Received: 8 November 2021 | Revised: 27 January 2023 | Accepted: 31 January 2023 | Published on: 15 February 2023

DOI: 10.1002/smj.3491

RESEARCH ARTICLE

SMS | Strategic Management Journal

WILEY

# Managing exploration in organizations: The effect of superior monitoring on subordinate search behavior

**Valentina Richter**[1] | **Robert Janjic**[1] | **Helge Klapper**[2] | **Steffen Keck**[1] | **Markus Reitzig**[1]

[1]Strategic Management Subject Area, University of Vienna, Vienna, Austria

[2]Krannert School of Management, Purdue University, West Lafayette, Indiana, USA

**Correspondence**
Valentina Richter, Strategic Management Subject Area, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria.
Email: valentina.richter@univie.ac.at

**Abstract**

**Research Summary:** In this article, we explore the effects of managerial monitoring on the behavior of subordinates tasked with the search for alternatives in a complex environment. We argue that managerial monitoring will lead subordinates to exhibit more search than they would engage in otherwise as they try to impress their superiors by exerting more effort. We test and confirm our hypothesis in four laboratory studies with a total of 444 participants. Our findings show that search distance and duration are highly susceptible to managerial monitoring, whereas similar interventions from peers and subordinates are ineffective.

**Managerial Summary:** A key task for employees is to find new solutions to corporate problems when and where they occur. But what determines how long these individuals search and how far they venture before settling on one solution eventually? And how can and should managers steer this process, particularly if they are as unknowing about potential solutions as their subordinates? Here, we show that leaders can bring their staff to explore complex solution spaces longer and more remotely by regularly appraising their subordinates' efforts. This is because employees will feel an

enhanced need to demonstrate their industriousness to their bosses, and long and distant search is easily justified. Whether this increases corporate performance depends on the complexity of the solution space and the opportunity costs of search.

**KEYWORDS**

laboratory experiments, managerial appraisal, monitoring, NK model, search

## 1 | INTRODUCTION

The process of search (Becker, Knudsen, & March, 2006; Denrell, Fang, & Winter, 2003; Fleming, 2001; Ghemawat & Levinthal, 2008; Porter, 1991; Rivkin, 2000; Schumpeter, 1934) is omnipresent in organizations (Angus, 2019; Dahlander, O'Mahony, & Gann, 2016; Tandon & Toh, 2022; Yayavaram, Srivastava, & Sarkar, 2018; Yu, Minniti, & Nason, 2019). Companies aiming to launch novel business models, for example, have to find new profitable combinations of bringing customer needs and products together (Markides, 1999; Zott & Amit, 2010). The more interdependent these different activities or choices are, the more complex the search problem becomes (Levinthal, 1997; Simon, 1962). The more complex a search problem, however, the less feasible it will be to evaluate how to spend effort to explore the unknown. This poses a challenge for organizations in which managers seek to monitor the individual search processes of their direct reports in the hierarchy. Notably, in truly complex situations managers will be unable to tell how well these subordinates, tasked with searching for new opportunities, pursue their assignments. How will subordinates behave in the presence of superiors who cannot compare their search progress to an objective benchmark? Will they ignore feedback from management, or will it still have an impact on subordinate behavior? If so, does supervision increase or decrease exploration? These are the questions we study in this article.

In his seminal paper, March (1991) suggested that successful firms must not solely exploit extant business models but also explore novel opportunities. Such corporate exploration can best be conceived of as the process of aggregating individual search behavior through organizational structure. In line with this suggestion, a variety of prior theoretical studies have shown that when firms adapt their structure, they simultaneously also shape the direction and intensity of their search activities (see the review by Baumann, Schmidt, & Stieglitz, 2019). For example, previous work suggests that when firms divide their search into multiple subunits (Ethiraj & Levinthal, 2004) and promote decentralized decision-making (Rivkin & Siggelkow, 2003), they will increase the exploration of novel opportunities.

Notwithstanding the importance of these contributions, we suggest that they may not be conclusive in explaining how corporate hierarchies affect search in organizations; the reason being that, in addition to the hierarchy's overall impact on information aggregation, it may also affect individual exploration in other ways. Of particular importance in that regard is the dyadic authoritative superior-subordinate relationship—a constituting element of each hierarchical structure. Notably, subordinates are known to adjust their behaviors when interacting with superiors—irrespective of whether the latter reward or sanction their subordinates explicitly. In particular, when giving their staff tasks that can more easily be evaluated against known

benchmarks, such as allocating budgets across initiatives or preselecting alternative proposals for innovative projects, supervisors often instill a sense of evaluation apprehension or lack of control among their subordinates, which leads them to avoid more uncertain actions whose consequences would be visible to their superiors (Detert & Treviño, 2010; Fang, Kim, & Milliken, 2014; Glass & Singer, 1972; Keum & See, 2017; Milliken, Morrison, & Hewlin, 2003; Nai, Kotha, Narayanan, & Puranam, 2020; Reitzig & Maciejovsky, 2015).

In complex search settings, so we argue, that situation changes. More specifically, such environments hamper managers in assessing the quality of their subordinates' search actions as the superiors lack intuitive benchmarks *ex-ante*. In the absence of a prior, they only form their expectations for what qualifies as good search performance over time. Once their subordinates' average search performance no longer improves, managers may expect their employees to stop engaging in uncertain actions. However, up to that point, superiors will not be inclined to punish subordinates for engaging in distant and extended search. To the extent that subordinates are aware of the above difficulties, they may actually react quite differently to the monitoring of superiors than they would in other situations, by taking less certain actions over an extended period of time. As we argue more generally in this paper, in the absence of clear performance benchmarks, individuals will be driven by the motivation to positively shape the perceptions of their supervisor through other means. They may take actions that signal higher effort (e.g., Kyl-Heku & Buss, 1996; Lund, Tamnes, Moestue, Buss, & Vollrath, 2007). They may also take actions which indicate a willingness to improve on the current status quo and can thus be more readily justified to a superior (e.g., Baron & Ritov, 2004; Lerner & Tetlock, 1999; Zeelenberg, van de Bos, van Dijk, & Pieters, 2002). When tasked with exploration specifically, individuals will thus engage in search for longer and pursue more distant search than in the absence of monitoring supervisors, believing that such actions may serve as a substitute yardstick for superiors to assess search performance.

Motivated by the tension between known effects of the superior-subordinate relationship on subordinate behavior in other task settings and potential deviations thereof in the context of complex search, we conducted four controlled laboratory experiments with a total sample of 444 participants, who completed a complex search task in the presence and absence of monitoring superiors. This approach allows us to provide clear causal evidence for our main hypothesis that subordinates will engage in more distant and extended search when they feel they are being monitored by their superior.

Monitoring describes the process by which managers obtain information about the performance of their subordinates and then use this information to influence their actions and motivation (Niehoff & Moorman, 1993). It can thus incorporate a number of practices such as providing simple feedback about a subordinate's current work performance, an open discussion of this feedback between manager and subordinate, the provision of contingent rewards, or communicating the importance of accomplishing particular tasks and goals (see, e.g., Larson & Callahan, 1990; Niehoff & Moorman, 1993; Townley, 1993). From the subordinate's perspective, one of the most common ways monitoring by managers is likely to manifest itself is for them to receive regular performance appraisals by their superiors about the consequences of their specific actions (see e.g., Fletcher, 2001; Townley, 1993). Such appraisals typically combine two related actions—the provision of feedback by a superior, and the request to the subordinate to explain her actions as part of a feedback discussion (e.g., Elicker, Levy, & Hall, 2006; Pichler, 2012; Taylor, Tracy, Renard, Harrison, & Carroll, 1995). Guided by this prior work, in our empirical study we thus focus our attention on the effects of such performance appraisals of

subordinates by their supervisors, which we operationalize in our experimental studies by manipulating both the provision of managerial feedback and the need for subordinates to explain their actions. We then test the joint effect of these two manipulations, but also disentangle the independent contributions they may have on individual search behavior. Given the nature of search problems, superior feedback does not contain objectively novel information for the subordinate, but merely signals the manager's observation of her subordinate's actions. Nevertheless, we find that managerial appraisal increases individual search activity, both by prolonging search duration and increasing search distance. Moreover, this effect does not depend on the framing of the feedback; that is, whether it highlights past improvements or deteriorations in the search process. Additional follow-up experiments also show that the combination of feedback and justification requests causes distinctly different reactions than the mere solicitation of justifications from subordinates, which has, on its own, no effect on their search behavior. In fact, the most powerful effect we obtain appears to be caused by the mere provision of feedback by superiors to subordinates, despite it reiterating known information only. Finally, our results also show that appraisal is ineffective if it is provided by a peer or subordinate.

Our findings bear insights for scholars studying search in organizations in the Carnegie tradition. In particular, we contribute to the ongoing discussion about the relationship between search and structure (Chang, 2022; Ethiraj & Levinthal, 2004; Knudsen & Srikanth, 2014; Lee & Puranam, 2015; Rivkin & Siggelkow, 2003; Siggelkow & Rivkin, 2005) and to the emerging literature on the role of behavioral factors in search tasks (Billinger, Srikanth, Stieglitz, & Schumacher, 2021; Billinger, Stieglitz, & Schumacher, 2014; Døjbak Håkonsson et al., 2016; Tracy, Markovitch, Peters, Phani, & Philip, 2017; Vuculescu, 2017). We add to these literatures by examining the key question of how the manager–subordinate relationship, characteristic of hierarchies within most companies (Simon, 1951), shapes the subordinate's tendency to explore or exploit. Our findings appear particularly relevant against the backdrop of the known problem of under-exploration in many settings (Piao & Zajac, 2016)—providing ideas to practicing executives on how to leverage their authority to have their subordinates engage in more distant search.

## 2 | SEARCH IN ORGANIZATIONS

Search is a central activity in every organization. More often than not (organizational) decision-makers first need to search for the alternatives from which they can eventually choose (Simon, 1957). Search tasks may reach from the creation of novel products or technologies (Angus, 2019; Dahlander et al., 2016; Fleming & Sorenson, 2001; Kiss, Libaers, Barr, Wang, & Zachary, 2020; Mihm, Loch, & Huchzermeier, 2003; Simon, 1962; Tandon & Toh, 2022; Yu et al., 2019), through the development of new business strategies (Porter, 1991; Rivkin, 2000) and search for alliance partners (Yayavaram et al., 2018), to the creation of organizational capabilities (Bruderer & Singh, 1996) and the resolution of organizational design issues (Ethiraj & Levinthal, 2004; Raveendran, 2020). Organizational search is often coordinated by managers (Rivkin & Siggelkow, 2003) who integrate different search actions by their subordinates. A central question is therefore how the monitoring that superiors exert in this process affects their subordinates' search behavior. In particular, there are two related aspects of search behavior that could be affected by monitoring: individuals' propensity to search locally—for marginal adjustments from extant solutions—or distantly—trying out

significantly different recombinations of choice variables (Fleming, 2001), and the overall time spent on search compared to simple exploitation of already existing solutions (e.g., Billinger et al., 2021).

