Technovation 129 (2024) 102895

Contents lists available at ScienceDirect

Technovation

journal homepage: www.elsevier.com/locate/technovation

Initial conditions and path dependence in explorative and exploitative learning: An experimental study

Luigi Mittone [a,c], Azzurra Morreale [b,c], Paavo Ritala [c,*]

[a] Cognitive and Experimental Economics Laboratory (CEEL), Department of Economics and Management, University of Trento, Italy
[b] Jönköping International Business School, Jönköping University, 553 18 Jönköping, Sweden
[c] Business School, LUT University, Finland

A B S T R A C T

The question of why individuals choose to explore or exploit as their learning accumulates remains largely unexplored in organizational literature in a strictly causal sense. To bridge this gap, we conducted an experimental laboratory study of individual decision-making sequences using a real-effort task that involved a training phase and an active phase. In the training phase, the participants used their skills to solve the same task in eight rounds to simulate the development of individual-level learning. In the active phase, we observed sequential choices over exploring or exploiting. The participants were financially incentivized to abandon a familiar task (that they learned in the training phase) by providing higher performance-related payoffs for exploring novel task environments. Interestingly, we not only found that different kinds of performance feedback affected the exploration-exploitation choice, but that the feedback-choice linkage is contingent upon the initial conditions of the task environment in terms of its simplicity or complexity. We found that when individuals are initially exposed to simpler tasks, they are more likely to continue exploiting a familiar task; and when they are initially exposed to more complex tasks, they are more likely to explore new and more profitable tasks and then continue exploiting the new tasks they learned. These findings contribute to the literature on individual search by demonstrating the important role of initial conditions and path dependence in exploration and exploitation behavior.

1. Introduction

Organizational decision-makers constantly make choices to explore or exploit (Lavie et al., 2010). These decisions are embedded in a process of individual search, where the decision-makers choose to continue exploiting their knowledge of a previously known task or choose to explore new tasks (Billinger et al., 2014). The exploration-exploitation choice is ultimately driven by the expected differences in the performance of these alternative models of learning, considering that the returns of exploration are uncertain while the returns of exploitation are more predictable (March, 1991). Therefore, given the lower cognitive requirements and the higher certainty of expected performance, decision-makers are often more prone to exploit than to explore, which sometimes leads to “overexploitation” and the “competency trap” (Denrell and Le Mens, 2020; Levinthal and March 1993). On the other hand, too much emphasis on exploration also takes place and can be harmful to decision performance (Billinger et al., 2014). Therefore, the exploration-exploitation choice remains a critical subject of study for organizational scholars, especially because according to Puranam et al. (2015), “The conditions under which people switch between exploration and exploitation are not yet fully understood” (p. 352). Furthermore, individuals’ search behavior has been shown to affect organizational-level explorative and exploitative innovation (Enkel et al., 2017), driven by organizational structures and practices (Ali et al., 2022), and ultimately embedded in organizational-level explorative and explorative intents and logics (Chadwick and Raver, 2015; Ko et al., 2021). Therefore, examining the microfoundations of exploration-exploitation choices represents an important, untapped research opportunity for scholarship and practice of innovation management and organization studies.

The features of exploration-exploitation choice have been theorized in the growing body of literature on individual search (Puranam et al., 2015). Empirically, the most prominent causal evidence of the antecedents of individual search comes from two strands of experimental literature. The first strand builds on the premises of behavioral literature and aspiration adaptation (Cyert and March 1963; Greve 1998; Simon 1955) and views the exploration-exploitation choice as a function of performance feedback. This literature includes experimental designs where the decision problem is modeled as a decision-maker’s choice between previously unknown probabilistic properties, such as choosing among the arms of a multi-armed bandit (Puranam et al., 2015), and searching for unknown features on rugged landscapes, including the NK

* Corresponding author.
E-mail addresses: luigi.mittone@unitn.it (L. Mittone), azzurra.morreale@ju.se (A. Morreale), ritala@lut.fi (P. Ritala).

https://doi.org/10.1016/j.technovation.2023.102895
Received 24 June 2023; Received in revised form 27 September 2023; Accepted 9 October 2023
Available online 11 November 2023

models (Billinger et al., 2014, 2021; Levine et al., 2019). The main findings of this suggest that underperformance related to aspirations often leads to exploration, and performing according to aspirations often leads to continuing exploitation. The second strand of literature focuses on path dependency in learning processes and, more specifically, on the nature of the decision landscape in which the decision-makers are trained to perform a given task (Luchins, 1942; Egidi and Narduzzo, 1997; Egidi, 2015; Hoeffler et al., 2006). This literature typically uses tasks that require individual learning and apply a set of rules aimed at building sequences of "actions" that translate into the desired outcome. The main findings from this literature imply that once subjects learn a task, they tend to stick to it even if it is inefficient (Luchins 1942), which is recognized later in the literature as a variety of biases that make exploitation more likely than it would ideally be (Denrell and Le Mens, 2020; Levinthal and March 1993).

Importantly, there is increasing evidence that exploration-exploitation decisions are prone to various preconditions and triggers that alter how this choice is being made. In the current study, we argue that the complexity of the task environment under which learning occurs is a key condition in the latter choices to deviate from or stick to the initially learned task. This intuition is supported by literature that has demonstrated that exploration-exploitation choice involves a lot of inertia, stickiness, and attachment to past success (Hoeffler et al., 2006; Brusoni et al., 2020) and that such choice is affected by the task environment (MacLeod and Pingle 2005; Rahmandad et al., 2021). Therefore, to model the exploration-exploitation choice that takes into account these features, in this study we put forward an experimental design that accommodates both the qualitative features of learning (i.e., the complexity of the task environment under which the initial task learning occurs) and the situations in which the choice is made to stick to or deviate from a specific task.

Empirically, we examine *what triggers an individual's adherence to or abandonment of a particular task that has been learned initially in more or less complex task environments*. To that end, we devised a laboratory experiment with 240 participants, internal conceptual replication, two abstract experimental designs (visual and numeric), and four different tasks that allowed the participants to explore and exploit. Following previous experimental literature, our experiment utilized a "real-effort task" (Brüggen and Strobel 2007). The real-effort task setting mimicked a process of *learning by doing* as a function of repetition and task success (Benndorf et al., 2019; Weiss and Ilgen 1985, p. 59), thereby capturing the "incremental refinements achieved by internalizing or combining existing knowledge" (Laureiro-Martínez et al., 2015, p. 321). An important feature of our experimental design is the *training phase,* which forced the participants to perform the same task multiple times to manipulate the build-up of learning and the related path-dependency, "imprinting" (Levine et al., 2019) or "attachment" (Brusoni et al., 2020) to particular tasks. During the training phase, the participants were randomly assigned to two tasks: a relatively simple memory task and a more difficult building task to simulate the establishment of learning in diverse task environments. Then, in the second phase, namely, the *active phase*, the participants had to decide whether to continue exploiting the training task, explore an unknown and different task in subsequent rounds, and potentially continue to exploit such task or retreat to the training task. To further simulate an environment in which exploring new alternatives can deliver higher rewards albeit with greater uncertainty (March, 1991), our experimental design offered increased payoffs if the newly explored task was successful.

Our main findings demonstrate the important role of initial conditions in exploration-exploitation choice. When individuals are initially exposed to simpler tasks, they are more likely to continue exploiting a familiar task. When they are initially exposed to more complex tasks, they are more likely to explore new and more profitable tasks and then continue exploiting the new tasks they learned. In the following, we briefly review the relevant literature, discuss the experimental design, and follow up with results, implications, and future research suggestions.

2. Theoretical background

2.1. Exploration-exploitation at the individual level: bounded adaptive rationality model

The seminal literature on exploration and exploitation is grounded on the organizational level, and it was initially concerned with organizational learning (March, 1991; Levinthal and March 1993). While much of this research builds upon March's work, primarily at the organizational level (e.g., Hoang and Rothaermel, 2010; Smith and Tushman 2005; Junni et al., 2013), March acknowledged the role of individual decision-makers, influenced by Simon's early notions of adaptive aspirations and exploration choices (Simon 1955; March, 1991; Lee and Meyer-Doyle 2017).

Recent studies in organizational literature emphasize variations in individual decision-makers' exploration and exploitation capabilities (Mom, Van den Bosch, and Volberda 2007; Mom, Van den Bosch, and Volberda 2009). Experimental designs investigating the impact of past performance on these choices (Billinger et al., 2014; Levine et al., 2019) and micro-level behavior patterns under different performance incentives (Lee and Meyer-Doyle 2017) support the importance of individual-level examination. Neuroimaging research also underscores cognitive disparities between exploration and exploitation (Laureiro-Martínez et al., 2010; 2015). However, the precise conditions governing switches between these modes remain incompletely understood (Puranam et al., 2015, p. 352).

Building on March's original definition (1991, p. 71), we conceptualize exploration-exploitation choice at the individual level. With exploitation, the key issue is that the decision-maker knows the task based on previous experience and, thus, a process of learning and refinement can take place (Ericsson and Lehmann 1996). With exploration, the opposite is true: the decision-maker chooses to abandon their current task to find a new, potentially more profitable task (March, 1991; Laureiro-Martínez et al., 2015). These choices serve performance goals: learning through task repetition or experimenting with new tasks. This dynamic process allows exploration to transition into exploitation and vice versa, mirroring changing task environments (Posen and Levinthal 2012; Lee and Meyer-Doyle 2017). Research suggests a tendency to overexploit due to a preference for existing competencies and risk-aversion, but overexploration can also occur, cutting short performance improvements (Levinthal and March 1993; Wiseman and Gomez-Mejia 1998; Greve 1998; Billinger et al., 2014). Therefore, individual exploration-exploitation patterns vary based on performance and task environment.

Our approach draws from Herbert Simon's concept of bounded rationality with adaptive aspirations (Simon 1955) and behavioral theory on problemistic search (Cyert and March 1963; Gavetti et al., 2012). The trade-off between exploration and exploitation relates to dynamic sequential decision-making processes. This process aligns with reinforcement learning theory (Sutton and Barto, 1998; Cohen et al., 2007) and Thorndike's "Law of Effect." Together, these assumptions can be conceptualized under the *bounded adaptive rationality* model (Puranam et al., 2015).

Our experimental approach follows the bounded adaptive rationality model's main components. The task environment represents a decision-maker's objective reality with various actions leading to different performance outcomes. The decision-maker aims to reach an acceptable performance level (Cyert and March 1963; Posen et al., 2018). The choice process is often incomplete, with limited information, and may prioritize "satisficing" over maximizing performance (Simon 1955; Cyert and March 1963). Performance feedback, varying in availability and quality, informs the decision-maker's representation of the task environment. This feedback triggers changes in the choice process, such as seeking alternatives in response to poor performance or repeating behaviors linked to strong performance (Greve 1998; March, 1988). Our application of the bounded adaptive rationality model views exploration-exploitation choice as dynamic, with the decision-maker comparing options and adjusting aspirations based on both recent and past experiences.

2.2. The role of initial conditions on learning and exploration-exploitation choice

Based on the bounded adaptive rationality model, we move to discuss how individual exploration-exploitation choice is potentially constrained by path dependency of learning and a variety of behavioral biases. At the individual level, when a problem solution is experienced through a process of learning via repetitions, it is ultimately applied *automatically* (Narduzzo and Warglien, 2008). According to Egidi (2015; see also Narduzzo and Warglien, 2008), this result is related to the cognitive characteristics of individual learning and more precisely, to the *mechanization of thought* introduced by Luchins (1942), one of the first and few scholars who investigated the relationship between initial conditions (problems framing) and solutions-search behavior, induced by being exposed repeatedly to the same class of problems. In Luchins's experiment, once the participants were required to solve a similar problem for a set of repetitions, they started to develop a procedure that they then mechanically applied to new problems, although the new problems required abandoning the original procedure to look for new and more efficient ones. This phenomenon of "habituation to a repeatedly used procedure" (Luchins 1942, p. 3) to solve a given problem is called the *Einstellung effect*. As Narduzzo and Warglien (2008) noticed, the situation does not change even when other factors are added to the experimental setting (e.g., incentives to reward efficiency or more realistic experimental conditions aimed at increasing the reflexivity of the participants; Luchins and Luchins 1950).

