BIASES IN THE SELECTION STAGE OF BOTTOM-UP STRATEGY FORMULATION

MARKUS REITZIG1* and OLAV SORENSON2
1 University of Vienna, Strategic Management, Vienna, Austria
2 Yale School of Management, Yale University, New Haven, Connecticut, U.S.A.

We propose that the failure to adopt an idea or innovation can arise from an in-group bias among employees within an organizational subunit that leads the subunit’s members to undervalue systematically ideas associated with members of the organization outside their subunit. Such biases in internal selection processes can stymie organizational adaptation and therefore depress the performance of the firm. Analyzing data on innovation proposals inside a large, multinational consumer goods firm, we find that evaluators are biased in favor of ideas submitted by individuals that work in the same division and facility as they do, particularly when they belong to small or high-status subunits. Copyright © 2013 John Wiley & Sons, Ltd.

INTRODUCTION

One of the primary goals of strategy research has been to develop a better understanding of how firms come to occupy advantaged positions within the market (Porter, 1980) and to develop valuable capabilities and resources that allow them to differentiate themselves from competitors (Wernerfelt, 1984; Barney, 1991). In understanding the origins of these competitive advantages, two distinct, but not unrelated, schools of thought have made important contributions. On the one hand, an evolutionary perspective, intertwined with the literature on organizational learning, has emphasized the fact that firms can benefit from processes of internal variation, selection, and retention (Baum and McKelvey, 1999; Lovás and Ghoshal, 2000; Zollo and Winter, 2002). For example, Sorenson (2000) demonstrated that firms that introduced and culled product variants at higher rates performed better than those that invested their resources in narrower product lines. Similarly, an extensive literature on economies of experience has found that productivity, profitability, and survival tend to increase as firms gain experience, presumably as trial-and-error processes engender the refinement of ever more efficient routines (for a recent review, see Thompson, 2010).

On the other hand, the process school of strategy has called attention to the fact that the decisions made within firms, and ultimately firms’ overall strategies, emerge not only from the actions of the chief executive but also from the cumulative decisions of numerous line employees and middle managers (Bower, 1970; Burgelman, 1983, 1994; Mintzberg and Waters, 1985; Noda and Bower, 1996; Huy, 2011). For example, Bower’s (1970) pioneering research on the resource allocation process highlighted the fact that strategic decisions often emerge from the lower levels of the organization. Burgelman’s (1994) famous case study of Intel similarly revealed that information gathering

Copyright © 2013 John Wiley & Sons, Ltd.
and decisions at middle levels of management led to Intel’s prescient abandonment of memory production and entry into microprocessors. And more recently, Huy (2011) documented the importance of middle managers even to the implementation of strategies introduced from the top of the organization.

Despite the fact that these literatures have drawn attention to the importance of internal selection processes to firm strategy and performance and to the critical role of middle managers in these selection processes, the field’s understanding of the role of middle managers and internal selection remains incomplete. Though the process school has posed a strong challenge to the top-down view of strategy and has called attention to a range of internal processes that influence planning, these insights have largely been based on qualitative research; and therefore it has been difficult to assess whether the influence of middle managers to project selection represents more the norm or exceptions as well as to examine the extent to which the selection decisions of middle managers might vary within the organization. Evolutionary perspectives on strategy, meanwhile, have analyzed organizational variation, selection, and retention systematically and quantitatively but have generally relied on externally observable correlates of these processes, such as cumulative output or product introductions, and have therefore had to treat the organization itself—and internal selection processes—as a black box.

This paper represents an early step in opening that black box. We build on the existing literature in two ways. Theoretically, we develop a more nuanced account of how middle managers might vary in their evaluations of ideas and projects originating from within the organization. To do so, we draw on three, somewhat disjoint, literatures in social psychology as well as research on social networks: One line of largely experimental research has found that individuals display a strong in-group bias in evaluations and choice decisions (e.g., Brewer, 1979; Mullen, Brown, and Smith, 1992); another examines the antecedents of identity (e.g., Tajfel et al., 1971; Doise et al., 1972). Building on these ideas, we argue that middlelevel managers in large, multunit organizations will come to identify psychologically with subparts of the organization and, as a result, they will favor ideas and projects forwarded by members of the subgroup(s) with which they have come to identify.1 A third line of experimental psychology has demonstrated that simply being exposed to another individual leads people to associate that individual with positive feelings and, as a result, to evaluate that individual and his ideas as being of unduly high quality (Zajonc, 1968; Lawler, 1992). Meanwhile, the literature on social structure has demonstrated that organizations strongly influence who interacts with whom, with those belonging to the same organization or part of the organization interacting most frequently (Feld, 1981, 1982; Lynch, 2006). When combined, these processes similarly suggest that middle managers would prefer projects forwarded by members of their own subunits of the organization.

Though past research has suggested that the professional and organizational identities of top-level managers might generate biases in executives’ evaluations of problems and solutions at the firm level (Dearborn and Simon, 1958; Hambrick and Mason, 1984; Tripsas and Gavetti, 2000; Zhang and Rajagopalan, 2010), the fact that a parallel set of processes and biases might operate at the level of middle management within the firm has been underappreciated. The choices made by these middle managers may nonetheless differ from those one might expect from a chief executive for a variety of reasons. They may wish to promote their subunit at the expense of the firm or feel constrained by personal connections to those they manage (Bower, 1970; Guth and MacMillan, 1986). Their positions within the firm may constrain their access to information or orient their attention toward a small subset of the information available (Ren and Guo, 2011). Or, they may internalize the effects of their decisions on those who report to them (Huy, 2011). Here, we call attention to the fact that their affiliations with and identification with a subunit of the firm—rather than with the firm as a whole—might also bias their evaluations of ideas emerging from across the organization.

Empirically, we exploit an unusual proprietary dataset that documents one large, multinational

---

1 Such a bias could also arise from rational calculation. Bower (1970: 302), for example, portrays such favoritism as a price paid to obtain loyalty from subordinates. Though undoubtedly an issue in many situations, the specifics of our context allow us to exclude such quid pro quo calculations as the mechanism underlying our results.
M. Reitzig and O. Sorenson

firm’s attempt to increase its rate of learning. Because the data include information on more than ten thousand ideas forwarded for evaluation, the outcomes of those evaluations, and the identities of the proposers and evaluators, they allow us to explore several aspects of intraorganizational selection processes. Using these data, we estimated the degree to which the evaluations of innovation proposals within the firm depended on the identities of the evaluators and submitters of those ideas. We found strong evidence that evaluators had an in-group bias, preferring ideas from employees who belonged to the same subunit of the organization—in this case, the same broad product division and physical site. Consistent with this explanation, the strength of these biases declined with the (1) size of the subunit to which the evaluator and submitter belonged—where subunit identity becomes less salient and the probability of direct personal connections becomes lower—and (2) stature of the submitter’s subunit—which mitigates the odds of the other unit being considered an outsider. These effects remained robust to the inclusion both of fixed effects for the people proposing these ideas and of controls for within-individual heterogeneity in the quality of these ideas.

