Choose <input type="checkbox"/> # 2
<input type="button" value="NEXT CHOICE"/>		

Panel B:

TASK 2

Please click on the check-box under the 'modification package' that you prefer for this creature.

Note that, regardless of the picture shown, the modification package you select represents the final set of abilities the creature will have. Also, assume that both modification packages cost the same, and that you have sufficient DNA points for one package, but not both.

<u>Modification #1</u>		<u>Modification #2</u>
Sing: 5		Sing: 1
Dance: 2		Dance: 4
Charm: 4		Charm: 5
Pose: 2		Pose: 3
Choose <input type="checkbox"/> # 1		Choose <input type="checkbox"/> # 2
<input type="button" value="NEXT CHOICE"/>		

FIG. 2.—Sample Modification-Package Choices. Above are two samples of the 14 pairings used to generate participants' surrogation scores, the primary dependent variable. Panel A (panel B) is a sample pairing shown to participants who implemented the attack (socialize) strategy. Each pairing was the same across strategies, with the exception of the measure-labels. For each

We use a different subset of six pairings to construct a *no opportunity cost* surrogation score. For these pairings, selecting the surrogation-consistent modification package does not involve sacrificing a higher total ability level, because the total ability level is equal for both modification packages. An example of a pairing used to construct the *no opportunity cost* surrogation score is displayed in panel B of figure 2. While modification package choice #1 maximizes *sing* ($sing = 5$ for modification package #1; $sing = 1$ for modification package #2), both modification package pairings have the same total social-ability level (the measures sum to 13 for both modification packages). We report the opportunity cost status for each modification-package pairing in the last column of table 1.

In addition to participants' surrogation scores, we also measure surrogation using survey responses collected approximately one week after the experimental session. As explained in greater detail later, the survey tested both recall of the strategy itself, as well as the strength of word associations that provide additional insights into the mental representations participants have of both the strategy they implemented and of measures related to their strategy.

Importantly, we use neither the "time to dominate" measure nor the related tournament-based compensation described earlier as our dependent variable, because these measures may simply capture participants' desire to maximize wealth. Participants' surrogation scores and their survey responses allow us to measure the extent to which participants surrogated the strategy with measures without the confounding influence of wealth maximization.

3.4 SESSION TIMELINE

As participants arrived at a session, they were randomly assigned to one of the experiment conditions. Participants signed a consent form, then read computer-based instructions, then watched a video tutorial on using the *Creature Creator* module of *Spore*. After the tutorial, participants

FIG. 2.—*Continued.* pairing, participants viewed a picture of a creature, along with two modification packages. Participants were asked to select the modification package that they believed, if added to the creature, would best implement the strategy. After making their selection, participants clicked "Next Choice" to move on to the next creature and modification-package pairing. Note that the choice displayed in panel A speaks to a participant's propensity to surrogate with the *bite* measure. That is, a choice of modification package #2 is consistent with favoring a higher *bite* measure of 3 (versus 2 for modification package #1). This modification-package pairing also represents an opportunity cost of surrogating with *bite*, as the surrogation-consistent modification package does not maximize the total value of the attack-ability measures (*bite*, *charge*, *spit*, and *strike* levels sum to 12 for modification package #1, but only sum to 11 for modification package #2). The choice displayed in panel B above speaks to a participant's propensity to surrogate with the *sing* measure. However, this pairing does not entail an opportunity cost as the surrogation-consistent modification package has the same total value of the social-ability measures (*sing*, *dance*, *charm*, and *pose*) as does the alternative modification package (levels sum to 13 for both modification packages).

TABLE 1
Modification-Package Pairings

Pairing	Surrogation-consistent modification package ^a				Alternative modification package ^a				Pairing used to evaluate opportunity cost? ^b
	Bite/ Sing	Charge/ Dance	Spit/ Charm	Strike/ Pose	Bite/ Sing	Charge/ Dance	Spit/ Charm	Strike/ Pose	
1	4	3	5	1	3	5	3	2	No
2	5	2	4	2	1	4	5	3	No
3	3	1	5	2	1	4	1	5	No
4	5	2	4	1	2	4	1	5	No
5	3	2	5	1	1	5	3	4	Yes
6	5	2	3	1	4	1	4	3	Yes
7	4	4	2	2	1	3	5	5	Yes
8	3	2	4	2	2	4	1	5	Yes
9	4	2	3	4	1	3	5	4	No
10	3	1	4	1	2	5	1	1	No
11	3	3	2	1	2	1	5	3	Yes
12	5	2	1	3	2	4	2	5	Yes
13 ^c	5	3	4	4	4	2	2	3	N/A
14 ^c	3	4	5	3	2	3	2	2	N/A

This table lists the 14 modification-package pairings provided to participants. Though numbered in the table, pairings were not numbered in the study. Pairings were presented in four, predetermined, random orders (order is fully crossed with the compensation manipulation).

^aThe column labeled “Surrogation-consistent modification package” reports the measure realizations for *bite* (*sing*), *charge* (*dance*), *spit* (*charm*), and *strike* (*pose*) of the modification package within each pairing that is consistent with surrogation on *bite* (*sing*). The column labeled “Alternative modification package” reports the alternative to the surrogation-consistent option. Though not shown in the table, modification packages are labeled for each pairing as “Modification #1” or “Modification #2” (see figure 2). In addition, the label on the surrogation-consistent modification package is randomly determined, and is held constant for all participants.

^bThe last column indicates whether or not there is an opportunity cost to selecting the surrogation-consistent modification package. In *no opportunity cost* pairings, the total level of the four ability measures is equal across the two modification packages. Thus, a participant who chooses the surrogation-consistent modification package does not sacrifice a higher total level in order to do so. In contrast, in *opportunity cost* pairings, the surrogation-consistent modification package has a lower total level than the alternative modification package.