Based on the experimental evidence discussed above, and the theoretical intuition of aspiration adaptation, decision makers are likely to favor continuing exploitation once they have learned to perform a certain task, *ceteris paribus*. However, whether and how different initial experiences in (exploitative) learning affect search behavior has been less examined (Billinger et al., 2021). Understanding better the role of initial conditions is crucial, as learning takes place within a context, and this context will matter in the later exploration-exploitation choice. The reason for this gap in the literature is that most of the experimental studies did not use a training phase or a "learning phase" to simulate the stabilization of exploitative learning (and exploitation choice) under different task environments. A rare exception is the study of Betsch et al. (2001), where micro-world computer simulation was used to investigate the role of the "routine strength" in the search for information and in the decision to stick to the learned task. The main finding is that the strong-routinization participants relied on the beliefs built in the training phase much more than the weak-routinization participants. This translates to a greater reluctance to give up the exploitation and demonstrates that initial conditions matter for exploration-exploitation choice.

For the current study, we are particularly interested in the complexity of decision making in the context in which learning occurs, i.e. the *task environment* (Puranam et al., 2015). Task complexity is a crucial feature of decision-making as it affects the goal setting of actors and is related to the different strategies for solving a task (Campbell 1988). While some task environments are perceived as more challenging (due to their complexity), others are perceived as easier. A complex task environment includes many interacting variables, making it difficult to understand, which leads to a particular level of task performance; and the inverse is true for simple task environments where the sources of performance are more visible and where there are fewer interacting decision variables (Rahmandad et al., 2021).

In the context of individual search, the task complexity regulates the process of aspiration adaptation. MacLeod and Pingle (2005) demonstrated that with complex tasks, the decision-maker experiences uncertainty when facing a problem and, relatedly, experiences uncertainty as to where the aspiration should be set. Conversely, they found that with simple problems, aspirations are easier to set, which further encourages individual search behavior. The question remains that while complex task environments seem to be initially more challenging for a decision-maker, what happens when learning takes place over multiple repetitions in such an environment as opposed to a simple task environment? In this regard, we expect that learning taking place in simpler task environments may offer less cognitive flexibility or room to maneuver later on when exploring new tasks. Multiple repetitions in a simple task environment lead to learning that helps the decision-maker to conserve cognitive effort and thus may discourage the decision-maker from exploring other tasks even with incentives to do so. Indeed, there is evidence that favorable initial task experiences may limit explorative search. Hoeffler et al. (2006) found that especially individuals who enjoy favorable results of particular tasks, can become dependent on those initial experiences, and lead to a "biased search process in which entire regions of potentially attractive alternatives are relatively unlikely to be discovered" (p. 218). Therefore, since we expect a simple task environment to be more likely to provide such favorable experiences (in terms of performance), there is an inherent bias to continue exploitation. For learning taking place initially in complex environments, we expect the inverse to be true. Decision-makers need to learn to perform under a state of complexity, which imposes initially high cognitive demands on them (Rahmandad et al., 2021). When the situation stabilizes, however, the decision-makers will have experienced a learning process that conditioned them to face higher levels of uncertainty and complexity, which we expect to be carried over to the subsequent exploration-exploitation choice. Experiences in facing uncertainty will prime the decision-maker to encounter further uncertainty, and in our context, to make more likely choices to explore new tasks.

3. The experiment

This section outlines the experimental design employed to examine how individuals navigate between exploitation and exploration in their decision-making. Our research design consists of two key elements: first, a computer-based decision-making game, which we discuss in Section 3.1, serves as the central experiment. Second, we utilize the Bomb Risk Elicitation Task (BRET; Crosetto and Filippin, 2013) to assess participants' risk preferences, detailed in Section 3.2. Section 3.3 elaborates on our approach to conceptualizing exploration-exploitation choices. Finally, Section 3.4 provides an overview of the study's participants and the experimental procedures.

3.1. The main experiment

The main experiment utilized a computer game based on a real-effort task (Brüggen and Strobel 2007). This allowed us to analyze decision-makers performance and learning, leading to choices between exploitation and exploration during the experiment. The game unfolded over 18 rounds and was divided into two main phases: the training phase (eight rounds) and the active phase (10 rounds). During the training phase, the participants were randomly assigned to what we call a "building task" (learning a technique for constructing a visual puzzle or a number), or to what we define as a "memory task" (learning a technique for memorizing a figure or a number). This procedure forced the subjects to gain practical experience and establish a learning path in two distinct task environments with different difficulty levels. This experimental expedient was introduced to verify the effect of different degrees of complexity of the training environment on the subsequent exploration and exploitation decisions of the participants in the active phase of the experiment. The building task proved to be more challenging than the memory task because the subjects who were required to train on it achieved lower average performances compared to those who were trained on the memory task (see section 4.1).

The training phase plays a crucial role in inducing learning by doing. This phase is compulsory, forcing the decision-makers to repeat the same task for 8 rounds. In the subsequent active phase, the decision-makers can choose to exploit the same task in multiple rounds or explore another task from a total of four tasks (*Memory* – labeled in the instructions "Modality Y" –, *Building* – labeled in the instructions "Modality X" –, *Combine* – labeled in the instructions "Modality Z" –, and *Destroy* – labeled in the instructions "Modality K" –), including the training phase task. The decision to choose neutral rather than descriptive terminology to present the tasks is aimed at avoiding the risk of uncontrollable factors that can influence the participants' choices. Regardless of the specific task chosen, participants aimed to approximate a "target" as closely as possible, which served as a proxy for various types of goals in decision-making scenarios.

The tasks in the main game were framed in two ways: visually or numerically.[1] In the *visual* frame, the participants were asked to select or construct an abstract target figure. As in Mittone and Papi (2017), all the figures used in this experiment were grids with red or beige cells. Each cell is called a *pixel*, and a set of four adjacent cells is called a *block*. The *numeric* frame corresponds closely to the visual frame regarding task design, but in this case, the participants were asked to select or construct a target number. Thus, our experiment relies on 4 tasks: one learned in the training phase and the other three unknown tasks that, in addition to the learned one, could be chosen by subjects during the active phase. Table 1 provides a detailed description of each task.

When examining the task descriptions, it becomes evident that the "Destroy" task mirrors the "Building" task, but it compels the subjects to view the game solution as a simplification of the initial frame rather than a progressive complexification of it. As for the "Combine" task, it is derived from a fusion of the "Building" and "Memory" tasks. Notably, each round gave each participant a different figure or number.

3.1.1. Information on task performance: availability of performance updates versus No updating

Alongside the task presentation format, which was either numerical or visual, we also manipulated the provision of performance feedback to participants: participants either received continual updates on their performance after each round or did not receive any performance feedback at all.

Informed decision-making about whether to explore or exploit depends on performance feedback (Greve 2003; Puranam et al., 2015; Simon 1996). Decision-makers typically receive some form of performance feedback, but the availability of useful performance updates (feedback) varies. While updating performance data is sometimes immediate (e.g., in stock markets), the organizational literature confirmed that in many situations, individuals make decisions in the absence of complete or even partial information about the immediate consequences of their actions (e.g., Posen and Levinthal 2012; Ritala et al., 2016; Weick 1993). For that reason, our experiment tested two distinct scenarios: decision-making with constant explicit updates on performance and decision-making in the absence of any such updates. Feedback is generally thought to contribute positively to individual performance (Larson 1984). Extensive organizational and economic experiments that analyzed the role of feedback (Diehl and Sterman 1995; Roth and Malouf 1979; Smith 1962) confirmed that the information provided often supports the participant's ability to choose the best option or task-performing strategy. On that basis, we expect the availability of explicit feedback updates on performance to improve task performance when other things remain constant.

To test the influence of the availability of updates on performance, we created two versions of the game. In the version with explicit performance updates, the participants received information on their performance in each of the 18 rounds. Specifically, the participants were updated about their performance regarding the extent to which they had fulfilled the requirements of the given task (expressed as a percentage). These participants received an update on their performance but not on the associated payoffs; they were informed about the payoff scheme at the beginning of the game but not after each round. This design sought to reproduce the incomplete information environment in which managers often operate.[2] The performance-update (feedback) version of the experiment was replicated in both numeric and visual scenarios.

In the no-performance-update (no-feedback) version, the participants received no information about their performance in any round of the game (although they were informed about the payoff scheme at the beginning of the experiment). As in the performance-update version of the experiment, we replicated the no-performance-update sessions for both the numeric and visual scenarios.

Thus, within the main experiment, four separate sub-studies were conducted, each distinguished by two factors: whether participants receive feedback on their performance or not, and the framework—either numerical or visual—used for the tasks. Therefore, the experiment includes four distinct settings: numeric feedback study (NF), numeric no-feedback study (NNF), visual feedback study (VF), and visual no-feedback study (VFN) (see Table 2).

3.1.2. Performance and payoff computation

In our experimental setting, realization of the decision-maker's objective function depends both on the decision-maker's performance level and the associated monetary reinforcement (payoff). For all versions of the experiment, we assessed performance in terms of the resulting distance from the target figure (or number) using the following formula:

$$Fitness(\%) = 100 - \left(100 * \frac{d_i(c,s)}{T}\right)\%. \qquad [3.1]$$

In the visual frame, $d_i(c,s)$ is the number of mistakes made by participant i for choice problem c in scenario $s = visual$. Following Mittone and Papi (2017), our design recorded a mistake each time a pixel of the chosen (or constructed) figure differed from the corresponding pixel of the target figure. In the numeric frame, $d_i(c,s)$ is the difference between the absolute value of the target number and that of the number selected (or constructed) by participant i for choice problem c in scenario $s = numeric$. In the visual frame, T is equal to $d_{max}(c,\ s)$, which is the maximum number of mistakes a participant can make for choice problem c in scenario $s = visual$. In the numeric frame, T is equal to the target number.[3]

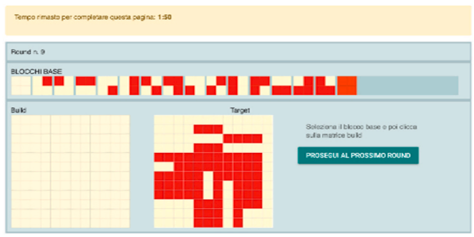
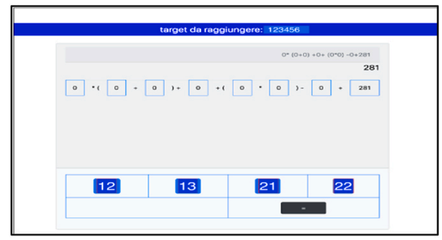
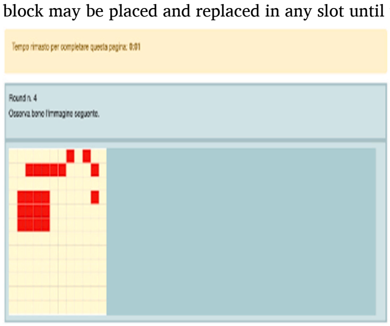
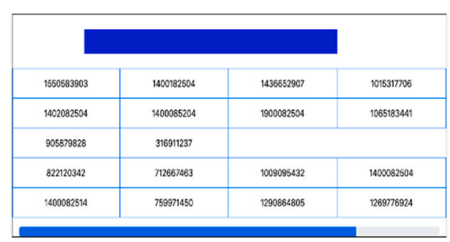
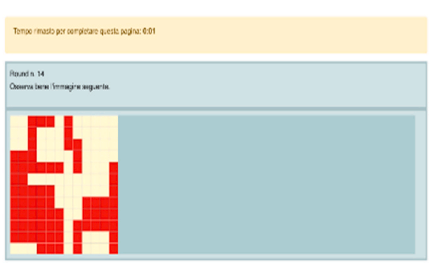
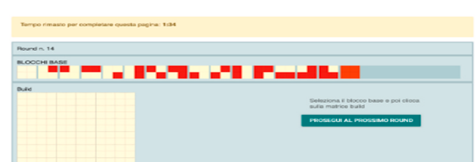
The participants received monetary incentives based on the standard paradigm in experimental economics (Smith 1976). To pay the participants, we computed their performance in terms of their correct approximation of the target, after which we valued their performance with experimental currency units (ECU) using the conversion formula in Table 3. We fully informed the participants about this approach.

[1] For a better understanding of our design, detailed instructions provided to the participants during the experiment can be accessed at https://osf.io/8j3gz/?view_only=6e7c37d42a1d4873983fcec74f28f14b.

[2] For example, describing the task of a manager who is responsible for a complex set of organizational resources, Kantola concluded that "In a way, managers in many cases are forced to work and make decisions with partial information and in violation of right decision" (2016, p. 8).