## 2.1 | How organizational hierarchy affects search—What we (do not) know

Hierarchical structures and their effects on corporate performance have been subject to study across a wide range of earlier studies (see, e.g., Joseph & Gaba, 2020). Prior research has, among others, studied incentive conflicts in principal–agent relationships (Aghion & Tirole, 1997; Alonso, Dessein, & Matouschek, 2008), and the influence of hierarchical structures on organizational decision making from an informational perspective (Csaszar & Eggers, 2013; Sah & Stiglitz, 1986). Our work differs from these earlier works in that we focus on the purely behavioral consequences of hierarchies for the actions of subordinates.

A number of prior investigations have documented that such subtler behavioral effects in hierarchies exist and may have serious implications for firm performance. For example, Fang and colleagues highlight that subordinates "may [...] omit negative feedback when reporting to their bosses" (Fang et al., 2014, p. 1187). Reitzig and Maciejovsky (2015) provide related field and experimental evidence. They demonstrate that subordinates, when fearing negative feedback from their direct superiors, would pass on fewer proposals to their bosses the more uncertain they were about the quality of the proposals. The authors also show that a perceived lack of control on the part of the subordinates might lead the latter to invest less effort in proposing uncertain alternatives that would require a lot of explanation to their superiors, irrespective of the fear of being evaluated negatively. Finally, Nai et al. (2020) demonstrate that subordinates can form different deservingness expectations towards their superiors, depending on the secrecy of the environment in which they operate. What these behavioral contributions thus jointly suggest is that subordinates adjust their behaviors towards their superiors based on the monitoring they are exposed to, and the broader feedback culture that surrounds them. It, therefore, stands to reason that subordinates, tasked with searching for new alternative solutions to problems, will also react to their superiors' monitoring in some form. Just how exactly it will affect subordinates' search actions is an open question to date, however.

## 2.2 | Organizational hierarchies—Search in complex environments

At the most basic level, transferring the results from prior studies to the search context would suggest that subordinates, in the presence of superiors, should avoid taking uncertain search actions. At first glance, one might thus assume that individuals, when observed by managers, should favor local as opposed to distant search, and should stop search sooner rather than later, if their opportunity costs of time matter. Upon reflection, however, this inference seems problematic, as it falls short of acknowledging the different boundary conditions under which coordinated search unfolds within organizations; most importantly, that much of the search organizations engage in is complex, meaning that nontrivial interdependencies exist among relevant choice variables (Levinthal, 1997; Simon, 1962). Experimenting with one variable might affect other variables, and potentially change the overall outcome and obfuscate inference (Billinger et al., 2014). Where and for how long a subordinate should search is not clear *ex ante*.

Even worse, when maneuvering uncharted territory and facing a complex search problem, superiors may not even be able to benchmark their subordinates' search against any reasonable prior. As such, they initially cannot compare subordinates' search performance to anything other than what is being discovered in the actual search process. For as long as this benchmark moves with every new solution the subordinates unearth, there is no reason to punish employees for either search strategy they adopt—be it local or distant search; it is to be expected that taking uncertain search actions may be objectively just as valuable as others, and prolonging the period of search may even help reduce the uncertainty around what qualifies as good search performance and when search should stop eventually. Subordinates who see through their superiors' challenge will thus not feel inhibited from taking uncertain actions for as long as no clear benchmark exists[1]—such as making long jumps, searching for remote solutions, and searching longer than usual, instead of merely incrementally modifying current approaches and locking in early on solutions (Cyert & March, 1963; March & Simon, 1958; Mihm, Loch, Wilkinson, & Huberman, 2010; Nelson & Winter, 1982). Different from a selection or allocation scenario, they need not assume that their managers can and will evaluate their performance against initially known benchmarks, so that evaluation apprehension would make them avoid taking uncertain actions from the beginning (Detert & Treviño, 2010; Milliken et al., 2003). Similarly, it seems hard to imagine why a lack of control would create a behavioral bias against subordinates engaging in prolonged or distant search.

## 2.3 | How managerial monitoring should affect subordinates' search

More likely, so we argue, will the superior–subordinate relationship in complex environments create a different type of bias which may, in fact, lead subordinates to engage in more distant and extended search when being monitored, compared to the ones they might take in isolation. Specifically, we suggest that in the absence of objective performance metrics, individuals will instead consider more distant and prolonged search as actions that signal effort and willingness to strive for further gains, hoping that their superiors may appreciate them against alternative yardsticks to assess search performance. Here, two closely related mechanisms would appear to reinforce one another in creating this overall effect.

First, research in organizational behavior suggests that subordinates frequently strive to impress their supervisor through high levels of so-called "industriousness," which is comprised of various work-related effortful behaviors such as hard work on a given task, or the acquisition of new relevant knowledge (Kyl-Heku & Buss, 1996). For example, Kyl-Heku and Buss (1996) report results from both laboratory experiments and survey-based field data, showing that individuals frequently engage in behavior designed to signal their willingness to work hard and to obtain novel knowledge necessary to complete important tasks, in an attempt to advance professionally. Surveys conducted by Lund et al. (2007) with professionals in two different industries resulted in very similar findings, again showing that individuals frequently attempt to positively shape their superiors' perceptions by engaging in high levels of work-related effort. Taking these findings seriously, we argue that subordinates, when having to report to their superiors who are not able to objectively assess their performance, should tend to engage in

[1]Note that when an agent explores a complex landscape, benchmarks for each individual search result will only be formed over time. During the initial stage of the exploration phase it is safe to assume that these benchmarks do not exist at all. The longer the agent searches, the better the benchmark becomes, but it will never be perfect.

more distant and extended searches than they would otherwise. This is because searching for remote solutions signals higher levels of search effort, and thus exemplifies a willingness to spend effort on something better than merely pursuing incremental changes of existing solutions or not searching at all. Similarly, they should extend their active search period, irrespective of how far they search. By doing so they signal that they actively seek to reduce the uncertainty around the threshold of what qualifies as a good search result, and do not stop exploring prematurely.

Second, more distant and longer search may also appear more easily justifiable to managers who are not able to directly observe objective search performance. As frequently shown in the large literature on the effects of accountability, subordinates, especially when they feel that their actions are being observed by others, will seek to take actions for which they can more easily provide justifications than for others (see e.g., Lerner & Tetlock, 1999). Importantly, prior research suggests that decisions that involve being active and signal a willingness to strive for further gains are fundamentally easier to justify to others—and especially to a manager—than decisions that signal contentment with the status quo and a desire to settle for a potentially suboptimal outcome, all else being equal (Baron & Ritov, 2004; Zeelenberg et al., 2002). Subordinates should thus perceive more distant search to be more easily justifiable to superiors than passive exploitation, causing them to increase the distance of their search in the presence of monitoring managers. Along the same lines, prolonging their active search phase, when given the chance, may help employees justify their actions. Irrespective of how far employees intend to search, delaying their decision to stop this process should be more easily justified to a superior.

In general, although related, these two mechanisms are different from one another, as they could in principle occur independently, and the behaviors they produce might differ in settings different from ours.[2] However, importantly, as we argued previously, it is a special feature of the challenge that comes with the search for new alternatives that both of these effects that might be triggered by monitoring—the desire to signal industriousness and the motivation to take more easily justifiable actions—are likely to create a unidirectional "action bias" among subordinates to engage in more distant and extended search, when they are subject to monitoring by superiors. Based on these arguments we therefore posit the following:

> **Hypothesis.** Subordinates will engage in more search activity, in terms of increasing both distance and duration, when being monitored by managers, all else being equal.

## 3 | METHODS

We conducted four experimental studies[3] to test our main hypothesis and to provide insights into the underlying drivers and boundary conditions of our findings.

[2]To appreciate our point, think of an alternative scenario in which subordinates are not tasked with search, but with allocating budgets to known alternatives. Justification would likely lead subordinates to stick close to historical allocation patterns; industriousness, on the other hand, might lead them to argue for reallocations.

[3]We complied with all institutional and national ethical guidelines for experiments involving human subjects. All participants gave their consent to participate in the study.

As we outlined previously, across all studies our operationalization of monitoring is strongly guided by the types of measures managers are likely to employ in practice in order to monitor the search activities of their subordinates. In particular, a typical form of monitoring in organizational practice is an appraisal of subordinates' current actions in the form of summarized information about their relative performance progress, which managers might use to steer subordinates' actions in a certain direction (e.g., Taylor et al., 1995). Moreover, such reflection is rarely provided on its own but is instead accompanied by an opportunity for subordinates to explain—or justify—the reasons for their specific actions, thereby providing important feedback to their superiors (e.g., Elicker et al., 2006; Pichler, 2012). Based on these observations, we thus operationalize monitoring in our experimental studies as a recurring managerial appraisal consisting of summarized performance feedback and justification requests.

The main goal of Study 1 is to establish evidence for our main hypothesis. We also explore how the framing of the feedback that is part of the appraisal might influence our results, as the framing of the manager's feedback may signal subjective satisfaction or dissatisfaction by the superior. The results of this analysis can be found in the Technical Appendix (Section 3.1).

In Study 2, we focus on demonstrating that the source of the appraisal is driving the predicted increase in search activity. For this purpose, we directly compare the effects of appraisals—again consisting of feedback and justification requests—issued by superiors with those issued by a colleague or a lower-ranked co-worker.

Next, having explored only the joint effects of feedback and justification requests on individual search behavior in Studies 1 and 2, in Study 3, we attempt to disentangle the effects arising from each of these two components of managerial appraisal. Finally, in Study 4, we attempt to rule out the possibility that managerial feedback could merely cause participants to think more actively about the task and the optimal level of search, which would potentially provide an alternative explanation of our findings.

## 3.1 | Experimental task

Our studies build on the "alien game," an experiment based on the NK model (Levinthal, 1997) first deployed by Billinger et al. (2014), which has been established as a useful way to investigate the search behavior of individuals in a controlled environment (e.g., Billinger et al., 2021). The NK model offers an appropriate test bed for our hypothesis for several reasons. First, the combinatorial nature of NK models allows depicting intra-organizational interdependencies well (Ganco, Kapoor, & Lee, 2020; Rivkin & Siggelkow, 2003; Siggelkow & Rivkin, 2005, 2009). Second, there is empirical support for the validity of the main search tenets expressed in that type of model (Fleming & Sorenson, 2001; Ganco, 2017; Lee & Alnahedh, 2016; Lenox, Rockart, & Lewin, 2010). Third, the key variable of interest in this article—search activity (capturing both search distance and search duration, see below for more details)—features prominently in this setup and can be conceptualized more easily than in other models, notably multi-armed bandits (Sutton & Barto, 1998). Fourth and related, our approach is thus more amenable to comparisons with prior experimental work in this domain (e.g., Billinger et al., 2014; Billinger et al., 2021; Gavetti & Levinthal, 2000; Posen & Martignoni, 2018; Rivkin, 2000; Rivkin & Siggelkow, 2003; Siggelkow & Rivkin, 2005, 2009; Tracy et al., 2017; for an overview see Baumann et al., 2019).

In its very essence, the NK model opens up a complex performance landscape as a space of interdependent combinatorial alternatives, within which search agents must uncover the value

of alternative solutions. Here, the parameter N represents the number of choice attributes that an alternative can have. Attributes are binary—they can be chosen or not—yielding a total of $2^N$ possible combinations, and their contribution, or combination-specific payoff, to the performance landscape, is determined by a random number drawn from a uniform distribution ranging from 0 to 1. The NK algorithm thus creates an environment where sets of possible attribute combinations are mapped onto their respective payoff values (Levinthal, 1997).