^cPairings #13 and #14 are not used in any test of surrogation. These pairings are included in the experiment as distracters. In both pairings, the modification package with higher *bite/sing* also has a higher total level, confounding surrogation and optimal modification-package selection.

practiced building a virtual creature for 10 minutes. Next, participants read additional instructions describing the two global domination strategies and how participants would be compensated. Participants next took a short quiz to ensure they understood the instructions, and proceeded to the creature-design phase of the experiment. Participants then viewed the 14 modification-package pairings and made their selections for each pairing. Finally, participants completed a questionnaire eliciting process-related and demographic information. Approximately one week after all sessions were completed, participants filled out a brief postexperiment survey, were debriefed, and received payment.

3.5 DIFFERENCES BETWEEN EXPERIMENTS

Experiment 1 includes our full experiment design. Experiment 2 replicates two of the conditions from Experiment 1 (the no-involvement

condition and the choice condition), but asks two additional questions as part of debriefing intended to help better understand our results. Specifically, we asked participants in Experiment 2 to rate 1) the relative effectiveness of the two global-domination strategies, and 2) how committed they were to the strategy they implemented. Experiment 3 also replicates two of the conditions from Experiment 1 (the flat-wage condition and the no-involvement condition), but includes additional wording as part of the compensation manipulation intended to help rule out an alternative explanation for the results from Experiment 1.

3.6 ADDITIONAL DISCUSSION OF THE EXPERIMENT DESIGN

Multiple issues relating to our experiment design warrant further discussion. First, we explicitly chose *not* to provide any form of compensation for participants' modification-package selections. Doing so would introduce wealth maximization as a potential confound, and preclude us from observing the phenomenon of interest. Second, we used the computer game *Spore*, which is far removed from most participants' practical and educational experiences, to minimize the effects of participants' priors about potential surrogates for strategic constructs. Importantly, the setting still allows participants to "get their hands dirty" selecting and implementing the strategy and to deal with various associated measures of strategic constructs. Third, we took several steps to maximize the internal validity of our study. To ensure that the strategy statement was salient, we stated the strategy multiple times, with the last instance occurring immediately before the creature-design phase of the experiment. We also included in the task instructions an explicit statement that all four measures of attack-ability or social-ability are equally important in implementing the strategy and achieving global domination. Fourth, participants had to demonstrate thorough understanding of the various aspects of our research design via a quiz. Specifically, participants could not proceed to the creature-design phase of the experiment until they had demonstrated their knowledge regarding the strategy they were implementing (attack or socialize) and the fact that all attack (social) abilities were equally important to their ability to implement the attack (socialize) strategy. Finally, we use flat-wage participants' modification-package selections as a benchmark against which we compare the level of surrogation in the incentive compensation conditions. Participants' choices in the flat-wage condition capture participants' natural tendency to exhibit surrogation in our setting. Further, by including the flat-wage condition in our experimental design, we are able to draw inferences about the effects of our independent variables on surrogation using a manageable number of modification-package selections. Absent the flat-wage condition, drawing such inferences would require us to counter-balance myriad combinations of measure-values within modification-package selections, which would impose a substantial burden on participants.

4. Results

4.1 PRELIMINARY ANALYSES AND MANIPULATION CHECK

Before reporting the results of our hypothesis tests and related supplemental analyses, we examine three issues. First, we examine participants' prior experience with *Spore*. As previously noted, we conduct our investigation using a setting unfamiliar to participants in order to reduce their tendency to apply preconceived notions about potential surrogates for strategic constructs. Participants' prior experience with *Spore* threatens the internal validity of our study, as their behavior is likely influenced by factors other than our independent variables of interest. For this reason, we ask participants in the postexperimental questionnaire if they had ever played *Spore*. We drop participants reporting previous *Spore* experience, leaving a total of 144 participants in Experiment 1, 70 participants in Experiment 2, and 83 participants in Experiment 3.

Second, because Experiments 2 and 3 replicate conditions in Experiment 1, we investigate whether it is appropriate to collapse across the three experiments for purposes of our analyses. As expected, comparisons of the conditions in Experiments 1 and 2 do not indicate differences across experiments (two-tailed $p > 0.10$ for all primary dependent measures). Thus, for simplicity of our discussion of analyses, we collapse across these two experiments and present results at this aggregate level. However, we control for experiment session to minimize noise due to any experiment-level differences. This control is not significant in any analysis, and all results reported below are inferentially identical if we exclude this control. A comparison of Experiments 1 and 3 indicates only a few differences across experiments for some of our dependent measures.¹ However, because Experiment 3 includes substantial changes to the wording of the compensation manipulation, we report the results of Experiment 3 separately (see section 4.2.2). A summary of our experimental design and the conditions included in each experiment is depicted in figure 3.

Finally, we assess whether our compensation manipulation was successful. First, regarding the initial creature-design task, we asked each participant to "indicate how much attention you gave to the [*bite (sing)*, *charge (dance)*, *spit (charm)*, and *strike (pose)*] abilities, by allocating 200 points among the four abilities. Allocate more points to abilities to which you gave more attention." Table 2 reports the mean number of points allocated to each ability across conditions. Consistent with expectations, pairwise comparisons indicate that participants earning incentive compensation on the initial creature-design task allocated more points (i.e., paid more attention) to the incentivized measure than participants in the flat-wage condition

¹ For example, we find that participants' *comprehensive* and *no opportunity cost* surrogation scores are different across Experiments 1 and 3 (one-tailed p -values are 0.07 and 0.10, respectively).

Condition:	Flat-Wage	No-Involvement	Choice	Deliberation
Incentive Compensation in Initial Creature-Design Task? ^a	No	Yes	Yes	Yes
Participant Chooses Strategy? ^b	No	No	Yes	No
Participant Provides List of Pros and Cons of Implemented Strategy? ^c	No	No	Yes	Yes
Number of Participants in Experiment 1	39	40	39	39
Number of Participants in Experiment 2 ^d	0	39	39	0
Number of Participants in Experiment 3 ^e	44	48	0	0

FIG. 3.—Experiment Design. ^aIn the creature-design phase of the experiment, participants received either flat-wage compensation or incentive compensation on a measure of their creature’s ability. In the flat-wage condition, participants received \$10 regardless of their creature’s ability. In the no-involvement, choice, and deliberation conditions, participants received \$2 per level of either *bite* or *sing*, depending on the strategy implemented (attack or socialize). For example, if a participant implementing the attack (socialize) strategy designed a creature with *bite* (*sing*) level 4, then she received \$8.

