Discovering Alternative Strategies: Experimental Evidence on the Impact of Frameworks *

Hyunjin Kim and Nety Wu

INSEAD

June 2025

Abstract

How can decision-makers discover strategic alternatives? Using randomized experiments across 820 MBA students and executives, we investigate how frameworks influence the generation of strategic alternatives. Our findings suggest that frameworks significantly expand the set of options visible to decision-makers and the ultimate choices they make. Participants who were randomly provided with a framework on strategic options were more likely to generate and select mutually exclusive strategic alternatives rather than operational improvements, and to explore options beyond the current strategy. Interview and survey data suggest that frameworks appear to impact how participants generate alternatives by broadening how they formulate the underlying problem. Treatment effects are muted when multiple frameworks are provided, suggesting that focused attention may be essential for frameworks to effectively stimulate the generation of strategic alternatives. However, we find suggestive evidence that this cognitive bottleneck may be less binding in AI-assisted strategy making, as LLMs can readily process and integrate multiple frameworks to uncover more strategic options.

Keywords— Decision Making and Theory of the Firm, Managerial and Organizational Cognition, Laboratory research

*We are grateful to Angus Mcquillan-Gordon, Ronak Mogra, Si Yi Ng, and Zhe Chong Tang for their excellent research assistance. This project would not have been possible without the support of Vicky Lee, Phanish Puranam, Daniel Simonovich, Ithai Stern, Thomaz Teodorovicz, and Nghi Truong. We also sincerely thank participants at the INSEAD seminar and the 2024 Organization Science Winter Conference for their valuable comments. The experiments were approved by the INSEAD IRB and pre-registered on the AEA RCT registry. Authors contributed equally and are listed alphabetically.

1 Introduction

Generating multiple viable strategic alternatives is critical to effective strategic decision-making. It allows organizations to avoid prematurely committing to suboptimal courses of action, especially in complex and uncertain business environments (Gans et al., 2019). Recent empirical evidence shows that experimenting with alternatives leads to more pivots and terminations (Camuffo et al., 2020) and is positively associated with product changes and better performance conditional on survival (Koning et al., 2022).

However, less is known about how managers can overcome the challenges involved in discovering multiple viable strategic alternatives. Managers hold entrenched mental models (Levinthal, 2011; Tripsas & Gavetti, 2000) that constrain problem formulation and information processing (Baer et al., 2013; Csaszar, 2018; Simon, 1947). These cognitive constraints lead managers to overweigh certain aspects of a strategic problem while neglecting others, hampering their ability to generate alternatives (Lane et al., 2024). While extensive literature exists on ideation and creativity (for example, Anderson et al., 2014; Berg et al., 2023; Girotra et al., 2010), there has been limited insight on the unique challenges of generating strategic options—a process characterized by a high level of uncertainty, substantial resource commitments, and long-term implications. Due to its complexity and high stakes, this process often discourages the exploration of alternatives, underscoring the need to understand how managers can overcome cognitive barriers to generate multiple strategic options.

One possible approach to addressing these challenges is the use of strategic frameworks. Widely adopted among managers (Csaszar, Hinrichs, & Heshmati, 2024) and central to management education (Acemoglu et al., 2022; Heshmati & Csaszar, 2024; Yang et al., 2020), frameworks aim to structure complex strategic problems and stimulate strategic thinking. As Simon (1996, p. 132) noted, "solving a problem simply means representing it so as to make the solution transparent." Through this structured representation, frameworks may enhance managers' ability to generate strategic alternatives.

However, *whether* and *how* frameworks aid alternative generation remains theoretically ambiguous and empirically untested. If managers' mental models are so deeply ingrained, the question remains whether frameworks can induce even a temporary cognitive shift to enable the generation of alternatives. Moreover, while frameworks may stimulate idea generation (Csaszar, 2018; Kirsh, 2010), their structured nature—designed to "make the solution transparent"—might also narrow attention (Scaife & Rogers, 1996; Wilson, 2002), which might limit the exploration of alternative strategies and reduce the number and diversity of options considered. The specific mechanisms through which frameworks either facilitate or hinder the generation of multiple strategic alternatives remain unclear.

In this paper, we investigate how providing a framework impacts the set of strategic alternatives considered and ultimately selected using a series of randomized controlled trials across 820 MBA students and executives enrolled in a leading business school. Participants were tasked with developing as many strategic options as possible for a venture-backed online review platform that observed tremendous user growth without reaching profitability eight years after its inception. Those randomly assigned to treatment received a framework designed to help them generate strategic options, while those in the control group received only general instructions. We then collected and analyzed all the alternatives the participants generated using human coders and transformer-based text analysis algorithms.

Our findings suggest that frameworks can expand the set of strategic alternatives visible to decision-makers and change their ultimate choice. While both groups generated a similar number of options, participants with the framework generated 20% more strategic alternatives and 21% fewer options that merely continued the existing strategy. They were more likely to consider options like entering underserved markets or positioning for strategic acquisition, rather than incremental adjustments that the control group focused on, such as repricing advertising and restructuring sales incentives. Treated participants were also

17 percentage points more likely to generate a set of mutually exclusive options, suggesting that they generated genuine alternatives. Most importantly, these expanded alternatives translated into meaningfully different decisions: participants with the framework were 16 percentage points more likely to select strategic rather than operational solutions when choosing their best option, and 10 percentage points less likely to default to maintaining current strategies.

Surveys and interviews suggest that a key mechanism through which the framework helped generate strategic alternatives was by impacting how participants represented the problem they were tasked with. Participants who received the framework found it easier to formulate and represent the problem at hand. They also tended to adopt a more comprehensive perspective that included multiple viewpoints, which led to a richer understanding of the issues. In contrast, those in the control group tended to focus on straightforward, easily identifiable business issues.

One might expect that if one framework enhances alternative generation, multiple frameworks would provide even greater benefits by offering diverse perspectives through which to view the problem. To examine this possibility, we conducted a follow-up experiment in which treated participants received a set of classic strategic frameworks, including our focal one. Surprisingly, instead of broadening the range of strategic alternatives identified through varied representations, providing a larger set of frameworks muted the treatment effect. This suggests that for frameworks to effectively stimulate the generation of strategic alternatives, focused attention may be crucial. Given the difficulty of shifting entrenched mental models, presenting many frameworks may overwhelm rather than redirect thinking—ultimately diluting, rather than enhancing, their impact.

However, this cognitive bottleneck may be less binding in AI-assisted strategy making. To explore, we simulated our experiments using Large Language Models (LLMs). We generated a comparable sample of virtual participants and randomly assigned them to different treatments to consider the same problem as our human participants, replicating both our main experiment and the multi-framework follow-up. We find that when prompted with the framework, LLMs produced significantly more strategic alternatives, consistent with our main findings. More strikingly, LLMs demonstrated the ability to process and integrate multiple frameworks to simultaneously uncover additional strategic options, suggesting that this cognitive bottleneck may diminish as AI becomes increasingly integrated into strategic decision-making (Boussioux et al., 2024; Csaszar, Ketkar, & Kim, 2024; Dell'Acqua et al., 2023; Doshi et al., 2025).

This paper sheds light on how frameworks influence the generation of multiple strategic alternatives—a critical yet understudied component of strategy formation. Theoretically, we identify the mechanisms through which strategic frameworks may shape alternative generation, offering insight into both their enabling and constraining effects. While we do not focus on experimentation per se, our findings speak to how frameworks may aid in the experimentation process, as the ability to generate diverse alternatives is a prerequisite for effective learning from experimentation. Empirically, we show that prompting decision-makers with a well-chosen strategic framework can shift the set of alternatives they consider and nudge them toward more strategic choices, but only when it can focus their attention. Practically, these insights inform the design and deployment of strategic frameworks in both organizational settings and management education, as well as how managers can overcome cognitive blinders in generating alternative strategies.

2 Theoretical Motivation

2.1 The importance of multiple alternatives

Strategic decision-making requires decision-makers to navigate the inherent complexity and uncertainty embedded in strategic problems. These decisions are characterized by their high-stakes nature, long-term implications, and the involvement of multiple stakeholders (Csaszar, 2018; Leiblein et al., 2018). The interdependence of various organizational elements and external environmental factors complicates the decision-making process, challenging managers to fully understand and consider all relevant aspects of a strategic issue. Furthermore, these decisions involve substantial uncertainty, as the values of alternatives are not known to the decision-maker ex-ante. The interplay between complexity and uncertainty means that the full set of alternatives and their potential values are often not immediately clear (Gans et al., 2019).

Scholars have increasingly highlighted the critical role of discovering and evaluating multiple viable strategic alternatives, rather than searching for a single solution. This shift acknowledges that in complex and dynamic business environments, there may not be a single "best" strategy, but rather multiple potentially effective approaches (Gans et al., 2019). By generating multiple alternatives, decision-makers can avoid committing prematurely to a strategy without gathering available information about possible alternatives (Gans et al., 2019). Recent empirical evidence suggests that experimenting with alternatives results in more pivots and terminations (Camuffo et al., 2020) and is positively correlated with product changes and better performance conditional on survival (Koning et al., 2022).

2.2 Strategic alternative generation and mental representations

The complexity and uncertainty inherent in strategic problems make the generation of multiple viable alternatives a significant challenge. Central to this challenge are managers' mental representations, which play a key role in how strategic problems are framed and potential solutions conceived.

Prior research reveals several key insights about the role of mental representations in the process of generating strategic options. First, mental representation can serve as a powerful cognitive tool in facilitating the process of generating options. Mental representations allow individuals to simplify complex real-world issues into manageable "small world" scenarios (Craik, 1943; Cyert & March, 1963; Levinthal, 2011). Research has shown that leveraging these mental representations can facilitate "long jump" searches, enabling managers to consider options beyond their immediate, familiar terrain (Gavetti & Levinthal, 2000). They also facilitate navigating novel situations through analogy (Gavetti et al., 2005). Empirical evidence suggests that mental representations provide substantial help to both the search and the evaluation stages (Csaszar & Laureiro-Mart'inez, 2018; Heshmati & Csaszar, 2024), eventually leading to better decisions (Gary & Wood, 2011). Scholars have further theorized that mental representations can drive strategic renewal (Helfat & Peteraf, 2015), shape market entry decisions (Eggers & Kaplan, 2009), and enable coupled learning despite initial misrepresentations (Puranam & Swamy, 2016), leading to performance heterogeneity among firms (Csaszar & Levinthal, 2016; Felin & Zenger, 2017).

While mental representations can facilitate the generation of strategic options that fit within a given representation, they have also been theorized to constrain the quality and diversity of the alternatives considered by managers. Mental representations shape how managers frame strategic problems and process information (Baer et al., 2013; Csaszar, 2018; Simon, 1947), which can lead to an overemphasis on certain cues while inadequately accounting for others (Hanna et al., 2014). This selective attention narrows the scope of information considered in decision-making. Recent research has further shown that the way managers attend to information significantly influences their choices (Lane et al., 2024). Consequently, managers may become fixated on certain perspectives, inadvertently overlooking other viable strategic options. This might hinder their ability to generate and consider multiple alternatives, potentially limiting the effectiveness of their strategic decision-making.

Furthermore, mental representations themselves are deeply ingrained and difficult to change (Tripsas & Gavetti, 2000). Prior research suggests that managerial beliefs exhibit strong resistance to change (Argyris & Schön, 1978; Levitt & March, 1988). Managers often underreact to new data due to entrenched beliefs (Tripsas & Gavetti, 2000) or adopt models that overfit the past (Schwartzstein & Sunderam, 2021). This resistance is further complicated by psychological discomfort from altering deeply held beliefs, which can manifest as cognitive dissonance in individuals when confronted with conflicting information (Festinger, 1957). Furthermore, people tend to seek, interpret, and recall information in a way that confirms their preexisting beliefs or hypotheses, making it difficult to consider information that contradicts their current mental models (Nickerson, 1998; Rabin & Schrag, 1999). This tendency to confirm pre-existing mental representations further reinforces the difficulty of altering them, leading to too few changes in the long run (Augenblick & Rabin, 2021). This entrenchment of mental representations raises a challenge for strategic decision-making, as these representations may constrain and shape the alternatives decision-makers consider and ultimately choose.