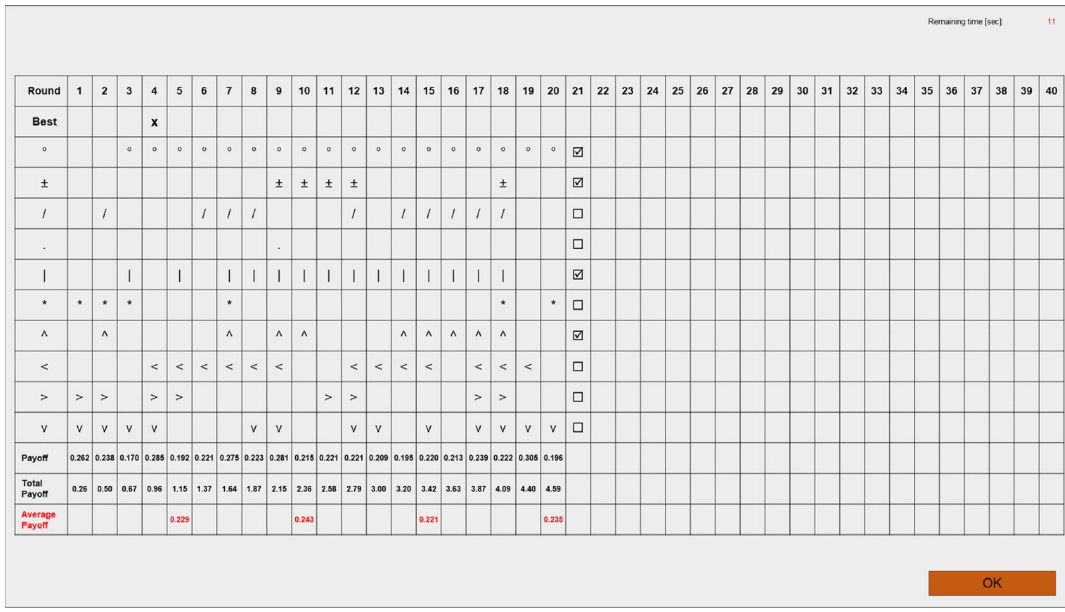
Attributes can be more or less interdependent with one another, as captured by the parameter K. The higher the value of K, the more attributes depend on one another: at one extreme, K can take the value 0, indicating that decision attributes are completely independent of one another. Maximum interdependence between choice variables is given where $K = N - 1$, indicating that the contribution of any attribute is determined by its own value, as well as by the value of all the other attributes within the performance landscape. A high-parameter $K$ contributes to more local peaks, that is, to more choice combinations that can be considered distinctively more successful than others, and thus to a more "rugged" performance landscape. The more rugged a performance landscape, the harder it gets to distinguish local peaks from the global maximum, which is the overall best-performing combination of choices (Baumann et al., 2019), of the performance landscape, and thus the harder the identification of truly successful choice combinations (Levinthal, 1997).

In all studies, participants are asked to combine 10 interdependent binary design features (parameter N) into an artistic illustration for aliens. In all experimental groups, participants select different combinations over 40 separate rounds and are informed of the payoff resulting from their selected combination at the end of each round. Setting N equal to 10 opens up a performance landscape with 1,024 ($2^{10}$) possible combinations, providing enough opportunity to search the topography without finding exceptionally successful combinations too quickly.

All participants play the alien game twice. Across the two games, we vary the complexity parameter $K$ to create two nontrivial search tasks, albeit with different levels of complexity. We choose complexity parameter $K = 9$ and $K = 2$ to represent the upper and lower end of a so-called complexity spectrum. As the number of interdependent binary design features $N$ is equal to 10 in our experimental task, a complexity parameter of $K = 9$ opens up a maximally rugged performance landscape on which the search for successful choice combinations can be considered extremely complex due to interdependencies between all choice variables.

A performance landscape with a complexity parameter of $K = 2$, on the other hand, creates a performance landscape with fewer peaks and fewer interdependencies between choice variables, but still represents a search task complex enough that simple search strategies such as "hill-climbing" in the neighborhood of recently uncovered solutions (Levinthal, 1997) will not guarantee the discovery of successful choice combinations. As opposed to a complexity parameter of $K = 0$ or $K = 1$, $K = 2$ thus creates a performance landscape with enough peaks that search can still be considered nontrivial and complex for individual agents.

To inform participants of the task's complexity level, that is, whether a participant is about to search on a performance landscape where $K = 2$ or $K = 9$, we describe to them their respective alien customers' taste in artistic illustrations prior to each game. When providing this customer-specific information to participants, we stress the level of interdependence between choice parameters to give participants an impression of the complexity level of the search task. To foreshadow that participants will be dealing with a less complex, yet nontrivial search task ($K = 2$), customers are described to "favor single design features over facetted combinations in an artistic illustration" and to "have 'favorite' symbols which are valued higher than others" and that "constructing artwork around such symbols is a good idea". Therefore, customers "do




| Round | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Best | | | | x | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ○ | | | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ☑ | | | | | | | | | | | | | | | | | | |
| ± | | | | | | | | | ± | ± | ± | ± | | | | | | ± | | | ☑ | | | | | | | | | | | | | | | | | | |
| / | | / | | | | / | / | / | | | | / | | / | / | / | / | / | | | ☐ | | | | | | | | | | | | | | | | | | |
| . | | | | | | | | | . | | | | | | | | | | | | ☐ | | | | | | | | | | | | | | | | | | |
| \| | | | \| | | \| | | \| | \| | \| | \| | \| | \| | \| | \| | \| | \| | | | ☑ | | | | | | | | | | | | | | | | | | |
| * | * | * | * | | | | * | | | | | | | | | | | * | | * | ☐ | | | | | | | | | | | | | | | | | | |
| ^ | | ^ | | | | | ^ | | ^ | ^ | | | | ^ | ^ | ^ | ^ | ^ | | | ☑ | | | | | | | | | | | | | | | | | | |
| < | | | | < | < | < | < | < | < | | | < | < | < | < | | < | < | < | | ☐ | | | | | | | | | | | | | | | | | | |
| > | > | > | | > | > | | | | | | > | > | | | | | > | > | | | ☐ | | | | | | | | | | | | | | | | | | |
| v | v | v | v | v | | | | v | v | | | v | v | | v | | v | v | v | v | ☐ | | | | | | | | | | | | | | | | | | |
| Payoff | 0.262 | 0.238 | 0.170 | 0.285 | 0.192 | 0.221 | 0.275 | 0.223 | 0.281 | 0.215 | 0.221 | 0.221 | 0.209 | 0.195 | 0.220 | 0.213 | 0.239 | 0.222 | 0.305 | 0.196 | | | | | | | | | | | | | | | | | | | | |
| Total Payoff | 0.26 | 0.50 | 0.67 | 0.96 | 1.15 | 1.37 | 1.64 | 1.87 | 2.15 | 2.36 | 2.58 | 2.79 | 3.00 | 3.20 | 3.42 | 3.63 | 3.87 | 4.09 | 4.40 | 4.59 | | | | | | | | | | | | | | | | | | | | |
| Average Payoff | | | | | 0.229 | | | | | 0.243 | | | | | 0.221 | | | | | 0.235 | | | | | | | | | | | | | | | | | | | | |

FIGURE 1 Schematic screenshot of the search mask

not care that much about overall looks of the artwork" but "prefer to have certain symbols in the piece of art they purchase." To prepare participants that they will be dealing with a highly complex search task ($K = 9$), they are informed that it is rather complicated to identify customers' tastes in art and that customers "do not strictly prefer certain symbols but care more about the 'overall' look of the artwork.". Participants are told that "concentrating on specific symbols might not be the best strategy" with these customers and that "a certain combination of seemingly unpopular symbols" can yield a high payoff. The performance landscapes created with complexity parameters $K = 2$ and $K = 9$ remain unchanged for all treatment groups and studies to ensure comparability across the experiment.

## 3.2 | General procedure

We use the programming software z-Tree (Fischbacher, 2007) to create a version of the alien game and to conduct the experiment. Before beginning the main study, all participants complete a mandatory introduction session that includes detailed descriptions of the task and five trial rounds to familiarize themselves with the game. We do not determine a fixed starting position in any of our treatments. The combination of symbols submitted by a participant in the first round thus constitutes one's individual starting position (for a detailed analysis of the starting position, see Section 2.3 in the Technical Appendix). Participants of both the treatment and control groups can see all previously submitted combinations and their respective payoffs. A new search round is started with the previous combination already implemented. Figure 1 illustrates the alien game's universal search mask. The search interface resembles a large table where rows represent the symbols available to the participant and columns indicate the configuration rounds of the alien game. In each round, participants choose the symbols they want to use for their artistic illustration by ticking the corresponding boxes placed in each row next to

the configuration element. Artistic illustrations are submitted by clicking on the "OK" button at the bottom right of the screen. Previously submitted combinations, their respective payoffs, and a running total payoff are visible at all times in all experimental conditions. Additionally, participants are informed of their best-uncovered combination so far with an X in the respective column of the row "Best" right below the round display. After every fifth round, the average payoff of the previous five round payoffs is displayed in the last row of the search mask.

Each experimental session lasts approximately 90 min, giving participants 15 min to complete the mandatory introduction and 30 min each to complete the two consecutive types of the alien game. Participants are paid in the remaining 15 min. Participants are told that their aim is to maximize their final payoffs over all 40 rounds of one game, and that they will receive their cumulative payoffs over the total of 80 rounds as experimental remuneration.[4] Thus, the search task in our experiment is similar to the one used in Billinger et al. (2021), but differs from the setup in Billinger et al. (2014), where participants are required to find the single highest combination-specific payoff. We find the quest for maximizing overall payoffs to be more appropriate for mimicking managerial monitoring of subordinate search processes in an organizational context. That is because, in contrast to search problems concerned with identifying one individual superior solution, rewarding searchers based on cumulative payoffs introduces additional opportunity costs for search, similar to real-world situations such as the development of novel products or new business models (Billinger et al., 2021; Markides, 1999), which, in turn, creates an incentive to avoid costly (distant) exploration by exploiting known alternatives. The experimental procedure is illustrated in the Technical Appendix (Figure S2).

## 3.3 | Variables

### 3.3.1 | Dependent variable

*Search activity* (*varies at participant-round level*). Central to all observations is the search behavior of individuals. Participants search the underlying NK landscape to find attractive combinations, either locally or more distantly, and some participants may search longer than others. Measuring this search activity thus renders the search process of any search agent quantifiable and their search behavior measurable. Search activity is measured as the number of changed geometrical elements per round (Billinger et al., 2014), which corresponds to the Hamming (1950) distance frequently employed to measure organizational search (e.g., Billinger et al., 2014; Levinthal, 1997; Rivkin, 2000). Given that participants are working with 10 geometrical elements, search activity takes a discrete value between 0 and 10, 0 representing no change at all and 10 constituting maximum change of all attributes. To observe a searcher's willingness to move away from a *successful* combination, potential deviations from prior maximum-payoff submissions are measured and the round-specific Hamming distance is computed with the best-performing combination submitted so far.[5]

[4]A post-study manipulation check, conducted with 49 participants, confirmed that the vast majority of subjects (>90%) correctly understood the experimental instructions as regards compensation—notably, that the amount a subject would receive in exchange for participating in the experiment would solely depend on her performance in the game and not on any actions by the superior.