Our paper contributes to the literature in multiple ways: For practitioners, it demonstrates the importance of appropriately assigning individuals to evaluation roles within the firm. Even in the absence of financial incentives for favoritism, connections between those generating and those evaluating ideas can introduce noise—and hence inefficiency—into the selection of projects and ideas. For students of strategy, our paper grounds the idea of in-group bias more firmly on a set of well-established psychological and sociological mechanisms and uses that grounding to develop a more nuanced set of predictions about when these biases should operate most strongly. With respect to evolutionary perspectives on strategy, it reveals that biases in beliefs can lead the selection stage of the process to deviate substantially from the identification of the “best” ideas. Finally, with respect to the process school, though this school has sometimes been seen as more about the “art” of management, our paper demonstrates that many of its insights nonetheless lend themselves to general theorization and systematic quantitative testing.

THE EVOLUTION OF ORGANIZATIONS

Campbell (1965) has been widely credited with introducing the variation-selection-retention model of evolution to the study of social systems. One sees direct descendants of his influence in the current crop of evolutionary perspectives on strategy formulation and organizational learning. Burgelman (1991), for example, describes the strategy formulation process as one in which individuals and groups within the firm lobby for the firm to engage in activities currently outside the scope of the firm (variation), management chooses among these proposals (selection), and then devotes organizational resources to their implementation (retention). Most models of organizational learning similarly have elements of variety creation, either through errors or experiments, selection based on environmental feedback, and retention through the diffusion of improved routines throughout the organization (Levitt and March, 1988; March, 1991).

Most of the intraorganizational research to date, however, has been on either the variation or the retention stage of this model. March and his students, for example, have done a great deal of (mostly theoretical) research to understand better the ways in which firms might differ in the degree of variety available in their internal sampling systems (e.g., March, 1991; Levinthal and March, 1993; Denrell and March, 2001). They have demonstrated that firms can fail to learn effectively because they do not sample the space of possible actions sufficiently or at random. Meanwhile, a set of strategy scholars has studied the processes of retention, the diffusion of selected practices within the organization (e.g., Szulanski, 1996; Hansen, 1999; Mors, 2010). That stage is not without its own difficulties. For instance, retained ideas must often be modified to fit settings different from the one in which they originated (Szulanski, 1996; Sorenson, Rivkin, and Fleming, 2006).

Though it has received less attention, between the stages of variation and retention sits another crucial to the innovation adoption process: the screening or selection stage. Engaging in the

---

2 A number of authors have acknowledged this problem in principle (e.g., Szulanski, 1996: 31). The existing literature has nonetheless generally treated the screening stage as informational (simply getting accurate information on which to base a decision).
adoption of an idea consumes significant organizational resources. Managers must therefore decide which ones to pursue.

To the extent that existing research on evolutionary processes has considered selection processes, it has often treated the organization as a unitary actor. Empirical investigations of the selection of more versus less risky organizational actions, for instance, have primarily examined these decisions at the level of the organization as a whole (e.g., Bromiley, 1991; Greve, 1998; Souder and Bromiley, 2012). Studies of product introductions and culling similarly have been at the level of the firm (e.g., Sorenson, 2000; Eggers, 2012). In part, this focus stems from data limitations; in part, it reflects a bias in favor of a top-down view of strategy. A vibrant case-based literature on the strategy formation process nevertheless suggests that many, if not most, of these selection decisions emerge not from the corner suites but rather from the middle ranks of management (Bower, 1970; Burgelman, 1994). Once one couples this fact with the idea that these midlevel managers might vary systematically in their biases, it opens the possibility that the acceptance or adoption of an idea might depend not just on its inherent quality but also upon its place of origin within the organization. Here, we explore one form of bias that may impede the selection and retention of good ideas.

In-group evaluation biases

“Each group nourishes its own pride and vanity, boasts itself superior, exalts its own divinities, and looks with contempt on outsiders.” (Sumner, 1906: 13)

Social scientists have long recognized the tendency for individuals to identify psychologically with groups and to favor other members of those groups in a variety of ways (for reviews of the psychology literature, see Brewer, 1979; Mullen et al., 1992). Early experiments, for example, demonstrated that subjects who had been randomly assigned to groups of individuals and then had time to socialize with those individuals would later treat the members of other (randomly assigned) groups with hostility (Sherif et al., 1961) and would evaluate their performances less favorably than those by members of their own group (Ferguson and Kelley, 1964). Consistent with this relationship-based effect, experimental studies of dyadic evaluation and exchange have found that even brief prior exposure to others leads to positively biased evaluations of them and of their products and ideas (Zajonc, 1968; Kroll, 1994). Given that relationships tend to concentrate within groups (Feld, 1981, 1982), this relationship-based bias could account for the favoritism towards others in the same group.

Subsequent experiments, however, have demonstrated that these biases in favor of the members of one’s own group emerge even in the absence of interaction—the mere act of randomly assigning two subjects to the same group seems sufficient to induce in-group preferences (Tajfel et al., 1971; Doise et al., 1972). Some have suggested that individuals come to adopt the values and goals of their group and therefore to promote it and its members’ interests (e.g., Ashforth and Mael, 1989). They may. But the experimental evidence suggests that, at least in part, this in-group bias stems either from innate cognitive processes or from social rules learned at a very early age. The minimal manipulations that involve neither contact with other members of the group nor even beliefs of future membership in the group would seem to preclude the internalization of values and goals as an explanation for their effects (Tajfel et al., 1971; Doise et al., 1972). And experiments on children as young as eight years of age have found in-group biases (Sherif et al., 1961).

But do these in-group biases have purchase in the field, on decisions with real economic consequences? We propose that these in-group biases will influence the screening function in the acceptance of ideas within the firm. To the extent that those involved in the selection of ideas and projects see themselves as part of some group—rather than as members of the organization as a whole—and interact more intensely with others in that group, a natural extension of these lab-based in-group biases would be for those screeners to exhibit a subconscious preference for ideas proposed by other members of their nominal group relative to those coming from outside.

An underlying assumption of our premise is that employees, in particular midlevel managers, see themselves as members of groups both at the firm and at the subunit level. At the firm level, such a proposition seems uncontroversial. Indeed, many have suggested that individuals come to identify with the organizations that employ them
Identification at the level of the organization as a whole combined with in-group biases may well account for the widely discussed NIH (“Not Invented Here”) syndrome, whereby organizations systematically undervalue innovations from outside the firm (e.g., Katz and Allen, 1982).

At the subunit level, however, this idea appears more novel, yet several factors jointly seem to justify it. First, in terms of identities, most organizations label employees and middle managers according to their subunit of the organization. Hence, a middle-level manager might hold an appointment such as senior marketing manager for Clorox in Colombia. Far from being meaningless, these titles not only assign prestige to individuals but also affiliate them with a particular part of the organization. Subunit memberships also shape dyadic interactions for at least two reasons. First, firms intentionally organize themselves and their reporting relationships—whether by product, function, geography, market, or some combination thereof—to encourage interaction within those subunits (Lawrence and Lorsch, 1967; Galbraith, 1973). Second, most organizations also arrange their offices physically along these same lines. The microgeography of buildings therefore tends to reinforce these functional patterns (Festinger, Schachter, and Back, 1950; Newcomb, 1956; Allen, 1977). Concomitant to these interactions, middle managers as well as other employees likely develop an affinity for and identification with the employees closest to them on the organization chart. We therefore expect middle managers to favor ideas from their own subunits.