Together, these insights highlight that while mental representations can facilitate the search for options, the constraints and persistence of entrenched mental models create substantial barriers to effectively generating multiple strategic alternatives. Yet despite widespread recognition of these challenges, less is known about how to overcome these barriers effectively. Moreover, while there is extensive research on ideation and creativity more broadly (for example, Anderson et al., 2014; Berg et al., 2023; Girotra et al., 2010), it has generally not focused on the unique challenges of generating options that are strategic in nature. The constrained and high-stakes environment of strategic decision-making may differ significantly from general creative processes, as decision-makers must discover not only novel ideas but also ones that provide viable strategic paths forward.

2.3 The role of frameworks on strategic alternative generation

One possible solution to overcoming these cognitive challenges lies in strategic frameworks, which are widely used in practice. Leading consulting firms such as McKinsey (McKinsey, 2008) and Boston Consulting Group (Reeves et al., 2014) have developed and extensively deployed these frameworks in their work with clients. Frameworks also form the backbone of management education in business schools, where core MBA strategy courses center on teaching frameworks as a method to enhance strategic thinking. This ubiquity underscores the importance of understanding their impact on strategic decision-making.

Initial evidence suggests that management education broadly facilitates deeper mental representations (Heshmati & Csaszar, 2024) with measurable effects on firm performance (Yang et al., 2020) and key operational decisions like cost-cutting (Acemoglu et al., 2022). However, these studies examine the broader educational impact rather than the specific role of external representations in decision-making. Similarly, research on entrepreneurial experimentation emphasizes the value of scientific approaches in strategy formulation (Agrawal et al., 2021; Camuffo et al., 2020; Gans et al., 2019), but has not explored how managers generate multiple alternatives—a prerequisite for effective experimentation.

Whether frameworks expand the strategic options considered by managers remains theoretically ambiguous. If managers' mental models are deeply entrenched, frameworks may fail to induce even temporary cognitive shifts. Moreover, frameworks' structured nature might actually narrow attention to a predefined set of attributes and limit the exploration of diverse alternatives (Scaife & Rogers, 1996; Wilson, 2002). If "solving a problem simply means representing it so as to make the solution transparent" (Simon, 1996, p. 132), frameworks might inadvertently funnel thinking toward singular strategies.

Conversely, frameworks could stimulate strategic option generation by enhancing cognitive functions. Csaszar, Hinrichs, and Heshmati (2024) theorize that visual frameworks facilitate critical cognitive processes in strategic decision-making, potentially yielding more diverse alternatives. Frameworks may reshape how decision-makers conceptualize problems, expanding the "opportunity space" by reorganizing and reinterpreting information (Stenning & Oberlander, 1995; Zhang & Norman, 1994). For instance, Porter's Five Forces framework dissects industry analysis into distinct components, enabling systematic evaluation of competitive dynamics that might remain invisible when relying solely on internalized thought processes.

Different frameworks may also activate distinct cognitive processes (Kleinmuntz & Schkade, 1993; Zhang, 1997), altering how strategists approach problems. Metaphoric diagrams can inspire decision-makers to view familiar challenges through alternative industry lenses, prompting analogies that remap problems onto different conceptual domains (Gavetti et al., 2005). By actively guiding cognition (Csaszar, Hinrichs, & Heshmati, 2024; Giere, 2004; Zhang & Norman, 1994), frameworks may fundamentally transform how managers understand strategic challenges, highlighting previously overlooked attributes and relationships.

In the subsequent sections, we empirically investigate this question of whether and how frameworks influence the generation of strategic alternatives and through what mechanisms they operate in practice.

3 Experimental Design

3.1 Sample, experimental task, and randomization

We ran our main experiment across 340 MBA students enrolled in a core strategy course at a leading business school during 2022-2023. Table 1 presents the summary statistics of the sample: participants were on average 30 years of age, 38% female, and 30% employed in management consulting prior to enrolling in the MBA program.[1]

The experimental task focused on developing strategic options for a company based on a pre-distributed case study. This case described Rated (a pseudonym to maintain confidentiality), a venture-backed online reviews platform that was reassessing its strategic direction

[1] We also collected a binary indicator of their education level (equaling one if the participant holds a master's or above degree), the number of years since they obtained their last degree, the number of years of work experience, and GMAT scores. For students without a GMAT score but with a GRE score, we transformed the GRE scores into a predicted GMAT score based on a formula developed by the Educational Testing Service (ETS): https://www.ets.org/content/dam/ets-org/pdfs/gre/gre-bschool-comparison-tool-faq.pdf.

after eight years of substantial growth without reaching profitability (Kim, 2021). The case presented a snapshot of the company and its performance at the time, along with viewpoints from key executives who pointed out various issues, such as outdated pricing of the advertising feature and high turnover in their salesforce. These elements, while significant, were primarily concerned with fixing operational issues in their existing processes to improve efficiency,[2] rather than identifying solutions that could fundamentally alter the strategic direction of the company. Furthermore, the case captured the complexity of real organizational challenges, where surface-level problems often mask deeper underlying issues. The case ended with a question: what strategic options should the CEO consider to improve Rated's survival and competitiveness in an evolving industry landscape. Participants were given the case several days in advance to understand the issues at hand and analyze financial and market share figures.

Participants were randomly assigned to two groups, control and treatment, stratified on their section, gender, and a binary indicator of whether they were employed in consulting prior to enrolling in the MBA program. Each group was allocated to work in separate rooms to prevent potential contamination. The experiment took place within the classroom, where participants were asked to identify all strategic alternatives that they saw as being available for the CEO of Rated. They worked individually and were provided with unique links to an online survey interface to submit their responses.

The experimental treatment changed the instructions that each group received in the online interface. All participants received general instructions for the task, along with a visual graphic (Figure 1). The treatment group additionally received a visual representation of a strategic framework designed to guide the generation of strategic alternatives (Figure 2) (see Kim et al., 2024 for further background on the framework). In contrast, the control group received no additional instructions beyond the general task instructions. In follow-

[2] Such changes may also be categorized as tactical rather than strategic, focusing on short-term gains without addressing the need for long-term sustainability and growth.

up experiments detailed in Section 4.3, we replaced the single framework with a set of frameworks (Figures 3a and 3b). After identifying all strategic alternatives available to the company, participants were asked to indicate which one they assessed to be the best from their set of alternatives.

To further explore mechanisms, additional questions were provided on the online interface for a subsample of 127 participants. Participants were asked to explain their thought processes and, on a 5-point scale, rate their perceived task difficulty and confidence in their responses. The treatment group was also asked how helpful they found the framework to be and why. Additionally, we conducted short interviews with 23 participants who volunteered to share their experiences in completing the task. Appendix C shows further details on the interview questions.

The experiments were pre-registered at the AEA RCT Registry and conducted in compliance with IRB approval from the authors' institution.[3] Table A.1 in the Appendix shows the timeline of experimental interventions. Table 1 shows that control and treatment groups were balanced on all baseline attributes.

3.2 Outcome measures

In total, participants generated 2,253 options in the main experiment. To construct outcome measures, we employed both hand-coding by independent human coders and natural language processing algorithms.

3.2.1 Human-coded measures

Each alternative was coded by two independent coders blind to the experimental groups using a predefined rubric (see Appendix B for further details). In instances of discrepancies in coding, the disputed alternatives were assigned to a third independent coder for resolution.

[3]We describe all differences between the paper and pre-registration in detail in Appendix F.

We constructed three count-related measures at the participant level: (1) *the total number of options*, (2) *the number of strategic options*, and (3) the number of alternatives related to continuing the current strategy (henceforth called "*the number of continue options*").

Variables (2) and (3) were based on the manual coding of the alternatives generated by the participant. For (2), each alternative was coded as strategic or not (Cohen's kappa: 0.604, $p < 0.001$). Strategic alternatives were defined as high-level plans that are difficult to reverse and require substantial resource commitment, following prior literature (Csaszar, 2018; Leiblein et al., 2018; Van den Steen, 2017). According to this definition, launching a new product or expanding into a new market would be categorized as strategic, while increasing sales team salaries would not.

For variable (3), each alternative was coded as "continuing" or "not continuing" (Cohen's kappa: 0.636, $p < 0.001$) based on whether it aligned with the current strategy. This coding generally represented options that did not fundamentally alter the firm's value proposition, target market, or key activities, keeping the broader strategy largely unchanged.

In addition to these count-related variables, we constructed two binary indicators on the entire set of alternatives that a given participant generated: (4) whether the set was mutually exclusive, and (5) whether the set included an alternative related to exit (i.e., selling or closing the company). *Mutually exclusive* equals one if choosing one option would prevent the firm from choosing any of the others, assuming some resource constraints (Cohen's kappa: 0.505, $p < 0.001$). *Exit* equals one if the set of options contains at least one choice that indicates exit such as selling the business (Cohen's kappa: 0.884, $p < 0.001$).

Similarly, we constructed three binary indicators of whether the best alternative chosen by the participant is (1) strategic; (2) continues the current strategy; and (3) involves exit.

Appendix B provides more details and examples of the entire coding process and rubrics for each variable.

3.2.2 Text analysis-based measures

We also leveraged text analysis to analyze differences in the content of the alternatives generated by participants, using topic modeling based on transformer models. We preprocessed the text by removing special characters and lemmatizing words, then applied topic modeling to identify distinct themes. To do this, we used BERTopic, a topic modeling framework that leverages transformer models, which can handle large amounts of text and identify nuanced topics better than traditional topic modeling techniques (Grootendorst, 2022).[4] This analysis provided us with both the topic representations and the texts associated with each to evaluate how the content of alternatives differed between control and treatment groups (see Appendix E for further details).

In addition, we constructed a measure of how distinctive each alternative was relative to others. We used the sentence transformer model "all-mpnet-base-v2,"[5] one of the best-performing pre-trained models available based on the quality of embedded sentences, search queries, and paragraphs, to generate embedding—vector representations of each alternative generated by participants. Using the embedding vectors, we constructed the measure, average similarity, as the cosine similarity between the embedding $e_{i,p}$ of alternative i generated by participant p and the average embedding of alternatives generated by all other participants assigned to the same experimental group excluding participant p, averaged across all N alternatives generated by participant p:

$$S_{\text{avg},p} = \frac{1}{N}\sum_{i=1}^{N}\cos(e_{i,p}, e_{\text{avg},-p})$$

This measure assesses the distinctiveness of the focal alternative in relation to the broader set of alternatives generated by participants from the same experimental group. As a ro-

[4] This model maps sentences to a 768-dimensional vector space that captures the semantic information, and it is usually used for clustering or sentence similarity tasks (Hugging Face, 2022).

[5] https://www.sbert.net/docs/pretrained_models.html.

bustness check, we also computed the cosine similarity between each alternative's embedding and the average embedding across all alternatives, which is strongly correlated (correlation: 0.986).

3.3 Empirical specification

To study the impact of providing frameworks on generating strategic alternatives, we estimate the following model:

$$y_{i,s} = \beta_0 + \beta_1 \cdot \text{Treatment}_{i,s} + \boldsymbol{X}_i + \gamma_s + \epsilon_{i,s} \tag{1}$$

where $y_{i,s}$ represents any of our dependent variables for participant i in randomization strata s. β_1 measures the treatment effect of the framework. $Treatment_{i,s}$ is an indicator that takes the value 1 for participants assigned to treatment and 0 otherwise. $\boldsymbol{X}_i$ is the vector of control variables including individual demographic attributes. γ_s controls for strata fixed effects. $\epsilon_{i,s}$ is an idiosyncratic error term.

We define each randomization stratum as a combination of section, gender, and whether their last job was in consulting. This enables us to improve the precision of results by accounting for any section-specific factors or variation in treatment effects by gender or prior experience (Cox & Reid, 2000). Models are estimated with robust standard errors at the participant level, which is the level of randomization (Abadie et al., 2023).