[5]In addition to measuring search distance relative to the best-performing solution found so far, we also compute distances to the last position visited and to the closest combination submitted. Our results remain robust (See Technical Appendix, Table S3).

*Duration of active search* (*varies at participant level*). Duration of active search is the sum of all periods in which search activity is different from zero. Conceptually, this allows capturing for how many periods an individual attempts to uncover new solutions.

*Search distance* (*varies at participant-round level*). Search distance is defined as search activity conditional on individuals actively searching (i.e. search activity being greater than zero).

### 3.3.2 | Independent variables

*Treatment dummies* (*varies at participant level*). In all studies, we code the treatment variable as 1 to indicate a treatment group, and 0 to indicate the control group.

*Task experience* (*varies at session level*). Participants always play two versions of the alien game, once on a less rugged performance landscape, with $K = 2$, and once on a maximally rugged search landscape, with $K = 9$. The order in which participants are exposed to the two landscapes is determined randomly. The variable is coded 1 for participants in the second round of the experiment, and 0 for participants in the first round.

*Payoff feedback* (*varies at participant-round level*). This variable indicates the objective configuration-specific payoff achieved in the prior round, which is ultimately determined by the underlying NK algorithm.

*Complexity* (*varies at session level*). This binary variable indicates the ruggedness of the underlying performance landscape and thus represents search-task complexity. A value of 1 indicates the highly complex search task, where $K = 9$, whereas a value of 0 represents a less complex search task, where $K = 2$.

*Round number*. This variable captures the specific round of the search task and ranges from 1 to 40.

## 3.4 | Estimation strategy

Primarily, we use the nonparametric Wilcoxon rank-sum test (WRT) to investigate differences between treatments using participants as units of analysis. Since the variable *Search activity* is non-normally distributed (Shapiro–Wilk test: $p = .000$), a WRT is considered more efficient (Conover, 1998) and we use the *Average search activity* per participant per condition. As additional analysis, we employ Poisson regressions with a participant-round panel with standard errors clustered at the individual level to examine the treatment effects on *Search activity*. In the Technical Appendix, we also replicate the regressions using an ordinary least squares model with random effects, as some treatments are stable for participants, at the round and individual levels (Tables S1 and S2). The results remain robust.

# 4 | RESULTS

## 4.1 | Study 1: Managerial appraisal

Participants are randomly allocated to the treatment or control group. The treatment group exposes participants to the following manipulations: (1) periods of search are regularly

interrupted by managerial interventions; (2) participants receive managerial appraisal of their search performance to regularly remind them of their attempts to improve their search outputs—information of which they are already objectively aware; and (3) participants are required to regularly justify (changes in) search outcomes and their choice of search parameters to a virtual manager, which gives them a chance to explain their past actions to their virtual boss. Thus, our treatment creates the perception of the general presence of a supervisor for participants and exposes them to appraisal by this supervisor. Participants randomly allocated to the control group, on the other hand, play the alien game without managerial interventions and do not have to provide justifications for past search actions.

### 4.1.1 | Participants

We recruited 148 participants[6] (72 male, 74 female, 2 nonbinary; mean age = 24.8 years) from a subject pool of a large research-based university. All experiments were conducted at an experimental laboratory equipped with 28 computers. For Study 1, which tests the phenomenon-driven effect of managerial appraisal, we conducted two independent testing sessions for each experimental group with the aim of recruiting roughly 100 participants for the treatment group and 50 for the control group. All participants were undergraduate or graduate students from a broad variety of fields of study. Participants received their cumulative payoffs over all rounds (mean payoff = €21.97) as experimental remuneration.

### 4.1.2 | Procedure

After five search rounds, participants of the treatment group are interrupted by their virtual alien boss. The Technical Appendix provides a schematic timeline of treatment in the experimental groups (Figure S3). Subsequently, they are given a summary of their round-specific combinations and combination-specific payoffs over the previous five rounds. As of round 10, that is during the second managerial intervention, they receive appraisal of their search performance over the previous five rounds from their virtual manager. Additionally, the average payoff achieved during the preceding five rounds is displayed and serves as the point of reference for the virtual manager to provide managerial appraisal.

Participants are presented with the number of combination-specific payoff increases or decreases over the previous five rounds compared with the previous average payoff. Instructions read "Compared with your previous average, your payoffs in this section were [*above/below*][7] average X out of 5 times. Please explain and justify your search process using the textbox below". Participants are then required to justify the number of changes and their choice of search parameters to their virtual manager. The Technical Appendix contains illustrations of the two types of managerial intervention (Figures S4 and S5). Note that the managerial appraisal

[6]Note that for subsequent studies, members of the subject pool who previously took part in an alternate version of the alien game were no longer eligible for participation.

[7]Half of the participants in the treatment group received the positively framed feedback (above) and half the negatively framed feedback (below). The results showed that there were no differences between the two groups. See Technical Appendix, Section 3.1 for a more detailed analysis.

provided by the virtual manager represents a mere observation of the subordinate's past search performance and includes no new or subjective information.

Rather than aggregating successes or failures to a single measure, we introduced this frequency score to direct participants' attention to understanding how often recent search improved the prior status quo. Our experimental task is specific, as we opted for an experiment that specifically contains opportunity costs for more search [e.g., in contrast to Billinger et al. (2014)]. In real-world settings, costs for (more distant) search are often implicit (e.g., more time and more material to be invested). Our frequency score should thus adequately reflect the way in which managers illustrate past performance of innovation pipelines that yield discrete ideas for novel business models or products. A number of other corporate scenarios also come to mind that map onto our representation. Marketing campaigns, for example, where managers care about how often new campaigns outperform previously successful strategies (Kitchen, Brignell, Li, & Jones, 2004). Similarly, procedural legal strategies, where lawyers focus on how often a newly developed legal strategy employed in court beats their earlier standard approach, might be another setting to which our setup maps well (Taylor & Walsh, 2002). Yet, aggregated comparisons of current with past search progresses might still be more representative of some other organizational settings. Therefore, we tested the robustness of our results with a different representation of feedback. A detailed description of an adapted version of the experiment can be found in Section 4.6.2 of this article. Importantly, the results we obtain support our main findings.

As all treatments include 40 rounds, participants experience a managerial intervention eight times in total. We increase the number of configuration rounds of the alien game from 25 (Billinger et al., 2014) to 40 rounds. This way, (1) participants have more time to develop a search strategy, and (2) sufficient managerial interventions can be conducted. Participants in the control group can explore or exploit the landscape for 40 rounds without managerial interventions.

To ensure comparability among experimental conditions, the search mask employed by participants of the control group still includes previously submitted combinations and their respective payoffs, a running total payoff, the best-performing combination uncovered so far, and averaged payoffs of previous combinations in a five-round interval (Figure S6). As described previously, participants play the alien game twice in all experimental groups: once on a maximally rugged performance landscape ($K = 9$) and once on a less rugged, but still complex performance landscape ($K = 2$). A counterbalanced experimental design controlled for potential learning effects across sessions.

### 4.1.3 | Results

Figure 2 presents an overview of the average search activity across treatment and control groups and rounds over time. In line with previous findings, the average search activity of all experimental groups decreases as the game progresses (Fleming, 2001; March, 1991). Importantly, and consistent with our hypothesis, it is apparent that, throughout the experiment, participants who receive managerial appraisal engage in more search than those in the control group.

We compare the search activity between treatment and control groups and find that managerial appraisal leads to more search throughout the game. For each comparison, we show the results of a WRT and the coefficient of a Poisson regression analysis with robust standard errors (Table **1**).

In particular, the comparison of search activity over all rounds between pooled treatment groups and the control group reveals higher search activity for participants in either of the treatment groups (WRT: $N = 148$, $z = -2.374$, $p = .018$). The average search activity of participants

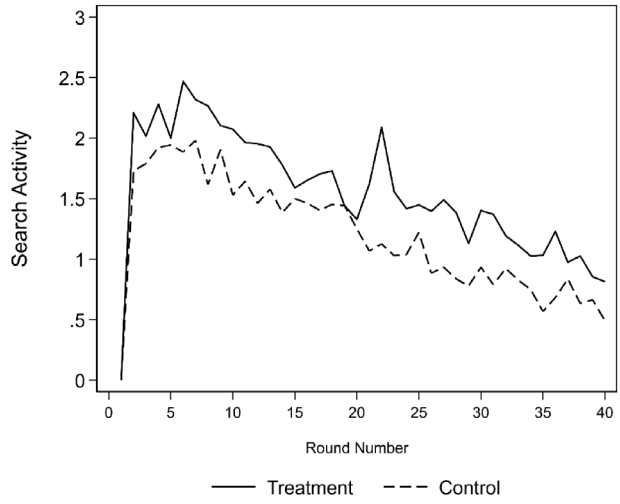



FIGURE 2 Study 1: Search activity of treatment and control groups over time

in the treatment groups is 1.56, meaning that participants change three choice variables for every two rounds, whereas participants in the control group display an average search activity of 1.20, which represents six changes in five rounds. The results of the Poisson regression are highly consistent with this analysis (Poisson: $\beta = .256$, $p = .008$, CI = [0.066, 0.447]). Overall, the results provide clear support for our hypothesis.

## 4.2 | Study 2: Sources of appraisal

How important is being a superior to the effectiveness of search appraisal? Peers or even lower-ranked staff can also provide appraisal of a focal subordinate's individual performance (Morrison, 1993) and can ask for explanations. However, we have argued that regular appraisal and justification demands will be particularly effective if they originate from a superior. In the following, we test whether the monitoring coming from a superior is a strictly necessary condition for our observed effect of appraisal to emerge, or whether interventions by peers or even subordinates might be sufficient to affect individuals' search behavior.

### 4.2.1 | Procedure

For Study 2, we recruited 103 participants[8] (40 male, 61 female, 2 nonbinary; mean age = 22.7 years) in which we vary the hierarchical position of the person providing appraisal and justification demands in the fictitious alien company. In the two treatments, we replace the virtual manager of our main study with (1) a virtual colleague, to mimic a peer-to-peer relationship; and (2) a virtual intern, to observe the effect of communication with an individual on a lower hierarchical level. Even though regular interventions originate from the (equally or

[8]Refer to the Technical Appendix (Section 1.3) for treatment-specific sample sizes. We performed a power analysis based on Study 1 and calculated that we needed about 80 participants per study to achieve a power of 0.80.