^bParticipants were all presented with two possible strategies to implement in the game of *Spore*: the “attack strategy” and the “socialize strategy.” In the choice condition, participants chose which strategy they would implement. In all other conditions, participants implemented an assigned strategy.

^cParticipants in the choice and deliberation conditions provided a list of pros and cons of the implemented strategy.

^dExperiment 2 replicates the no-involvement and choice conditions from Experiment 1 with changes to debriefing questions (see section 3).

^eExperiment 3 replicates the flat-wage and no-involvement conditions from Experiment 1 with changes to the explanation given to participants regarding their compensation (see section 3).

($p < 0.01$ in all three comparisons). Second, we examine the abilities of the creatures that participants designed. Participants earning incentive compensation on the initial creature-design task designed creatures with higher levels of the compensated measure than participants in the flat-wage condition ($p < 0.05$ in all three comparisons). Collectively, these results suggest that our compensation manipulation was successful.

4.2 PRIMARY TESTS OF HYPOTHESIS I: INCENTIVE COMPENSATION

Table 3 presents mean surrogation scores across conditions (panel A) and hypothesis test results (panel B). Hypothesis 1 states that tying compensation to a measure of a strategic construct increases managers’ propensity to surrogate the strategic construct with the measure, relative to when flat-wage compensation is provided. To test this hypothesis, we compare

TABLE 2
Responsiveness to Incentive Compensation Manipulation
Experiments 1 and 2 (Pooled)

Panel A: Mean Creature Design Point Allocations^a [Standard Deviation]				
Condition ^b				
Ability Measures	Flat-Wage (<i>n</i> = 39)	No-Involvement (<i>n</i> = 79)	Choice (<i>n</i> = 78)	Deliberation (<i>n</i> = 39)
<i>Bite (Sing)</i>	54.6 [32.2]	85.7 [35.5]	74.3 [27.9]	87.1 [29.4]
<i>Charge (Dance)</i>	46.5 [22.7]	36.4 [20.0]	42.4 [14.5]	37.9 [17.3]
<i>Spit (Charm)</i>	51.5 [38.9]	39.2 [24.0]	37.0 [20.1]	37.3 [19.2]
<i>Strike (Pose)</i>	47.4 [25.2]	38.8 [19.2]	36.4 [20.4]	46.2 [20.0]

Panel B: Statistical Tests				
	Estimate	<i>t</i>	<i>p</i> -Value ^c	
Points Allocated to <i>bite/sing</i> : Flat-Wage vs. No-Involvement	31.1	4.89	<0.001	
Points Allocated to <i>bite/sing</i> : Flat-Wage vs. Choice	19.7	3.24	0.001	
Points Allocated to <i>bite/sing</i> : Flat-Wage vs. Deliberation	32.5	4.39	<0.001	

^aCreature Design Point Allocations represent participants' allocations of exactly 200 points across the four ability measures (*bite*, *charge*, *spit*, and *strike* for the attack strategy, or *sing*, *dance*, *charm*, and *pose* for the socialize strategy), when asked to "indicate how much attention you gave to the [each ability], by allocating 200 points among the four abilities. Allocate more points to abilities to which you gave more attention."

^bPlease see figure 3 (and section 3) for a detailed description of the four experiment conditions. Reported conditions and statistics combine data from Experiments 1 and 2.

^cAll *p*-values are reported on a one-tailed basis, given the expected directional influence of incentive compensation.

participants' surrogation scores in the flat-wage condition to the surrogation scores of participants in the no-involvement condition.

We begin with an examination of the *comprehensive* surrogation scores. The mean *comprehensive* surrogation score is 0.36 in the flat-wage condition, and is 0.53 in the no-involvement condition. The difference of 0.17 is significant ($t = 3.73$, $p < 0.001$).

These results are robust to differentiating surrogation scores according to opportunity cost. The mean *opportunity cost* surrogation score is 0.20 in the flat-wage condition, and is 0.38 in the no-involvement condition. The difference of 0.18 is significant ($t = 3.00$, $p = 0.002$). In addition, the mean *no opportunity cost* surrogation score is 0.52 in the flat-wage condition, and is 0.69 in the no-involvement condition. The difference of 0.17 is significant ($t = 3.00$, $p = 0.002$).

These results support our hypothesis that tying compensation to a measure of a strategic construct increases the propensity to surrogate that strategic construct with the compensated measure. Importantly, a simple propensity toward wealth-maximization cannot explain these results because participants were not paid to maximize specific measures during the phase of the experiment in which surrogation scores were generated.

4.2.1. Additional Tests of Hypothesis 1: Experiment 3 Analyses. An alternative explanation for results relating to H1 is that, despite instructions to the contrary, participants who received incentive compensation inferred that

TABLE 3
Tests of Hypotheses: Surrogation Score
Experiments 1 and 2 (Pooled)

Panel A: Cell Means [Standard Deviations]	Condition ^a			
	Flat-Wage (<i>n</i> = 39)	No-Involvement (<i>n</i> = 79)	Choice (<i>n</i> = 78)	Deliberation (<i>n</i> = 39)
Surrogation Scores				
Surrogation Score: <i>Comprehensive</i> ^b	0.36 [0.15]	0.53 [0.25]	0.48 [0.22]	0.51 [0.25]
Surrogation Score: <i>No Opportunity Cost</i> ^c	0.52 [0.23]	0.69 [0.28]	0.68 [0.25]	0.68 [0.29]
Surrogation Score: <i>Opportunity Cost</i> ^d	0.20 [0.19]	0.38 [0.34]	0.28 [0.27]	0.34 [0.34]

Panel B: Hypothesis Tests	Difference	<i>t</i>	<i>p</i> -Value ^e
H1: Flat-Wage vs. No-Involvement			
<i>Comprehensive</i>	0.17	3.73	< 0.001
<i>No Opportunity Cost</i>	0.17	3.00	0.002
<i>Opportunity Cost</i>	0.18	3.00	0.002
H2: No-Involvement vs. Choice			
<i>Comprehensive</i>	-0.05	-1.49	0.069
<i>No Opportunity Cost</i>	-0.01	-0.36	NS
<i>Opportunity Cost</i>	-0.10	-1.96	0.026
H3: No-Involvement vs. Deliberation			
<i>Comprehensive</i>	-0.02	-0.76	NS
<i>No Opportunity Cost</i>	-0.01	-0.31	NS
<i>Opportunity Cost</i>	-0.04	-0.89	NS

^aPlease see figure 3 (and section 3) for a detailed description of the four experiment conditions. Reported conditions and statistics combine data from Experiments 1 and 2.