4 Results

4.1 Main results

We find that the framework leads participants to generate more strategic alternatives, which are also more likely to be mutually exclusive and less likely to continue with the current strategy. These effects carry over to their ultimate choices. However, the framework also leads participants to generate options that are more similar to those of others who also use the framework.

Column 1 in Table 2 shows that having access to the framework did not significantly change the total number of options that participants generated. This is consistent with graphical evidence shown in Panel (a) in Figure 4, which plots the kernel density of the total number of options for both control and treatment groups (see histograms in Panel (a) in Appendix Figure A.1). The two distributions are not significantly different from each other (Kolmogorov–Smirnov test $p = 0.994$).

However, having access to the framework increased the number of strategic options that participants generated. Column 2 in Table 2 shows that on average, participants who were exposed to the framework generated 0.9 more strategic options ($p = 0.006$). This translates into a 20% increase in the number of strategic options generated. We observe that this treatment effect emerges across the entire distribution of participants. Panel (b) in Figure 4 shows the distribution of the number of strategic options generated by the treatment group, which is shifted to the right of that for the control group (Kolmogorov–Smirnov test $p = 0.039$), showing that participants working with the framework generated more strategic alternatives overall. Together, these results suggest that providing the framework helps decision-makers identify and consider additional strategic alternatives that they might not have considered in the absence of the framework.

We find that this result is also reflected in the topics of the options. Figure 5 highlights a difference in the distribution of topics between the treatment and control groups (Chi-squared test $p < 0.001$). The treatment group tended to come up with more choices related to expansion ($p = 0.002$) and exit ($p = 0.019$), while the control group focused more on advertising ($p = 0.013$) and sales team salaries ($p < 0.001$). This result is in line with our findings from the main analysis, providing additional evidence that the treatment group tended to generate more strategic alternatives compared to operational ones.

Using the framework also increased the likelihood of generating a mutually exclusive set of alternatives. Column 3 of Table 2 shows that on average, participants assigned to treatment were 17 percentage points more likely to generate a set of mutually exclusive alternatives ($p < 0.001$). This provides additional support that the framework appears to encourage the development of alternatives that are genuinely different from one another. The framework's effectiveness may stem from its clear articulation of key strategic dimensions, which encourages participants to explore truly different options.

Furthermore, using the framework decreased the number of alternatives that involved continuing with the current strategy. Column 4 in Table 2 shows that compared to the control group, participants who received the framework on average generated 0.6 fewer alternatives involving continuing with the firm's current direction ($p = 0.003$). This translates into a 21% decrease compared to the control group, suggesting that the framework may have nudged decision-makers toward exploring alternative strategic paths from the status quo. Panel (c) in Figure 4 shows that the distribution of the number of continuing alternatives is shifted to the left for the treatment group relative to the control group (Kolmogorov–Smirnov test $p = 0.019$), suggesting that the framework had treatment effects across the full distribution. In addition to this, treated participants were 12 percentage points more likely to consider exit ($p = 0.009$) on average compared to the control group.

While our results suggest that frameworks may help decision-makers develop more strategic alternatives, one possible challenge is that they may lead them to think similarly to others who use the same framework. To explore, we analyze the cosine similarity measures based on word embedding from the sentence transformer model, which assesses semantic closeness between the alternatives. Column 6 in Table 2 shows that using the framework increased the average similarity of alternatives relative to those generated by others in the same experimental group by 2% ($p < 0.001$). These results suggest that while decision-makers might consider more strategic alternatives when provided with a framework, they may end up considering more similar alternatives as others who also use the framework. This result echoes findings from cognitive science, which have shown that external representations may constrain cognitive processes, potentially leading to more similar decision-making perspectives across individuals (Zhang, 1997). However, it is important to note that this similarity is not necessarily a disadvantage, particularly if the strategic ideas require coordinated effort for implementation. In fact, external representations provided by frameworks could enhance shared understanding among decision-makers, potentially facilitating more effective collaboration and implementation of strategies (Csaszar, Hinrichs, & Heshmati, 2024). This shared cognitive framework may prove beneficial in organizational settings where alignment and coordinated action are crucial.

Table 3 shows that these treatment effects also carry over to the ultimate choice that participants select. We find that treated participants exposed to the framework were 17 percentage points more likely to choose strategic options as their best choices ($p < 0.001$), 10 percentage points less likely to choose to continue with the current strategy ($p = 0.05$), and 3 percentage points more likely to choose to exit ($p = 0.054$). These results suggest that the framework aided participants in charting out a diverse set of strategic options and influenced their perceptions of what constituted the best strategic path forward.

Together, these results suggest that frameworks may help decision-makers shift to a more strategic focus when considering and choosing what alternatives are available for the company. We find consistent results across different regression specifications and different ways to construct the measurements, such as using a natural log transformation of count variables (total number of options, number of strategic options, and number of continuing options), as well as using the raw count instead of a binary indicator for exit (see Appendix D for further details). However, we also find that frameworks may lead participants to generate options that are similar to others who also use the framework.

4.2 Mechanisms: The effect on problem formulation

To investigate potential mechanisms through which frameworks might influence the set of options considered by participants, we explore data from surveys and semi-structured interviews of approximately 15-30 minutes each (see Appendix C for detailed interview questions). Our analyses provide suggestive evidence that the framework may have helped generate more strategic options by impacting how participants formulated the problem at hand as well as the solution space.

Compared to the control group, participants assigned to treatment stated that they found it easier to formulate and represent the problem the company faced. For example, Interviewee 20 (Treatment) noted: “The biggest challenge was starting out. The chart provided was useful in sparking ideas.” Similarly, Interviewee 2 (Treatment) described the framework as “a lens through which I could look at the situation at hand.” Interviewee 16 (Treatment) expressed confidence in identifying problems. On the contrary, interviewee 6 (Control) commented on struggling with defining the problem at hand. Interviewee 14 (Control) found initiating solutions challenging and mentioned the tendency to stray: “Once you just get the task, you need to organize your chain of thoughts and it’s very easy to go in the wrong direction and then just one thing leads to the other.” Similarly, Interviewee 9 (Control) pointed out a lack of clear problem scope: “There was no specific scope of what the current problem we are trying to solve [is].” This contrast in experience between the two experimental groups suggests the potential effectiveness of the framework in aiding participants to identify and formulate problems more effectively.

Furthermore, participants across experimental groups appeared to represent the problem differently. Treated participants tended to include multiple viewpoints and consider a broader strategic context. For instance, Interviewee 2 (Treatment) described how the framework enabled them “to think about the problem from all the points of view.” Similarly, Interviewee 15 (Treatment) employed a comprehensive approach by focusing on the problem, the time variable, and the stakeholders involved. Interviewee 16 (Treatment) started by understanding the customer base and identifying the target segment. This broader perspective was echoed in survey responses, with participants noting, “[The framework] helped me look at the problem from different points of view,” and “provide a comprehensive process to look through the problem.” In contrast, participants in the Control Group tended to focus predominantly on discrete and visible business issues without integrating them into a broader strategic context. Interviewee 8 (Control) focused on the problem of leveraging volume to obtain increased revenue, while Interviewee 9 (Control) highlighted the problem of customer retention. Interviewee 10 (Control) described a more ad hoc approach: “I just picked up some obvious problems that I saw. And I brought up a solution to that problem.” Overall, the framework appeared to enable a broader representation of problems, while the control group appeared to focus more on straightforward, easily identifiable business issues and adopt a one-by-one problem-solving approach.

The framework also structured participants’ thinking processes and refocused their attention. Interviewee 21 (Treatment) noted the framework’s role in organizing their thoughts, and Interviewee 23 (Treatment) suggested that the framework “gave me some structure.” Survey responses further illustrated this point. One participant described their method as “Structured the options around the three building blocks (continue, expand and exit), taking into account the constraints currently facing the company.” Another participant noted their thought process as “go[ing] through the structural decision map.” Other responses included “Very structured approach so you can create a more comprehensive/exhaustive option set,” “It allowed me to structure my thinking,” and “It helps to guide thinking and be MECE (Mutually Exclusive, Collectively Exhaustive).”

Participants in the treatment group also used the framework as a reference to add and prune alternatives. In this way, the framework influenced how the participants perceived their solution space. For example, Interviewee 1 (Treatment) described using the framework to first brainstorm ideas and then adjust them to fit the framework better. Relatedly, others highlighted that the framework led them to “anchor” on certain alternatives, potentially overlooking other viable alternatives like forming alliances (Interviewee 4 (Treatment)). Interviewee 23 (Treatment) described an iterative process: “For my second option, I revisited the framework. This process reminded me of our lessons on blue oceans, which prompted me to think of unconventional strategies. In the end, I had two more ideas using the framework.” Interviewee 4 (Treatment) recognized the framework’s role in inspiration: “[The framework] helps remind you of options that you might not have in the first place.” A survey respondent also highlighted the inspirations from the framework: “It also triggered potential ideas that I would have never thought of.” This demonstrates how the framework was instrumental in both generating and pruning alternatives.

We further quantitatively analyzed survey responses and found consistent evidence that control and treatment groups used different approaches to formulate the problem at hand and develop their alternatives. We coded text responses from the participants on how they developed the alternatives into different categories using a self-defined dictionary approach.[6] Figure 6 shows that the treatment group leaned towards external analysis and expansion, suggesting that the framework may have influenced decision-makers towards a more outward-looking approach to consider the external environment and identify strategic opportunities and threats. Conversely, the control group was more likely to use internal analysis and a problem-solving approach, indicating a more introspective and potentially reactive approach to address immediate issues. These results suggest that frameworks may help shift decision-makers from an introspective, reactive mode to a more external and expansion-centric approach.

These findings suggest that frameworks may influence not only the alternative generation stage but also the problem formulation stage in strategic decision-making. This is consistent with extant research in cognitive science, which argues that external representations

[6]The categories we landed on were devised based on recurring themes and terminologies that emerged from the participants’ descriptions of how they developed their options.

can influence internal mental interpretation by simplifying, omitting, adding, or distorting information (Bryant & Tversky, 1999).

4.3 Follow-up experiments: A set of frameworks versus a specific framework

The findings from the main experiment suggest that the treatment framework in Figure 2 influenced how participants identified strategic alternatives. This raises the question of whether the observed outcomes may be specific to this particular framework, or whether any frameworks might have similar effects simply by perturbing the decision environment. To address this question, we conducted follow-up experiments that varied the frameworks provided to participants.

The follow-up experiments involved 480 participants, including MBA students and executives enrolled in a strategy program at a top-tier business school. The experimental design and tasks closely mirrored those of the main study, with the primary variation being that participants in the treatment group received a set of strategic frameworks rather than a single framework. We varied the set of strategic frameworks in two ways: either including the framework itself or incorporating its key components (see Figures 3a and 3b). This design allows us to test whether the mere presence of strategic frameworks enhances strategic thinking, and whether our framework's effectiveness persists when presented alongside other frameworks. In the follow-up experiments, a total of 2,842 alternatives were generated by the participants, which were coded using a combination of human coders and a fine-tuned GPT-3.5 Turbo model trained on previously human-coded data (see Appendix B.2).

Our results show that providing participants with multiple frameworks mutes the treatment effect observed in the main experiment. When participants receive a set of frameworks rather than a single framework, both the magnitude and precision of the treatment effects decrease substantially. This pattern holds across our key measurements—from option generation to selection—including the number of strategic options, their mutual exclusivity, and whether participants select strategic options as their best choice (Appendix Tables G.2 and G.4). The results remain consistent whether we analyze the follow-up experiments separately or pool them together.

One might question whether this lack of impact stems from participants being presented with too many frameworks, potentially leading them to disregard any specific framework. However, analysis of participants' descriptions of their thought processes confirmed that they recognized and paid substantial attention to the frameworks (Appendix Figure G.1). These results suggest that for frameworks to effectively stimulate the generation of strategic alternatives, focused attention may be crucial. Given the difficulty of shifting entrenched mental models, presenting many frameworks may overwhelm rather than redirect thinking—ultimately diluting, rather than enhancing, their impact.