**TABLE 1** Poisson regression with robust standard errors.

| Dependent variable: Search activity | Model 1 |
|---|---|
| Appraisal | 0.256 |
| | (0.097) |
| | [0.008] |
| Task experience | −0.279 |
| | (0.054) |
| | [0.000] |
| Complexity | 0.121 |
| | (0.054) |
| | [0.026] |
| Payoff feedback | −6.794 |
| | (0.490) |
| | [0.000] |
| Round number | −0.014 |
| | (0.003) |
| | [0.000] |
| Constant | 2.379 |
| | (0.146) |
| | [0.000] |
| Log-likelihood | −22,016 |
| Pseudo $R^2$ | 0.0784 |
| Number of observations | 11,544 |
| Number of total searchers | 296 |
| Subject cluster | Yes |

*Note*: Robust standard errors (clustered at the participant level) in parentheses; $p$ values in brackets.

lower-ranked) co-worker, participants still receive appraisal and are then required to justify their search performance over the five preceding rounds to the person in charge.[9] The rest of the setup remains identical to Study 1.

### 4.2.2 | Results

We do not find support (WRT: $N = 104$, $z = -0.423$, $p = .673$) for the idea that average search activity of participants who receive appraisal from a colleague differ from those of the control group. Likewise, appraisal issued by an intern does not influence search activity levels of

[9]Participants in the treatment conditions still receive positively or negatively framed managerial appraisal but, for the sake of simplicity, we subsequently report only findings that compare the pooled treatment groups with the control group. The individual treatments were still tested separately during our analysis and their results can be accessed in the Technical Appendix in Section 2.6.

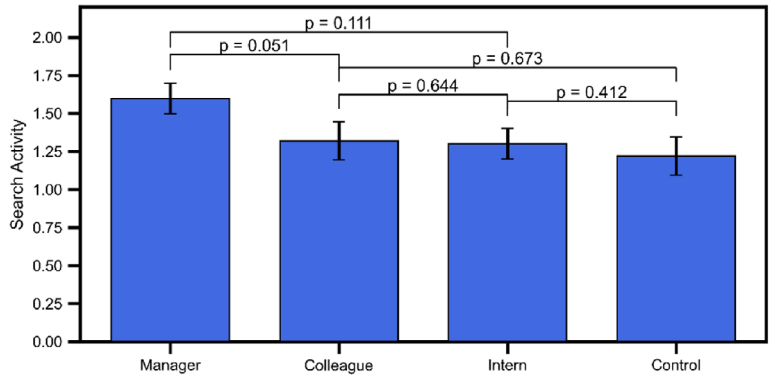



**FIGURE 3** Mean search activities, standard errors, and *p* values for experimental treatments of Study 2

participants (WRT: $N = 103$, $z = -0.821$, $p = .412$). Whereas the intern and colleague treatments do not differ from each other (WRT: $N = 103$, $z = -0.465$, $p = .642$), both treatments lead to lower search activity than the treatment (Manager) condition from Study 1 (Colleague: WRT: $N = 148$, $z = 1.950$, $p = .051$; Intern: WRT: $N = 147$, $z = 1.595$, $p = .111$). In summary, for appraisal to be effective, it needs to originate from a manager.

The findings reported above bear interesting insights for two reasons: first, we can reject a common appraisal effect that leads to participants searching more under external control as an alternative explanation. Appraisal given by a manager encourages the subordinate to search more for successful combinations, appraisal provided by other subordinates does not. Second, our results identify the managerially initiated intervention to be a determinant of subordinate search behavior. One could argue that this finding is due to the mere interruption that forces participants in the treatment groups to stop their search process and to reposition themselves on the performance landscape. Yet, treatments in which the appraisal comes from colleagues or interns do not show the same effects.[10] Thus, we can conclude that it is mainly the asymmetry inherent in the superior–subordinate relationship that contributes to participants searching more in the appraisal treatment of Study 1 and that the presence of a superior is thus necessary for our observed effects of appraisal to emerge. Figure 3 depicts an overview of mean search activities, standard errors, and *p* values of experimental treatments for Studies 1 and 2.

## 4.3 | Study 3: Disentangling the two mechanisms of managerial appraisal

Managerial feedback is rarely provided on its own but is instead typically accompanied by the opportunity for subordinates to justify the reasons for their specific actions (e.g., Elicker et al., 2006; Pichler, 2012) in order to give them a voice in the overall appraisal of their performance (e.g., Lizzio, Wilson, & MacKay, 2008; Taylor et al., 1995). Representing this typical organizational practice in our previous experimental treatments, participants both received managerial feedback on their prior performance and were asked to explain or justify their actions to their superior. This poses the question if only one of these two components of appraisal is mainly driving our previously observed results, or if both of them need to be present

[10]Note that this criticism is also refuted by the results of the pure-justification treatment from Study 3 below.

simultaneously to affect individual search behavior. The goal of Study 3 is to address this question by empirically disentangling the separate effects of feedback and justification requests.

### 4.3.1 | Procedure

To that end, we conduct a study with 120 participants (38 male, 81 female, 1 nonbinary; mean age = 24.8 years), in which participants were exposed to only one part of the managerial intervention during their search process. In a pure-justification treatment, participants were regularly asked by their virtual alien boss to provide an explanation for their search performance over the preceding five rounds, while a pure-feedback treatment repeatedly exposed subordinates to managerial appraisal of their search results over the preceding five rounds. Otherwise, the setup from Study 1 remains unchanged.

### 4.3.2 | Results

Study 3 shows that there is no indication that search behavior between the pure-justification treatment and the control group differs (WRT: $N = 100$, $z = 0.480$, $p = .632$), demonstrating that it is not the process of explaining one's search performance alone that drives search levels when exposed to regular managerial intervention.

As for the pure-feedback treatment of the alien game, we can observe higher levels of search activity (WRT: $N = 124$, $z = -2.491$, $p = .012$) with average values of 1.39 among participants receiving managerial appraisal, compared to their peers in the control group and in the pure justification treatment (WRT: $N = 120$, $z = 3.265$, $p = .001$).

In summary, the previous results indicate that asking subordinates to explain their behavior without providing managerial feedback does not affect individual search behavior, whereas providing managerial feedback without demands for an explanation from the subordinate has an effect similar to the combination of justification demands and managerial feedback explored in the main part of our study. See Figure 4 for comparisons of mean search activities, standard errors, and $p$ values of all experimental treatments of Study 3.

One possible explanation for the former result is that asking the subordinate for an explanation without specifying what exactly this explanation should refer to was not perceived as meaningful by our study participants. Providing some supportive evidence for this suggestion, previous research emphasizes the need to link demands for explanations or justification to specific outcomes or actions in order for such interventions to be effective in changing subordinate behavior (e.g., Wood Jr. & Winston, 2005). Similarly, our result concerning the effectiveness of managerial feedback even in the absence of explanation requests is generally consistent with previous work on performance appraisal and feedback that has demonstrated the strong independent effects of such feedback provided to subordinates in the absence of any other intervention on various outcomes, such as subordinates' learning behavior, satisfaction, and performance improvement (see, e.g., Levy & Williams, 2004).

## 4.4 | Study 4: Exploring the effects of landscape size

We previously argued that the observed increase in subordinates' individual search activities caused by managerial appraisal results from the subordinates' desire to show effort, thereby to

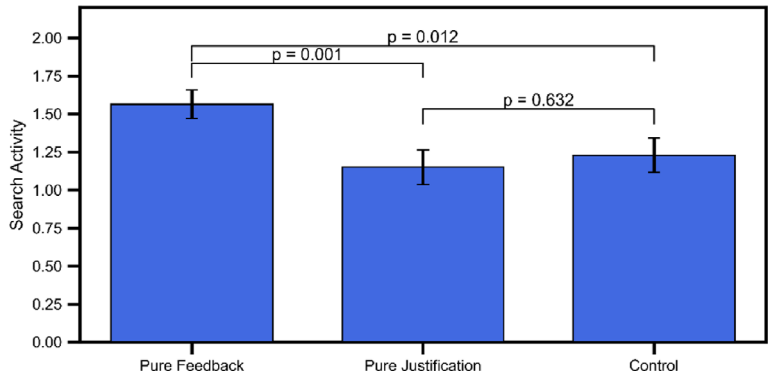



**FIGURE 4** Mean search activities, standard errors, and *p* values for experimental treatments of Study 3

please their superiors, and to engage in actions that are subjectively easier to justify. A possible *alternative* explanation might be that appraisal causes participants to think more thoroughly about the search task and thus realize that—given the relatively large landscape they conduct their search on—more search would likely lead to a higher expected performance. To test for this possibility, in Study 4, we presented participants with a considerably smaller landscape in which the optimal degree of search is therefore lower. If managerial appraisal mostly activated more thorough thinking about the task at hand, we should thus observe a smaller effect of appraisal on search activity. If, however, appraisal created a bias to engage in actions that signal effort and enable easier justification, as per our theory, its effects should be independent of landscape size.

### 4.4.1 | Procedure

We conduct an additional study with 73 participants (34 male, 39 female; mean age = 22.8 years) in which participants search on a landscape where the number of symbols is reduced from $N = 10$ to $N = 6$. In this alternate version of the alien game, participants face a total of 64 ($2^6$) possible attribute combinations. We vary the complexity parameter $K$ to create two landscapes with different levels of ruggedness where $K = 2$ and $K = 5$, respectively.

### 4.4.2 | Results

To derive the (average) normalized search activity of our main study for comparison, the (average) absolute search activity must be divided by 10. Evidently, individual search behavior with and without managerial interventions remains comparable across the different landscapes ($N = 10$ and $N = 6$): participants in the treatment groups engage in substantially more search (WRT: $N = 73$, $z = -4.641$, $p = .000$) with an average value of 1.04; for participants in the control group, we observe average search activity of 0.53. Yet, the difference between relative search activity of the treatments in the large landscape (Study 1) and small landscape are relatively minor (WRT: $N = 147$, $z = 1.585$, $p = .113$), indicating that managerial appraisal leads to the same reaction independent of landscape size. Figure 5 depicts

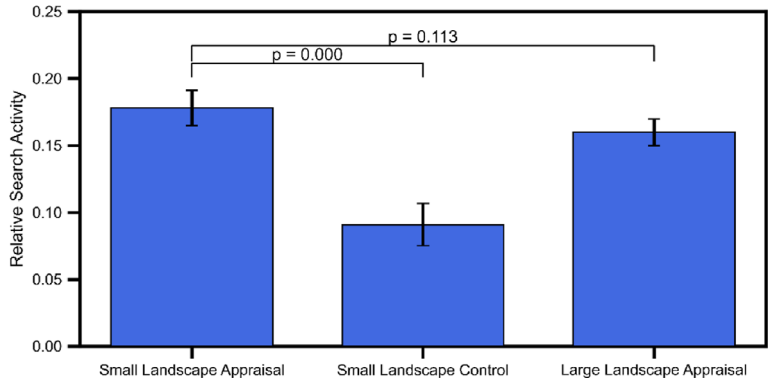



**FIGURE 5** Mean search activities, standard errors, and *p* values for experimental treatments of Studies 1 and 4

mean search activities, standard errors, and *p* values for experimental treatments of Studies 1 and 4.

If participants were simply thinking more thoroughly about the task due to the managerial appraisal and thereby obtained a better understanding of it, we should have observed participants on the small landscape displaying a faster and earlier decrease in search activity than their peers searching a larger landscape. Our findings, therefore, suggest that the main mechanism driving our results is indeed a desire to engage in easily explainable actions that signal high effort, which is independent of landscape size, as we proposed.