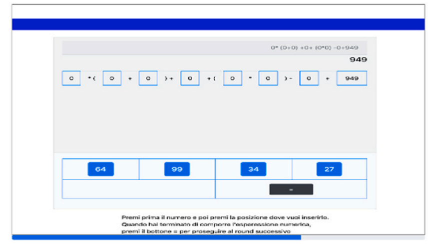
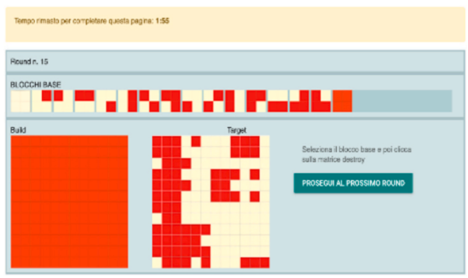
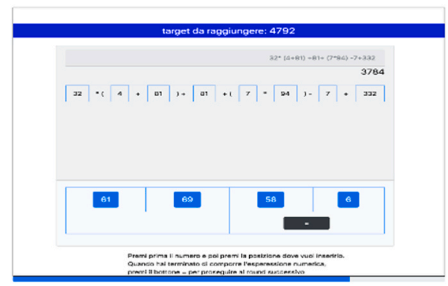
[3] It is important to note that in the Building (Destroy) task in the visual frame, when the participants did not use the available blocks but left the figure as it was (i.e., partitioned into blank or red spots), the performance was, by definition, higher than zero because some pixels were already correctly positioned in relation to the corresponding pixels in the target figure. To make the numeric frame comparable, we assumed that the initial performance was higher than zero by placing a given number in the provided formula.

Table 1
Task descriptions: visual and numeric frames.

Task description	Visual Game	Numeric Game
Building task. To the best of their ability, the participants must construct a target figure (visual) or number (numeric), using a set of blocks in the visual frame or numbers in the numeric frame, as displayed on their computer screen for the entire duration of each round.	 The participants are shown several white and red blocks (along the top), a target figure (made up of red and white blocks), and a blank matrix (totally white) partitioned into blank spots with the same dimensions as the blocks. Starting from the figure provided, the participants are asked to construct a figure by placing the blocks in the slots within 120 s. Any block may be placed and replaced in any slot until the time expires.	 The participants are shown a target number (at the top) and four numbers (along the bottom). Using a predetermined formula, they are asked to construct a number by assigning the available numbers to empty cells within 60 s. Any number may be assigned and reassigned to any cell until the time expires.
Memory task. The participants are shown a target figure or number for a predetermined time (20 s). The target is then hidden, and the participants are presented with a set of 18 figures (including the target) or 18 numbers (including the target). They must choose whichever item they believe to be the target within 120 s in the visual frame or 60 s in the numeric frame.	 	 
The *Combined task* includes elements of both the memory and building tasks. The participants are presented with a target figure or number for a given time (20 s). The target is then hidden, and the participants must build it (as in the *building* task); the difference is that on this occasion, the target figure or number is no longer visible while building. The participants must construct the target figure or number within	 	

(*continued on next page*)

Table 1 (*continued*)

Task description	Visual Game	Numeric Game
120 s (visual frame) or 60 s (numeric frame).		
The *Destroy task* is the mirror-image of the *Building* task. The only difference is that the participants are shown a figure partitioned into red rather than white spots in the visual frame. In the numeric frame, they see a target formula.	 The participants are shown several blocks, including a target figure (on the right) and a figure partitioned into red spots (on the left) with the same dimensions as the blocks. Starting from the figure on the left, they are asked to complete a figure by placing the blocks in the slots within 120 s. Before the time expires, they are allowed to place and replace any block in any slot.	 The participants are shown a target number (at the top) and four numbers (along the bottom). Using a predetermined formula, they are asked to compute a number by placing the available numbers into filled cells within 60 s. Before the time expires, they are allowed to place and replace any number in any cell.

During the active phase, a participant who chose to abandon the training task could obtain a higher payoff (see Table 4).

A higher-performance ECU ratio was introduced to compensate for the costs of learning a new task. That is, the payoff was calibrated to make it less profitable to exploit the familiar task learned in the training phase and more profitable to explore a new task (and to continue to exploit this new task). This feature was designed to motivate exploration and to mimic the issue at the core of the exploitation-exploration dilemma (March, 1991): that exploitation generates more certain low-cost benefits but may generate inertia, while exploration offers potentially higher benefits but also entails higher risks (Greve 2007, Laureiro-Martínez et al., 2015). The higher potential monetary reward for abandoning a familiar was also aimed at counterbalancing the effects of performance-based incentives, which have been shown to reduce exploration and increase exploitation of previously acquired knowledge (Lee and Meyer-Doyle 2017). In our analyses, we will exclusively focus on the payoff.[4]

3.2. The risk elicitation task

After completing the main experiment (the training and active phases), the participants were asked to play the BRET (bomb risk elicitation task) to determine their risk preferences (Crosetto and Filippin 2013). In this task, they were presented with 64 boxes and were informed that one box contained a bomb while each of the remaining 63 boxes contained one ECU. The participants were asked to collect whatever number of boxes they wished, knowing that if they collected the box that contained the bomb, it would "explode" and nullify their earnings for this part of the experiment. The participants' earnings increased each time they opened a box, provided they did not collect the box that contained the bomb; and if they selected the box with the bomb, their earnings were wiped out. This procedure provided an index of risk-taking; the more boxes the participants collected, the higher their propensity for taking risks.

3.3. Conceptualization of exploration-exploitation choices

To conceptualize exploration-exploitation choice at the individual level, we build on March's original definition (1991, p. 71). *Exploitation* refers to the accumulation of decision-making experience through task repetition and related processes of learning and refinement (Ericsson and Lehmann 1996, Laureiro-Martínez et al., 2015). *Exploration* involves the opposite, as decision-makers choose to abandon their current task to find a new and potentially more profitable task (Laureiro-Martínez et al., 2015; March, 1991, Smith and Tushman 2005). This disengagement facilitates "the pursuit of new knowledge, of things that may come to be known" (Levinthal and March 1993, p. 105). According to this definition, in our study's context, we categorize *exploration* as behavior that involves switching to a task that differs from the task performed in the previous round or in the training phase. *Exploitation*, in contrast, is sticking with the same task. Specifically, we define two types of exploitation:

1. "*Exploitation of a familiar task environment*" (ExploitFT) occurs when individuals stick with the task learned during the training phase.

[4] Here, the decision-maker is not interested in improving performance as such but in reaching a higher level of achievement of an objective function associated with a monetary reward. In this regard, we follow the Induced Value Theory of Vernon Smith (1976), which is a standard technique in experimental economics.

Table 2
Overview of the four studies in the main experiment.

		Frame	
		Numeric	**Visual**
Update about performance in each round	**No**	numeric no-feedback study (NNF)	visual no-feedback study (VNF)
	Yes	numeric-feedback study (NF)	visual-feedback study (VF)

Table 3
Conversion scheme: Performance incentives if the chosen task is the one learned in the training phase.

Performance	ECU
100%–95%	400
94%–80%	200
79%–65%	70
64%–50%	20
49%–0%	0

Note. ECU stands for Experimental Currency Units.

Table 4
Conversion scheme: Performance incentives if the chosen task differs from the one learned in the training phase.

Performance	ECU
100%–95%	600
94%–80%	300
79%–65%	100
64%–50%	30
49%–0%	0

2. "*Exploitation of a novel task environment*" (ExploitNT) takes place when individuals opt to continue a new task that was initially explored (i.e., a task that differs from the one they learned in the training phase).

During the training phase (i.e., during the first eight rounds), participants gained confidence by repeatedly performing the same task, which aligned with March (1991)'s definition of exploitation as: "the refinement and extension of existing competences, technologies, and paradigms … returns are positive, proximate, and predictable" (p. 81). In the active phase, beginning with the ninth round, participants faced the choice of either continuing the familiar task from the training phase or exploring a new and unknown one (whose outcome is largely uncertain). Subsequently, from the tenth round on, participants had three options:

1. Persist with the well-known familiar task from the Training Phase (ExploitFT), which assumes a well-known environment (Laureiro-Martinez 2014; Oehler et al., 2019).
2. Maintain the course with the recently explored task (ExploitNT). This essentially means exploiting a new task that was recently explored in the active phase and chosen as an alternative to the one from the training phase.
3. Decide to abandon the current task for another task that differs from the one learned in the training phase or performed in the previous round. We refer to this as the "*Exploration of a task environment*" (*ExploreT*). This entails a choice to switch to a task that deviates from ExploitFT or ExploitNT.

Fig. 1 summarizes this temporal framework, considering a participant trained on the memory task who, for example, had chosen the following sequence of tasks.

In the initial stage, decision-makers are induced to learn a task. Then, they must decide whether to continue exploiting the same task (ExploitFT) or to switch to a new task (ExploreT). If they choose to explore a new task, they can decide to continue with that task to learn and exploit this task (ExploitNT), to go back to the training task (Selten, 1998), or to explore another task.

3.4. Participants and experiment procedures

The experiments were conducted at the Cognitive and Experimental Economics Laboratory (CEEL) of the University of Trento (Italy). The participants were recruited using customized software implemented at the laboratory, and the experiment was programmed and administered using oTree software (Chen et al., 2016, Holzmeister and Pfurtscheller, 2016). To facilitate replication and extensions of our study, we shared the software and study data in an open-access repository.[5] In total, 240 people participated in the experiment; 132 were female and 108 were male.

The average age of the participants was 21.80 years (sd = 2.36). Most were students of economics (54.17%), while the remainder included students of law (21.25%), engineering (7.50%), mathematics and natural sciences (7.50%), social sciences (4.58%), humanities (3.33%), and psychology (1.67%).

Each participant participated in only one study—81 in the numeric-feedback study (NF); another 81 in the visual-feedback study (VF); 39 in the numeric no-feedback study (NNF); and 39 in the visual no-feedback study (VNF). Upon arriving at the laboratory, the participants were randomly assigned to a computer, where they read the instructions for the first part of the experiment (the training phase). An experimenter also read the instructions aloud, and the participants were invited to ask questions to ensure that they understood the instructions.

The participants were told that the experiment would involve two main phases (the training phase and the active phase). They knew that they would participate in both phases, but none of them knew in advance the purpose of the second phase. After completing the training phase, they received instructions regarding the active phase. Finally, they were asked to play the BRET (Crosetto and Filippin 2013) and to answer a short demographic questionnaire. The experiment lasted approximately 40 min. Besides being given a show-up fee of €3, the subjects received a certain amount of experimental currency units (ECU) depending on the choices they made in the two parts of the experiment and the BRET. They were informed that their payment would be based on the results that they would obtain in two rounds, which would be randomly extracted at the end of the experiment—one from the training phase and one from the active phase. Based on a conversion rate of 100 ECU = 1 euro, the average payment was €11.07 (including the show-up fee).

4. Results

In this section, we begin by presenting our analysis of the performance results and the response time evolution during the first phase of the game, that is, during the training phase (section 4.1). Then we present our analysis of the data during the second phase of the game, that is, during the active phase (section 4.2). This is followed by a multivariate analysis of several determinants of explorative-exploitative behavior in the active phase (section 4.3).

4.1. Training phase with two different initial conditions

Table 5 reports the means and standard deviations of the

[5] The software and data can be accessed at https://osf.io/8j3gz/?view_only=6e7c37d42a1d4873983fcec74f28f14b.

participants' performance in each experimental condition during the training phase. Regardless of the frame, their average performance increased significantly when they were asked to play the memory task rather than the building task. A post-hoc statistical power analysis conducted with G*Power (Faul et al., 2009) revealed an achieved power (1-β) greater than 0.9 at the standard 0.05 alpha error probability. Only in the Numeric (with feedback) scenario, we achieved a power greater than 0.59 at the standard 0.05 alpha error probability.[6]

Thus, we can deduce that the two training tasks had different degrees of complexity.