This favoritism might manifest in multiple ways. For example, managers might allocate less attention to ideas that originate outside their subunit and therefore evaluate them less intensely. Even if they do evaluate ideas carefully, in-group biases may lead them to prefer subconsciously those submitted by members of their own subunit, even if an outsider would consider the two ideas of equivalent promise (quality). Regardless of the precise processes, we expect these in-group biases to influence idea selection. In particular, we predict:

**Hypothesis 1 (H1): The probability of an idea being approved for further consideration increases when the idea submitter belongs to the same subunit as the evaluator.**

Students of strategy, particularly those from the process school, have noted this phenomenon in the allocation of resources within the firm (e.g., Bower, 1970: 302). Strategy scholars have nonetheless often interpreted this tendency as a pragmatic solution on the part of the manager, consciously weighing the value of avoiding poor projects against the costs of upsetting underlings and potentially losing their support. Although able managers must balance these concerns, we contend that this phenomenon also stems in part from a pathological and subconscious bias in the evaluation process (and empirically explore a context that allows us to separate the two).

Although the main expected effect of these in-group biases is for middle managers to favor ideas from their own subunits of the organization, the fact that this prediction depends on the degree to which individuals identify with particular subunits and on their interactions with others from their own as well as from other subunits allows us to derive several additional predictions. In particular, let us consider the importance of the characteristics of both evaluators and subunits as moderators of the basic in-group effect. As we will argue below, the moderating effects not only shed additional light on the baseline effect, but also help us to isolate the mechanisms involved, and therefore help us to identify our core effect.

**Evaluator characteristics**

Although experimental evidence suggests that one can induce in-group bias with minimal manipulations, the strength with which an individual comes to identify with an employer or one particular subunit of the organization should increase over time (Huy, 2011). Unlike experiments in which individuals enter de novo, except in startups, the typical hire or transfer in an established firm enters a pre-existing community. He or she must therefore undergo a period of assimilation into the group. Even in startups, it requires time for individuals to become comfortable with their fellow employees and to perceive the organization as an ongoing concern (Stinchcombe, 1965). Given that this process of assimilation and identification unfolds over time, one would expect in-group biases to grow stronger with organizational and subunit tenure.
Consistent with this idea, at the firm level, the resistance to ideas from outside the firm has been found to increase with organizational tenure (Katz and Allen, 1982).

**Hypothesis 2 (H2):** The bias of evaluators in favor of ideas from their own subunits increases with their tenure within those subunits.

### Subunit characteristics

Characteristics of the subunit may also influence these biases. We see two factors, in particular, as potentially moderating the effect of in-group bias: group size and standing. Consider first the size of the subunit to which the evaluator and submittor belong. Experiments that involve the assignment of individuals to larger groups tend to produce lower levels of bias (Mullen et al., 1992). At least in the minimal manipulations used in many of the laboratory experiments—that do not even involve having subjects interact with these groups but simply telling them the size of the group—this effect probably stems from the fact that people find larger groupings less salient. One might therefore expect middle managers from larger subunits to identify less strongly with them, perhaps even perceiving themselves as belonging first and foremost to some smaller group within the subunit.

Bias based on dyadic relationships between managers and employees would also lead one to expect a weaker level of prejudice in larger units. Holding constant the total number of contacts that any particular individual has, the probability that two individuals have a direct relationship decreases with group size. Hence, to the extent that in-group favoritism stems from these direct connections, the probability of a connection and therefore also of the mean expected magnitude of the in-group bias should decline with the number of members in the group.

**Hypothesis 3 (H3):** The bias of evaluators in favor of ideas from their own subunits decreases with the size of their subunits.

Note that this prediction may depend, however, to some extent on the absence—or at least a relatively low level—of perceived competition between groups in the selection process. Though this assumption appears consistent with our empirical context, in cases where evaluators feel that subunits compete with other groups for resources, a countervailing force could exist: Middle managers from larger subunits may feel a greater need to favor proposals from within their subunit either to foster loyalty (Bower, 1970) or because they internalize the feelings of their fellow group members (Huy, 2011). Which force would prevail under competition—the declining salience and probability of personal connection or the increasing demand for resources—remains an open empirical question for future research.

Finally, subunit standing should also matter. The sociological literature has suggested and found substantial evidence for the notion that people treat status as a signal of quality (Podolny, 1993; Stuart, Hoang, and Hybels, 1999). Ideas from higher status sources might therefore be perceived as being of higher potential value. Indeed, consistent with this idea, at the interorganizational level, one of the primary exceptions to the bias against ideas from outside the organization has been the tendency for firms to adopt the practices of high-status firms (Strang and Macy, 2001). Translating this tendency to the intrafirm level, one might expect evaluators to have more favorable beliefs about the probable quality of ideas and innovations emerging from parts of the organization with higher standing within the firm.

**Hypothesis 4 (H4):** The bias of evaluators against ideas from other subunits decreases with the status of those subunits.

### METHOD

We tested these hypotheses at a large, multinational firm with operations in both consumer products and manufacturing. The company, which manufactures and sells thousands of products, employs more than 50,000 employees at 318 (traceable) facilities located across 66 countries. It is now structured as a matrix, with three broad product divisions as its primary dimension of organization and 66 countries as a secondary dimension of organization.3

---

3 The corporation underwent organizational restructuring—involving business unit mergers—during our data collection
In 2006, the CEO led the firm on an initiative to spur innovation throughout the firm. As part of the initiative, management offered gift bonuses, as well as (promotion-unrelated) recognition, to employees who submitted ideas that the firm decided to adopt. In part because the firm wanted to study the process but also in part to determine who should receive payouts, the firm developed an extensive database to track the ideas proposed by employees as part of this initiative.

From a research perspective, this database affords rare access to the screening process. Most studies of the adoption of innovations simply observe successful outcomes and therefore cannot determine whether nonadoption stems from not being exposed to the idea (variation), from objectively poor quality (accurate selection), from a bias in the screening function (biased selection), or from difficulty in implementing the idea following a decision to adopt it (retention). As our data provides complete information on all of the ideas considered, they allowed us to control for exposure. They also provide insight into idea quality to some extent. And, because they include direct information on the decision to consider the idea further, they allowed us to distinguish screening processes from implementation processes.

Any employee could submit an idea through an online interface. The submitter would give the idea a title, type in a short description of it (usually around 50 words), and classify it as applying to one of 54 product categories (e.g., body care or packaging adhesives), and possibly also to one of 18 idea subcategories.4

The system would then automatically assign the idea to a set of potential evaluators (middle managers). Although submitters could nominate potential evaluators, the company did not consider these nominations in their assignments. Instead, matching occurred on the basis of the business unit to which the idea appeared to apply (not necessarily the business unit of the submitter), the idea product category, and the idea subcategories.

When evaluators logged into the system, they would see those ideas assigned to them that had yet to receive an evaluation. Upon reading a proposal, they would then either deem it unworthy of further consideration or pass it on to either a business unit manager or technical expert for a secondary evaluation. Roughly half (42.3%) of the proposals received favorable evaluations (contingent on being evaluated at all) at this first stage.