However, this cognitive bottleneck may be less binding in AI-assisted strategy making. To explore, we replicated our experiments using Large Language Models (LLMs). We generated a comparable sample of virtual participants, randomly assigned them to different treatments, and instructed them to consider the same problem faced by human participants. This procedure was used to replicate both our main experiment and the multi-framework follow-up (see Appendix H for more details).

We find that when prompted with the framework, LLMs produced significantly more strategic alternatives, consistent with our main findings (Table H.2). More strikingly, when presented with multiple frameworks, LLMs were able to process and integrate them to identify additional strategic options, although the effect size was smaller than in the main simulation (Table H.3). These results suggest that the cognitive bottleneck observed in human participants may be mitigated as AI becomes increasingly integrated into strategic decision-making (Boussioux et al., 2024; Csaszar, Ketkar, & Kim, 2024; Dell'Acqua et al., 2023; Doshi et al., 2025).

5 Discussion and conclusion

Using randomized controlled trials involving 820 MBA students and executives, this study explores how frameworks influence the strategic alternatives decision-makers generate and ultimately select. We provide empirical evidence on the effectiveness of frameworks in strategic decision-making, showing that they significantly shift the set of options visible to decision-makers and the choices they make. Frameworks increase the likelihood that decision-makers consider more strategic alternatives, develop a mutually exclusive set of options, and select a strategic choice as the best option.

Moreover, we provide suggestive evidence that the mechanism through which frameworks might affect strategic decisions is by shaping how decision-makers formulate the problem. Our findings suggest that frameworks may facilitate problem formulation and nudge decision-makers to think from a broader perspective, compared to a more internal and reactive approach. However, frameworks also introduce a potential tradeoff: while they may help generate more strategic alternatives, they also appear to increase convergence in how individuals frame the problem, resulting in greater similarity across the alternatives generated.

The follow-up experiments using multiple frameworks reveal an important boundary condition: frameworks enable alternative generation only when they focus attention rather than fragment it. Contrary to expectations that multiple frameworks would broaden perspectives, providing multiple frameworks *muted* the treatment effects observed in the main experiment. The results suggest that providing multiple frameworks may overwhelm rather than redirect thinking, especially given the difficulty in shifting decision-makers' entrenched mental models.

However, this cognitive bottleneck may be less binding in AI-assisted strategy making. In our LLM-based simulations of the experiments, we find that LLMs were able to integrate both the single framework and the multi-framework, suggesting the cognitive bottleneck of processing multiple frameworks may diminish as AI becomes increasingly integrated into strategic decision-making (Boussioux et al., 2024; Csaszar, Ketkar, & Kim, 2024; Dell'Acqua et al., 2023; Doshi et al., 2025). Moreover, AI systems themselves can serve as powerful tools for creating new types of external representations, from textual narratives to data visualizations and simulations, expanding the representational repertoire available to strategic decision-makers. These AI systems are distinctly valuable because they are interactive, allowing decision-makers to rapidly iterate and examine strategic issues from diverse stakeholder viewpoints. This interactivity and fluidity could counter tendencies toward cognitive entrenchment, prompting a continuous re-evaluation of assumptions and dominant mental models governing strategic issues. Future research could explore how strategic issues are framed and understood using AI-enabled external representations.

Our findings also speak to the scientific method of strategy making (Camuffo et al., 2020; Koning et al., 2022), which posits that generating viable alternatives is a prerequisite for learning through experimentation (Gans et al., 2019). While prior work emphasizes the value of experimentation, it overlooks how managers generate the diverse options for experimentation in the first place. We bridge this gap by demonstrating that frameworks enable decision-makers to produce more strategic and mutually exclusive alternatives, which are critical attributes for meaningful experimentation. By structuring problem representation, frameworks help managers move beyond incremental adjustments (e.g., cost-cutting) to consider divergent paths (e.g., market entry or acquisition), thereby creating the necessary "portfolio" of alternatives to test and iterate. This suggests frameworks are not merely analytical tools but generative mechanisms for strategic experimentation.

Our work also has its limitations. First, the empirical setting of our experiment, involving MBA students and executives enrolled in a program, leaves open questions about its generalizability to real-world corporate scenarios. The context in which MBA students make decisions during an experiment is significantly different from the high-stakes, high-pressure environments in which corporate decisions are often made. Real-world scenarios involve a complex interplay of factors such as organizational constraints and incentives, stakeholder expectations, regulatory environments, and competitive pressures, which might not be adequately simulated in an academic setting. While these students represent those who already occupy key organizational and entrepreneurial positions (Levine et al., 2023), they may perceive and respond to risk differently in these contexts relative to practice.

In addition, our study highlights the immediate effects of frameworks, but important questions remain about their long-term impact. Our findings show that frameworks can nudge decision-makers to approach problems differently, suggesting that it may lead to a short-term shift in their mental models. However, sustaining these effects over time may require repeated exposure and reinforcement. Achieving lasting cognitive change is likely more challenging, though emerging technologies such as AI agents may offer new opportunities for continuous engagement. Exploring how prolonged interaction with frameworks influences strategic thinking over time, and how technological tools can support this process, might be a promising direction for future research.

References

Abadie, A., Athey, S., Imbens, G., & Wooldridge, J. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, *138*(1), 1–35.

Acemoglu, D., He, A., & Maire, D. L. (2022). Eclipse of rent-sharing: The effects of managers' business education on wages and the labor share in the us and denmark [Available at: https://www.nber.org/papers/w29874]. *National Bureau of Economic Research*, (w29874).

Agrawal, A., Gans, J., & Stern, S. (2021). Enabling entrepreneurial choice. *Management Science*, *67*(9), 5510–5524.

Anderson, N., Potočnik, K., & Zhou, J. (2014). Innovation and Creativity in Organizations: A State-of-the-Science Review, Prospective Commentary, and Guiding Framework [Publisher: SAGE Publications Inc]. *Journal of Management*, *40*(5), 1297–1333. https://doi.org/10.1177/0149206314527128

Argyris, C., & Schön, D. (1978). *Organizational learning: A theory of action perspective*. Addison-Wesley Publishing Company.

Augenblick, N., & Rabin, M. (2021). Belief movement, uncertainty reduction, and rational updating*. *The Quarterly Journal of Economics*, *136*(2), 933–985.

Baer, M., Dirks, K., & Nickerson, J. (2013). Microfoundations of strategic problem formulation. *Strategic Management Journal*, *34*(2), 197–214.

Berg, J. M., Duguid, M. M., Goncalo, J. A., Harrison, S. H., & Miron-Spektor, E. (2023). Escaping irony: Making research on creativity in organizations more creative. *Organizational Behavior and Human Decision Processes*, *175*, 104235. https://doi.org/10.1016/j.obhdp.2023.104235

Boussioux, L., Lane, J. N., Zhang, M., Jacimovic, V., & Lakhani, K. R. (2024). The Crowdless Future? Generative AI and Creative Problem-Solving [Publisher: INFORMS]. *Organization Science*. https://doi.org/10.1287/orsc.2023.18430

Bryant, D., & Tversky, B. (1999). Mental representations of perspective and spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(1), 137–156.

Camuffo, A., Cordova, A., Gambardella, A., & Spina, C. (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science*, *66*(2), 564–586.

Cox, D., & Reid, N. (2000). *The theory of the design of experiments*. CRC Press.

Craik, K. (1943). *The nature of explanation*. University Press, Macmillan.

Csaszar, F. (2018). What makes a decision strategic? strategic representations. *Strategy Science*, *3*(4), 606–619.

Csaszar, F., Hinrichs, N., & Heshmati, M. (2024). External representations in strategic decision-making: Understanding strategy's reliance on visuals [Forthcoming. Available at: https://onlinelibrary.wiley.com/doi/abs/10.1002/smj.3613]. *Strategic Management Journal*.

Csaszar, F., Ketkar, H., & Kim, H. (2024). Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors. *SSRN*.

Csaszar, F., & Laureiro-Mart'inez, D. (2018). Individual and organizational antecedents of strategic foresight: A representational approach. *Strategy Science*, *3*(3), 513–532.

Csaszar, F., & Levinthal, D. (2016). Mental representation and the discovery of new strategies. *Strategic Management Journal*, *37*(10), 2031–2049.

Cyert, R., & March, J. (1963). *A behavioral theory of the firm*. Prentice-Hall.

Dell'Acqua, F., McFowland III, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023, September). Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. Retrieved May 30, 2024, from https://papers.ssrn.com/abstract=4573321

Doshi, A. R., Bell, J. J., Mirzayev, E., & Vanneste, B. S. (2025). Generative artificial intelligence and evaluating strategic decisions. *Strategic Management Journal*, *46*(3), 583–610. https://doi.org/10.1002/smj.3677

Eggers, J., & Kaplan, S. (2009). Cognition and renewal: Comparing ceo and organizational effects on incumbent adaptation to technical change. *Organization Science*, *20*(2), 461–477.

Felin, T., & Zenger, T. (2017). The theory-based view: Economic actors as theorists. *Strategy Science*, *2*(4), 258–271.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Gans, J., Stern, S., & Wu, J. (2019). Foundations of entrepreneurial strategy. *Strategic Management Journal*, *40*(5), 736–756.

Gary, M., & Wood, R. (2011). Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal*, *32*(6), 569–594.

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

Gavetti, G., Levinthal, D., & Rivkin, J. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, *26*(8), 691–712.

Giere, R. (2004). The problem of agency in scientific distributed cognitive systems. *Journal of Cognition and Culture*, *4*(3–4), 759–774.

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea Generation and the Quality of the Best Idea [Publisher: INFORMS]. *Management Science*, *56*(4), 591–605. https://doi.org/10.1287/mnsc.1090.1144

Grootendorst, M. (2022). Bertopic: Neural topic modeling with a class-based tf-idf procedure [Available at: http://arxiv.org/abs/2203.05794]. *arXiv*.

Hanna, R., Mullainathan, S., & Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment*. *The Quarterly Journal of Economics*, *129*(3), 1311–1353.

Helfat, C., & Peteraf, M. (2015). Managerial cognitive capabilities and the microfoundations of dynamic capabilities. *Strategic Management Journal*, *36*(6), 831–850.

Heshmati, M., & Csaszar, F. A. (2024). Learning strategic representations: Exploring the effects of taking a strategy course. *Organization Science*, *35*(2), 453–473. https://doi.org/10.1287/orsc.2023.1676

Hugging Face. (2022, November). Sentence-transformers/all-mpnet-base-v2 · hugging face [Available at: https://huggingface.co/sentence-transformers/all-mpnet-base-v2].

Kim, H. (2021). Rated: In search of strategic options. *INSEAD Case*.

Kim, H., Simonovich, D., & Stern, I. (2024). Crafting strategic options. *Working Paper*.

Kirsh, D. (2010). Thinking with external representations. *AI & SOCIETY*, *25*(4), 441–454.

Kleinmuntz, D., & Schkade, D. (1993). Information displays and decision processes. *Psychological Science*, *4*(4), 221–227.

Koning, R., Hasan, S., & Chatterji, A. (2022). Experimentation and start-up performance: Evidence from a/b testing. *Management Science*, *68*(9), 6434–6453.

Lane, J., Cai, T., Menietti, M., Weber, G., & Guinan, E. (2024). Greenlighting innovative projects: How evaluation format shapes the perceived feasibility of novel ideas. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*.

Leiblein, M., Reuer, J., & Zenger, T. (2018). What makes a decision strategic? *Strategy Science*, *3*(4), 558–573.

Levine, S. S., Schilke, O., Kacperczyk, O., & Zucker, L. G. (2023). Primer for experimental methods in organization theory. *Organization Science*, *34*(6), 1997–2025. https://doi.org/10.1287/orsc.2023.18093

Levinthal, D. (2011). A behavioral approach to strategy-what's the alternative? *Strategic Management Journal*, *32*(13), 1517–1523.

Levitt, B., & March, J. (1988). Organizational learning. *Annual Review of Sociology*, *14*(1), 319–338.

McKinsey. (2008, July). Enduring ideas: Classic mckinsey frameworks that continue to inform management thinking [Available at: https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/enduring-ideas-classic-mckinsey-frameworks-that-continue-to-inform-management-thinking].

Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.