## 4.5 | Disentangling search activity—Discerning the effects of search distance and search duration

Jointly, our studies clearly show that managerial supervision increases the search activity of participants. This activity, as we argued earlier (see Section 3.3.1), comprises both effects of search distance, that is, the length of jumps while searching, and search duration—the overall number of periods during which subjects attempt to uncover new solutions at all. How the aggregate increase in search activity is driven by either one of the two partial effects is a question that seems interesting to answer, when seeking to understand the underlying mechanics of how superior monitoring affects subordinate search behavior.

We address this question by taking a close look at two related phenomena we can observe in our data. To determine search duration, we measure the average number of rounds per individual per condition where search activity equals zero. We then compare search duration between treated individuals and those in the control group using a WRT. The results we obtain on the entirety of our sample reveal that participants which are subject to regular appraisals search longer than those, who do not receive any feedback ($N = 444$, $z = 3.179$, $p = .002$). To determine search distance, we average search distances per individual per condition, not counting those rounds in which search activity is zero. Again, we test differences in search distance between treated and other subjects using a WRT. In line with our reasoning, we find that appraisals lead participants to search more distantly compared to their peers without such an impetus ($N = 444$, $z = -5.834$, $p = .000$).

We corroborate our WRT results using an additional way to identify the partial effects of managerial appraisal on search duration and search distance, respectively. To that end, we run a two-step Heckman regression model (Billinger et al., 2021; Heckman, 1979), in which the first stage depicts whether individuals actively search or not. In Stage 2, we model how distantly they search. We find that managerial appraisal increases search duration, as well as search distance (see Technical Appendix, Table S10).

## 4.6 | Robustness checks

### 4.6.1 | Observing the effect of managerial appraisal during a noncomplex search task

In our paper, we argue that subordinates will engage in more search when management can no longer compare individuals' search progress to an objective benchmark, that is, when the search environment becomes increasingly more complex. We therefore classify search task complexity as our most important boundary condition. As our main studies contained complex landscapes ($K = 2$ and $K = 9$), we can test whether our results would hold on a noncomplex landscape where $K = 0$. Thus, we conducted an additional experiment where we recruited 54 participants (15 male, 39 female; mean age = 24.1 years). In contrast to our main studies, we find no difference in search activity (Poisson: $\beta = .021$, $p = .914$, CI = [−0.362, 0.404]) between participants with and without managerial appraisal. This result reported in Table S14 in Section 3.2 of the Technical Appendix thus supports our theoretical framework in which we argue that managerial appraisal will increase search activity only in the context of a complex search task, but not in other settings.

### 4.6.2 | Alternating the framing of the superior's performance feedback message

The superior's performance feedback message we chose for our experiment ("Compared with your previous average, your payoffs in this section were [*above/below*] average X out of 5 times") directs the participants' attention to understanding how often recent search improved the prior status quo. However, feedback can also be given on the average performance increase or decrease ("Compared to your best payoff so far, your current best payoff was X% [*higher/lower*]"). Therefore, to test for the robustness of our results, we conducted an additional experiment with 54 participants (22 male, 31 female, 1 nonbinary; mean age = 23.2 years) with the alternative feedback. The rest of the experiment remained identical to Study 1. We find that our treatment effect is robust to the framing of the performance feedback message participants receive, as we observe higher search activity (Poisson: $\beta = .318$, $p = .006$, CI = [0.091, 0.544]) for participants in the appraisal treatment compared to participants in the control group. The results can be accessed in Table S15 in the Technical Appendix.

## 5 | DISCUSSION AND CONCLUSION

Across four studies involving 444 participants, we find that managerial monitoring encourages more distant as well as prolonged search among subordinates tasked with exploring a complex

landscape. Monitoring—in the form of managerial appraisal comprising both managerial feedback and requests for subordinates' justification of their actions—increases subordinates' overall search activity (Study 1). Importantly, our effects only hold when superiors—as opposed to peers or subordinates—provide appraisal to our subjects (Study 2). Moreover, we find that the key reason for why managerial appraisal is effective is that it entails performance feedback, whereas the request for subordinates to justify their actions seems to play only a minor role (Study 3). Finally, there are no indications that managerial appraisal causes participants to engage in more thorough thinking and thereby obtain a better understanding of their task (Study 4).

Our work contributes to several different streams of literature. First, we expand on the existing literature on organizational design and its behavioral implications. The results presented here highlight the importance of organizational structure and hierarchies as posited by prior work (e.g., Fang, Lee, & Schilling, 2010; Rivkin & Siggelkow, 2003; Schoonhoven & Jelinek, 1990; Siggelkow & Levinthal, 2003) and thus contribute to the reintegration of formal structure into the behavioral foundations of the Carnegie tradition (Gavetti, Levinthal, & Ocasio, 2007). More specifically, we experimentally complement theoretical research pertaining to the effects of managerial appraisal on the aggregation of individual search results and organizational learning. We do so by eliciting the effects of such appraisal on subordinates' search behaviors. Through this, we also complete an emerging picture of the effect of managerial interventions across different types of task environments, and we demonstrate how the effect of managerial intervention differs in search environments as opposed to other settings. Notably, whereas in selection or aggregation tasks subordinates display a tendency to avoid uncertain actions because of fear of negative feedback or perceived lack of control (Detert & Treviño, 2010; Fang et al., 2014; Glass & Singer, 1972; Keum & See, 2017; Milliken et al., 2003; Reitzig & Maciejovsky, 2015), search environments produce different reactions. This is because "pleasing" a superior in a search environment is rather achieved by making bold moves—whose potentially negative consequences can be assessed by neither the manager nor the searching individuals.

Viewed from a different perspective, our findings also complement earlier research on corporate search in the field of strategy more broadly. Notably, we extend the list of contingencies—such as individual career concerns (Tandon & Toh, 2022), individual cognitive flexibility (Kiss et al., 2020), and corporate competitive pressure (Morandi Stagni, Fosfuri, & Santaló, 2021)—known to affect search distance by authoritative appraisal. Importantly, with such appraisal being manipulable by leaders, we offer managers a lever for steering explorative activities within their companies as they see fit.

Finally, from a methodological angle, we also add to the burgeoning literature that seeks to bring laboratory studies to the field of strategic management where useful (Cain, Moore, & Haran, 2015; Elfenbein, Knott, & Croson, 2017; Flammer & Kacperczyk, 2019). More specifically, we add to the prior studies that experimentally explored and tested search behavior on a performance landscape (Billinger et al., 2021; Tracy et al., 2017; Vuculescu, 2017). Our paper complements these studies by examining the behavioral implications that organizational structure adds to the mix. In doing so, our paper responds to the call of neo-Carnegie scholars for a "renewed behaviorally plausible, decision-centered perspective on organization" (Gavetti et al., 2007, p. 525), and our findings may help construct more realistic models of search behavior in the future (see Baumann et al., 2019).

Practitioners may find our research relevant because it offers an avenue to foster more search among their subordinates. Prior work on organizational search shows that firms often struggle to engage in search sufficiently distant to arrive at truly novel solutions that differ from their status quo approaches because subordinates are biased against more distant search (Piao & Zajac, 2016). Managerial appraisal, so it appears, might be a means to overcoming such bias. Additionally, regular appraisal benefits managers by providing them with insights into how subordinates conceive of their search process and what knowledge they have accumulated.

## 5.1 | Limitations and future research

Our work is not without limitations and points to the need for further (experimental) research on effectively and efficiently structuring search in complex, uncertain, and dynamic organizational settings where learning about optimal solutions is virtually impossible—or at least impracticable.

Our paper is limited by the type of performance landscapes we have created for the experiments. It is not clear which kinds of landscapes best represent realistic environments, or to what extent firms would want to encourage more search. Although there is extant research showing that firms are biased against exploration (Leonard-Barton, 1995; Levinthal & March, 1993; March, 1991; Piao & Zajac, 2016; Tripsas & Gavetti, 2000), the degree of "under"-exploration depends heavily on the complexity of the landscape.

We also focus our experiments on individual search. In particular, the effect of formal organizational structure on search groups is yet to be understood. Organizations often rely on groups rather than individuals as decision-makers when facing complex problems (Bernstein, Shore, & Lazer, 2018; Hill, 1982; Laughlin & Ellis, 1986). As demonstrated in a large body of research, people see themselves as less personally accountable for their group's outcomes than when they work individually (Latané & Darley, 1970, 1976; Mynatt & Sherman, 1975; Whyte, 1991), because they feel able to "hide" from individual responsibility behind the collective (Wallach, Kogan, & Bem, 1964). Applying this finding to our empirical setting, we would suspect that groups engage in even more active search than individuals when exposed to managerial control.

Related to the above, future research might also investigate the impact of informal—or social—control mechanisms on individual search behavior more closely. The immediate tendency to engage in social comparison with one another has been shown to drive individual behavior (Lant & Hewlin, 2002) in a variety of contexts. Previous studies show that the attempt to get close to (more) successful peer performance positively affects new product introductions and the contribution of new knowledge (Giachetti & Lampel, 2010; Greve, 2003). In the context of search, this could translate to elevated levels of search activity for subordinates who are in direct comparison with their peers.

Finally, exploring the link between managerial appraisal and overall search performance seems like a fruitful endeavor for future research. To provide some initial insights into this issue, we have investigated the effect of our treatments on performance by pooling the data from all studies and then conducted a regression estimating to what extent managerial appraisal increases or decreases overall search payoffs.[11] For our specific setting, we find a positive effect of appraisal on payoff which, however, diminishes during the later rounds of the experiment. This indicates that the increased search due to the treatment is beneficial for participants in the

[11]Refer to Section 2.5 of the Technical Appendix (Table S11) to review the results.

beginning, but extensive search at the end of the experiment erases that advantage as participants do not fully exploit their best combinations in order to maximize payoffs. While these findings show that appraisal, via increasing search distance, has an impact on search performance, the exact boundary conditions under which this effect will be positive or negative require further analysis and a different setting than the one we chose for our research study. Notably, scholars would want to vary the relative and absolute length of exploration versus exploitation in an experiment for landscapes of different size in order to get a sense when appraisal will benefit or hamper search performance. Similarly, whether bosses are being perceived as facilitators—governing largely self-controlled structures (Klapper & Reitzig, 2018; Reitzig, 2022a, 2022b)—or traditional controllers may determine the effects of appraisal.

## 6 | CONCLUSION

The results of four laboratory experiments consistently showed that individuals increased the duration and distance of their search activity when their actions were subject to managerial monitoring in the form of regular performance appraisals. By demonstrating this clear connection between the presence of monitoring and search behavior, our findings add to the literature on the interplay between organizational search and structure. Moreover, they should be of general value to scholars aiming to develop more behaviorally plausible assumptions about individual search behavior in organizations. Practitioners might use our results to develop better methods to steer the search behavior of their subordinates in a desired direction.

### ACKNOWLEDGEMENTS

We are grateful for the constructive feedback from two anonymous reviewers and the associate editor Martin Ganco. We would also like to thank Hart Posen, Phanish Puranam, and audience members at seminars at the University of Zurich and the University of Munich for helpful comments. All authors would like to acknowledge funding from the Austrian Science Fund (FWF) [Grant P 26574].