^b*Comprehensive* surrogation scores capture the percentage of times participants choose the modification package that maximizes the incentivized measure (*bite* or *sing*).

^c*No Opportunity Cost* surrogation scores capture the percentage of times participants choose the modification package that maximizes the incentivized measure (*bite* or *sing*) when the total level does not differ between modification packages.

^d*Opportunity Cost* surrogation scores capture the percentage of times participants choose the modification package that maximizes the incentivized measure (*bite* or *sing*) when such a choice means sacrificing a higher total level.

^eAll *p*-values are presented on a one-tailed basis, given the directional predictions for H1, H2, and H3.

the incentivized measure was particularly helpful in achieving global domination. Accordingly, surrogation scores may simply capture a response to the incentive compensation scheme for the creature-design task and not surrogation.

To address this alternative explanation, we conducted a third experiment in which we modify the wording of our compensation manipulation. Specifically, rather than simply telling participants how they would be compensated based on what condition they were randomly assigned to (as in Experiment 1), we told all participants in Experiment 3 the following regarding the manipulated compensation:

Some individuals will receive \$10 for designing their creature. For other participants, the amount of this additional compensation will depend on

the competence level of a specific, compensated ability. They will receive \$2 for each incremental level of the compensated ability their creature possesses. Different participants will have a different compensated ability; but the compensated ability will be one of the following: sing, bite, dance, charge, charm, spit, pose, or strike.

Similar to Experiment 1, we then told participants in the flat-wage condition “You have been randomly selected to receive \$10 for designing your creature” and participants in the no-involvement condition “You have been randomly selected to receive \$2 for each incremental ‘[*measure*]’ level your creature possesses.” To avoid deception, we randomly assigned six participants to be compensated based on the *dance*, *charm*, *pose*, *charge*, *spit*, and *strike* measures. However, because our dependent variable is structured to measure surrogation on either the *sing* or *bite* measures, we drop these six participants from our analyses. Telling participants about other forms of incentive compensation in the study helps to alleviate concerns that participants believed that their personal form of incentive compensation was informative as to the best way to achieve domination in *Spore*. Further, providing flat-wage participants with the same information as participants who received incentive compensation ensures that information is held constant across all conditions, and that significant differences are driven solely by the form of compensation (i.e., whether it is flat-wage or measure-based) and not by its information content.

Table 4 presents mean surrogation scores across conditions (panel A) and Hypothesis 1 test results (panel B). We begin our additional analysis of H1 with an examination of the *comprehensive* surrogation scores in Experiment 3. The mean *comprehensive* surrogation score is 0.36 in the flat-wage condition, and is 0.46 in the no-involvement condition. The difference of 0.10 is significant ($t = 2.30$, $p = 0.012$).

These results are robust to differentiating surrogation scores according to opportunity cost. The mean *opportunity cost* surrogation score is 0.17 in the flat-wage condition, and is 0.28 in the no-involvement condition. The difference of 0.11 is significant ($t = 2.02$, $p = 0.023$). Finally, the mean *no opportunity cost* surrogation score is 0.55 in the flat-wage condition, and is 0.64 in the no-involvement condition. The difference of 0.09 is significant ($t = 1.61$, $p = 0.056$). Ultimately, these results suggest that our support for H1 is robust to clarifications in the incentive compensation manipulation, and thus, not a function of unintentional signals related to incentives.

4.3 PRIMARY TESTS OF HYPOTHESIS 2: STRATEGY CHOICE

Hypothesis 2 predicts that allowing managers to choose strategy will decrease their tendency to surrogate the strategy with associated measures. Operationally, this suggests that surrogation scores will be lower in the choice condition than in the no-involvement condition.

TABLE 4
Test of Hypothesis 1
Experiment 3

Panel A: Cell Means [Standard Deviations]		Condition ^a	
	Flat-Wage ($n = 41$)	No-Involvement ($n = 42$)	
Surrogation Scores			
Surrogation Score: <i>Comprehensive</i> ^b	0.36 [0.16]	0.46 [0.23]	
Surrogation Score: <i>No Opportunity Cost</i> ^c	0.56 [0.22]	0.64 [0.27]	
Surrogation Score: <i>Opportunity Cost</i> ^d	0.17 [0.19]	0.29 [0.31]	

Panel B: Test of Difference			
Flat-Wage vs. No-Involvement	Difference	t	p -Value ^e
<i>Comprehensive</i>	0.10	2.31	0.012
<i>No Opportunity Cost</i>	0.17	2.02	0.023
<i>Opportunity Cost</i>	0.18	1.61	0.055

^aThe manipulated conditions in this additional experiment map to the flat-wage condition and the no-involvement condition in our main experiment, with one difference: in our supplemental experiment, all participants were made aware that other participants were compensated on different measures or on a fixed basis. Please see figure 3 and section 3 for a description of our experiment conditions, and see section 3.5 for a description of the difference between Experiment 1 and Experiment 3.

^b*Comprehensive* surrogation scores capture the percentage of times participants choose the modification package that maximizes the incentivized measure (*bite* or *sing*).

^c*No Opportunity Cost* surrogation scores capture the percentage of times participants choose the modification package that maximizes the incentivized measure (*bite* or *sing*) when the total level does not differ between modification packages.