Based on the arguments outlined in Section 2, we expected that the training phase would trigger the participants' learning processes. As their learning deepens, participants are likely to become more efficient, that is, their performance should improve over time, and/or they should be able to complete their tasks sooner (Laureiro-Martinez 2014). Therefore, we computed the average round-by-round performance during the training phase, and we also measured the average round-by-round response time needed to perform the task (Cohen and Bacdayan, 1994; Laureiro-Martinez, 2014; Oehler et al., 2021). As shown in Fig. 2, the average performance first increased quickly and gradually stabilized after the first three rounds among the participants who performed the building task; and for those who performed the memory task, the average performance remained relatively stable. This finding highlights the tasks' different degrees of cognitive complexity; the memory task was simpler, and the participants could make better use of their pre-existing skills. Fig. 3 shows that the average response time needed to accomplish the task diminished during the training phase, regardless of the task (building or memory) or frame (visual or numeric). The only exception to this general tendency is the experimental condition of numeric memory no-feedback. It is important to note that although the memory task was simpler, the participants completed it with increasing efficiency as the rounds progressed—that is, the average time needed to complete the task decreased.

Figs. 2 and 3 also reveal differences in the average performance and response time with and without performance updates. These results support our expectations based on the literature. Specifically, the participants who received updates on their performance in each round performed better throughout the building task than those who played the no-feedback version. The t-tests revealed significant differences in average performance for all rounds other than round 4 of the numeric frame ($p < 0.05$ to $p < 0.0005$); and in contrast, the visual frame exhibited significant differences only in rounds 1, 3, and 4 (p-values 0.099, 0.035, and 0.067, respectively). On the other hand, feedback seemed to have played no role among the participants in the memory task, with significant differences noted only in rounds 4 and 7 in the numeric frame ($p = 0.09$ and $p = 0.079$, respectively).

In terms of response times, the participants who received updates on their performance in both the visual and numeric frames were faster than those who did not receive any performance updates. The t-tests revealed significant differences in the majority of the rounds for both the building and memory tasks in the numeric frame, with the p-values ranging from <0.1 to <0.005. In the visual frame, significant differences were detected only in rounds 2 and 5 (the memory task) and in rounds 4 and 5 (the building task), with the p-values ranging from <0.1 to < 0.05.

4.2. Path-dependent behaviors that emerged from the experimental data

In the previous section, we revealed that individual responses to learning in the training phase differ significantly based on the complexity of the task environment—whether the task is simple (memory) or complex (building). A compelling question then arises:

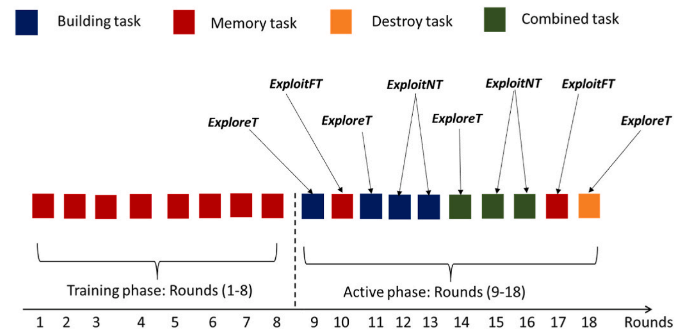


Fig. 1. Temporal framework: Exploitation and Exploration behaviors. ExploreT (exploration of a task environment); ExploitFT (exploitation of a familiar task environment); ExploitNT (exploitation of a novel task environment).

How do the two distinct groups of participants, each exposed to different training sets (memory and building), perform in the subsequent active phase of the tournament? Specifically, we are interested in understanding how learning in two different sets affected the decision to exploit or to explore. We operationalized ExploreT, ExploitFT, and ExploitNT as dummy variables (see Table 7).

Fig. 4 illustrates how exploitation (both ExploitFT and ExploitNT) and exploration (ExploreT) decisions alternated across successive rounds during the active phase (i.e., from round 9 to round 18).

We have evidence that the initial conditions of the game trigger different strategies in the active phase: players who had originally learned a simple task (the memory task) decided more frequently to choose the same task in the second tournament of the game. On the other hand, players trained on the complex task (the building task) opted to exploit a new task more often, that is, a task learned in the second tournament of the game. Fig. 5 reports how many times each of the four available tasks was chosen during the active phase (i.e., from round 9 to 18).

The results reported in Fig. 5 confirm that the majority of the participants trained on the memory task were stuck on it in the active phase, and at the same time, that the memory task was also the most frequently chosen for *exploiting a novel task environment/ExploitNT*. Indeed, it could be observed that the participants trained on the building task, once they discovered the memory task, prevalently got stuck on it in the active phase. However, the most frequent choice of the memory task for both groups is the outcome of two completely distinct decision processes. For those trained on the memory task, the fact that it was the most frequent choice was merely the consequence of a lock-in effect of the task that they were forced to learn during the training phase. For those trained on the building task, it was the consequence of the "discovery" of a simpler task and the ensuing decision to invest the necessary effort to exploit a novel task environment. The results allow us to say that the lock-in effect of the memory task was very robust, that is, it was applied independently from having been exogenously imposed by the experimenters for a set of repetitions.

A natural question that emerges is why those who were trained on the simpler memory task repeatedly employed the same task even in the presence of explicit incentives to explore, while the participants trained on the more complex building task were able to learn new ones. The answer might be found in the role played by the previously mentioned mechanization of thought (the *Einstellung* effect), which enabled "individuals to pass from effortful mental activity to partially automatic, unconscious, and effortless mental operations" (Egidi 2105, p. 198). Our results in section 4.1 show that the participants who played the memory task immediately reached the high-performance level. It is, therefore, possible to argue that the cognitive process involved in solving the memory task becomes automatic and effortless quite soon. According to Egidi (2015), such an automatic process can direct players' attention towards a familiar task. In other words, the strength of the learning

[6] We share this analysis in the same open-access repository where we share our software and data (https://osf.io/8j3gz/?view_only=6e7c37d42a1d4873983fcec74f28f14b).

Table 5
Performance across experimental conditions (training phase): Mean (standard deviation).

Frame	Average performance (Memory task)	Average performance (Building task)	Difference (Memory-Building) Mann-Whitney *U* test)
Visual (with feedback)	98.20 (8.92)	88.26 (12.41)	($p < 0.0001$)
Numeric (with feedback)	98.67 (9.8)	88.09 (22.19)	($p < 0.0001$)
Visual (no feedback)	98.82 (7.54)	84.83 (14.38)	($p < 0.0001$)
Numeric (no feedback)	98.64 (9.89)	70.26 (39.37)	($p < 0.0001$)
All	98.53 (9.18)	84.78 (23.21)	($p < 0.0001$)

experience in capturing participants' attention has a key role in triggering the search for a new alternative (Egidi 2015; see also Bilalić et al., 2008, 2010).

In terms of ex-post speculation, our experimental results can be interpreted through the framework of cognitive strategy, understood in terms of well-structured versus ill-structured problems, as well as Type 1 and Type 2 mental processes (Laureiro-Martínez and Brusoni, 2018). The literature mainly refers to Type 1 and Type 2 as "System 1" and "System 2" respectively (Kahneman, 2003). The key distinction between decision-making tasks performed using System 1 or System 2 lies in the cognitive processes involved. System 1 is cognitively effortless, fast, and automatic, whereas System 2 is cognitively demanding, slow, and rule-governed.

In our specific case, the difference between memory and building tasks lies in the cognitive process required to solve the decision problem. Memorization is a well-known cognitive tool, often employed by experimental subjects in their everyday lives, classifying the 'memory task' as a well-structured cognitive problem, thereby allowing subjects to be more likely to activate System 1 (Type 1). The *Einstellung* effect facilitated participants in swiftly reaching high-performance levels by automating the cognitive process. This automation induced a form of cognitive inertia, deterring them from venturing into new tasks even when incentivized, making them overly reliant on a particular mental framework.

Conversely, the 'building' task, which necessitates understanding the game's rules and applying them to achieve different objectives, belongs to the ill-structured cognitive problems. This necessitated the activation of System 2 (or Type 2) mental processes, which are slower and demand greater cognitive control for thoughtful deliberation. This implies that when subjects were engaged in the unfamiliar building task, they were more likely to activate System 2 (or Type 2) and start a costly learning process to find the most efficient strategy to reach high-performance levels. This translated into the acquisition of a problem-solving method that subjects could exploit in the active stage of the experiment, progressively shifting from the use of System 2 to System 1, exploiting the acquired knowledge of the task. This ultimately translates into the rise of a tendency towards sticking to the learned task. This phenomenon is more prevalent in those subjects trained in the memory task because this task implies using a cognitive skill already owned by the subjects without the need to learn it.

In summary, participants engaged in the well-structured memory task often found their learning to be a double-edged sword: it solidified existing knowledge but also led to cognitive inertia, inhibiting the exploration of novel tasks. Conversely, the ill-defined nature of the building task demanded a more flexible cognitive approach, promoting adaptability and encouraging the exploration of new tasks. It is crucial to acknowledge that these are speculative interpretations, and further empirical work is needed.

4.2.1. *Person-by-person analysis of the results*

We have analyzed the results of the overall merged data. We have identified the existence of cognitive traps (the *Eistellung* effect) in the learning process, which in turn affected the decision to explore-exploit. We now concentrate on the individual respondents and analyze the results by focusing on the choice patterns of each participant (see also Egidi and Narduzzo 1997; Laureiro-Martínez et al., 2015).

We start checking two extreme types of behaviors: the "strong ExploitFT-participants" and the "strong ExploitNT-participants". The former are participants who never switched from the task learned during the training phase. The latter, in contrast, explored only one or two tasks at the very start of the active phase (in rounds 9 and 11) and then slacked on that task to the end of the experiment.[7]

Additionally, we individuate Retreaters, that is, participants who, after the training phase, explored a new task only once at the very start of the active phase (between rounds 9 and 12) before deciding to revert to the training task for the entire duration of the game. Their tendency to avoid tasks with a lower outcome on the first trial can be interpreted as the "hot stove effect" (Denrell and March 2001), which entails retreating to the training task because of their inherent resistance to exploring new alternatives.

Table 6 presents descriptive statistics related to the game design, training task type, and which kind of exploitation (ExploitFT or ExploitNT) they engaged in (for their demographic statistics, see Appendix A2).

The effects of the training task show clearly that exposure to the memory task is strongly associated with Retreaters, 87% of whom received training on the memory task. Table 6 reports a similar result for the strong ExploitFT-participants, 76% of whom were trained on the memory task. Interestingly, 85% of the strong ExploitNT-participants (who explored once or twice before quickly choosing to exploit a new task) were trained on the more complex building task. Taken together, these findings further confirm that the type of learned task in the training phase fundamentally influences the decision to explore or exploit.

The broader implication is that any individual who achieves very high performance on a simple task is more likely to repeat the same task, even when there are strong incentives to abandon it. Conversely, an individual who learns a more challenging procedure (reflecting the greater intrinsic difficulty of the task), where low initial performance gradually improves with practice, is less likely to be trapped in the established task and is more likely to explore new options. Indeed, searching for new alternatives implies facing a dilemma: on the one hand, a short-term desire to experience the familiar and favorable outcomes, and on the other hand, a long-term desire to explore new and potentially more promising tasks that may increase utility in the future (Hoeffler et al., 2006). We argue that the participants who had a more favorable experience were more likely to be trapped in the

[7] This category includes one subject who switched in the last round and one subject who switched once in round 14 before immediately returning to the training task.

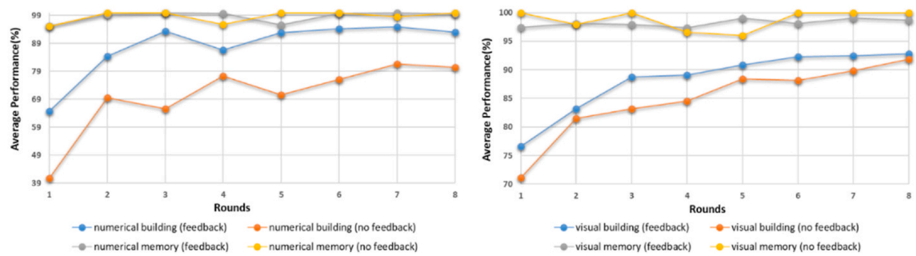


Fig. 2. Average performance during the training phase (rounds 1–8): (left) Numeric frame; (right) Visual frame.