In terms of the awareness of identities, one can consider the system a form of single-blind review. Evaluators could see who submitted an idea in the information available to them on the online system. But evaluators remained anonymous to the submitters, so they did not need to worry about reciprocity, future interpersonal interactions, or potential frictions, with those who submitted the ideas that they assessed. Importantly, this feature of the evaluation system removed the possibility of a quid pro quo trading of favoritism in exchange for employee loyalty as an explanation for in-group biases (cf. Bower, 1970).

**Data**

After removing identifying information, the company provided us with data on 22,958 valid idea proposals.5 To facilitate the analysis, we narrowed this sample in two ways. First, as process ideas accounted for less than 20 percent of proposals and appeared to differ on a number of dimensions from product ideas, we decided to limit our analysis to product innovations. Second, some of the control variables that we developed for idea quality depended on linguistic analysis of the proposals. Although the firm received proposals in several languages, we could only reliably develop these metrics for the two primary languages used in the firm: English and German.6 Eliminating process innovations and ideas in the other four languages left us with 16,758 cases.

From this set, we lost some additional cases due to missing data. First, not all of the ideas

---

4 Evaluators could change these categories, but the system did not track these reclassifications. Such changes, however, would influence the second-stage evaluations rather than the first-stage evaluations analyzed here.

5 To be considered valid, the online form had to be filled out in full. Incomplete entries, those submitted anonymously, or those submitted using invalid e-mail addresses therefore do not appear in our records.

6 We must infer the language of the proposal from the language chosen on the interface, from a menu of six, by the submitter.
had been evaluated by the time the company discontinued the program in May of 2008. Since both our theory and our dependent variable depend on these evaluations, we excluded 2,485 (right-censored) cases from the analysis. Second, incomplete biographic information on the attributes of submitters eliminated another 2,289 cases, leaving us with 11,984 cases for our main analyses. Third, incomplete information on the attributes of submitters, required to determine the exact size of their subunits, precluded the use of an additional nine cases, leaving us with 11,975 cases in which to examine interaction effects (Hypotheses 2 through 4).7

**Dependent variable**

Our dependent variable is a dichotomous variable, *approval*, with a value of one indicating that the initial evaluator decided that the idea merited further consideration and a value of zero denoting that the evaluator rejected the idea.8 Although a somewhat subjective and uncertain evaluation, as noted above, the sponsorship of the program came from the highest level of the company, and all evidence suggests that evaluators took this task seriously.

**Independent variables**

As noted above, the firm organizes itself primarily by broad product divisions and then by geography. We therefore considered the division-site as the relevant subunit, and calculated our primary measure of interest, *same unit*, as an indicator variable with a value of one if both the submitter and evaluator of an idea belonged to the same division of the firm (of three) and worked at the same site (of 318).9 Our expectation is that a shared identity between two individuals will lead the evaluator to have a biased assessment in favor of the idea and therefore to increase the probability that he or she recommends it for further evaluation.

*Evaluator tenure* Human Resources did not maintain longitudinal data in its central office on employees’ geographic locations. We therefore could only measure the evaluator’s tenure in his or her current division (as opposed to tenure in his or her division-site or division-country). Our variable therefore measures the (logged) time, in years (with precision at the level of the number of days), for which the evaluator worked in the division prior to assessing the idea (plus one to avoid logging zeroes).

*Subunit size* The company also does not maintain centralized records on the number of coworkers employed at each site in every country. Our measure of the subunit size therefore uses the information from the idea database to determine it. In particular, our measure counts the (logged) number of different individuals who submitted an idea and who worked within the same division and site as the evaluator (plus one to avoid logging zeroes).

The need to rely on the submission data itself to develop a size measure raises concerns. Obviously, it adds noise to our measure, potentially rendering our assessments of statistical significance more conservative. More worryingly, it could lead to biased estimates of the interaction effects if the probability of idea submission varied systematically as a function of subunit size. To explore this possibility, we created a second measure of subunit size based not on the affiliations of submitters but rather on the subunit memberships of evaluators. Since the company selected evaluators before opening the submission system, the distribution of those evaluators across subunits could not have been based on the probability of receiving submissions from a particular subdivision. Estimates using that alternative measure nevertheless produced qualitatively and statistically equivalent results. Hence, though we cannot dismiss the possibility of bias in our estimates, we feel reasonably confident of the direction of the effects.

*Submitter subunit stature* Discussions with the firm have made it clear that management sees the firm’s strength as engineering and design and that it reveres its historical success in research and development. As one of our contacts, a senior officer in the company, related to us, “The R&D function, formerly called applied technologies and...
product development, enjoys a high status for historical reasons.” Consistent with this stature, senior managers in the company have most commonly risen from the R&D ranks.

Although it did not seem feasible to ask senior management to assess the within-firm stature of all 318 sites, given the importance of the R&D function to the corporate identity, we believe that the strength of a local subunit’s research and development operations provided a reasonable proxy for its standing within the firm. For each site, our measure calculates the proportion of all submitted ideas classified as R&D related that came from the same site as the submitter. For example, if employees of Alpha Beta Aerospace (a disguised subunit name) forwarded four ideas and the company classified two of them as being R&D related, then we would compute a status measure for Alpha Beta equal to 0.5 (= 2/4).

Senior management affirmed the face validity of this measure for the highest- and lowest-ranked offices. Note that our models included controls at the idea level for whether or not the focal idea itself had an R&D component. Hence, one can interpret this effect as a spillover, or halo effect, from the submission of R&D ideas on other ideas at the site level (i.e., the degree to which even non-R&D ideas benefit from coming from a site that generated a large number of R&D ideas).

We also included several control variables to address potential heterogeneity in the underlying quality of the ideas submitted. Our first set of controls dealt with potential differences across submitters. *Same country* indicated whether or not the submitter and the evaluator worked in the same country. Most importantly, this measure captures the linguistic difficulties arising in the communication between idea originator and idea evaluator (e.g., Marschan-Piekkari, Welch, and Welch, 1999; Huy, 2011). But also, it controls for the potential effects of nationalism. In many of the models, moreover, we included submitter-level fixed effects to control for all observed and unobserved characteristics of these individuals.

Even with submitter-level fixed effects, however, one might still worry about heterogeneity in the quality of the ideas submitted by any particular individual. Our second set of controls therefore attempted to address this issue in two ways. First, we included indicator variables for the proposed application of the innovation. Submitters could assign their ideas to one of 54 different idea categories (though only 20 of them received valid submissions). We therefore entered into the model a vector of dichotomous variables for each of these idea categories. We also included indicator variables for each of the three divisions (treating the largest one as the baseline) and for whether the idea involved either the marketing or R&D function (the two most common functional classifications of the proposals).

Second, we developed a set of measures based on the content of each idea description. The first stage in developing these measures involved reading through 250 of the proposals—half in English and half in German—and noting factors that appeared to distinguish accepted proposals from rejected ones. The measures used in the models nevertheless relied on automated scripts, given the large amount of content to code. Two measures in particular appeared important: First, we counted the *number of words* in the description. Short descriptions often appeared terse on details and therefore received negative evaluations. Idea quality peaked at intermediate levels of description length (the optimum being approximately 250 words). We therefore included both the number of words and its square in the models. Second, we coded the tenor of the idea. Some descriptions focused positively on the benefits of the solution, while others highlighted the negatives of the current situation. To capture these differences, we created two word count variables—one tallying terms bearing positive connotations (e.g., “success,” “efficiency”) and another those with negative connotations (e.g., “frustration,” “complicated,” “confusing”).