Puranam, P., & Swamy, M. (2016). How initial representations shape coupled learning processes. *Organization Science*, *27*(2), 323–335.

Rabin, M., & Schrag, J. (1999). First impressions matter: A model of confirmatory bias*. *The Quarterly Journal of Economics*, *114*(1), 37–82.

Reeves, M., Moose, S., & Venema, T. (2014, June). Bcg classics revisited: The growth share matrix [Available at: https://www.bcg.com/publications/2014/growth-share-matrix-bcg-classics-revisited].

Scaife, M., & Rogers, Y. (1996). External cognition: How do graphical representations work? *International Journal of Human-Computer Studies*, *45*(2), 185–213.

Schwartzstein, J., & Sunderam, A. (2021). Using models to persuade. *American Economic Review*, *111*(1), 276–323.

Simon, H. (1947). *Administrative behavior; a study of decision-making processes in administrative organization*. Macmillan.

Simon, H. (1996). *The sciences of the artificial, 3rd edition*. MIT Press.

Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, *19*(1), 97–140.

Tripsas, M., & Gavetti, G. (2000). Capabilities, cognition, and inertia: Evidence from digital imaging. *Strategic Management Journal*, *21*(10–11), 1147–1161.

Van den Steen, E. (2017). A formal theory of strategy. *Management Science*, *63*(8), 2616–2636. https://doi.org/10.1287/mnsc.2016.2468

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9*(4), 625–636.

Yang, M.-J., Christensen, M., Bloom, N., Sadun, R., & Rivkin, J. (2020). How do ceos make strategy? [Available at: https://www.nber.org/papers/w27952]. *National Bureau of Economic Research*, (w27952).

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science*, *21*(2), 179–217.

Zhang, J., & Norman, D. (1994). Representations in distributed cognitive tasks. *Cognitive Science*, *18*(1), 87–122.

Tables

Table 1: Summary statistics

	(1) Full sample					(2) Treatment			(3) Control			(4) Difference (3)-(2)		
	N	Mean	SD	Min	Max	N	Mean	SD	N	Mean	SD	b	se	p
Baseline Variables														
Female	340	0.38	0.49	0	1	169	0.37	0.48	171	0.40	0.49	0.03	0.05	0.560
Consulting	340	0.30	0.46	0	1	169	0.31	0.47	171	0.29	0.45	-0.03	0.05	0.587
Age	340	29.62	2.58	24	36	169	29.47	2.58	171	29.77	2.58	0.30	0.28	0.286
Master's or higher	340	0.39	0.49	0	1	169	0.37	0.48	171	0.41	0.49	0.04	0.05	0.491
Years since Grad	340	6.49	2.41	0	14	169	6.44	2.42	171	6.53	2.41	0.09	0.26	0.736
Years of Work Experience	340	5.86	2.15	2	13	169	5.86	2.14	171	5.87	2.17	0.01	0.23	0.962
GMAT (Standardized)	340	-1.97	5.57	-17	2	169	-2.04	5.79	171	-1.90	5.37	0.14	0.61	0.817
Outcome Variables														
Number of Options	340	6.63	3.64	2	23	169	6.63	3.76	171	6.62	3.53	-0.01	0.40	0.973
Number of Strategic Options	340	4.62	2.99	0	17	169	5.05	3.11	171	4.19	2.81	-0.85	0.32	0.008
Mutually Exclusive	340	0.19	0.39	0	1	169	0.27	0.44	171	0.11	0.31	-0.16	0.04	0.000
Number of Continue Options	340	2.67	1.95	0	9	169	2.35	1.79	171	2.98	2.06	0.63	0.21	0.003
Exit	340	0.24	0.42	0	1	169	0.30	0.46	171	0.18	0.38	-0.12	0.05	0.009
Best Option is Strategic	340	0.78	0.41	0	1	169	0.86	0.34	171	0.70	0.46	-0.16	0.04	0.000
Best Option is Continue	340	0.28	0.45	0	1	169	0.23	0.42	171	0.33	0.47	0.10	0.05	0.036
Best Option is Exit	340	0.02	0.14	0	1	169	0.04	0.19	171	0.01	0.08	-0.03	0.02	0.054

Notes: This table reports the summary statistics and the balance of variables across experimental groups. GMAT includes the conversion of GRE scores into a predicted GMAT score based on a formula developed by the Educational Testing Service (ETS). The standardized GMAT was computed by subtracting the sample mean from each individual score and dividing by the sample standard deviation.

Table 2: The impact of frameworks on alternative generation

VARIABLES	(1) Number of Options	(2) Number of Strategic Options	(3) Mutually Exclusive	(4) Number of Continue Options	(5) Exit	(6) Average Similarity
Treatment	-0.01	0.86***	0.17***	-0.62***	0.12***	0.02***
	(0.39)	(0.31)	(0.04)	(0.21)	(0.05)	(0.00)
Constant	6.42**	2.62	-0.23	2.54	-0.16	0.37***
	(3.12)	(2.53)	(0.40)	(1.84)	(0.40)	(0.04)
Observations	340	340	340	340	340	340
Gender-Consulting-Section FE	YES	YES	YES	YES	YES	YES
Control Variables	YES	YES	YES	YES	YES	YES
Adjusted R-squared	0.03	0.06	0.06	0.09	0.03	0.05
Control Group Mean	6.62	4.19	0.11	2.98	0.18	0.32

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports regression results for equation (1). All regression models include fixed effects for the randomization strata (gender-consulting-section) and control for age, whether the participant has a master's degree or above, the number of years since they obtained their last degree, the number of years of working experience, and standardized GMAT scores.

Table 3: The impact of frameworks on the best alternative chosen

VARIABLES	(1) Best Option is Strategic	(2) Best Option is Continue	(3) Best Option is Exit
Treatment	0.17***	-0.10*	0.03*
	(0.04)	(0.05)	(0.01)
Constant	0.68*	0.54	-0.02
	(0.37)	(0.44)	(0.11)
Observations	340	340	340
Gender-Consulting-Section FE	YES	YES	YES
Control Variables	YES	YES	YES
Adjusted R-squared	0.040	-0.007	-0.023
Control Group Mean	0.702	0.333	0.006

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports regression results for equation (1). All regression models include fixed effects for the randomization strata (gender-consulting-section) and control for age, whether the participant has a master's degree or above, the number of years since they obtained their last degree, the number of years of working experience, and standardized GMAT scores.

Figures

Figure 1: Instructions for all participants

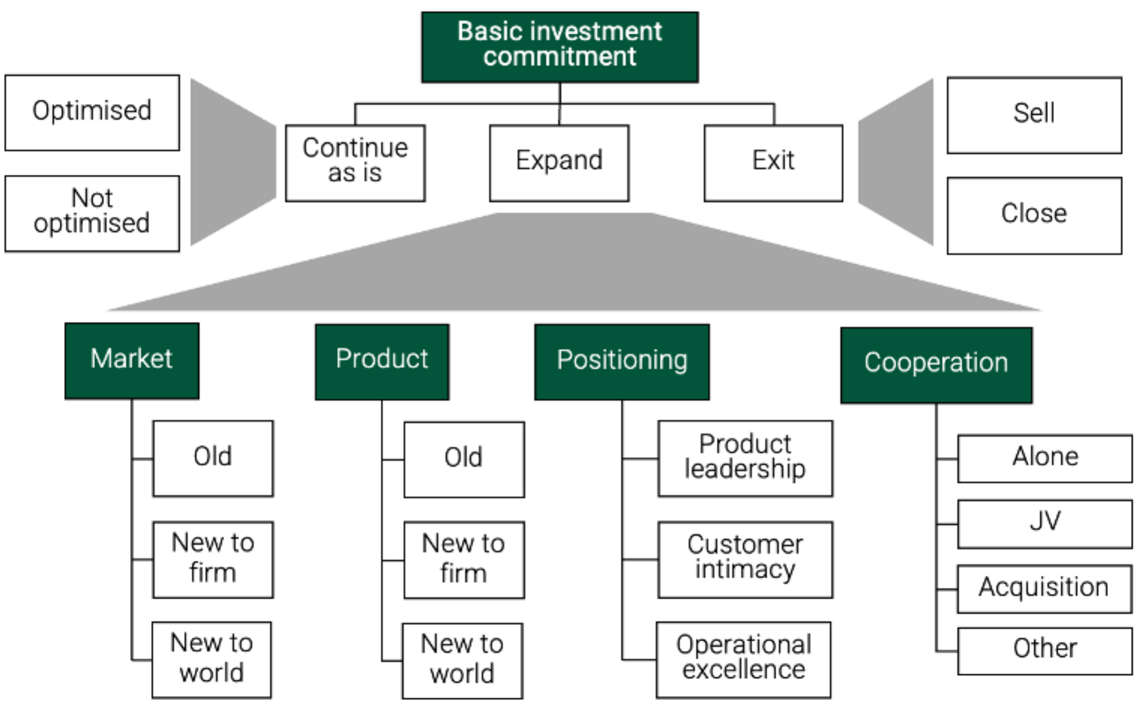
CRAFTING STRATEGIC OPTIONS

Strategy development includes the articulation of strategic options available to a business before making a final choice for execution.

Your task is to articulate what strategic options are available for Rated, Inc. Please input only one option at a time.

Notes: This figure shows the general instructions provided to all participants at the online interface.

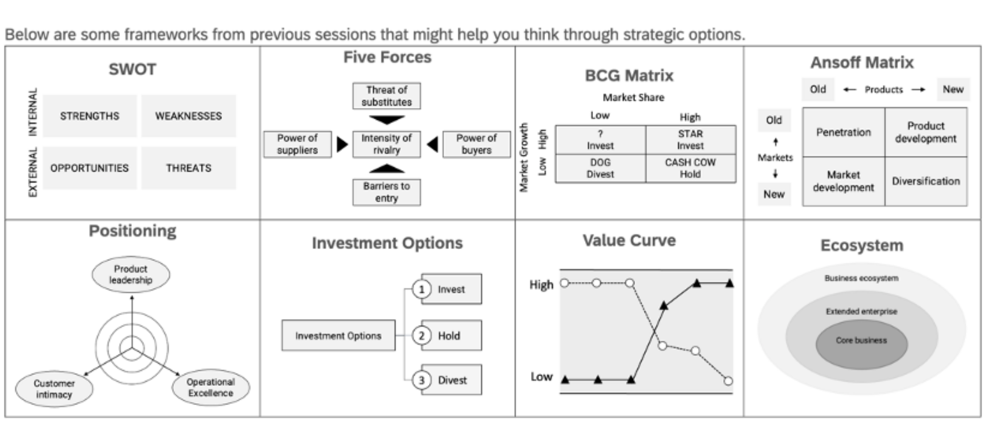
Figure 2: Design of treatment - Main experiment



Notes: This figure shows the framework designed to guide the generation of strategic options (Kim et al., 2024). It was provided to the treatment group in the main experiment.

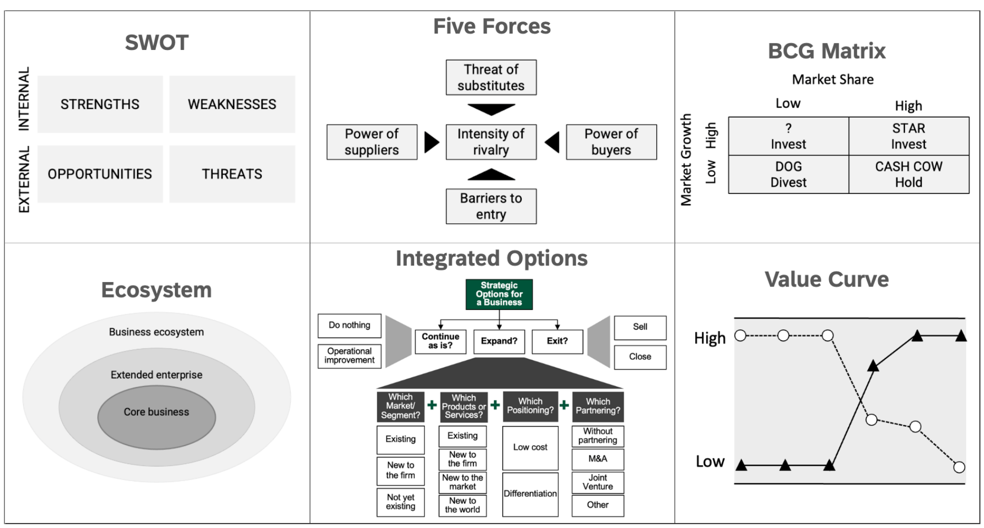
Figure 3: Design of treatments - Follow-up Experiments

(a) A set of frameworks provided in the follow-up experiment 1



Notes: This figure shows a summary of well-known frameworks. They were provided to the treatment group in the follow-up experiment.