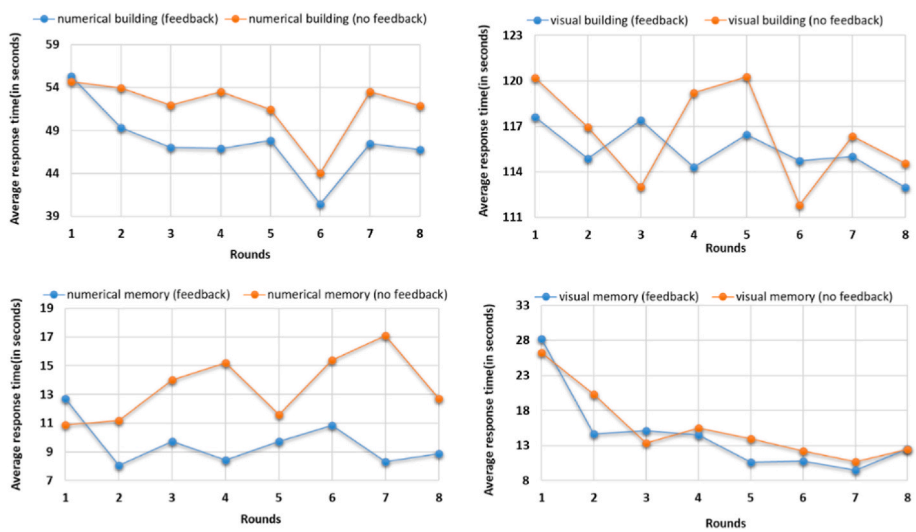


Fig. 3. Average time response (sec) during the training phase (rounds 1–8): (left) Numeric frame; (right) Visual frame.

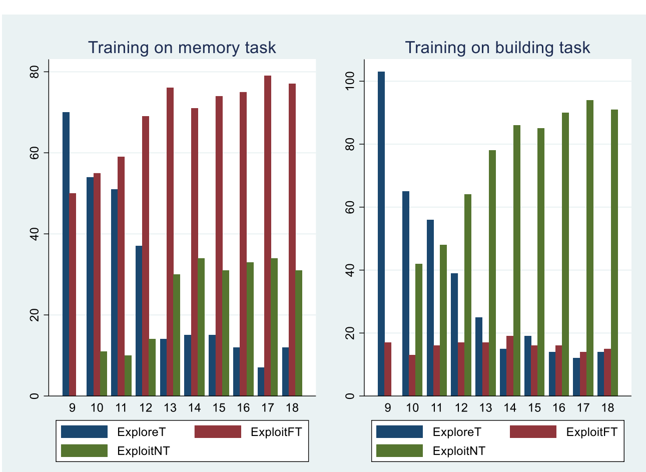


Fig. 4. Frequency distribution of the exploitation-exploration decisions in the active phase.

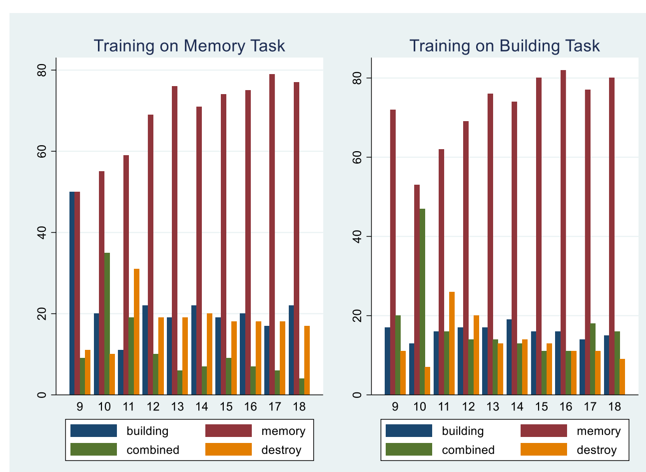


Fig. 5. Frequency distribution of the tasks chosen in the active phase.

Table 6
Composition of types.

Choice pattern	Training (*Memory*)	Game (*Numeric*)	Type of Exploitation
Strong ExploitFT-participants (N = 29)	76%	41%	ExploitFT
Retreaters (N = 23)	87%	39%	ExploitFT
Strong ExploitNT-participants (N = 54)	15%	35%	ExploitNT

mechanization of thought, which in turn triggered a more myopic search.[8]

4.3. Determinants of exploitation and exploration behavior

Here, we investigate the effect of initial conditions and performance feedback as antecedents of the decision to explore or exploit. That is, we assume that decision-makers engage in satisfying behavior, adapting their aspiration level based on the results of past experiences. Before introducing our regression analyses, we will explain how the independent and dependent variables were operationalized (as summarized in Table 7).

Dependent Variables. Our main variables of interest are exploration of a task environment (ExploreT),[9] exploitation of a familiar task environment (ExploitFT), and exploitation of a novel task environment (ExploitNT). As anticipated in the previous section, we operationalized these variables as dummy variables. Based on our observations during the active phase, we ran a set of logit regressions (including individual random effects) to investigate the determinants of these three variables.

Independent Variables. We start by considering the effect of the kind of training task played from rounds 1 to 8 (the *Training* variable). Then, we analyze the role of feedback on exploration-exploitation behavior. Specifically, inspired by the most recent experimental literature (Billinger et al., 2021), we consider different feedback variables. First, the literature highlights the role of feedback based on recent experiences (e.g., Billinger et al., 2014; Van Rijnsoever et al., 2012), as there was a progressive decline in the impact of older experiences. In particular, if the outcomes of the decision-makers' most recent choices are perceived as failures, the decision-makers are less likely to make those choices again and are more likely to explore a different choice. Conversely, if the decision-makers perceive the outcomes of their most recent choices as successes, they are likely to repeat those choices (Hoeffler et al., 2006; Lave and March 1993; Sitkin and Pablo 1992; Van Rijnsoever et al., 2012). To test the effects of recent experience, we measured the payoff from the previous round (*Last Payoff Experienced*). Second, while the decision-makers' most recent experience arguably matters the most, the behavioral research stream also suggests a role for longer-term aspiration levels (Cyert and March 1963, Greve 1998, Lant TK Montgomery, 1987; Levitt and March 1988, March, 1988). This means that the decision-makers revise their aspiration level based on all the feedback that they have received relative to their recent performance (Billinger et al., 2014; MacLeod and Pingle 2005). This argument is supported by the prospect theory of Kahneman and Tversky (1979), which suggests that decision-makers evaluate their ongoing performance against a subjectively fixed reference point. On that basis, current performance is perceived as a failure if it falls below such a reference point, increasing the likelihood of exploration (Bromiley 1991; Greve 1998, March, 1988). Exceeding aspiration levels leads to risk aversion and more exploitative behavior (Billinger et al., 2014; Levinthal and March 1993). Accordingly, following Levine et al. (2019), we consider the difference between the current round payoff and the past best payoff as an assumed proxy of the variable *Aspiration-Payoff Gap*. Third, we also include the variable *Average Payoff*, as recent experimental literature suggests that the higher the average payoff is, the less likely exploration is (Billinger et al., 2021).

Our econometric analysis additionally controlled for the training task and its interactions with the three variables (*Aspiration-Payoff Gap*, *Last Payoff Experienced*, and *Average Payoff)* to check whether the effect of these three forms of feedback varied across the two different training experiences.

Finally, we added control variables, including the participants' demographics. Specifically, we included a set of dummy variables to isolate the presence or absence of updates on performance (the *Feedback* variable) and the different frames (the *Game* variable), and we controlled for the impact of risk preferences (*BRET*), *Age* (in years), a dummy variable for gender (*Male*), a dummy variable for economics students (Economics), and a dummy variable to capture participation in previous experiments (*Number of Experiments*). Regarding risk, it is well established in the decision-making literature that individuals exhibit different attitudes to risk (see, for example, Diamond and Stiglitz 1974), which may affect explorative behavior. Based on organizational data, Miller and Chen (2004) found empirical support for March and Shapira's (1987) model, which includes risk-taking when performance falls below the aspiration level. On that basis, it was also considered important to control for individual attitudes to risk. We also controlled for gender since the exploration-exploitation literature presents evidence of gender-related differences in behavior (Mehlhorn et al., 2015).

Regression Results. Table 8 displays the results of the logit regressions. Since the results of logit, that is, nonlinear models, cannot be easily interpreted (Hoetker 2007), we report the marginal effect of the explanatory variable of interest, that is "the effect of a unit change in an explanatory variable on the dependent variable" (Wiersema and Bowen 2009, p. 682), by keeping the other variables at observed values (results not shown in the exhibits).

Models 1, 5, and 9 controlled for the task learned during the training phase (simple or complex, where *simple* refers to memory and *complex* refers to building). We confirm the results we already discussed in the previous section: training on a simple task reduced exploration (ExploreT) by 6.24%, and exploitation of a novel task environment (ExploitNT) by 37.61%, while it increased exploitation of a familiar task environment (ExploitFT) by 42.66%. Models 2, 6, and 10 show how the independent variables *Aspiration-Payoff Gap*, *Last Payoff Experienced*, and *Average Payoff* influenced our dependent variables. A negative gap (i.e., payoff falls under aspirations) increased ExploreT and reduced

[8] Evidence of this effect has been demonstrated by Hoeffler et al. (2006) (p. 218): "*If people are overly focused on extracting immediate utility, they should be more likely to select an experience that is similar to the favorable one they recently enjoyed. However, people who have a neutral or moderately negative experience should not be affected to the same degree. Ironically, such dependency on the quality of earlier experiences also implies that favorable initial experiences may prevent people from experimenting with dissimilar other products, leading to a biased search process in which entire regions of potentially attractive alternatives are relatively unlikely to be discovered. The starting point may heavily influence which particular region people select from initially, and favorableness and myopic search are the mechanisms that limit their search in that particular region*".

[9] In Appendix A1, we include an alternative analysis of antecedents of exploration conceptualized as a unique instance of trying a new task. Our main exploration variable refers only in general to trying new tasks that deviate from an established task or from a previous task.

Table 7
Variables and measurement.

Dependent variables	Explanation of the measurement
Exploration of a task environment (ExploreT)	Binary variable [0,1]. If the participant changed tasks from the previous round but has not retreated to the task learned in the training task, the value is 1 (otherwise, 0).
Exploitation of a familiar task environment (ExploitFT)	Binary variable [0,1]. If the participant remained in or returned to the same task learned in the training task, the value is 1 (otherwise, 0).
Exploitation of a novel task environment (ExploitNT)	Binary variable [0,1]. If the participant repeated consecutively (i.e., for at least two trials) a task different from the task learned in the training phase task, the value is 1 (otherwise, 0).
Independent variables	
Average Payoff	Continuous variable: Average of accumulated payoffs from the start of the game to the current round (variable standardized)
Last payoff Experienced	Continuous variable: Payoff in the previous round (variable standardized)
Aspiration-Payoff Gap	Binary variable [0,1]. Computed as the difference between the payoff in the current round and the maximum payoff across all past rounds. If the difference is negative, the value is 1 (otherwise, 0).
BRET (risk-measure attitude)	Number of boxes collected in the bomb risk elicitation task (BRET). Higher values indicate a higher level of risk taking.
Game	Binary variable [0 = visual, 1 = numeric]
Gender	Binary variable [0 = female,1 = male]
Feedback	Binary variable [0 = no, 1 = yes] (received performance feedback or not)
Training Task	Binary variable [0 = build, 1 = memory] (task in previous 8 rounds)
Age	Continuous variable (years)
Economics	Binary variable [0, 1]. If participants have a background in Economics, the value is 1 (otherwise, 0).
Number of experiments	Continuous variable (previous experiments in which the subject participated)

both ExploitFT and ExploitNT. Specifically, falling below the aspiration level increased the likelihood of exploration (ExploreT) by 16.63% while reducing the likelihood of ExploitFT and ExploitNT by 6.43% and 4.14%, respectively. These findings align with recent experimental evidence (Levine et al., 2019) and offer strong support for March's proposition regarding the role of the aspiration-performance gap in boosting exploration and reducing exploitation (March, 1988). The results also indicate that a higher recent payoff (variable *Last Payoff Experienced*) reduces the probability of ExploreT and ExploitFT by 5.47% and 2.72%, respectively, while it increases the probability of ExploitNT by 8.24%. This is consistent with earlier findings that repetition of a choice rather than further search can be expected if the outcome of the most recent choice was successful (Lave and March 1993; Sitkin and Pablo 1992; Van Rijnsoever et al., 2012). The findings also suggest that the average payoff (*Average Payoff*) decreases the probability of ExploreT by 11.14%, while increasing the likelihood of ExploitNT by 15.93%.