^d*Opportunity Cost* surrogation scores capture the percentage of times participants choose the modification package that maximizes the incentivized measure (*bite* or *sing*) when such a choice means sacrificing a higher total level.

^eAll p -values are presented on a one-tailed basis, given the directional predictions for H1.

The mean *comprehensive* surrogation score is 0.53 in the no-involvement condition, and is 0.48 in the choice condition. The difference of 0.05 is marginally significant ($t = -1.49$, $p = 0.069$). These results are robust to considering the *opportunity cost* surrogation score; the mean *opportunity cost* surrogation score is 0.38 in the no-involvement condition, and is 0.28 in the choice condition. The difference of 0.10 is significant ($t = -1.96$, $p = 0.026$). *No opportunity cost* surrogation scores, however, are not statistically different across conditions (mean_{no-involvement} = 0.69; mean_{choice} = 0.68; $t = -0.36$, NS). The lack of statistical significance for the *no opportunity cost* comparison is not entirely surprising, given that participants do not incur an opportunity cost in selecting the surrogation-consistent choices in this case. Overall, these results provide support for our hypothesis that allowing managers to choose strategy will decrease their tendency to surrogate strategic constructs with performance measures.²

² Interestingly, while these results are robust when examining only the data from participants implementing the attack strategy, we do not find support for H2 when examining only the data from participants implementing the socialize strategy. We discuss these results, and their implications, in a supplemental analysis below.

4.4 PRIMARY TESTS OF HYPOTHESIS 3: STRATEGY DELIBERATION

Hypothesis 3 predicts that strategy deliberation alone will decrease the tendency to exhibit surrogation. Operationally, this suggests that surrogation scores will be lower in the deliberation condition than in the no-involvement condition. The mean *comprehensive* surrogation score is 0.53 in the no-involvement condition, and is 0.51 in the deliberation condition. The difference of 0.02 is not significant ($t = -0.76$, NS). Differences in the *opportunity cost* and *no opportunity cost* surrogation scores across conditions are also not statistically significant (*opportunity cost*: $\text{mean}_{\text{no-involvement}} = 0.38$; $\text{mean}_{\text{deliberation}} = 0.34$; $t = -0.89$, NS; *no opportunity cost*: $\text{mean}_{\text{no-involvement}} = 0.69$; $\text{mean}_{\text{deliberation}} = 0.68$; $t = -0.31$, NS). Thus, H3 is not supported. These results suggest that, in our setting, strategy deliberation is not sufficient to reduce the tendency to surrogate strategy with performance measures.

4.5 STRATEGY RECALL

Approximately one week after the experiment, participants completed a brief survey regarding their experience in the study. In the survey, we reminded participants that in the study they chose (in the choice condition) or were provided (in all other conditions) a strategy. We then presented participants with a list of 14 words and asked participants to “Please circle the word below that best represents that strategy.” The list included both strategic constructs (attack and socialize), all of the corresponding ability measures, and additional ability measures unrelated to either attack or socialize (e.g., coordinate, fly, etc.). We use participants’ responses to this question to create two additional measures of surrogation: 1) the proportion of participants who do not correctly select the strategy they implemented, and 2) the proportion of participants who select the compensated measure. The latter measure is an especially strong indicator of surrogation, as it directly assesses participants’ tendency to replace the strategy with a measure.

In addition to the strategy-recall question, we also asked participants who implemented the attack (socialize) strategy to “Please indicate the strength of the relationship between” the word “attack” (“socialize”) and the phrase “global domination,” as well as between the word “bite” (“sing”) and the phrase “global domination.” Participants answered the latter two questions using 101-point scales, where 0 = “Very weak relationship” and 100 = “Very strong relationship.” The ratio of these two assessments constitutes a third additional measure of surrogation (e.g., “socialize to achieve global domination” vs. “sing to achieve global domination”). To construct the ratio, we divide participants’ global domination-*measure* assessment by their global domination-*strategy* assessment; the higher this “surrogation ratio,” the stronger the indication of surrogation.

We repeat our hypothesis tests using these three additional measures. Specifically, we expect that the tendency to not select the correct strategy,

TABLE 5
Tests of Hypotheses: Recall Measures
Experiments 1 and 2 (Pooled)

Recall Measure ^b	Condition ^a			
	Flat-Wage (<i>n</i> = 39)	No-Involvement (<i>n</i> = 79)	Choice (<i>n</i> = 78)	Deliberation (<i>n</i> = 39)
<i>Strategy Not Selected</i> ^c	0.33 [0.48]	0.68 [0.47]	0.54 [0.50]	0.76 [0.43]
<i>Measure Selected</i> ^d	0.06 [0.24]	0.44 [0.50]	0.33 [0.47]	0.63 [0.49]
<i>Surrogation Ratio</i> ^e	0.68 [0.30]	1.16 [1.39]	0.83 [0.32]	0.92 [0.35]

Panel B: Hypothesis Tests	Difference	<i>t</i>	<i>p</i> -Value ^f
H1: Flat-Wage vs. No-Involvement			
<i>Strategy Not Selected</i>	0.35	4.53	<0.001
<i>Measure Selected</i>	0.38	4.13	<0.001
<i>Surrogation Ratio</i>	0.48	2.61	0.005
H2: No-Involvement vs. Choice			
<i>Strategy Not Selected</i>	-0.14	-1.90	0.029
<i>Measure Selected</i>	-0.11	-1.45	0.074
<i>Surrogation Ratio</i>	-0.33	-2.28	0.012
H3: No-Involvement vs. Deliberation			
<i>Strategy Not Selected</i>	0.08	0.43	NS
<i>Measure Selected</i>	0.19	1.48	NS
<i>Surrogation Ratio</i>	-0.24	-1.41	0.081

^aPlease see figure 3 (and section 3) for a detailed description of the four experiment conditions. Reported conditions and statistics combine data from Experiments 1 and 2.

^bRecall measures were collected from a brief survey administered approximately one week after the experiment.