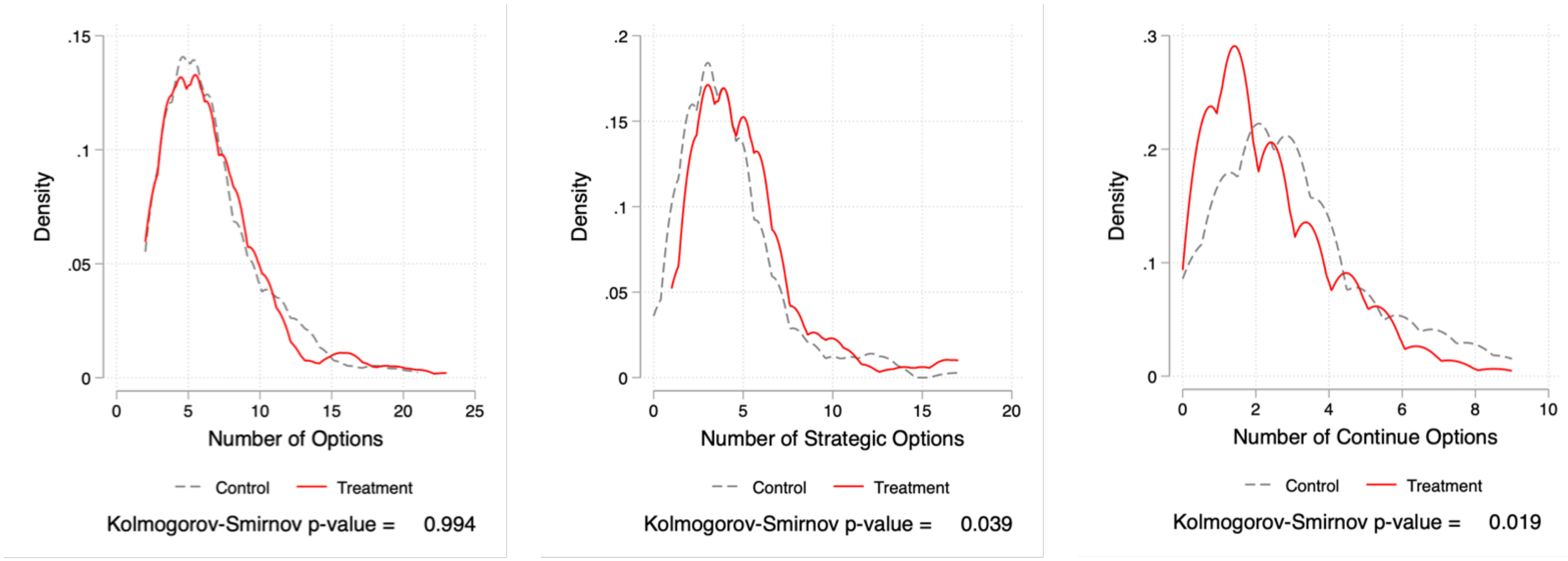
(b) A set of frameworks provided in the follow-up experiment 2



Notes: This figure shows a summary of well-known frameworks. They were provided to the treatment group in the second follow-up experiment. The "Integrated Options" framework was provided as the treatment in the main experiment.

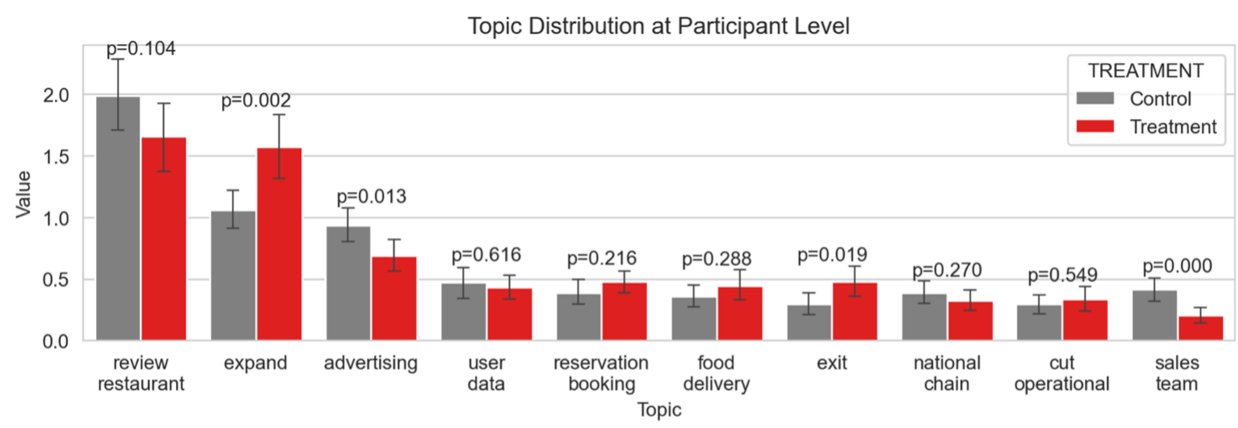
Figure 4: Distributions of outcome variables

(a) Total Number of Options (b) Number of Strategic Options (c) Number of Continue Options



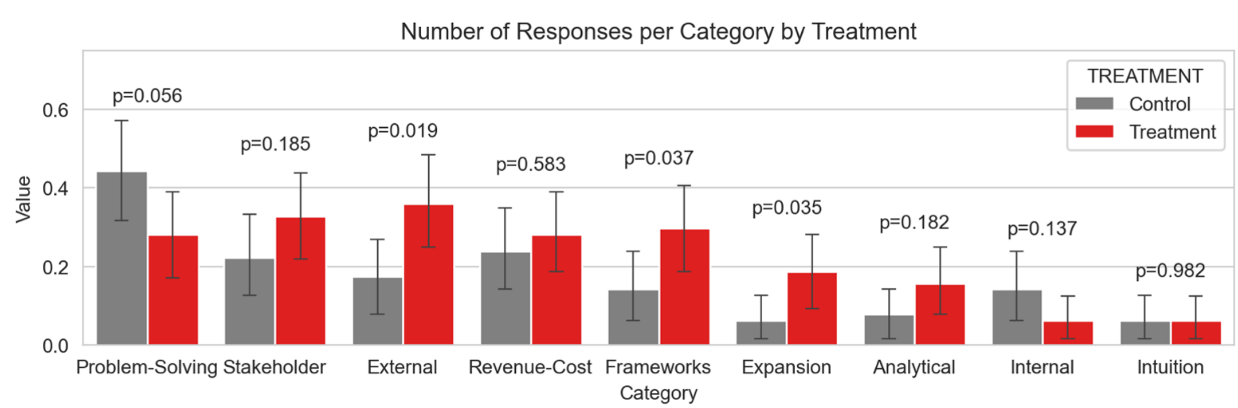
Notes: These figures show the distribution of outcome variables of the treatment group (in red, solid line) versus the control group (in grey, dotted line). Panel (a) shows the distribution of the number of alternatives generated. Panel (b) shows the distribution of the number of strategic options generated. Panel (c) shows the distribution of the number of continuing options generated. Histogram versions of the figures are shown in Appendix Figure A.1.

Figure 5: Comparison of topic distributions by treatment



Notes: This figure plots the result of the BERTopic Model analysis. The x-axis shows the topics. The y-axis shows the average number of alternatives per topic at the participant level. The p-value indicates the t-test results comparing the average outcomes between the control and the treatment groups.

Figure 6: Comparison of alternative generation approaches by treatment



Notes: This figure plots the result of the categorization of the responses to the survey question "Describe how you developed your alternatives." The x-axis shows the categories. The y-axis shows the average number of participants per category. The p-value indicates the t-test results comparing the average outcomes between the control and the treatment groups.

Appendices

A Experiment details

This appendix provides additional details on the experiment. Table A.1 shows the timeline of experimental interventions. Figure A.1 shows the distribution of outcome variables in histograms.

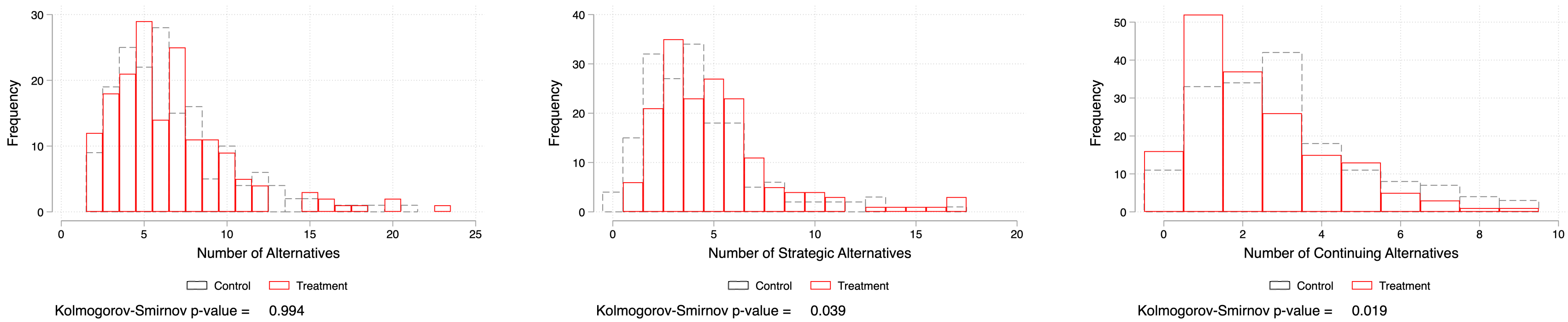
Table A.1: Timeline of experimental interventions

	Date of Experiment	Number of Sections	Number of Participants
1	30 September 2022	2	113
2	16 February 2023	2	104
3	15 September 2023	2	123
	Total	6	340

Notes: This table shows the timeline and the corresponding details for each experiment.

Figure A.1: Distributions of outcome variables

(a) Total Number of Options (b) Number of Strategic Options (c) Number of Continue Options



Notes: These figures show the distribution in histograms of outcome variables of the treatment group (in red, solid line) versus the control group (in grey, dotted line). Panel (a) shows the distribution of the number of alternatives generated. Panel (b) shows the distribution of the number of strategic options generated. Panel (c) shows the distribution of the number of continuing options generated.

B Coding process

This appendix provides additional details on the coding process.

Each option and each set of options were coded by two research assistants independently based on the coding rubric provided by the authors. Following this, their coding results were compared to identify discrepancies. In cases of disagreement, a third assistant served as the arbitrator to resolve the issue and finalize the coding decision. In total, four research assistants were involved in the coding process. To maintain objectivity, all of them were blind to the experiment details, the hypotheses, and the conditions.

Section B.1 presents the coding rubric, which was constructed by the authors based on the case study used in the experiment. Section B.2 describes the process of training a fine-tuned GPT-3.5 Turbo model trained on previously human-coded data.