[Correction made on 27 May 2023 after first online publication: Grant number was changed from P 25768-G16 to P 26574]

### DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### ORCID

*Valentina Richter* https://orcid.org/0000-0003-0064-356X
*Robert Janjic* https://orcid.org/0000-0001-8897-6478
*Helge Klapper* https://orcid.org/0000-0002-1253-9180
*Steffen Keck* https://orcid.org/0000-0003-0090-3456
*Markus Reitzig* https://orcid.org/0000-0002-8562-3754

### REFERENCES

Aghion, P., & Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy*, *105*(1), 1–29. https://doi.org/10.1086/262063

Alonso, R., Dessein, W., & Matouschek, N. (2008). When does coordination require centralization? *American Economic Review*, *98*(1), 145–179. https://doi.org/10.1257/aer.98.1.145

Angus, R. W. (2019). Problemistic search distance and entrepreneurial performance. *Strategic Management Journal*, *40*(12), 2011–2023. https://doi.org/10.1002/smj.3068

Baron, J., & Ritov, I. (2004). Omission bias individual differences and normality. *Organizational Behavior and Human Decision Processes*, *94*(2), 74–85. https://doi.org/10.1016/j.obhdp.2004.03.003

Baumann, O., Schmidt, J., & Stieglitz, N. (2019). Effective search in rugged performance landscapes: A review and outlook. *Journal of Management*, *45*(1), 285–318. https://doi.org/10.1177/0149206318808594

Becker, M. C., Knudsen, T., & March, J. G. (2006). Schumpeter, Winter, and the sources of novelty. *Industrial and Corporate Change*, *15*(2), 353–371. https://doi.org/10.1093/icc/dtl003

Bernstein, E., Shore, J., & Lazer, D. (2018). How intermittent breaks in interaction improve collective intelligence. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(35), 8734–8739. https://doi.org/10.1073/pnas.1802407115

Billinger, S., Srikanth, K., Stieglitz, N., & Schumacher, T. R. (2021). Exploration and exploitation in complex search tasks: How feedback influences whether and where human agents search. *Strategic Management Journal*, *24*(2), 361–385. https://doi.org/10.1002/smj.3225

Billinger, S., Stieglitz, N., & Schumacher, T. R. (2014). Search on rugged landscapes: An experimental study. *Organization Science*, *25*(1), 93–108. https://doi.org/10.1287/orsc.2013.0829

Bruderer, E., & Singh, J. V. (1996). Organizational evolution learning and selection: A genetic-algorithm-based model. *Academy of Management Journal*, *39*(5), 1322–1349. https://doi.org/10.5465/257001

Cain, D. M., Moore, D. A., & Haran, U. (2015). Making sense of overconfidence in market entry. *Strategic Management Journal*, *36*(1), 1–18. https://doi.org/10.1002/smj.2196

Chang, S. (2022). Two faces of decomposability in organizational search: Evidence from singles vs. albums in the music industry 1995-2015. *Strategic Management Journal*, 1–37. https://doi.org/10.1002/smj.3479

Conover, W. J. (1998). *Practical nonparametric statistics*. New York, NY: John Wiley & Sons.

Csaszar, F. A., & Eggers, J. P. (2013). Organizational decision making: An information aggregation view. *Management Science*, *59*(10), 2257–2277. https://doi.org/10.1287/mnsc.1120.1698

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall.

Dahlander, L., O'Mahony, S., & Gann, D. M. (2016). One foot in, one foot out: How does individuals' external search breadth affect innovation outcomes? *Strategic Management Journal*, *37*(2), 280–302. https://doi.org/10.1002/smj.2342

Denrell, J., Fang, C., & Winter, S. G. (2003). The economics of strategic opportunity. *Strategic Management Journal*, *24*(10), 977–990. https://doi.org/10.1002/smj.341

Detert, J. R., & Treviño, L. K. (2010). Speaking up to higher-ups: How supervisors and skip-level leaders influence employee voice. *Organization Science*, *21*(1), 249–270. https://doi.org/10.1287/orsc.1080.0405

Døjbak Håkonsson, D., Eskildsen, J. K., Argote, L., Mønster, D., Burton, R. M., & Obel, B. (2016). Exploration versus exploitation: Emotions and performance as antecedents and consequences of team decisions. *Strategic Management Journal*, *37*(6), 985–1001. https://doi.org/10.1002/smj.2380

Elfenbein, D. W., Knott, A. M., & Croson, R. (2017). Equity stakes and exit: An experimental approach to decomposing exit delay. *Strategic Management Journal*, *38*(2), 278–299. https://doi.org/10.1002/smj.2493

Elicker, J. D., Levy, P. E., & Hall, R. J. (2006). The role of leader-member exchange in the performance appraisal process. *Journal of Management*, *32*(4), 531–551. https://doi.org/10.1177/0149206306286622

Ethiraj, S. K., & Levinthal, D. A. (2004). Modularity and innovation in complex systems. *Management Science*, *50*(2), 159–173. https://doi.org/10.1287/mnsc.1030.0145

Fang, C., Kim, J. H., & Milliken, F. J. (2014). When bad news is sugarcoated: Information distortion, organizational search and the behavioral theory of the firm. *Strategic Management Journal*, *35*(8), 1186–1201. https://doi.org/10.1002/smj.2146

Fang, C., Lee, J., & Schilling, M. A. (2010). Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science*, *21*(3), 625–642. https://doi.org/10.1287/orsc.1090.0468

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178. https://doi.org/10.1007/s10683-006-9159-4

Flammer, C., & Kacperczyk, A. (2019). Corporate social responsibility as a defense against knowledge spillovers: Evidence from the inevitable disclosure doctrine. *Strategic Management Journal*, *40*(8), 1243–1267. https://doi.org/10.1002/smj.3025

Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, *47*(1), 117–132. https://doi.org/10.1287/mnsc.47.1.117.10671

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: Evidence from patent data. *Research Policy*, *30*(7), 1019–1039. https://doi.org/10.1016/S0048-7333(00)00135-9

Fletcher, C. (2001). Performance appraisal and management: The developing research agenda. *Journal of Occupational and Organizational Psychology*, *74*(4), 473–487. https://doi.org/10.1348/096317901167488

Ganco, M. (2017). NK model as a representation of innovative search. *Research Policy*, *46*(10), 1783–1800. https://doi.org/10.1016/j.respol.2017.08.009

Ganco, M., Kapoor, R., & Lee, G. K. (2020). From rugged landscapes to rugged ecosystems: Structure of interdependencies and firms' innovative search. *Academy of Management Review*, *45*(3), 646–674. https://doi.org/10.5465/amr.2017.0549

Gavetti, G., & Levinthal, D. A. (2000). Looking forward and looking backward: Cognitive and experiential search. *Administrative Science Quarterly*, *45*(1), 113–137. https://doi.org/10.2307/2666981

Gavetti, G., Levinthal, D. A., & Ocasio, W. (2007). Perspective-neo-Carnegie: The Carnegie School's past, present, and reconstructing for the future. *Organization Science*, *18*(3), 523–545. https://doi.org/10.1287/orsc.1070.0277

Ghemawat, P., & Levinthal, D. A. (2008). Choice interactions and business strategy. *Management Science*, *54*(9), 1638–1651. https://doi.org/10.1287/mnsc.1080.0883

Giachetti, C., & Lampel, J. (2010). Keeping both eyes on the competition: Strategic adjustment to multiple targets in the UK mobile phone industry. *Strategic Organization*, *8*(4), 347–376. https://doi.org/10.1177/1476127010391235

Glass, D. C., & Singer, J. E. (1972). *Urban stress: Experiments on noise and social stressors*. New York, NY: Academic Press.

Greve, H. R. (2003). A behavioral theory of R&D expenditures and innovations: Evidence from shipbuilding. *Academy of Management Journal*, *46*(6), 685–702. https://doi.org/10.5465/30040661

Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal*, *29*(2), 147–160. https://doi.org/10.1002/j.1538-7305.1950.tb00463.x

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), 153–161. https://doi.org/10.2307/1912352

Hill, G. W. (1982). Group versus individual performance: Are N+l heads better than one? *Psychological Bulletin*, *91*(3), 517–539. https://doi.org/10.1037/0033-2909.91.3.517

Joseph, J., & Gaba, V. (2020). Organizational structure, information processing, and decision-making: A retrospective and road map for research. *Academy of Management Annals*, *14*(1), 267–302. https://doi.org/10.5465/annals.2017.0103

Keum, D. D., & See, K. E. (2017). The influence of hierarchy on idea generation and selection in the innovation process. *Organization Science*, *28*(4), 653–669. https://doi.org/10.1287/orsc.2017.1142

Kiss, A. N., Libaers, D., Barr, P. S., Wang, T., & Zachary, M. A. (2020). CEO cognitive flexibility, information search, and organizational ambidexterity. *Strategic Management Journal*, *41*(12), 2200–2233. https://doi.org/10.1002/smj.3192

Kitchen, P. J., Brignell, J., Li, T., & Jones, G. S. (2004). The emergence of IMC: A theoretical perspective. *Journal of Advertising Research*, *44*(1), 19–30. https://doi.org/10.1017/S0021849904040048

Klapper, H., & Reitzig, M. (2018). On the effects of authority on peer motivation: Learning from Wikipedia. *Strategic Management Journal*, *39*(8), 2178–2203. https://doi.org/10.1002/smj.2909

Knudsen, T., & Srikanth, K. (2014). Coordinated exploration: Organizing joint search by multiple specialists to overcome mutual confusion and joint myopia. *Administrative Science Quarterly*, *59*(3), 409–441. https://doi.org/10.1177/0001839214538021

Kyl-Heku, L. M., & Buss, D. M. (1996). Tactics as units of analysis in personality psychology: An illustration using tactics of hierarchy negotiation. *Personality and Individual Differences*, *21*(4), 497–517. https://doi.org/10.1016/0191-8869(96)00103-1

Lant, T. K., & Hewlin, P. F. (2002). Information cues and decision making: The effects of learning, momentum, and social comparison in competing teams. *Group & Organization Management*, *27*(3), 374–407. https://doi.org/10.1177/1059601102027003004

Larson, J. R., & Callahan, C. (1990). Performance monitoring: How it affects work productivity. *Journal of Applied Psychology*, *75*(5), 530–538. https://doi.org/10.1037/0021-9010.75.5.530

Latané, B., & Darley, J. M. (1970). *The unresponsive bystander: Why doesn't he help?* New York, NY: Appleton-Century-Crofts.