Models 3, 7, and 11 investigated exploration-exploitation behavior in greater detail by also controlling for the interaction terms between the variable training and the three feedback variables (*Aspiration Payoff Gap*, *Average Payoff*, and *Last Payoff Experienced*). To reduce collinearity, the *Average Payoff* and *Last Payoff experienced* were standardized. As coefficients offer no satisfactory basis for inference of interaction effects in a logit model (e.g., Hoetker 2007; Zelner 2009), we adopted Zelner's (2009) approach to test the statistical significance of the marginal change in the likelihood of exploring and exploiting a familiar (new) task due to an increase in the *Average Payoff* and the *Last Payoff Experienced* at the two different values of the dummy variable *Training*, setting the remaining covariates as observed values. We also tested the statistical significance of the marginal change in the likelihood of exploring and exploiting a familiar (new) task when trailing aspiration at the two different values of the dummy variable *Training*. We rely on graphical analysis (see Figs. 6–8) to offer a better understanding of the interaction effects (The results are based on the estimates computed in the full models, columns 4, 8, and 12 of Table 8).

We notice that an increase in *Average Payoff* reduces the odds of ExploreT (the marginal effect of the average payoff is always negative and significant) and increases the likelihood of ExploitNT (the marginal effect of the average payoff is always positive) irrespective of the specific training task.[10] At the same time, regardless of the specific training task, an increase in the average payoff does not affect the probability of exploiting a familiar task (ExploitFT), as the confidence interval (i.e., the vertical bar) includes zero. Therefore, we conclude that the effect of the average payoff so far is not conditioned on the initial conditions, even if such effect is more sizeable in the case of the participants who were trained on the more complex building task.

Interestingly, we previously noticed that a recent high payoff (variable *Last Payoff Experienced*) increased the tendency to exploit a novel task environment. Our analysis of the marginal effects revealed that this effect is contingent upon the establishment of initial learning in the more complex building task. In fact, an increase in the *Last Payoff Experienced* reduces the odds of exploring and exploiting a familiar task and increases the probability of exploiting a novel task only in the building training task condition. On the other hand, an increase in the recent payoff does not affect the probability of exploring and exploiting a familiar task, while it slightly increases the probability of exploiting a new task in the memory task condition.[11]

On what is regarded as the role of the aspiration gap, we found that falling below the aspiration level increases the likelihood of ExploreT by 27.21% in the memory task, while no effect was found in the building task. On the other hand, trailing aspiration decreases the likelihood of ExploitFT and increases the likelihood of ExploitNT in the memory task, while the opposite is true in the building task: falling below the aspiration level increases the likelihood of ExploitFT and decreases the likelihood of ExploitNT.[12].

Finally, additional control variables were also incorporated in Models 4, 8, and 12, including the presence or absence of performance

[10] Specifically, the average change in the probability of exploration (ExploreT) when the average payoff increases is −30.69% (p = 0.000) in the complex task and −3.21% in the simple task (p = 0.015). The average change in the probability of ExploitNT when the average payoff increases is 27.17% (p = 0.000) in the complex task and 9.24% in the simple task (p = 0.000).

[11] Specifically, the average change in probability of exploration (ExploreT) and ExploitFT when last payoff experienced increases is respectively equal to −5.82% (p = 0.000) and −3.48% (p = 0.000) in the complex task. The average change in probability of ExploitNT when last payoff experienced increases is equal to 9.35% (p = 0.000) in the complex task and 2.7% in the simple task (p = 0.033).

[12] Specifically, the average change in the probability of ExploitFT when trailing aspirations is equal to 19.22% (p = 0.000) in the complex task and −23.93% in the simple task (p = 0.000). The average change in probability of ExploitNT when trailing aspirations is equal to −12.12% (p = 0.000) in the complex task and 4.7% in the simple task (p = 0.067).

Table 8
Determinants of ExploreT, ExploitFT, and ExploitNT.

	DP variable: ExploreT				DP variable: ExploitFT				DP variable: ExploitNT			
VARIABLES	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	model 9	model 10	model 11	model 12
Training (Memory = 1)	−0.357*	−1.556***	−2.891***	−2.891***	5.016***	4.888***	6.637***	6.628***	−2.886***	−1.594***	−0.720	−0.667
	(0.015)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.104)	(0.121)
Average Payoff		−0.785***	−2.527***	−2.512***		−0.007	0.169	0.138		1.692***	3.632***	3.635***
		(0.000)	(0.000)	(0.000)		(0.953)	(0.653)	(0.714)		(0.000)	(0.000)	(0.000)
Asp_Payoff_gap(neg. gap = 1)		1.063***	−0.342	−0.288		−0.853***	3.219***	3.184***		−0.440*	−1.521***	−1.590***
		(0.000)	(0.141)	(0.216)		(0.000)	(0.000)	(0.000)		(0.026)	(0.000)	(0.000)
Last Payoff Experienced		−0.386***	−0.467***	−0.477***		−0.350***	−0.747***	−0.725***		0.876***	1.246***	1.251***
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Training x Average Payoff			2.272***	2.243***			−0.284	−0.285			−2.615***	−2.598***
			(0.000)	(0.000)			(0.479)	(0.477)			(0.000)	(0.000)
Training x Asp_payoff_gap			2.096***	2.126***			−5.299***	−5.453***			1.964***	2.105***
			(0.000)	(0.000)			(0.000)	(0.000)			(0.000)	(0.000)
Training x Last payoff exp.ed			0.323*	0.324*			0.584**	0.587**			−0.932***	−0.950***
			(0.014)	(0.015)			(0.007)	(0.007)			(0.000)	(0.000)
Feedback (Yes = 1)				−0.055				−0.221				0.308
				(0.780)				(0.672)				(0.441)
Game (Numeric = 1)				0.783***				−0.077				−0.922*
				(0.000)				(0.879)				(0.016)
Gender (Male = 1)				0.017				−0.439				0.417
				(0.928)				(0.389)				(0.277)
BRET				0.001				−0.006				0.001
				(0.823)				(0.693)				(0.927)
Economics				0.283				−0.994+				0.594
				(0.144)				(0.051)				(0.130)
Age				−0.037				0.130				0.013
				(0.384)				(0.228)				(0.880)
Num_Experiments				−0.023+				−0.019				0.036
				(0.064)				(0.559)				(0.139)
Constant	−0.959***	−0.564***	0.695***	1.105	−4.559***	−4.170***	−5.513***	−6.944**	0.272	1.758***	−2.352***	−3.263+
	(0.000)	(0.000)	(0.000)	(0.254)	(0.000)	(0.000)	(0.000)	(0.006)	(0.179)	(0.000)	(0.000)	(0.091)
Observations	2400	2400	2400	2360	2400	2400	2400	2360	2400	2400	2400	2360
Number of codeid	240	240	240	236	240	240	240	236	240	240	240	236

Pval in parentheses.
***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, + $p < 0.1$.
Note. ExploreT stands for Exploration of a task environment; ExploitFT stands for Exploitation of familiar task environment; ExploitNT stands for Exploitation of a novel task environment.

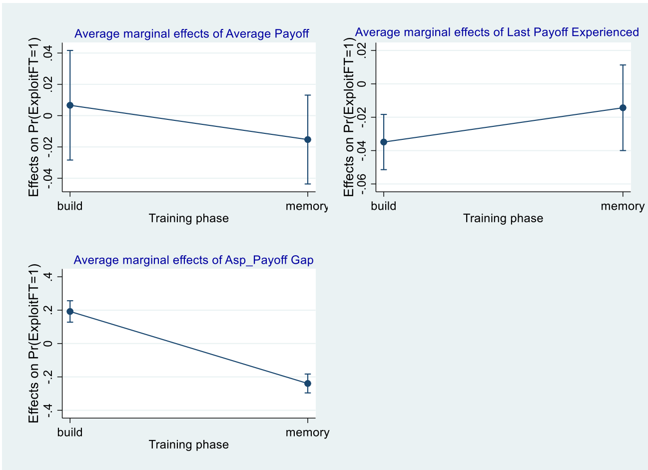


Fig. 6. Marginal effects of *Average payoff*, *Last payoff experienced*, and *Aspiration payoff* on exploitation of a familiar task environment (ExploitFT).

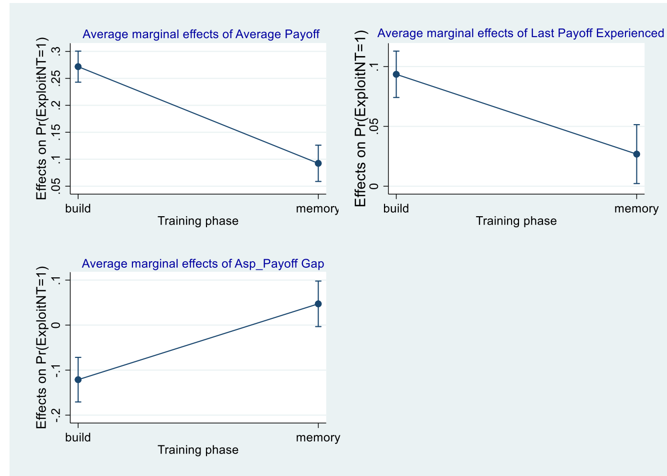


Fig. 7. Marginal effects of *Average payoff*, *Last payoff experienced*, and *Aspiration-payoff gap* on exploitation of a novel task environment (ExploitNT).

feedback, the framework (visual or numeric), the risk attitude (*BRET*), the participation in prior experiments (*Number of Experiments*), and demographic variables (*Gender*, *Age*, and *Economics*). These confirmed the results already discussed and indicate that the presence of performance feedback does not affect exploitation-exploration behavior[13] and that playing the numeric game is associated with a slight increase in exploration. Additional control variables play no explanatory role in relation to the dependent variables (except for *Economics* and *Number of Experiments*, which reduce the likelihood of ExploitFT and ExploreT, respectively).

5. Discussion and implications

Our study's main finding is that initial learning in task environments with different levels of complexity leads to different types of decision patterns in exploration-exploitation choice. We found that individuals who were trained on a simple task tend to continue exploiting the training stage task, either by directly continuing the exploitation after the training stage or later retreating to the initial task after experiencing exploration. Conversely, those who have built their learning on a more complex task are more likely to explore new task environments and, later, to continue exploiting a newly learned task.

An important takeaway is that being conditioned to initially simple task environments might lead to being "trapped" in them. Despite incentives to leave the training task, we found that many participants were unwilling to try alternative tasks after the training phase. These results imply the existence of an individual-level cognitive trap that we call the *easy training trap*. This bias can be seen as a special case of the competency trap (see Denrell and Le Mens, 2020; Levinthal and March 1993), which means that suboptimal behavior driven by initial conditions may persist for long periods despite financial incentives to switch to something different. Consequently, analyses of individual search behavior should consider, at the same time, the exploration-exploitation choice and the initial conditions that precede that choice.

Our study also sheds more light on how different types of performance feedback, in interaction with initial learning conditions, affect

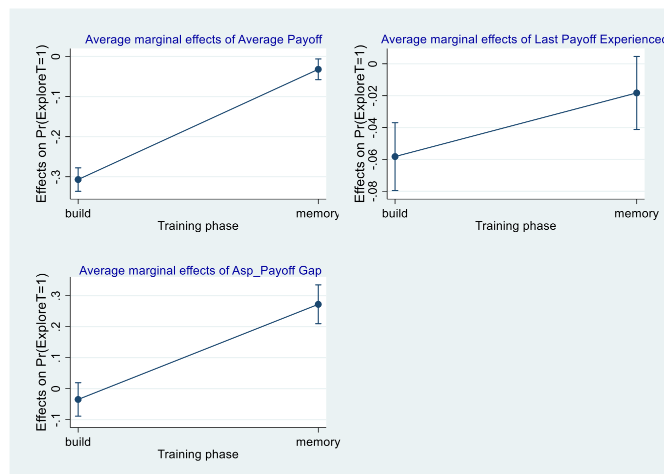


Fig. 8. Marginal effects of *Average payoff*, *Last payoff experienced*, and *Aspiration-payoff gap* on exploration (ExploreT).

individual search behavior. Interestingly, we found that not only do different kinds of feedback differently affect the decision to explore-exploit, but also that feedback differently affects the decision to explore-exploit contingent upon the task environment. First, we found that subjects with larger average cumulated payoffs are less likely to explore, regardless of whether they were trained in a simpler or more complex task. Second, favorable recent feedback (i.e., the payoff in the previous round) increases the tendency to exploit a novel task. However, this result is contingent upon establishing initial learning in the more complex building task. Simply put, after experiencing success in a new, simpler task, participants are more likely to focus on improving in that same task (Sydow et al., 2009). As a result, they become less inclined to return to the previously established, more complex task especially if they continue to perform well in the new task (i.e., as long as they continue to receive high payoffs). Third, in terms of the aspiration-performance gap, as suggested by the behavioral theory and by more recent experimental literature (Levine et al., 2019), we found that lagging behind long-term aspirations triggers exploration; that is, exploration is promoted by a poor payoff relative to the best previous payoff. Interestingly, this result is contingent upon the initial learning in the simpler task. In this case, falling below the aspiration level increases the likelihood of exploration. However, the opposite is true for subjects who had been trained in the more complex task. For those subjects,

[13] The results in section 4.1 show that the participants can improve their performance without explicit feedback over consecutive rounds when repeating the same task. On that basis, we can deduce that such repetitions enabled the participants to develop a sufficient approximation of their performance. This may explain why feedback does not significantly influence exploration-exploitation choices.

falling below the aspiration level increases the likelihood of exploiting the familiar task. These results highlight the contingency role of the task environment complexity in problemistic search, helping to clarify the surprisingly mixed empirical evidence to date regarding the effect of the aspiration-performance gap on exploration (Posen et al., 2018).