B.1 Coding rubric used by human coders

For **each option**, code as strategic versus operational option:

	Description	Examples
Strategic option	Long-term, high-level plan, not easy to reverse, involves significant resource allocation.	• Launch new product • Enter new market (e.g., geographic) • Mergers & Acquisitions (M&A) • Exit • A complete overhaul of the platform's user interface • Strategic partnerships
Operational option	Specific course of action with short-term goals, focusing on day-to-day activities and tasks.	• Cut costs • Promotion / marketing • Increase sales team pay • Training sessions

Then, code whether the option suggests that the company continue its current strategy:

	Examples
Continue as it is	• Explicitly mention "continue as it is" • Increase customer stickiness • Change pricing system — note as a special case as this could be seen as a change • Streamline operations, improve efficiency, or increase profit margins without fundamentally altering its value proposition, target market, or key activities. The broader business model or strategy remains largely unchanged. (e.g., keeping the focus on local businesses as the key target market would represent "continuing as is", while changing the target market from local businesses to national chains would not)
Not continue	• Launch new product • Enter new market

For **each student**, code whether the options were mutually exclusive:

- Mutually exclusive: Choosing one option would prevent the firm from choosing any of the other options.
- All the options need to be mutually exclusive with each other, so any presence of a non-mutually exclusive option would make the set as a whole not mutually exclusive
- Assume the company has limited resources, launching one strategic initiative would constraint the company from launching another

	Examples
Mutually exclusive	• 1. Continue; 2. Expand; 3. Exit • 1. Expand ecosystem; 2. Focus on ad sales; 3. Sell the business • (Refer to the figure)
Not mutually exclusive	• 1. Bring in new clients; 2. Automation to cut costs • 1. Expand service offering; 2. Optimize pricing; 3. Modify compensation; 4. Reduce admin costs

Output

1. Create one spreadsheet with the coding of all strategic options specifying response ID (with the actual options text):
   a. ResponseID
   b. OptionID
   c. Option: Actual option text
   d. Strategic: Dummy, whether the option is strategic or not, 1 means strategic, 0 otherwise
   e. Continue: Dummy, whether the option is "continue as it is" or not, 1 means continue as it is, 0 means otherwise
   f. Change Pricing of Ads: Dummy, 1 means the option indicates a change in pricing, 0 means otherwise
      i. Changing the pricing of the ads
      ii. Change in pricing would explicitly mention phrases like "change of pricing structure", "change the pricing model", etc.
   g. Exit: Dummy, 1 means the option indicates exit, and 0 otherwise
      i. Exit means selling the business or exiting the market
   h. Note
      i. Improve algorithm if the option mentions this
      ii. Special case of "continue as it is" if not fully sure
      iii. Multiple if the option seems to include multiple options
      iv. Partially exit if the option mentions exit of some market or some product
2. Create a spreadsheet with the following columns: response ID, whether options were mutually exclusive:
   a. ResponseID
   b. Mutually Exclusive: Dummy; NA if only 1 option is reported

B.2 Training a fine-tuned GPT-3.5 Turbo model

We employed a fine-tuned GPT-3.5 Turbo model [7], trained on a dataset of human-coded options (2,269 options), to automate the coding of the options. The training data consisted of instances coded by human coders, which we used to fine-tune the model to replicate the human coding process. This fine-tuning process has allowed us to significantly speed up the coding process while maintaining high accuracy, reducing the time and labor required for this task.

For each options, we asked the model to code whether it is 1) Strategic; 2) Continue; 3) Change Pricing of Ads; 4) Exit.[8] A system message was provided at both the training and deployment stages to further guide the model's behavior. This ensured consistency in how the model interpreted and coded the options, aligning its outputs with the predefined rubric. Below is the system message we provided when fine-tuning and deploying the model, which was adapted from the coding rubric in Section B.1:

"Code each option according to the following rubric. Return only the coding results. (1) Strategic: Dummy, whether the option is strategic or not, 1 means strategic, 0 otherwise. Strategic options are long-term, high-level plans, not easy to reverse, involve significant resource allocation. Examples include launching new products, entering new markets (e.g., geographic), M&A, exit, a complete overhaul of the platform's user interface, strategic partnerships. Operational options are specific courses of action with specific short-term goals, focus on day-to-day activities and tasks. Examples include cutting costs, promotion/marketing, increasing sales team pay, training sessions. (2) Continue: Dummy, whether the option is 'continue as it is' or not, 1 means continue as it is, 0 means otherwise. Examples for 'continue as it is' include explicitly mentioning 'continue as it is,' increasing customer stickiness, changing pricing system — note as a special case as in this instance as it could be seen as a change, streamlining operations, improving efficiency, or increasing profit margins without fundamentally altering its value proposition, target market, or key activities. The broader business model or strategy remains largely unchanged. (e.g., keeping the focus on local businesses as the key target market would represent 'continuing as is,' while changing the target market from local businesses to national chains would not). (3) Change Pricing of Ads: Dummy, 1 means the option indicates a change in pricing, 0 means otherwise. Changing the pricing of the ads. Change in pricing would explicitly mention phrases like 'change of pricing structure,' 'change the pricing model,' etc. (4) Exit: Dummy, 1 means the option indicates exit, and 0 otherwise. Exit means selling the business or exiting the market."

[7] https://platform.openai.com/docs/guides/fine-tuning/

[8] Based model: gpt-3.5-turbo-0613; Fine-tuned on: Feb 20, 2024

Similarly, for each set of options generated by the participant, we fine-tuned the model to code whether it is mutually exclusive.[9] This was trained on a set of 340 participants coded by human coders. Below is the system message we provided when fine-tuning the model, which was adapted from the coding rubric in Section B.1:

"Code whether the set of options were mutually exclusive. Mutually exclusive means choosing one option would prevent the firm from choosing any of the other options. For the set of options to be mutually exclusive, all the options need to be mutually exclusive with each other, so any presence of a non-mutually exclusive option would make the set as a whole not mutually exclusive. Assume the company has limited resources, launching one strategic initiative would constraint the company from launching another."

After training, we evaluated the model's performance by coding a new set of options and participants (835 options, 138 participants) and comparing the results with human coders.

Table B.1: Consistency between fine-tuned GPT and human coders

	Agreement	**Cohen's Kappa**
Strategic	87.90%	0.7350
Continue	87.07%	0.7373
Change Pricing	97.72%	0.8792
Exit	99.04%	0.8413
Mutually Exclusive	97.10%	0.6520

These results indicate that the fine-tuned GPT-3.5 model can be a reliable tool this specific coding task, with performance metrics that are comparable to human coders in many cases.

[9] Based model: gpt-3.5-turbo-0125; Fine-tuned on: Apr 8, 2024.

C Interview questions

This appendix provides additional details on the interview questions.

The interviews were conducted in a semi-structured way with the below questions:

1. How did you develop your options?
2. Are there any particular strategies or methodologies you employed?
3. How confident are you in the options you developed? Why?
4. How challenging did you find the task? Can you pinpoint where the specific challenges arose?
5. Was the instruction helpful with the option crafting process? How?
6. Are there any tools or resources you wish you had access to during this process?
7. How did the group discussion go? How was the process like?

D Full sets of regressions

This appendix reports the full regression results in table form for the main analyses in the paper.

Table D.1 to Table D.4 report the full regression results across different regression specifications. Table D.5 to Table D.10 report the full regression results for different construction of measures, including the log form of count outcome variables (total number of options, number of strategic options, and number of continue options), raw count instead of a binary indicator for change pricing and exit, and the proportion of strategic options and continue options.

Table D.1: Main Regression Results on Number of Options, Number of Strategic Options, and Mutually Exclusive

Dependent Variables	Number of Options				Number of Strategic Options				Mutually Exclusive			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.0133	-0.0444	-0.0271	-0.00832	0.854***	0.814**	0.822***	0.863***	0.161***	0.163***	0.163***	0.170***
	(0.395)	(0.392)	(0.390)	(0.387)	(0.322)	(0.317)	(0.317)	(0.313)	(0.0414)	(0.0408)	(0.0415)	(0.0424)
Age				-0.0323				0.0284				0.0190
				(0.122)				(0.0998)				(0.0164)
Master's or higher				0.305				0.259				-0.00587
				(0.412)				(0.339)				(0.0497)
Years since Grad				0.166				0.185*				-0.00461
				(0.133)				(0.112)				(0.0177)
Years of Work Experience				-0.0687				-0.173				-0.0201
				(0.171)				(0.137)				(0.0177)
GMAT (Standardized)				0.0202				0.0174				0.00261
				(0.0361)				(0.0298)				(0.00398)
Constant	6.620***	5.544***	6.318***	6.424**	4.193***	3.376***	3.761***	2.622	0.105***	0.123*	0.168	-0.225
	(0.270)	(0.580)	(1.118)	(3.123)	(0.215)	(0.451)	(0.811)	(2.531)	(0.0235)	(0.0668)	(0.116)	(0.401)
Observations	340	340	340	340	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	-0.003	0.030	0.042	0.032	0.018	0.053	0.065	0.059	0.040	0.086	0.062	0.055
Control Group Mean	6.620	6.620	6.620	6.620	4.193	4.193	4.193	4.193	0.105	0.105	0.105	0.105

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports regression results regressing outcome variables (Number of Options, Number of Strategic Options, and Mutually Exclusive (binary indicator), definition in the main text). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.2: Main Regression Results on Number of Continue Options, Change Pricing, and Exit

Dependent Variables	Number of Continue Options				Change Pricing				Exit			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	-0.633***	-0.645***	-0.628***	-0.619***	-0.0829	-0.0895*	-0.0910*	-0.0818	0.120***	0.117**	0.114**	0.121***
	(0.209)	(0.206)	(0.205)	(0.206)	(0.0539)	(0.0541)	(0.0534)	(0.0543)	(0.0457)	(0.0457)	(0.0459)	(0.0462)
Age				-0.00581				0.0142				0.0154
				(0.0740)				(0.0183)				(0.0166)
Master's or higher				0.243				0.0319				-0.0623
				(0.226)				(0.0572)				(0.0508)
Years since Grad				0.0166				0.0221				0.0167
				(0.0644)				(0.0195)				(0.0161)
Years of Work Experience				0.0232				-0.0332				-0.0552***
				(0.0939)				(0.0247)				(0.0207)
GMAT (Standardized)				0.01000				0.00319				0.00271
				(0.0190)				(0.00488)				(0.00408)
Constant	2.982***	2.450***	2.703***	2.535	0.491***	0.417***	0.343**	-0.0315	0.175***	0.0962	0.0331	-0.161
	(0.157)	(0.287)	(0.465)	(1.842)	(0.0383)	(0.0850)	(0.140)	(0.467)	(0.0292)	(0.0636)	(0.0720)	(0.399)
Observations	340	340	340	340	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.024	0.082	0.098	0.089	0.004	0.004	0.055	0.049	0.017	0.030	0.020	0.030
Control Group Mean	2.982	2.982	2.982	2.982	0.491	0.491	0.491	0.491	0.175	0.175	0.175	0.175

Robust standard errors in parentheses:
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Notes: This table reports regression results regressing outcome variables (Number of Continue Options, Change Pricing (binary indicator), and Exit (binary indicator), definition in the main text). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.3: Main Regression Results on Best Option is Strategic and Best Option is Continue

Dependent Variables	Best Option is Strategic				Best Option is Continue			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.162***	0.167***	0.165***	0.169***	-0.103**	-0.101**	-0.0969**	-0.0977*
	(0.0439)	(0.0439)	(0.0438)	(0.0439)	(0.0486)	(0.0490)	(0.0490)	(0.0497)
Age				0.00499				-0.00783
				(0.0154)				(0.0180)
Master's or higher				-0.0197				0.0409
				(0.0541)				(0.0578)
Years since Grad				0.0231				-0.0143
				(0.0227)				(0.0215)
Years of Work Experience				-0.0198				0.00955
				(0.0246)				(0.0244)
GMAT (Standardized)				0.00267				0.00319
				(0.00434)				(0.00421)
Constant	0.702***	0.769***	0.860***	0.679*	0.333***	0.299***	0.268**	0.536
	(0.0351)	(0.0673)	(0.0785)	(0.368)	(0.0362)	(0.0756)	(0.120)	(0.436)
Observations	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.036	0.031	0.046	0.040	0.010	-0.001	0.001	-0.007
Control Group Mean	0.702	0.702	0.702	0.702	0.333	0.333	0.333	0.333

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports regression results regressing outcome variables (Best Option is Strategic, Best Option is Continue, both are binary indicators). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.4: Main Regression Results on Best Option is Change Pricing, Best Option is Exit, and Average Similarity