Latané, B., & Darley, J. M. (1976). *Help in crisis: Bystander response to an emergency*. Morristown, NJ: General Learning.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology*, *22*(3), 177–189. https://doi.org/10.1016/0022-1031(86)90022-3

Lee, E., & Puranam, P. (2015). The nature of expertise in organization design: Evidence from an expert–novice comparison. In G. Gavetti & W. Ocasio (Eds.), *Cognition and strategy* (pp. 191–209). Bingley, UK: Emerald Group. https://doi.org/10.1108/S0742-332220150000032006

Lee, G. K., & Alnahedh, M. A. (2016). Industries' potential for interdependency and profitability: A panel of 135 industries, 1988–1996. *Strategy Science*, *1*(4), 285–308. https://doi.org/10.1287/stsc.2016.0023

Lenox, M. J., Rockart, S. F., & Lewin, A. Y. (2010). Does interdependency affect firm and industry profitability? An empirical test. *Strategic Management Journal*, *31*(2), 121–139. https://doi.org/10.1002/smj.811

Leonard-Barton, D. (1995). *Wellsprings of knowledge*. Boston, MA: Harvard Business School Press.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*(2), 255–275. https://doi.org/10.1037/0033-2909.125.2.255

Levinthal, D. A. (1997). Adaptation on rugged landscapes. *Management Science*, *43*(7), 934–950. https://doi.org/10.1287/mnsc.43.7.934

Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic Management Journal*, *14*(S2), 95–112. https://doi.org/10.1002/smj.4250141009

Levy, P. E., & Williams, J. R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*, *30*(6), 881–905. https://doi.org/10.1016/j.jm.2004.06.005

Lizzio, A., Wilson, K., & MacKay, L. (2008). Managers' and subordinates' evaluations of feedback strategies: The critical contribution of voice. *Journal of Applied Social Psychology*, *38*(4), 919–946. https://doi.org/10.1111/j.1559-1816.2008.00332.x

Lund, O. C. H., Tamnes, C. K., Moestue, C., Buss, D. M., & Vollrath, M. (2007). Tactics of hierarchy negotiation. *Journal of Research in Personality*, *41*(1), 25–44. https://doi.org/10.1016/j.jrp.2006.01.002

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, *2*(1), 71–87. https://doi.org/10.1287/orsc.2.1.71

March, J. G., & Simon, H. A. (1958). *Organizations*. New York, NY: Wiley.

Markides, C. C. (1999). A dynamic view of strategy. *Sloan Management Review*, *40*(3), 55–63.

Mihm, J., Loch, C. H., & Huchzermeier, A. (2003). Problem-solving oscillations in complex engineering projects. *Management Science*, *49*(6), 733–750. https://doi.org/10.1287/mnsc.49.6.733.16021

Mihm, J., Loch, C. H., Wilkinson, D., & Huberman, B. A. (2010). Hierarchical structure and search in complex organizations. *Management Science*, *56*(5), 831–848. https://doi.org/10.1287/mnsc.1100.1148

Milliken, F. J., Morrison, E. W., & Hewlin, P. F. (2003). An exploratory study of employee silence: Issues that employees don't communicate upward and why. *Journal of Management Studies*, *40*(6), 1453–1476. https://doi.org/10.1111/1467-6486.00387

Morandi Stagni, R., Fosfuri, A., & Santaló, J. (2021). A bird in the hand is worth two in the bush: Technology search strategies and competition due to import penetration. *Strategic Management Journal*, *42*(8), 1516–1544. https://doi.org/10.1002/smj.3277

Morrison, E. W. (1993). Newcomer information seeking: Exploring types, modes, sources, and outcomes. *Academy of Management Journal*, *36*(3), 557–589. https://doi.org/10.5465/256592

Mynatt, C., & Sherman, S. J. (1975). Responsibility attribution in groups and individuals: A direct test of the diffusion of responsibility hypothesis. *Journal of Personality and Social Psychology*, *32*(6), 1111–1118. https://doi.org/10.1037/0022-3514.32.6.1111

Nai, J., Kotha, R., Narayanan, J., & Puranam, P. (2020). Transparency and fairness in organizational decisions: An experimental investigation using the paired ultimatum game. *Strategy Science*, *5*(1), 55–70. https://doi.org/10.1287/stsc.2019.0100

Nelson, R. R., & Winter, S. G. (1982). *Evolutionary theory of economic change*. Cambridge, MA: Harvard University Press.

Niehoff, B. P., & Moorman, R. H. (1993). Justice as a mediator of the relationship between methods of monitoring and organizational citizenship behavior. *Academy of Management Journal*, *36*(3), 527–556. https://doi.org/10.5465/256591

Piao, M., & Zajac, E. J. (2016). How exploitation impedes and impels exploration: Theory and evidence. *Strategic Management Journal*, *37*(7), 1431–1447. https://doi.org/10.1002/smj.2402

Pichler, S. (2012). The social context of performance appraisal and appraisal reactions: A meta-analysis. *Human Resource Management*, *51*(5), 709–732. https://doi.org/10.1002/hrm.21499

Porter, M. E. (1991). Towards a dynamic theory of strategy. *Strategic Management Journal*, *12*(S2), 95–117. https://doi.org/10.1002/smj.4250121008

Posen, H. E., & Martignoni, D. (2018). Revisiting the imitation assumption: Why imitation may increase, rather than decrease, performance heterogeneity. *Strategic Management Journal*, *39*(5), 1350–1369. https://doi.org/10.1002/smj.2751

Raveendran, M. (2020). Seeds of change: How current structure shapes the type and timing of reorganizations. *Strategic Management Journal*, *41*(1), 27–54. https://doi.org/10.1002/smj.3084

Reitzig, M. (2022a). *Get better at flatter—A guide to shaping and leading organizations with less hierarchy*. London, UK: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-89254-8

Reitzig, M. (2022b). How to get better at flatter designs: Considerations for shaping and leading organizations with less hierarchy. *Journal of Organization Design*, *11*(1), 5–10. https://doi.org/10.1007/s41469-022-00109-7

Reitzig, M., & Maciejovsky, B. (2015). Corporate hierarchy and vertical information flow inside the firm—A behavioral view. *Strategic Management Journal*, *36*(13), 1979–1999. https://doi.org/10.1002/smj.2334

Rivkin, J. W. (2000). Imitation of complex strategies. *Management Science*, *46*(6), 824–844. https://doi.org/10.1287/mnsc.46.6.824.11940

Rivkin, J. W., & Siggelkow, N. (2003). Balancing search and stability: Interdependencies among elements of organizational design. *Management Science*, *49*(3), 290–311. https://doi.org/10.1287/mnsc.49.3.290.12740

Sah, R. K., & Stiglitz, J. E. (1986). The architecture of economic systems: Hierarchies and polyarchies. *American Economic Review*, *76*(4), 716–727. https://doi.org/10.3386/w1334

Schoonhoven, D. B., & Jelinek, M. (1990). Dynamic tension in innovative high technology firms: Managing rapid technological change through organizational structure. In M. A. von Glinow & S. A. Mohrman (Eds.), *Managing complexity in high technology organizations* (pp. 90–118). New York, NY: Oxford University Press.

Schumpeter, J. A. (1934). *The theory of economic development*. Cambridge, MA: Harvard University Press.

Siggelkow, N., & Levinthal, D. A. (2003). Temporarily divide to conquer: Centralized, decentralized, and reintegrated organizational. *Organization Science*, *14*(6), 650–669. https://doi.org/10.1287/orsc.14.6.650.24840

Siggelkow, N., & Rivkin, J. W. (2005). Speed and search: Designing organizations for turbulence and complexity. *Organization Science*, *16*(2), 101–122. https://doi.org/10.1287/orsc.1050.0116

Siggelkow, N., & Rivkin, J. W. (2009). Hiding the evidence of valid theories: How coupled search processes obscure performance differences among organizations. *Administrative Science Quarterly*, *54*(4), 602–634. https://doi.org/10.2189/asqu.2009.54.4.602

Simon, H. A. (1951). A formal theory of the employment relationship. *Econometrica: Journal of the Econometric Society*, *19*, 293–305. https://doi.org/10.2307/1906815

Simon, H. A. (1957). *Models of man*. New York, NY: John Wiley & Sons.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, *106*(6), 467–482.

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning* (Vol. 135). Cambridge: MIT Press.

Tandon, V., & Toh, P. K. (2022). Who deviates? Technological opportunities, career concern, and inventor's distant search. *Strategic Management Journal*, *43*(4), 724–757. https://doi.org/10.1002/smj.3345

Taylor, C. R., & Walsh, M. G. (2002). Legal strategies for protecting brands from genericide: Recent trends in evidence weighted in court cases. *Journal of Public Policy & Marketing*, *21*(1), 160–167. https://doi.org/10.1509/jppm.21.1.160.17609

Taylor, M. S., Tracy, K. B., Renard, M. K., Harrison, J. K., & Carroll, S. J. (1995). Due process in performance appraisal: A quasi-experiment in procedural justice. *Administrative Science Quarterly*, *40*(3), 495–523. https://doi.org/10.2307/2393795

Townley, B. (1993). Performance appraisal and the emergence of management. *Journal of Management Studies*, *30*(2), 221–238. https://doi.org/10.1111/j.1467-6486.1993.tb00302.x

Tracy, W. M., Markovitch, D. G., Peters, L. S., Phani, B. V., & Philip, D. (2017). Algorithmic representations of managerial search behavior. *Computational Economics*, *49*(3), 343–361. https://doi.org/10.1007/s10614-015-9559-7

Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal*, *21*(10–11), 1147–1161. https://doi.org/10.1002/1097-0266(200010/11)21:10/11<1147::AID-SMJ128>3.0.CO;2-R

Vuculescu, O. (2017). Searching far away from the lamp-post: An agent-based model. *Strategic Organization*, *15*(2), 242–263. https://doi.org/10.1177/1476127016669869

Wallach, M. A., Kogan, N., & Bem, D. J. (1964). Diffusion of responsibility and level of risk taking in groups. *Journal of Abnormal and Social Psychology*, *68*(3), 263–274. https://doi.org/10.1037/h0042190

Whyte, G. (1991). Diffusion of responsibility: Effects on the escalation tendency. *The Journal of Applied Psychology*, *76*(3), 408–415. https://doi.org/10.1037/0021-9010.76.3.408

Wood, J. A., Jr., & Winston, B. E. (2005). Toward a new understanding of leader accountability: Defining a critical construct. *Journal of Leadership and Organizational Studies*, *11*(3), 84–94. https://doi.org/10.1177/107179190501100307

Yayavaram, S., Srivastava, M. K., & Sarkar, M. B. (2018). Role of search for domain knowledge and architectural knowledge in alliance partner selection. *Strategic Management Journal*, *39*(8), 2277–2302. https://doi.org/10.1002/smj.2791

Yu, W., Minniti, M., & Nason, R. (2019). Underperformance duration and innovative search: Evidence from the high-tech manufacturing industry. *Strategic Management Journal*, *40*(5), 836–861. https://doi.org/10.1002/smj.2988

Zeelenberg, M., van de Bos, K., van Dijk, E., & Pieters, R. (2002). The inaction effect in the psychology of regret. *Journal of Personality and Social Psychology*, *82*(3), 314–327. https://doi.org/10.1037/0022-3514.82.3.314

Zott, C., & Amit, R. (2010). Business model design: An activity system perspective. *Long Range Planning*, *43*(2–3), 216–226. https://doi.org/10.1016/j.lrp.2009.07.004

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Richter, V., Janjic, R., Klapper, H., Keck, S., & Reitzig, M. (2023). Managing exploration in organizations: The effect of superior monitoring on subordinate search behavior. *Strategic Management Journal*, *44*(9), 2226–2254. https://doi.org/10.1002/smj.3491