Overall, our findings support the argument that the complexity of the learning environment affects aspiration setting (MacLeod and Pingle 2005), and, therefore, has a major effect on choices to explore or exploit. Our results add nuance to the literature on the individual search that has thus far not considered the role of feedback contingent upon the initial learning conditions. Importantly, our arguments on learning in simple and complex task environments should be interpreted as relative changes in the exploration-exploitation tendencies. Regardless of the initial conditions, we still expect decision-makers to be driven by aspiration adaptation. However, we argue that individual search is conditioned by the underlying path-dependent learning, which changes the relative strength of the aspiration adaptation process.

5.1. Implications for experimental research on exploration-exploitation choice

Our experimental setting allowed us to investigate the degree of complexity of the task environment that the participants experienced early and the latter choices that were affected by those initial conditions. Furthermore, our setting allowed us to integrate into it the performance feedback with cognitive mechanisms and biases such as the competency trap (Denrell and Le Mens, 2020; Levinthal and March 1993). Unlike previous experiments on exploration-exploitation choice that typically started from a "blank slate" with no past reference point—as, for example, in the bandit problems (Puranam et al., 2015) or in the Billinger et al. (2014, 2021) alien game —our design builds a reference point artificially through path-dependent learning, revealing the connection between the complexity of initial conditions and individual search. The novelty of our study is that we introduce two dimensions of the decision-making process: at the first level, the learning process (i.e., the training phase) within a particular task environment; and at the second level, the search for different task environments (i.e., the active phase). By introducing these two dimensions, our setup captures the two fundamental elements of the definition of exploration by March (1991): "the novelty of alternatives and the uncertainty of outcomes" (Laureiro-Martínez et al., 2015, p. 322). Future experimental designs could take into account the relevance of the path-dependent learning that affects future exploration-exploitation choice to investigate different aspects of this phenomenon, such as the type of initial learning (e.g. the degree of difficulty, or nature of the task), or the time spent learning a particular task, and so on.

Our experimental results provide an outlook on individual search that demonstrates how stabilization of learning leads to a general persistency of exploitation that is regulated via performance feedback and how it is affected by the complexity of initial conditions of the task environment. Thus, our findings suggest the value of combining the insights from the literature on individual path-dependency learning (Luchins, 1942; Egidi and Narduzzo, 1997; Egidi, 2015; Hoeffler et al., 2006) and the exploration-exploitation literature based on bandit problems (Laureiro-Martínez et al., 2015; Puranam et al., 2015) or rugged landscapes of unknown features, including the NK models (Billinger et al., 2014, 2021; Levine et al., 2019). Going forward, we believe that cross-pollination of the different views continues to be a feasible way forward to address different aspects of individual search.

Finally, our study provides implications for experimental research utilizing real-effort task methodology. We utilized this methodology to get closer to the real-life processes that at the individual, as well as at the organizational level, are at the roots of the exploitation-exploration phenomenon. In real-life situations, exploration-exploitation choice is a matter of repeating a particular procedure, protocol, or routine versus deciding to learn new ways to reach our goals. This is in contrast to explaining the phenomena through abstracted lotteries, for instance, which might not capture the aspects of path-dependent learning on the task similarly as the real-effort task design. In reality, abandoning a mode of behavior (a product, a technology, etc.) to find a new one involves taking on new costs. Moreover, it especially involves psychological costs related to the sunk cost fallacy phenomenon (Thaler, 1980). Switching from an effortless choice to a new, equally effortless one (like switching from a two-arm bandit machine to another) is psychologically very different from switching from a well-established product or procedure to a new one.

5.2. Practical implications

Given that there is a link between individuals' search behavior and organizational explorative and exploitative innovation (Enkel et al., 2017), we believe our results can inform organizational practice. In particular, our results have implications for decision-making and task design. While stabilizing learning in a particular task with high performance can be beneficial in the short term, our results imply that, especially in simple task environments, this may lead to a suboptimal attachment to the previously built task. There are many ways to mitigate the competency traps and lock-in in organizations. For instance, it may be prudent to ensure that new recruits' tasks are challenging enough to prevent them from becoming unduly comfortable and locked into those initial tasks. Alternatively, classic organizational design features such as job rotation, job enrichment, and upskilling programs can help to prevent lock-in by exposing decision-makers to new task environments in which their skills remain applicable. Furthermore, as part of an upskilling program, individuals could be encouraged to accept short-term lower performance when sampling new tasks.

5.3. Limitations and future research opportunities

While our study addressed many important aspects of individual exploration-exploitation choices, it has certain limitations that indicate directions for further inquiry. First, although our experimental setting allowed for spontaneous learning in executing a real-effort task, we did not deeply investigate the dynamics of this process. Our analysis examines learning and related exploration-exploitation patterns but does not address, for example, the interaction between the participant type and learning. To illuminate this interaction, it would be useful to extend this experimental design to include the collection of fine-grained data on the selection process of participants' real-effort tasks during the experiment—for instance, data on decision times when making exploitation-exploration choices—and qualitative exploration of why individuals make certain choices, such as due to boredom, interest, and deliberate profit-seeking. It would also be useful to examine the dynamics of exploration within a specific task and in switching tasks. Such a two-level model could yield interesting results related to organizational complexity in decision-making.

Second, our experimental setting rendered the round-by-round performance level uncertain because it depended on the objective degree of difficulty of each real-effort task and the participants' individual skills. This uncertainty mimics real situations in which performance levels are determined by the nature of the task and by the managers' and employees' skills. In reality, however, the final performance outcome is also influenced by external elements that are often random, and incorporating both of these dimensions would help to clarify the nature of exploration-exploitation decisions.

Third, our participants made their choices in a social vacuum—that is, they received no feedback from other organizational or external agents, and they were not influenced by the decisions of other participants. As demonstrated in the previous literature (e.g., Ali et al., 2022) explorative and exploitative learning processes are driven and moderated by organizational structures and contexts. Thus, our experimental design is a simplification in comparison to real-world organizations.

However, we opted for this simplification because we were interested in understanding the interplay of individual cognitive processes in completing a real-effort task involving decisions about exploiting acquired knowledge and exploring new tasks. In future studies, introducing social interaction and different organizational contingencies should help to improve the external validity of the experimental results and further contextualize our findings.

Data availability

The research data is available open access (the links are provided in the endnotes of the article)

Acknowledgements

We would like to thank The Foundation for Economic Education (Liikesivistysrahasto) for providing funding for our experiment (grant number: 12-6686). We would also like to thank the Editor and the reviewers for their help in improving our manuscript.

APPENDICES

Appendix A1 Robustness checks (alternative measure of exploration)

To assess the sensitivity of our results to different conceptualizations of exploration, we used an alternative measure when performing the analysis in Table 8 (see Table 9). This measure conceptualizes exploration as a more extreme form of switching to a completely unique and previously untried task during the active phase. The binary variable is assigned a value of 1 if, during the active phase, the participant switches to a task never explored before; otherwise, it is assigned a value of 0. The results align with those reported in Table 8, with the exception that the variable *Last payoff experienced* is no longer significant. This suggests that, in relation to the Aspiration-Payoff Gap, the highest experienced payoff rather than recent experience (i.e., payoff in the previous round) is the key reference point.

Table 9
Determinants of exploration (alternative version)

	DP variable: Alternative exploration			
VARIABLES	Model 1	Model 2	Model 3	Model 4
Training (Memory = 1)	−0.145	−1.263***	−2.519***	−2.590***
	(0.163)	(0.000)	(0.000)	(0.000)
Average payoff		−0.873***	−2.995***	−3.029***
		(0.000)	(0.000)	(0.000)
Asp_payoff_gap (negative gap = 1)		1.080***	−0.129	−0.104
		(0.000)	(0.546)	(0.632)
Last payoff experienced		−0.072	0.029	0.015
		(0.201)	(0.722)	(0.859)
Training x Average payoff			2.769***	2.769***
			(0.000)	(0.000)
Training x Asp_payoff_gap			2.011***	2.048***
			(0.000)	(0.000)
Training x Last payoff experienced			−0.068	−0.073
			(0.586)	(0.565)
Feedback (Yes = 1)				−0.053
				(0.684)
Game (Numeric = 1)				0.362**
				(0.004)
Gender (Male = 1)				0.081
				(0.522)
bretbox				0.001
				(0.716)
Economics				0.128
				(0.320)
Age				−0.009
				(0.754)
Num_experiments				−0.011
				(0.199)
Constant	−1.360***	−1.096***	0.080	0.063
	(0.000)	(0.000)	(0.539)	(0.922)
Observations	2400	2400	2400	2360
Number of codeid	240	240	240	236

Pval in parentheses.
***$p < 0.001$, **$p < 0.01$, *$p < 0.05$, + $p < 0.1$.

Fig. 9 reveals that the average change in the probability of exploration when the average payoff increases is equal to −0.28% (p = 0.009) in the simple task and −34.54% (p = 0.000) in the complex task, thereby confirming a stronger effect in the complex task. Finally, we also confirmed the different effect of the aspiration-payoff feedback contingent upon the initial conditions built in the simple or complex tasks: that falling below the aspiration level increases the likelihood of exploration by 27.48% (p = 0.000) in the memory task, while no effect was found in the building task.

Fig. 9. Marginal effects of *Average payoff*, *Last payoff experienced*, and *Aspiration-Payoff gap* on exploration (alternative version).
Note: Results estimated using Model 4 (Table 9) by setting the other covariates at the observed values; 95% confidence interval reported.

Appendix A2. Demographic statistics: Exploration-exploitation patterns

	Age	Number of experiments		
	Mean	Std dev	Mean	Std dev
Strong ExploitFT-participants rowhead	22.17	2.83	8.07	5.69
Retreaters rowhead	21.4	2.6	10.7	9.2
Strong ExploitNT-participants rowhead	21.48	1.91	10.7	7.71

	Economics	Law	Engineering	Sociology	Faculty Literature	Computer Science	Psychology	Mathematics	Other	Gender Male
Strong ExploitFT-participants rowhead	41.38%	31.3%	3.45%	3.45%	13.79%				6.90%	41%
Retreaters rowhead	39.13%	13.04%	4.35%	17.39%	4.35%	4.35%	8.70%	4.35%	4.35%	48%
Strong ExploitNT-participants rowhead	59.26%	16.7%	7.41%	5.56%		1.85%	1.85%	1.85%	5.56%	48%