Table 1 provides descriptive statistics for the variables used in our analyses.

**RESULTS**

We began our analyses by estimating a linear probability model. The linear probability model has the advantage of being easy to interpret and provides valid inferences about factors associated

---

10 Author translation; original interview conducted in German.

11 We derived the initial list of terms through a manual inspection of ideas and then expanded the number of terms through the use of synonyms in the Microsoft Word English and German dictionaries.
### Table 1. Descriptive statistics and correlations

| Variable                                      | Number of observations | Mean   | S.D.  | Min  | Max  | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   |
|-----------------------------------------------|------------------------|--------|-------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1 Idea approved                               | 11,984                 | 0.42   | 1     | 1    |      | 1    |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 2 Same unit                                   | 11,984                 | 0.12   | 0     | 1    | 0.12 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 3 Same country                                | 11,984                 | 0.73   | 0     | 1    | 0.06 | 0.22 |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 4 Ln (evaluator tenure + 1)                   | 11,984                 | 2.34   | 0.63  | 0.28 | 3.73 | −0.01| −0.10|      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 5 Ln (submitter submit size + 1)              | 11,975                 | 4.42   | 1.50  | 0.69 | 6.05 | 0.03 | 0.14 | 0.20 | −0.26|      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 6 Submitter subunit R&D status                | 11,984                 | 0.15   | 0.12  | 0    | 1    | −0.03| −0.05| 0.23 | 0.24 | 0.08 |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 7 Same gender                                 | 11,984                 | 0.52   | 0     | 1    | −0.01| −0.01| 0.05 | 0    | −0.04| 0.06 |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 8 Age difference (absolute)                   | 11,984                 | −9.01  | 6.69  | −41.3| 0    | −0.03| 0.03 | 0.05 | −0.10| 0.08 | −0.05| 0.02 |      |      |      |      |      |      |      |      |      |      |      |      |
| 9 Idea subcategory marketing/advertising (dummy)| 11,984                 | 0.55   | 0     | 1    | −0.15| −0.06| −0.17| −0.18| −0.05| −0.24| −0.01| 0.01 |      |      |      |      |      |      |      |      |      |      |      |      |
| 10 Idea subcategory R&D (dummy)               | 11,984                 | 0.17   | 0     | 1    | 0.10 | −0.02| 0.10 | 0.08 | 0.03 | 0.31 | 0.02 | −0.04| −0.51|      |      |      |      |      |      |      |      |      |      |      |      |
| 11 Dummy for largest business unit           | 11,984                 | 0.11   | 0     | 1    | 0.15 | 0.12 | 0.13 | 0.17 | −0.01| −0.08| −0.02| 0.06 | −0.40| −0.08|      |      |      |      |      |      |      |      |      |      |      |
| 12 Dummy for second largest business unit     | 11,984                 | 0.46   | 0     | 1    | −0.08| −0.18| −0.01| −0.08| −0.09| −0.04| 0.01 | 0.04 | 0.34 | −0.05| −0.33|      |      |      |      |      |      |      |      |      |      |
| 13 Dummy for third largest business unit      | 11,984                 | 0.43   | 0     | 1    | −0.02| 0.10 | −0.07| −0.03| 0.11 | 0.09 | 0    | −0.08| −0.08| 0.10 | −0.31| −0.79|      |      |      |      |      |      |      |      |      |
| 14 Ln (days between submission and evaluation + 1) | 11,984                 | 3.57   | 1.47  | 0    | 6.53 | −0.02| −0.03| 0.12 | 0.02 | 0.03 | 0.14 | 0    | −0.01| −0.02| 0.03 | 0.03 | 0.14 | −0.16|      |      |      |      |      |      |
| 15 Total number of words (/100)               | 11,984                 | 0.49   | 0.46  | 0.01 | 6.44 | 0.13 | 0.04 | 0.04 | 0.02 | 0.13 | 0.08 | 0    | 0.07 | −0.08| 0.05 | 0.05 | −0.03| −0.01| 0.05 |      |      |      |      |
| 16 Total number of words (/100) squared      | 11,984                 | 0.45   | 1.17  | 0.0001| 41.47| 0.08 | 0.01 | 0.02 | 0    | 0.07 | 0.03 | −0.01| 0.04 | −0.05| 0.03 | 0.03 | −0.01| −0.01| 0.03 | 0.03 | 0.86 |      |
| 17 Number of positive words (/10)            | 11,984                 | 0.05   | 0.10  | 0    | 1.40 | 0.07 | 0.01 | 0.01 | 0.08 | −0.02| 0.10 | 0    | 0.02 | −0.05| 0.03 | 0.03 | −0.03| 0.01 | 0.07 | 0.48 | 0.42 |      |
| 18 Number of negative words (/10)            | 11,984                 | 0.01   | 0.04  | 0    | 0.60 | 0.05 | 0.02 | 0.01 | 0.01 | 0.05 | 0.01 | 0.01 | 0.02 | −0.07| 0.03 | 0.04 | −0.04| 0.02 | 0.01 | 0.28 | 0.24 | 0.15 |      |

All means calculated at the idea level.
with differences in the mean acceptance rate (even with binary data). In the first model reported in Table 2, we can see that when the evaluator and submitter come from the same division and site, an idea had a 15.9 percentage point higher probability of being given further consideration. Consistent with H1 then, evaluators exhibited a bias in favor of ideas from their own subunits. As one might have anticipated, the control variables also suggest that ideas have a higher probability of receiving a favorable evaluation when they pertain to R&D.\textsuperscript{12}

The second and third models reported introduce controls for submitter and within-submitter heterogeneity. As one would expect, some submitters consistently had better ideas than others—the individual level fixed effects are jointly significant with greater than 99.9 percent confidence (Model 3: $F(4081, 7875) = 1.35$; $p > F = 0.000$). Even within submitters, however, some ideas still seemed better than others. As expected, evaluators more frequently passed proposals of intermediate length along for further consideration. But even with both submitter fixed effects and controls for within-submitter idea heterogeneity, ideas still had a 6.3 percentage point higher chance of being evaluated favorably when the submitter and the evaluator came from the same division and site. Note that because of the submitter fixed effects, one should interpret this effect as being within-submitter. In other words, for a given submitter, ideas evaluated by a middle manager from the same subunit had a 6.3 percentage point higher probability of receiving a favorable evaluation relative to those submitted.