^c*Strategy not selected* represents the proportion of participants who did not correctly select the strategy they implemented when asked to select from a list of 14 words the word that they believed best represented the strategy they implemented. The higher the *strategy not selected* variable, the stronger the indication of surrogation.

^d*Measure selected* represents the proportion of participants who selected the compensated measure (*bite* or *sing*) instead of the implemented strategy. Participants were asked to select from a list of 14 words the word that they believed best represented the strategy they implemented. The higher the *measure selected* variable, the stronger the indication of surrogation.

^e*Surrogation ratio* represents the ratio of participants' assessment of the strength of the relationships between 1) the phrase "global domination" and the incentivized measure (*bite* or *sing*), to 2) the phrase "global domination" and the implemented strategy (attack or socialize). Both assessments were provided using 101-point scales, with 0 = "Very Weak Relationship" and 100 = "Very Strong Relationship". The higher the *surrogation ratio*, the stronger the indication of surrogation.

^fAll *p*-values are presented on a one-tailed basis, given the directional predictions of H1, H2, and H3.

the tendency to select the compensated measure, and the surrogation ratio will be: i) greater in the no-involvement condition than in the flat-wage condition (H1); ii) lower in the choice condition than in the no-involvement condition (H2); and iii) lower in the deliberation condition than in the no-involvement condition (H3).

Table 5 presents the means across conditions (panel A) and the hypothesis test results (panel B). Consistent with our earlier tests of H1, all three

alternative surrogation measures are greater in the no-involvement condition than in the flat-wage condition (*strategy not selected*: $t = 4.53$, $p < 0.001$; *measure selected*: $t = 4.13$, $p < 0.001$; *surrogation ratio*: $t = 2.61$, $p = 0.005$). Also consistent with earlier tests of H2, all three measures are lower in the choice condition than in the no-involvement condition (*strategy not selected*: $t = -1.90$, $p = 0.029$; *measure selected*: $t = -1.45$, $p = 0.074$; *surrogation ratio*: $t = -2.28$, $p = 0.012$). Finally, we find that the *surrogation ratio* is marginally lower in the deliberation condition than in the no-involvement condition ($t = -1.41$, $p = 0.081$). However, the remaining two measures are not statistically different across those two conditions (*strategy not selected*: $t = 0.43$, NS; *measure selected*: $t = 1.48$, NS). Collectively, these results generally corroborate the findings of our main hypothesis tests. That is, involvement in strategy choice reduces surrogation, but mere strategy deliberation does not.

4.6 STRATEGY BUY-IN

We also test our hypotheses splitting our sample by whether participants implemented the attack or socialize strategy. Using surrogation scores as the dependent measure, results are inferentially identical to those reported earlier, with the exception of H2. Specifically, we find strong support for H2 for participants implementing the attack strategy. The mean *comprehensive* surrogation score is higher in the no-involvement condition than in the choice condition ($t = -1.90$, $p = 0.030$). These results are robust when considering the *opportunity cost* surrogation score ($t = -2.58$, $p = 0.006$), but not when considering the *no opportunity cost* surrogation score ($t = -0.56$, NS). However, we do not find support for H2 for participants implementing the socialize strategy. The difference in surrogation scores across the choice and no-involvement conditions is not statistically significant ($p > 0.10$ for all three surrogation scores).

To better understand these results, we examine participants' strategy "buy-in" (Cokins [2005], Tayler [2010]). We define buy-in as the extent to which an individual *believes in the effectiveness of and is committed to the strategy*. In Experiment 2, we run the no-involvement and choice conditions again, and elicit participants' responses to two questions. The first question asks participants to rate "the relative effectiveness of the two global-domination strategies." We collect responses using a 101-point response scale, with 0 = "Attack strategy is much more effective," 50 = "The two strategies are equally effective," and 100 = "Socialize strategy is much more effective." The second question asks participants to rate "how committed [they were] to the [implemented] strategy." Again, we use a 101-point scale, with 0 = "Not committed at all," 50 = "Indifferent," and "100 = Very committed."

For participants implementing the attack strategy, assessment of the strategy's relative effectiveness is higher in the choice condition than in the no-involvement condition ($t = 3.16$, $p = 0.002$). Additionally, participants' self-reported commitment to the attack strategy is higher in the choice

condition than in the no-involvement condition ($t = 1.87, p = 0.035$). These findings were expected, since those choosing the strategy are likely to buy into the strategy more than those assigned the strategy. However, neither participants' assessment of the strategy's relative effectiveness, nor their self-reported commitment, differ across the choice and no-involvement conditions for participants implementing the socialize strategy ($p > 0.10$ for both variables).

While exploratory, one potential explanation for these results involves the beliefs of participants in the no-involvement condition. Recall that participants in this condition implemented an assigned strategy. Consequently, the no-involvement condition includes participants who, if given the opportunity, would have chosen to implement the non-assigned strategy. One possibility is that participants who preferred the attack strategy but assigned the socialize strategy still bought into the socialize strategy (perhaps they deemed the strategy to be of high quality even though they did not choose it), while those who preferred the socialize strategy but assigned the attack strategy did not buy into the attack strategy. This differential buy-in of strategy by participants in the no-involvement condition could help explain the strategy-contingent results reported above. That is, if the decrease in surrogation associated with manager involvement in strategy choice is driven in part by manager buy-in, then we would not expect to see a strategy-choice effect in settings where buy-in is not affected by involvement in strategy choice. Thus, the differential results for H2 across strategies potentially speak to the influence of buy-in on surrogation. Further, our results show that, at least in some settings, choice is a driver of buy-in, and has a mitigating influence on surrogation in these settings.

5. Conclusion

This study investigates the influence of managers' strategy involvement on their propensity to surrogate strategic constructs with compensated measures. Our findings suggest that managers' involvement in the *choice* of strategy is necessary and sufficient to mitigate the surrogation-inducing effects of incentive compensation. However, in our experiment setting, mere deliberation of strategy is not sufficient to mitigate surrogation.