References

Ali, M., Shujahat, M., Ali, Z., Kianto, A., Wang, M., Bontis, N., 2022. The neglected role of knowledge assets interplay in the pursuit of organisational ambidexterity. Technovation 114, 102452.
Benndorf, V., Rau, H.A., Sölch, C., 2019. Minimizing learning in repeated real-effort tasks. Journal of Behavioral and Experimental Finance 22, 239–248.
Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., Fiedler, K., 2001. The effects of routine strength on information acquisition and adaptation in recurrent decision making. Organ. Behav. Hum. Decis. Process. 84, 23–53.
Bilalić, M., McLeod, P., Gobet, F., 2008. Why good thoughts block better ones: the mechanism of the pernicious Einstellung (set) effect. Cognition 108 (3), 652–661.
Bilalić, M., McLeod, P., Gobet, F., 2010. The mechanism of the Einstellung (set) effect: a pervasive source of cognitive bias. Curr. Dir. Psychol. Sci. 19 (2), 111–115.
Billinger, S., Stieglitz, N., Schumacher, T.R., 2014. Search on rugged landscapes: an experimental study. Organ. Sci. 25 (1), 93–108.
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. Strat. Manag. J. 42 (2), 361–385.
Bromiley, P., 1991. Testing a causal model of corporate risk taking and performance. Acad. Manag. J. 34 (1), 37–59.
Brüggen, A., Strobel, M., 2007. Real effort versus chosen effort in experiments. Econ. Lett. 96 (2), 232–236.
Brusoni, S., Laureiro-Martínez, D., Canessa, N., Zollo, M., 2020. Exploring exploration: the role of affective states as forces that hinder change. Ind. Corp. Change 29 (1), 207–223.
Campbell, D.J., 1988. Task complexity: a review and analysis. Acad. Manag. Rev. 13 (1), 40–52.
Chadwick, I.C., Raver, J.L., 2015. Motivating organizations to learn: goal orientation and its influence on organizational learning. J. Manag. 41 (3), 957–986.
Chen, D.L., Schonger, M., Wickens, C., 2016. oTree - an open-source platform for laboratory, online and field experiments. Journal of Behavioral and Experimental Finance 9, 88–97.
Crosetto, P., Filippin, A., 2013. The “bomb” risk elicitation task. J. Risk Uncertain. 47 (1), 31–65.
Cyert, R.M., March, J.G., 1963. A Behavioral Theory of the Firm. Prentice-Hall, Englewood Cliffs, NJ).
Denrell, J., March, J.G., 2001. Adaptation as information restriction: the hot stove effect. Organ. Sci. 12 (5), 523–538.
Denrell, J., Le Mens, G., 2020. Revisiting the competency trap. Ind. Corp. Change 29 (1), 183–205.
Diamond, P.A., Stiglitz, J.E., 1974. Increases in risk and in risk aversion. J. Econ. Theor. 8 (3), 337–360.
Diehl, E., Sterman, J.D., 1995. Effects of feedback complexity on dynamic decision making. Organ. Behav. Hum. Decis. Process. 62 (2), 198–215.
Egidi, M., 2015. Organizational decisions in the lab. In: Minds, Models and Milieux. Palgrave Macmillan, London, pp. 186–206.
Egidi, M., Narduzzo, A., 1997. The emergence of path-dependent behaviors in cooperative contexts. Int. J. Ind. Organ. 15 (6), 677–709.
Enkel, E., Heil, S., Hengstler, M., Wirth, H., 2017. Exploratory and exploitative innovation: to what extent do the dimensions of individual level absorptive capacity contribute? Technovation 60, 29–38.
Ericsson, K.A., Lehmann, A.C., 1996. Expert and exceptional performance: evidence of maximal adaptation to task constraints. Annu. Rev. Psychol. 47 (1), 273–305.
Faul, F., Erdfelder, E., Buchner, A., Lang, A.-G., 2009. Statistical power analyses using gpower 3.1: tests for correlation and regression analyses. Behavioral Research Methods 41 (4), 1149–1160.
Greve, H.R., 2003. Organizational Learning from Performance Feedback: A Behavioral Perspective on Innovation and Change. Cambridge University Press, Cambridge, UK).
Greve, H.R., 1998. Performance, aspirations, and risky organizational change. Adm. Sci. Q. 43, 58–86.

Greve, H.R., 2007. Exploration and exploitation in product innovation. Ind. Corp. Change 16 (5), 945–975.
Hoeffler, S., Ariely, D., West, P., 2006. Path dependent preferences: the role of early experience and biased search in preference development. Organ. Behav. Hum. Decis. Process. 101 (2), 215–229.
Hoetker, G., 2007. The use of logit and probit models in strategic management research: critical issues. Strat. Manag. J. 28 (4), 331–343.
Holzmeister, Felix, Pfurtscheller, Armin, 2016. oTree: the "bomb" risk elicitation task. Journal of Behavioral and Experimental Finance 10 (C), 105–108. Elsevier.
Junni, P., Sarala, R.M., Taras, V.A.S., Tarba, S.Y., 2013. Organizational ambidexterity and performance: a meta-analysis. Acad. Manag. Perspect. 27 (4), 299–312.
Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. Econometrica 47 (2), 263–291.
Kahneman, D., 2003. A perspective on judgement and choice. Am. Psychol. 58 (9), 697–720.
Kantola, J., 2016. Organizational Resource Management: Theories, Methodologies and Applications. CRC Press, New York.
Ko, Y.J., O'Neill, H., Xie, X., 2021. Strategic intent as a contingency of the relationship between external knowledge and firm innovation. Technovation 104, 102260.
Lant TK Montgomery, D.B., 1987. Learning from strategic success and failure. J. Bus. Res. 15, 503–517.
Laureiro-Martínez, D., Brusoni, S., Zollo, M., 2010. The neuroscientific foundations of the exploration− exploitation dilemma. Journal of Neuroscience, Psychology, and Economics 3 (2), 95–115.
Laureiro-Martinez, D., 2014. Cognitive control capabilities, routinization propensity, and decision-making performance. Organ. Sci. 25 (4), 1111–1133.
Laureiro-Martínez, D., Brusoni, S., Canessa, N., Zollo, M., 2015. Understanding the exploration–exploitation dilemma: an fMRI study of attention control and decision-making performance. Strat. Manag. J. 36 (3), 319–338.
Laureiro-Martínez, D., Brusoni, S., 2018. Cognitive flexibility and adaptive decision-making: evidence from a laboratory study of expert decision makers. Strat. Manag. J. 39 (4), 1031–1058.
Larson Jr., J.R., 1984. The performance feedback process: a preliminary model. Organ. Behav. Hum. Perform. 33 (1), 42–76.
Lave, C.A., March, J.G., 1993. An Introduction to Models in the Social Sciences, second ed. University Press of America, Inc., London.
Lavie, D., Stettner, U., Tushman, M.L., 2010. Exploration and exploitation within and across organizations. Acad. Manag. Ann. 4 (1), 109–155.
Lee, S., Meyer-Doyle, P., 2017. How performance incentives shape individual exploration and exploitation: evidence from microdata. Organ. Sci. 28 (1), 19–38.
Levinthal, D.A., March, J.G., 1993. The myopia of learning. Strat. Manag. J. 14 (Special Issue), 95–112.
Levine, S.S., Reypens, C., Riedl, C., 2019. What directs search? Experimental evidence on some triggers of exploration. Available online: https://ssrn.com/abstract=2816541.
Levitt, B., March, J.G., 1988. Organizational learning. Annu. Rev. Sociol. 14, 319–340.
Luchins, A.S., 1942. Mechanization in problem solving: the effect of Einstellung. Psychol. Monogr. 54 (6), i.
Luchins, A.S., Luchins, E.H., 1950. New experimental attempts at preventing mechanization in problem solving. J. Gen. Psychol. 42 (2), 279–297.
MacLeod, W.B., Pingle, M., 2005. Aspiration uncertainty: its impact on decision performance and process. J. Econ. Behav. Organ. 56 (4), 617–629.
March, J.G., Shapira, Z., 1987. Managerial perspectives on risk and risk taking. Manag. Sci. 33 (11), 1404–1418.
March, J.G., 1988. Variable risk preferences and adaptive aspirations. J. Econ. Behav. Organ. 9 (1), 5–24.
March, J.G., 1991. Exploration and exploitation in organizational learning. Organ. Sci. 2 (1), 71–87.
Mehlhorn, K., Newell, B.R., Todd, P.M., Lee, M.D., Morgan, K., Braithwaite, V.A., et al., 2015. Unpacking the exploration–exploitation tradeoff: a synthesis of human and animal literatures. Decision 2 (3), 191.
Miller, K.D., Chen, W.R., 2004. Variable organizational risk preferences: tests of the March-Shapira model. Acad. Manag. J. 47 (1), 105–115.
Mittone, L., Papi, M., 2017. Does inducing choice procedures make individuals better off? An experimental study. Theor. Decis. 83 (1), 37–59.
Mom, T.J., Van Den Bosch, F.A., Volberda, H.W., 2007. Investigating managers' exploration and exploitation activities: the influence of top-down, bottom-up, and horizontal knowledge inflows. J. Manag. Stud. 44 (6), 910–931.
Mom, T.J., Van Den Bosch, F.A., Volberda, H.W., 2009. Understanding variation in managers' ambidexterity: investigating direct and interaction effects of formal structural and personal coordination mechanisms. Organ. Sci. 20 (4), 812–828.
Narduzzo, A., Warglien, M., 2008. Conducting experimental research on organizational routines. In: Becker, M.C. (Ed.), Handbook of Organizational Routines. Edward Elgar Publishing, pp. 301–324.
Oehler, P.J., Stumpf-Wollersheim, J., Welpe, I.M., 2019. Never change a winning routine? How performance feedback affects routine change. Ind. Corp. Change 28 (5), 1101–1124.
Oehler, P.J., Stumpf-Wollersheim, J., Welpe, I.M., Obstfeld, D., 2021. Beyond the truce: how conflict affects teams' decisions whether to enact routines or creative projects. Ind. Corp. Change 30 (3), 799–822.
Posen, H.E., Levinthal, D.A., 2012. Chasing a moving target: exploitation and exploration in dynamic environments. Manag. Sci. 58 (3), 587–601.
Posen, H.E., Keil, T., Kim, S., Meissner, F.D., 2018. Renewing research on problemistic search—a review and research agenda. Acad. Manag. Ann. 12 (1), 208–251.
Puranam, P., Stieglitz, N., Osman, M., Pillutla, M.M., 2015. Modelling bounded rationality in organizations: progress and prospects. Acad. Manag. Ann. 9 (1), 337–392.
Rahmandad, H., DexploitNTell, J., Prelec, D., 2021. What makes dynamic strategic problems difficult? Evidence from an experimental study. Strat. Manag. J. https://doi.org/10.1002/smj.3254.
Ritala, P., Heiman, B., Hurmelinna-Laukkanen, P., 2016. The need for speed—unfamiliar problems, capability rigidity, and ad hoc processes in organizations. Ind. Corp. Change 25 (5), 757–777.
Roth, A.E., Malouf, M.W., 1979. Game-theoretic models and the role of information in bargaining. Psychol. Rev. 86 (6), 574.
Selten, R., 1998. Aspiration adaptation theory. J. Math. Psychol. 42 (2–3), 191–214.
Simon, H.A., 1955. A behavioral model of rational choice. Q. J. Econ. 69 (1), 99–118.
Simon, H.A., 1996. The Sciences of the Artificial, third ed. MIT Press, Cambridge, MA.
Smith, V.L., 1962. An experimental study of competitive market behavior. J. Polit. Econ. 70 (2), 111–137.
Smith, V.L., 1976. Experimental economics: induced value theory, *the American economic review*. Papers and Proceedings of the Eighty-eighth Annual Meeting of the American Economic Association 66 (2), 274–279.
Smith, W.K., Tushman, M.L., 2005. Managing strategic contradictions: a top management model for managing innovation streams. Organ. Sci. 16 (5), 522–536.
Sitkin, S.B., Pablo, A.L., 1992. Reconceptualizing the determinants of risk behavior. Acad. Manag. Rev. 17, 9–38.
Sydow, J., Schreyögg, G., Koch, J., 2009. Organizational path dependence: opening the black box. Acad. Manag. Rev. 34 (4), 689–709.
Thaler, R., 1980. Toward a positive theory of consumer choice. J. Econ. Behav. Organ. 1 (1), 39–60.
Van Rijnsoever, F.J., Meeus, M.T., Donders, A.R.T., 2012. The effects of economic status and recent experience on innovative behavior under environmental variability: an experimental approach. Res. Pol. 41 (5), 833–847.
Weick, K.E., 1993. The collapse of sensemaking in organizations: the Mann Gulch disaster. Adm. Sci. Q. 38, 628–652.
Weiss, H.M., Ilgen, D.R., 1985. Routinized behavior in organizations. J. Behav. Econ. 14 (1), 57–67.
Wiersema, M.F., Bowen, H.P., 2009. The use of limited dependent variable techniques in strategy research: issues and methods. Strat. Manag. J. 30 (6), 679–692.
Wiseman, R.M., Gomez-Mejia, L.R., 1998. A behavioral agency model of managerial risk taking. Acad. Manag. Rev. 23 (1), 133–153.
Zelner, B.A., 2009. Using simulation to interpret results from logit, probit, and other nonlinear models. Strat. Manag. J. 30 (12), 1335–1348.