\textsuperscript{12}One might potentially worry that correlated errors across repeated observations leads to inappropriately small standard errors. The results nevertheless remain robust to clustering the error at either the evaluator or the submitter level, even though efficiency decreases.

\begin{table}[h]
\centering
\begin{tabular}{llllll}
\hline
 & 1 & 2 & 3 & 4 & 5 \\
Model & OLS & OLS, fixed effect & OLS, fixed effects & Logit & Logit, conditional fixed effects \\
\hline
H1: same unit & 0.159 & 0.064 & 0.063 & 0.658 & 0.307 \\
 & (0.014)*** & (0.022)*** & (0.022)*** & (0.063)*** & (0.108)*** \\
Same country & 0.012 & 0.061 & 0.063 & 0.307 & 0.300 \\
 & (0.011) & (0.015)*** & (0.015)*** & (0.047) & (0.075)*** \\
Idea subcategory = marketing & −0.071 & −0.092 & −0.090 & −0.267 & −0.429 \\
 & (0.013)*** & (0.016)*** & (0.016)*** & (0.055)*** & (0.078)*** \\
Idea subcategory = R&D & 0.081 & 0.094 & 0.094 & 0.328 & 0.417 \\
 & (0.016)*** & (0.019)*** & (0.019)*** & (0.067)*** & (0.094)*** \\
Third-largest division & 0.056 & 0.086 & 0.082 & 0.244 & 0.392 \\
 & (0.045) & (0.062) & (0.061) & (0.197) & (0.295) \\
Second-largest division & −0.149 & 0.013 & 0.013 & −0.627 & 0.155 \\
 & (0.037)*** & (0.048) & (0.048) & (0.160)*** & (0.241) \\
Total number of words (/100) & 0.133 & 1.052 & 0.618 & & \\
 & (0.029)*** & (0.092)*** & (0.142)*** & & \\
Total number of words (/100) squared & −0.029 & −0.250 & −0.131 & & \\
 & (0.010)*** & (0.037)*** & (0.046)*** & & \\
Number of positive words (/10) & 0.059 & 0.253 & 0.277 & & \\
 & (0.065) & (0.216) & (0.314) & & \\
Number of negative words (/10) & −0.035 & −0.067 & −0.207 & & \\
 & (0.126) & (0.462) & (0.617) & & \\
Constant & 0.564 & 0.495 & 0.435 & −0.159 & \\
 & (0.022)*** & (0.031)*** & (0.033)*** & (0.100) & \\
Idea categories & Y & Y & Y & Y & Y \\
Submitter fixed effects & N & Y (4,082 groups) & Y (4,082 groups) & N & Y (1,389 groups) \\
Observations & 11,984 & 11,984 & 11,984 & 11,984 & 7,870 \\
$R^2$ or pseudo $R^2$ & 0.06 & 0.05 & 0.06 & 0.06 & \\
\hline
\end{tabular}
\caption{Idea approval by evaluator and submitter subunit membership}
\end{table}
by the same person but evaluated by a manager in a different division or working at a different site.

The final two columns in Table 2 report the models re-estimated with logistic regression, both with and without submitter fixed effects. Note that only cases in which individuals both submit more than one idea and have at least one idea favorably evaluated and one idea unfavorably evaluated can contribute to the estimates of the conditional logit. Model 5 therefore drops 2,693 individuals who either submitted only one idea or had all of their ideas accepted or rejected in the first stage. Not only do the coefficients mirror those of the linear regression in terms of their qualitative implications, but also they suggest essentially equivalent effect sizes. For example, in the logistic regression (Model 4), the marginal effect of having an evaluator from the same subunit as the submitter—setting all other variables to their means—increases the expected probability of a favorable evaluation by 16.3 percentage points.

Table 3 explores whether various evaluator and subunit characteristics influence the degree of in-group bias. Because we expect these characteristics to moderate the intensity of in-group biases, we implemented each as an interaction between the evaluator or subunit characteristic and the same unit variable (we also included the uninteracted values of these characteristics as controls). Model 6 introduces these interaction effects in a linear probability model and Model 7 estimates them in the more efficient—though more difficult to interpret—logit model. In terms of the coefficient for evaluator tenure, we cannot reject the possibility that this coefficient does not differ from zero (i.e., that no interaction exists). Our data, however, do not allow us to determine whether this failure to find support for H2 stems from the actual absence of tenure effects or from the measurement error introduced by not having longitudinal data on the geographical locations of evaluators.

Consistent with H3, however, the interaction of same unit and subunit size has a negative and significant coefficient. In other words, evaluators from larger subunits exhibited less intense in-group biases. A doubling, for example, in the size of the evaluator’s subunit predicts a 6.5 percentage point decline in the magnitude of the same-unit bias, reducing its effect by roughly one-third. In support of H4, submitters from high stature—R&D intensive—subunits also experienced less bias when being evaluated by someone from another division or site. Stature, however, had much larger effects than subunit size; a one standard deviation increase in subunit standing corresponded to a roughly 16 percentage point decline in the magnitude of the same-unit bias (an almost complete elimination of the effect). Note that we could not include submitter fixed effects in these models because submitter characteristics do not vary within submitter and hence their inclusion would preclude estimation of the interaction effects of interest.

Although we have controlled for a wide range of factors that might confound the interpretation of the results, one issue merits further discussion.

**Endogenous assignment of evaluators**

In order to interpret our effects as causal, one would need to believe that the company randomly assigned reviewers to ideas (at least with respect to the subunit membership of the submitter). Though reasonable from the perspective of the submitter—submitters had no control over this assignment—evaluators did have some influence over this process. In particular, evaluators could choose the order in which they considered ideas. We examined whether this selection process might influence our results in two ways. First, we calculated the time between submission and evaluation, in terms of the logged number of days (plus one to avoid logging zeroes), and included it in the estimates of the first-stage evaluation to see whether evaluators favored ideas that they considered more rapidly (Model 8). Not only does the inclusion of this variable not affect the other results, but also the time that an idea spends in the queue from submission to initial evaluation had no significant relationship to its likelihood of receiving a favorable evaluation.

Second, we explored the factors that determined the length of time that ideas spent in the evaluation queue. Table 4 reports a regression of the number of days between submission and final evaluation as a function of idea characteristics. Interestingly, lengthier ideas and more positive ideas stayed in the queue longer. Since these ideas had higher likelihoods of receiving favorable evaluations, it appears that evaluators may attempt to minimize their effort. They suspect that they can quickly reject short, negative ideas, leaving the more
Table 3. Idea approval by evaluator and submitter subunit characteristics