Our findings contribute to a variety of research streams, including the literature on firms' use of strategic performance measurement systems, and the related phenomenon of surrogation. We advance academics' understanding of factors influencing managers' propensity to surrogate strategic constructs with performance measures. Additionally, our study complements extant literature that has begun to open the "black-box" of *how* goal commitment, strategy involvement, and buy-in facilitate improved performance. Our evidence is especially important to practicing managers and managerial accountants responsible for developing and implementing strategic performance measurement systems, and to those who must determine the degree of strategy-related autonomy afforded to managers within

a firm. Finally, we contribute to extant literature on strategic performance measurement systems. Specifically, we focus on managers' role within the strategic performance measurement system, highlighting the potential for different types of autonomy to endogenously influence the evolution and effectiveness of the strategic performance measurement system.

Our study is subject to limitations. First, we use a rich, but nonbusiness, setting to investigate our research question. While our setting and related design choices strengthen internal validity, they also invite questions of boundary conditions and the generalizability of our results. Second, while our study focuses on strategy selection and deliberation, it ignores other dimensions of how strategic performance measurement systems evolve with firm strategy. For instance, performance-measure identification and selection likely run parallel to strategy formulation (Tayler [2010]). With respect to surrogation, selecting performance measures may induce a "measure-focus" that offsets the benefits of managers' involvement in strategy formulation. This interesting possibility represents a potentially fruitful avenue for future research.

Our study highlights many other avenues of future research. For instance, our recall measures of surrogation highlight the robust and persistent nature of surrogation. Future research could examine the implications of this phenomenon in more dynamic (e.g., multiperiod, multiperson) environments. Future research could also examine how performance measure attributes such as congruence affect surrogation. Further, the measure's status as a leading or lagging indicator of a firm's achievement of a strategic goal may influence its use as a surrogate for a construct. Finally, future research may examine the impact of surrogation in some of the many real-world settings in which managers work with relatively static performance measures while grappling with a constantly changing business environment.

REFERENCES

- ATHEY, S., AND J. ROBERTS. "Organizational Design: Decision Rights and Incentive Contracts." *American Economic Review* 91 (2001): 200–205.
- BAKER, G. "Incentive Contracts and Performance Measurement." *Journal of Political Economy* 100 (1992): 598–614.
- BAKER, G. "The Use of Performance Measures in Incentive Contracting." *American Economic Review* 90 (2000): 415–20.
- BANKER, R. D.; H. CHANG; AND M. J. PIZZINI. "The Balanced Scorecard: Judgmental Effects of Performance Measures Linked to Strategy." *The Accounting Review* 79 (2004): 1–23.
- BECKER, B. E., AND M. A. HUSELID. "Strategic Human Resource Management in Five Leading Firms." *Human Resource Management* 38 (1999): 287–301.
- BECKER, S., AND D. GREEN. "Budgeting and Employee Behavior." *Journal of Business* 35 (1962): 392–402.
- BESTER, H., AND D. KRÄMER. "Delegation and Incentives." *RAND Journal of Economics* 39 (2008): 664–82.
- BROMWICH, M. "The Case for Strategic Management Accounting: The Role of Accounting Information for Strategy in Competitive Markets." *Accounting, Organizations and Society* 15 (1990): 27–46.

- BUDDE, J. "Performance Measure Congruity and the Balanced Scorecard." *Journal of Accounting Research* 45 (2007): 515–39.
- BUSHMAN, R. M., AND R. J. INDJEKIAN. "Stewardship Value of "Distorted" Accounting Disclosures." *The Accounting Review* 68 (1993): 765–82.
- BUSHMAN, R. M.; R. J. INDJEKIAN; AND M. C. PENNO. "Private Predecision Information, Performance Measure Congruity, and the Value of Delegation." *Contemporary Accounting Research* 17 (2000): 561–87.
- CAMPBELL, D.; S. M. DATAR; S. L. KULP; AND V. G. NARAYANAN. "Testing Strategy with Multiple Performance Measures: Evidence from a Balanced Scorecard at Store24." Harvard Business School Working Paper, 2008, No. 08–081.
- CHEN, C. C., AND K. JONES. "Are Employees Buying the Balanced Scorecard?" *Management Accounting Quarterly* 11 (2009): 36–44.
- CHENHALL, R. H. "Management Control Systems Design Within Its Organizational Context: Findings from Contingency-Based Research and Directions for the Future." *Accounting, Organizations and Society* 28 (2003): 127–68.
- CHOI, J.; G. W. HECHT; AND W. B. TAYLER. "Lost in Translation: The Effects of Incentive Compensation on Strategy Surrogation." *The Accounting Review* 87 (2012): 1135–63.
- COKINS, G. "Performance Management: Making It Work: The Promise and Perils of the Balanced Scorecard." *DM Review Online* (2005). Available at <http://www.information-management.com/news/1034304-1.html?zkPrintable=1&nopagination=1>.
- COLEMAN, H. J. "Why Employee Empowerment Is Not Just a Fad." *Leadership & Organization Development Journal* 17 (1996): 29–36.
- COVALESKI, M. A.; J. H. EVANS; J. L. LUFT; AND M. D. SHIELDS. "Budgeting Research: Three Theoretical Perspectives and Criteria for Selective Integration." *Journal of Management Accounting Research* 15 (2003): 3–49.
- DEVARO, J., AND F. A. KURTULUS. "An Empirical Analysis of Risk, Incentives and the Delegation of Worker Authority." *Industrial and Labor Relations Review* 63 (2010): 641–61.
- DRAKE, A. R.; J. WONG; AND S. B. SALTER. "Empowerment, Motivation, and Performance: Examining the Impact of Feedback and Incentives on Nonmanagement Employees." *Behavioral Research in Accounting* 19 (2007): 71–89.
- ELECTRONIC ARTS. *Sponz*. Redwood City, CA, 2008.
- GILES, W. D. "Making Strategy Work." *Long Range Planning* 24 (1991): 75–91.
- IJIRI, Y. *The Foundations of Accounting Measurement: A Mathematical, Economic, and Behavioral Inquiry*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1967.
- IJIRI, Y. *Theory of Accounting Measurement*. Sarasota, FL: American Accounting Association, 1975.
- ITTNER, C. D., AND D. F. LARCKER. "Innovations in Performance Measurement: Trends and Research Implications." *Journal of Management Accounting Research* 10 (1998): 205–38.
- ITTNER, C. D., AND D. F. LARCKER. "Assessing Empirical Research in Managerial Accounting: A Value-Based Management Perspective." *Journal of Accounting and Economics* 32 (2001): 349–410.
- KAHNEMAN, D., AND S. FREDERICK. "Representiveness Revisited: Attribute Substitution in Intuitive Judgment" in *Heuristics and Biases: The Psychology of Intuitive Judgment*, edited by T. Gilovich, D. Griffin, and D. Kahneman. New York: Cambridge University Press, 2002: 49–81.
- KAHNEMAN, D.; A. B. KRUEGER; D. SCHKADE; N. SCHWARZ, AND A. A. STONE. "Would You Be Happier if You Were Richer? A Focusing illusion." *Science* 312 (2006): 1908–10.
- KAPLAN, R. S., AND D. P. NORTON. "The Balanced Scorecard—Measures that Drive Performance." *Harvard Business Review* 70 (1992): 71–79.
- KAPLAN, R. S., AND D. P. NORTON. *The Balanced Scorecard: Translating Strategy into Action*. Boston, MA: Harvard Business School Press, 1996.
- KAUFFMAN, C. "Employee Involvement: A New Blueprint for Success." *Journal of Accountancy* (2010): 46–49.
- KEARNEY, R. C., AND S. W. HAYS. "Labor-Management Relations and Participative Decision Making: Toward a New Paradigm." *Public Administration Review* 54 (1994): 44–51.