<table>
<thead>
<tr>
<th>Model</th>
<th>6 OLS</th>
<th>7 Logit</th>
<th>8 Logit</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1: same unit</td>
<td>0.652*** (0.096)***</td>
<td>3.181*** (0.480)***</td>
<td>3.148*** (0.487)***</td>
</tr>
<tr>
<td>Same country</td>
<td>0.028 (0.011)**</td>
<td>0.125 (0.051)**</td>
<td>0.147 (0.100)***</td>
</tr>
<tr>
<td>H2: same unit \times \text{Ln (evaluator tenure + 1)}</td>
<td>0.005 (0.022)</td>
<td>0.062 (0.099)</td>
<td>0.069 (0.100)***</td>
</tr>
<tr>
<td>H3: same unit \times \text{Ln (submitter subunit size + 1)}</td>
<td>-0.094 (0.012)***</td>
<td>-0.462 (0.060)***</td>
<td>-0.478 (0.061)***</td>
</tr>
<tr>
<td>H4: same unit \times \text{submitter subunit R&amp;D status}</td>
<td>-0.356 (0.279)</td>
<td>-2.436 (1.378)*</td>
<td>-2.559 (1.394)*</td>
</tr>
<tr>
<td>Same unit \times \text{number of words (/100)}</td>
<td></td>
<td>0.249 (0.161)</td>
<td></td>
</tr>
<tr>
<td>\text{Ln (Evaluator tenure + 1)}</td>
<td>0.027 (0.008)***</td>
<td>0.116 (0.037)***</td>
<td>0.096 (0.037)***</td>
</tr>
<tr>
<td>\text{Ln (Submitter subunit size + 1)}</td>
<td>0.007 (0.003)**</td>
<td>0.029 (0.015)*</td>
<td>0.028 (0.015)*</td>
</tr>
<tr>
<td>\text{Submitter subunit R&amp;D status}</td>
<td>-0.354 (0.044)***</td>
<td>-1.574 (0.203)***</td>
<td>-1.525 (0.204)***</td>
</tr>
<tr>
<td>Same gender</td>
<td>— —</td>
<td>— (0.003)***</td>
<td>— (0.003)***</td>
</tr>
<tr>
<td>Age difference (absolute) \times (-1)</td>
<td>-0.071 (0.013)***</td>
<td>-0.313 (0.056)***</td>
<td>-0.317 (0.056)***</td>
</tr>
<tr>
<td>Idea subcategory = Marketing</td>
<td>0.098 (0.016)***</td>
<td>0.418 (0.068)***</td>
<td>0.409 (0.068)***</td>
</tr>
<tr>
<td>Idea subcategory = R&amp;D</td>
<td>0.059 (0.044)***</td>
<td>0.282 (0.200)***</td>
<td>0.281 (0.201)***</td>
</tr>
<tr>
<td>Third largest division</td>
<td>-0.084 (0.037)***</td>
<td>-0.336 (0.165)***</td>
<td>-0.369 (0.166)***</td>
</tr>
<tr>
<td>Second largest division</td>
<td>— —</td>
<td>— (0.014)***</td>
<td>— (0.014)***</td>
</tr>
<tr>
<td>\text{Ln (days between submission and evaluation + 1)}</td>
<td>0.245 (0.019)***</td>
<td>1.114 (0.094)***</td>
<td>1.145 (0.096)***</td>
</tr>
<tr>
<td>Total number of words (/100)</td>
<td>-0.057 (0.007)***</td>
<td>-0.268 (0.037)***</td>
<td>-0.282 (0.038)***</td>
</tr>
<tr>
<td>Total number of words (/100) squared</td>
<td>0.070 (0.049)***</td>
<td>0.299 (0.220)***</td>
<td>0.338 (0.221)***</td>
</tr>
<tr>
<td>Number of positive words (/10)</td>
<td>-0.038 (0.104)</td>
<td>-0.168 (0.465)</td>
<td>-0.164 (0.466)</td>
</tr>
<tr>
<td>Number of negative words (/10)</td>
<td>0.391 (0.035)***</td>
<td>-0.493 (0.156)***</td>
<td>-0.439 (0.165)***</td>
</tr>
<tr>
<td>Constant</td>
<td>— —</td>
<td>— (0.015)***</td>
<td>— (0.015)***</td>
</tr>
<tr>
<td>Idea categories</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Observations</td>
<td>11,975</td>
<td>11,975</td>
<td>11,975</td>
</tr>
<tr>
<td>\text{R}^2 or pseudo \text{R}^2</td>
<td>0.09</td>
<td>0.07</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Standard errors in parentheses; p-values are two-tailed t-tests.
***p < 0.01; **p < 0.05; *p < 0.1

difficult evaluations for the next evaluator logging into the system. Although ideas from the same unit do leave the queue faster, this effect is smaller than that of writing quality. The average idea stayed in the evaluation queue for 77 days. Those from the same unit as the evaluator received decisions, on average, 9.5 days earlier than those from other units. Though interesting, the results from Model 8 suggest that this deviation from random matching cannot account for our results.
Table 4. Days between idea submission and evaluation by idea characteristics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>OLS</td>
</tr>
<tr>
<td>Total number of words (/100)</td>
<td>13.822</td>
</tr>
<tr>
<td></td>
<td>(3.515)***</td>
</tr>
<tr>
<td>Total number of words squared (/100)</td>
<td>−3.901</td>
</tr>
<tr>
<td></td>
<td>(1.315)***</td>
</tr>
<tr>
<td>Number of positive words (/10)</td>
<td>62.149</td>
</tr>
<tr>
<td></td>
<td>(8.945)***</td>
</tr>
<tr>
<td>Number of negative words (/10)</td>
<td>6.194</td>
</tr>
<tr>
<td></td>
<td>(19.097)</td>
</tr>
<tr>
<td>Same unit</td>
<td>−9.462</td>
</tr>
<tr>
<td></td>
<td>(2.544)***</td>
</tr>
<tr>
<td>Constant</td>
<td>69.467</td>
</tr>
<tr>
<td></td>
<td>(1.436)***</td>
</tr>
<tr>
<td>Observations</td>
<td>11,984</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Standard errors in parentheses; $p$-values are two-tailed $t$-tests. ***$p < 0.01$

DISCUSSION

Though others have noted the potential for in-group biases at the level of the firm—for example, in the so-called “Not Invented Here” syndrome—we have argued here that in-group biases can also operate within firms. Middle managers, and other employees below the level of the senior executives, often come to identify with subunits of the firm. When placed in a role that requires them to evaluate ideas and projects emerging from various parts of the organization, they then tend to evaluate these ideas unequally, favoring those from their own subunits.

Using a unique, proprietary dataset that captures information on more than ten thousand ideas submitted and evaluated at one large multinational firm, we found evidence supporting the existence of these within-firm biases in the field. On average, managers accepted ideas from their own subunits—defined as the intersection of a product division and a geographic location—at a rate nearly 16 percentage points higher than the baseline acceptance rate of 42 percent. This bias in screening, moreover, appeared robust to a wide range of controls. Though it declined by roughly half in terms of its estimated magnitude, the bias persisted even with the inclusion of submitter-level fixed effects and textual-analysis-based controls for the quality of the proposal. In other words, even for ideas coming from the same individual and of similar written quality, ideas reviewed by a manager from the same subunit passed the screen at a higher rate. Further analyses suggested that these selection biases appeared most acutely against ideas emerging from smaller and from lower-status subunits of the organization, consistent with what one would expect if subunit identification on the part of evaluators drives these biases.

Either of two similar, though distinct, mechanisms could account for our results. On the one hand, in-group biases may emerge from the aggregation of dyadic relationships: Having positive affect for the person forwarding an idea can lead an evaluator to consider it more favorably, a finding documented in both prior experimental and field research (Kollock, 1994; Sorenson and Waguespack, 2006). On the other hand, experimental research finds that in-group biases can also emerge even in situations in which subjects have had no contact with each other and in which they do not even know each other’s identities (Tajfel et al., 1971; Doise et al., 1972); in other words, the evaluation bias might emerge either from dyadic relationships or from the psychological identification of the evaluator with the group of the submitter. Our data, however, do not allow us to separate these mechanisms.13

Limitations

Despite finding robust evidence that supports our hypotheses about in-group biases within the firm, our study is not without its weaknesses. We see two issues, in particular, as most potentially problematic. First, our data came from a single organization. In-group biases appeared to operate strongly and systematically within this firm but they may play less of a role in other contexts. For example, we would expect subunit identities to become more salient in large, multinational firms where employees only rarely move or interact across offices. Our research site thus represents just the sort of firm that one might expect to anchor the end of the spectrum where these in-group biases would most strongly influence idea and project selection.

13 We could not feasibly gather social network data for all of the 50,000 employees of the firm, or even for the 4,082 employees in the sample that we analyzed.
Though we hope to see future studies assess the strength of these processes in other contexts, we nonetheless suspect that progress on in-group biases at the intrafirm level will proceed largely through quantitative case studies; the need for detailed internal data poses a daunting challenge to the recruitment of even one organization. Attempting comparative research that would require such data across multiple sites seems untenable without a clever means of observing these processes from publicly available data.

Second, we only examined a single stage of the variation-selection-retention process. Even within this selection stage, many firms, including the one studied here, have more than one filter. Our firm, for example, passes the projects accepted in this initial screening on to a second expert for further evaluation. To the extent that these secondary experts have different group affiliations, they may help to mitigate some of the biases introduced in the first screening (or they could introduce their own). More broadly, the process of moving from idea to implementation involves numerous steps. Each one could interact with earlier steps, either mitigating or exacerbating biases. We therefore cannot determine the extent to which the individual-level biases observed in a single stage of the process might manifest as firm-level inefficiencies in product selection.

Practical and theoretical implications

Though a better understanding of the mechanisms involved would allow for more nuanced managerial implications, the core finding itself nevertheless informs at least three managerial decisions. The first involves the assignment of individuals to evaluation roles. At first blush, given our results, one might assume that managers should always assign the review and selection process to individuals from parts of the organization unrelated to the one from which the idea being evaluated originates. Though that prescription would remove in-group biases from the process, it would also frequently relegate decisions to those with little relevant expertise, introducing more error into the evaluation process. In many cases, a firm might prefer the illness to the cure. A more effective and practical solution might therefore take the opposite tack, ensuring that individuals from the same subunit of the organization as the submitter always evaluate the ideas in question. Selection processes will then involve bias but the bias will operate similarly across ideas. From the perspective of the firm, the problems in biased screening arise not so much from an inappropriate average level of selection but rather from differential application of the selection threshold, accepting bad ideas while rejecting better ones.

The second concerns the design of screening processes within the organization. As the discussion above suggests, managers face a vexing problem: Should they choose the most informed or the least biased evaluators? Evaluators within a group bring high quality information to the assessment but potentially with a distorted lens. Evaluators from outside a group, meanwhile, though free from bias, potentially have limited expertise and access to information relevant to the decision. As a result, neither using only one nor the other seems optimal. Firms may therefore screen more effectively and efficiently by adopting hybrid processes that systematically introduce opinions from both biased and unbiased parties. One example of such a system is tenure review: Most research-active universities have adopted a model where an internal committee of (biased) evaluators decides on whether to promote a candidate but does so with extensive input from an external set of (unbiased) experts from other universities. Another is the Olympic scoring system: In many sports such as gymnastics, the judging panel includes both biased and unbiased reviewers. After dropping the top and bottom scores, the average score provides a relatively objective assessment of the quality of the performance.

The third decision involves the design of the organization itself. Because these biases stem from the interactions of psychological and sociological processes with the structure of the organization, firms may vary systematically in the effectiveness of their internal evaluation regimes. Flatter organizational structures, for example, may lead employees and managers to identify more with the organization as a whole than just with its parts (Reitzig, 2011). Engendering a stronger identification with the firm could, in turn, inhibit ingroup biases. Similarly, the rotation of employees through different divisions within the firm might mitigate in-group biases both by creating cross-subunit personal relationships and by limiting the degree to which individuals perceive themselves as members of a subunit rather than of the firm as a whole.
Our results also have a number of theoretical implications: Most directly, they suggest that individual identity may matter for a variety of behaviors relevant to firm performance (Powell, Lovallo, and Fox, 2011). Here, we have highlighted the potential for identification at the sub-unit level to influence idea screening and project selection. But identities can also influence other processes: For example, subunit identity might also influence motivation. They can also exist at other levels. Many management researchers, for example, have suggested that employees identify with their employers at the level of the firm (e.g., Ashforth and Mael, 1989; Dutton et al., 1994). Firm-level identification—rather than information isolation—may therefore account for the “Not Invented Here” syndrome observed in many organizations (cf. Katz and Allen, 1982).

More broadly, evolutionary perspectives on strategy have argued that firms learn by generating variety—product ideas, potential process improvements, alternative positioning vis-à-vis the competition—selecting among these variants for those that appear most beneficial to the firm, and then rolling out the selected variants across the firm. The effectiveness with which firms learn depends on the confluence of these variation, selection, and retention routines. To date, the strategy literature has been most cognizant of problems and the potential for firm-level competitive advantages in the first and final stages of this process. James G. March and his collaborators, for example, have called attention to the fact that firms often generate too little variety, or variety too similar to their current offerings and practices (e.g., March, 1991; Levinthal and March, 1993; Denrell and March, 2001). Strategy scholars, meanwhile, have highlighted many of the problems that can arise as firms attempt to implement the ideas selected (retention), particularly with the difficulty of transferring the tacit information involved in processes and in the design of products (Szulanski, 1996; Hansen, 1999). Here, we have demonstrated that difficulties may also arise in the selection stage of the learning process.

CONCLUSION

Many of the selection processes of firms emerge not from the deliberations of the CEO and his senior staff but rather from the aggregated decisions of a multitude of middle managers (Bower, 1970; Burgelman, 1983; Noda and Bower, 1996). The choices made by these middle managers may nonetheless differ from those one might expect from the chief executive for a variety of reasons. For example, they may have incentives to promote their subunit at the expense of the firm or they may feel more constrained by their personal connections to those they manage (Bower, 1970). Here, we call attention to the fact that middle managers’ affiliations with and identification with a subunit of the firm—rather than with the firm as a whole—might bias their evaluations. In particular, middle-level managers might exhibit something that we think of as “intraorganizational provincialism”—a tendency to favor ideas and projects forwarded by members of the subgroups with which they have come to identify.

Organizations have become quite sensitive to the ways in which employees might inappropriately favor particular people and projects. To date, however, the attention has been mostly on situations where those biases emerge from purposive action, such as when employees might favor a project for personal financial gain. The solution in these situations has often been to try to eliminate the conflict of interest. But such favoritism can also emerge from subconscious processes. Indeed, those forms of partisanship could prove even more pernicious as conscientious employees do not even notice them and therefore find them impossible to avoid. However, the solution here, rather than attempting to avoid the problem, requires embracing it.

ACKNOWLEDGEMENTS

We thank Rodrigo Canales, Phanish Puranam, Will Mitchell, and two anonymous reviewers for their useful comments on earlier drafts of this paper. All errors remain our own.

REFERENCES


Lawrence P, Lorsch J. 1967. Organization and Environment: Managing Differentiation and Integration. Division of Research, Graduate School of Business Administration, Harvard University: Boston, MA.


Biases in Selection