- KREN, L. "Budgetary Participation and Managerial Performance: The Impact of Information and Environmental Volatility." *The Accounting Review* 67 (1992): 511–26.
- LANGFIELD-SMITH, K. "Management Control Systems and Strategy: A Critical Review." *Accounting, Organizations and Society* 22 (1997): 207–32.
- LAWTON, C. "Electronic Arts Bets Big on New Game." *Wall Street Journal* (September 2, 2008): B1.
- LEVINE, D. I., AND L. D. TYSON. "Participation, Productivity, and the Firm's Environment" in *Paying for Productivity: A Look at the Evidence*, edited by A. Blinder. Washington, DC: The Brookings Institution, 1990: 183–237.
- LIDEN, R. C.; S. J. WAYNE; AND R. T. SPARROWE. "An Examination of the Mediating Role of Psychological Empowerment on the Relations Between the Job, Interpersonal Relationships, and Work Outcomes." *Journal of Applied Psychology* 85 (2000): 407–16.
- LIPE, M. G., AND S. E. SALTERIO. "The Balanced Scorecard: Judgmental Effects of Common and Unique Performance Measures." *The Accounting Review* 75 (2000): 283–98.
- MALINA, M. A., AND F. H. SELTO. "Communicating and Controlling Strategy: An Empirical Study of the Effectiveness of the Balanced Scorecard." *Journal of Management Accounting Research* 13 (2001): 47–90.
- MALINA, M. A., AND F. H. SELTO. "Choice and Change of Measures in Performance Measurement Models." *Management Accounting Research* 15 (2004): 441–89.
- MALINA, M. A.; H. S. O. NØRREKLIT; AND F. H. SELTO. "Relations Among Measures, Climate of Control, and Performance Measurement Models." *Contemporary Accounting Research* 24 (2007): 935–82.
- MALSHE, A., AND R. S. SOHL. "Sales Buy-in of Marketing Strategies: Explorations of Its Nuances, Antecedents, and Contextual Conditions." *Journal of Personal Selling & Sales Management* 29 (2009): 207–25.
- MENON, S. T. "Employee Empowerment: An Integrative Psychological Approach." *Applied Psychology: An International Review* 50 (2001): 153–80.
- MILANI, K. "The Relationship of Participation in Budget-Setting to Industrial Supervisor Performance and Attitudes: A Field Study." *The Accounting Review* 50 (1975): 274–84.
- MILLER, K. I., AND P. R. MONGE. "Participation, Satisfaction, and Productivity: A Meta-Analytic Review." *Academy of Management Journal* 29 (1986): 727–53.
- NAGAR, V. "Delegation and Incentive Compensation." *The Accounting Review* 77 (2002): 379–95.
- NOURI, H., AND R. J. PARKER. "The Relationship Between Budget Participation and Job Performance: The Roles of Budget Adequacy and Organizational Commitment." *Accounting, Organizations and Society* 23 (1998): 467–83.
- PRENDERGAST, C. "The Tenuous Trade-Off Between Risk and Incentives." *Journal of Political Economy* 110 (2002): 1071–102.
- SCHKADE, D.A., AND D. KAHNEMAN. "Does Living in California Make People Happy? A Focusing Illusion in Judgments of Life Satisfaction." *Psychological Science* 9 (1998): 340–46.
- SHIELDS, J. F., AND SHIELDS, M. D. "Antecedents of Participative Budgeting." *Accounting Organizations and Society* 23 (1998): 49–76.
- STERLING, J. "Translating Strategy into Effective Implementation: Dispelling the Myths and Highlighting What Works." *Strategy & Leadership* 31 (2003): 27–34.
- STRACK, F.; L. L. MARTIN; AND N. SCHWARZ. "Priming and Communication: The Social Determinants of Information Use in Judgments of Life-Satisfaction." *European Journal of Social Psychology* 18 (1988): 429–42.
- TAYLER, W. B. "The Balanced Scorecard as a Strategy-Evaluation Tool: The Effects of Responsibility and Causal-Chain Focus." *The Accounting Review* 85 (2010): 1095–117.
- WEBB, R. A. "Managers' Commitment to the Goals Contained in a Strategic Performance Measurement System." *Contemporary Accounting Research* 21 (2004): 925–58.