Dependent Variables	Best Option is Change Pricing				Best Option is Exit				Average Similarity			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	-0.0461	-0.0470	-0.0487	-0.0451	0.0297*	0.0285*	0.0281*	0.0278*	0.0172***	0.0173***	0.0171***	0.0167***
	(0.0298)	(0.0303)	(0.0300)	(0.0295)	(0.0154)	(0.0152)	(0.0152)	(0.0144)	(0.00440)	(0.00442)	(0.00455)	(0.00464)
Age				0.00920				0.000711				-0.00192
				(0.0111)				(0.00467)				(0.00154)
Master's or higher				-0.0211				-0.0128				-0.00197
				(0.0326)				(0.0134)				(0.00490)
Years since Grad				0.00597				-0.000588				0.00259*
				(0.0101)				(0.00255)				(0.00137)
Years of Work Experience				-0.0113				-0.000353				-0.00200
				(0.0139)				(0.00435)				(0.00196)
GMAT (Standardized)				0.00222				-0.000292				-0.000134
				(0.00236)				(0.00156)				(0.000393)
Constant	0.105***	0.110**	0.0187	-0.215	0.00585	0.0194	-0.0108	-0.0216	0.317***	0.316***	0.319***	0.371***
	(0.0235)	(0.0445)	(0.0134)	(0.255)	(0.00585)	(0.0235)	(0.00700)	(0.114)	(0.00289)	(0.00671)	(0.0114)	(0.0389)
Observations	340	340	340	340	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.004	-0.004	0.065	0.057	0.008	0.016	-0.009	-0.023	0.040	0.059	0.043	0.045
Control Group Mean	0.105	0.105	0.105	0.105	0.006	0.006	0.006	0.006	0.317	0.317	0.317	0.317

Robust standard errors in parentheses:
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table reports regression results regressing outcome variables (Best Option is Change Pricing, Best Option is Exit, both are binary indicators, and Average similarity, definition in the main text). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.5: Main Result – Number of Options

Dependent Variables	Number of Options				Log(Number of Options)			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.0133	-0.0444	-0.0271	-0.00832	-0.00970	-0.0226	-0.0188	-0.0183
	(0.395)	(0.392)	(0.390)	(0.387)	(0.0569)	(0.0561)	(0.0563)	(0.0564)
Age				-0.0323				-0.00813
				(0.122)				(0.0181)
Master's or higher				0.305				0.0533
				(0.412)				(0.0592)
Years since Grad				0.166				0.0160
				(0.133)				(0.0185)
Years of Work Experience				-0.0687				-0.00618
				(0.171)				(0.0247)
GMAT (Standardized)				0.0202				-0.000122
				(0.0361)				(0.00478)
Constant	6.620***	5.544***	6.318***	6.424**	1.760***	1.573***	1.653***	1.799***
	(0.270)	(0.580)	(1.118)	(3.123)	(0.0393)	(0.0935)	(0.180)	(0.467)
Observations	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	-0.003	0.030	0.042	0.032	-0.003	0.041	0.049	0.036
Control Group Mean	6.620	6.620	6.620	6.620	1.760	1.760	1.760	1.760

Robust standard errors in parentheses:
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Notes: This table reports regression results regressing outcome variables (Number of Options, Log of Number of Options). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.6: Main Result – Number of Strategic Options

Dependent Variables	Number of Strategic Options				Log(Number of Strategic Options)			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.854***	0.814**	0.822***	0.863***	0.186***	0.176***	0.178***	0.191***
	(0.322)	(0.317)	(0.317)	(0.313)	(0.0662)	(0.0651)	(0.0657)	(0.0659)
Age				0.0284				0.0180
				(0.0998)				(0.0212)
Master's or higher				0.259				0.0896
				(0.339)				(0.0708)
Years since Grad				0.185*				0.0295
				(0.112)				(0.0223)
Years of Work Experience				-0.173				-0.0438
				(0.137)				(0.0284)
GMAT (Standardized)				0.0174				0.00542
				(0.0298)				(0.00627)
Constant	4.193***	3.376***	3.761***	2.622	1.268***	1.062***	1.053***	0.565
	(0.215)	(0.451)	(0.811)	(2.531)	(0.0488)	(0.114)	(0.217)	(0.560)
Observations	340	340	340	340	336	336	336	336
Section FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.018	0.053	0.065	0.059	0.020	0.056	0.063	0.063
Control Group Mean	4.193	4.193	4.193	4.193	1.268	1.268	1.268	1.268

Robust standard errors in parentheses:
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table reports regression results regressing outcome variables (Number of Strategic Options, Log of Number of Strategic Options). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.7: Main Result – Number of Continue Options

Dependent Variables	Number of Continue Options				Log(Number of Continue Options)			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	-0.633***	-0.645***	-0.628***	-0.619***	-0.219***	-0.229***	-0.224***	-0.218***
	(0.209)	(0.206)	(0.205)	(0.206)	(0.0718)	(0.0710)	(0.0711)	(0.0720)
Age				-0.00581				-0.000501
				(0.0740)				(0.0251)
Master's or higher				0.243				0.111
				(0.226)				(0.0792)
Years since Grad				0.0166				0.00978
				(0.0644)				(0.0252)
Years of Work Experience				0.0232				-0.0108
				(0.0939)				(0.0320)
GMAT (Standardized)				0.01000				0.00195
				(0.0190)				(0.00666)
Constant	2.982***	2.450***	2.703***	2.535	0.970***	0.864***	1.002***	0.977
	(0.157)	(0.287)	(0.465)	(1.842)	(0.0500)	(0.105)	(0.173)	(0.617)
Observations	340	340	340	340	313	313	313	313
Section FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.024	0.082	0.098	0.089	0.026	0.087	0.109	0.101
Control Group Mean	2.982	2.982	2.982	2.982	0.970	0.970	0.970	0.970

Robust standard errors in parentheses:
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Notes: This table reports regression results regressing outcome variables (Number of Continue Options, Log of Number of Continue Options). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.8: Main Result – Change Pricing

Dependent Variables	Change Pricing (Binary Indicator)				Change Pricing (Raw Count)			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	-0.0829	-0.0895*	-0.0910*	-0.0818	-0.106	-0.116*	-0.126*	-0.116*
	(0.0539)	(0.0541)	(0.0534)	(0.0543)	(0.0677)	(0.0688)	(0.0684)	(0.0695)
Age				0.0142				0.0160
				(0.0183)				(0.0206)
Master's or higher				0.0319				0.0328
				(0.0572)				(0.0660)
Years since Grad				0.0221				0.00977
				(0.0195)				(0.0287)
Years of Work Experience				-0.0332				-0.0245
				(0.0247)				(0.0343)
GMAT (Standardized)				0.00319				0.00723
				(0.00488)				(0.00518)
Constant	0.491***	0.417***	0.343**	-0.0315	0.567***	0.562***	0.433**	0.0449
	(0.0383)	(0.0850)	(0.140)	(0.467)	(0.0498)	(0.115)	(0.188)	(0.534)
Observations	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.004	0.004	0.055	0.049	0.004	0.007	0.056	0.048
Control Group Mean	0.491	0.491	0.491	0.491	0.567	0.567	0.567	0.567

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports regression results regressing outcome variables (Change Pricing, as binary indicator, and as raw count). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.9: Main Result – Exit

Dependent Variables	Exit (Binary Indicator)				Exit (Raw Count)			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.120***	0.117**	0.114**	0.121***	0.174***	0.169***	0.165***	0.172***
	(0.0457)	(0.0457)	(0.0459)	(0.0462)	(0.0570)	(0.0573)	(0.0575)	(0.0571)
Age				0.0154				0.0162
				(0.0166)				(0.0189)
Master's or higher				-0.0623				-0.0891
				(0.0508)				(0.0597)
Years since Grad				0.0167				0.0137
				(0.0161)				(0.0212)
Years of Work Experience				-0.0552***				-0.0599**
				(0.0207)				(0.0234)
GMAT (Standardized)				0.00271				0.00219
				(0.00408)				(0.00563)
Constant	0.175***	0.0962	0.0331	-0.161	0.187***	0.123	0.0135	-0.142
	(0.0292)	(0.0636)	(0.0720)	(0.399)	(0.0321)	(0.0817)	(0.0724)	(0.455)
Observations	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.017	0.030	0.020	0.030	0.024	0.031	0.021	0.029
Control Group Mean	0.175	0.175	0.175	0.175	0.187	0.187	0.187	0.187

Robust standard errors in parentheses:
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: This table reports regression results regressing outcome variables (Number of Continue Options, Log of Number of Continue Options). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

Table D.10: Main Result – Proportion of Strategic Options and Continue Options

Dependent Variables	Proportion of Strategic Options				Proportion of Continue Options			
Models	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Treatment	0.145***	0.144***	0.145***	0.151***	-0.100***	-0.0986***	-0.0974***	-0.100***
	(0.0219)	(0.0218)	(0.0222)	(0.0226)	(0.0236)	(0.0233)	(0.0235)	(0.0240)
Age				0.0139*				-0.00856
				(0.00747)				(0.00907)
Master's or higher				0.0120				0.0171
				(0.0257)				(0.0269)
Years since Grad				0.00488				-0.00371
				(0.00754)				(0.00775)
Years of Work Experience				-0.0160*				0.0109
				(0.00948)				(0.00988)
GMAT (Standardized)				0.00179				-0.000318
				(0.00196)				(0.00212)
Constant	0.621***	0.615***	0.597***	0.252	0.450***	0.452***	0.444***	0.649***
	(0.0173)	(0.0383)	(0.0682)	(0.197)	(0.0173)	(0.0409)	(0.0711)	(0.227)
Observations	340	340	340	340	340	340	340	340
Section FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting FE	NO	YES	NO	NO	NO	YES	NO	NO
Gender-Consulting-Section FE	NO	NO	YES	YES	NO	NO	YES	YES
Adjusted R-squared	0.112	0.122	0.123	0.125	0.048	0.077	0.081	0.072
Control Group Mean	0.621	0.621	0.621	0.621	0.450	0.450	0.450	0.450

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports regression results regressing outcome variables (Proportion of Strategic Options, defined as the number of strategic options over the total number of options, and Continue Options, defined as the number of continue options over the total number of options). Model 1 examined the effect of treatment on the outcome variables. Model 2 added the section fixed effects and separate fixed effects for gender-consulting strata. Model 3 added a combined fixed effect for the gender-consulting-section strata. Model 4 adds controls for age, whether the participant has a master's degree or above, number of years since graduation, number of years of working experience, and standardized GMAT score on top of Model 3. Robust standard errors are in parentheses.

E Text analysis

This appendix provides additional details on the transformer-based text analyses. The preprocessing of text data is crucial for enhancing the performance of natural language processing (NLP) models. Our preprocessing pipeline involved the following steps. First, all words are converted to lowercase. Then, special characters were removed. Subsequently, the text was split into individual words, or tokens, a process known as tokenization. Lemmatization, the transformation of words into their base or dictionary form, was then applied to consolidate various forms of the same word. The final step involved reconstructing the tokens into sentences, preserving the flow of ideas for further analysis.

For the task of topic modeling, we used the BERTopic package in Python for topic modeling. BERT (Bidirectional Encoder Representations from Transformers) topic modeling represents a state-of-the-art technique in natural language processing (NLP) that leverages deep learning to understand and categorize textual data. This method extends traditional topic modeling approaches by employing a transformer-based architecture, specifically designed to capture the contextual relationships between words in a text. We followed the default choices suggested in the documentation of the BERTopic package to build the model.

We used the sentence transformer model ”all-mpnet-base-v2” for embedding. This model maps sentences to a 768-dimensional vector space that captures the semantic information, and it is usually used for clustering or sentence similarity tasks (Hugging Face, 2022). It is one of the best-performing pre-trained models based on evaluations for the quality of embedded sentences and embedded search queries and paragraphs.

The UMAP (Uniform Manifold Approximation and Projection) algorithm was used for dimensionality reduction. This algorithm was selected for its ability to preserve both the global and local structure of the data when reducing dimensionality. The clustering of topics was then performed using the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) algorithm. The choice of HDBSCAN was motivated by its effectiveness in identifying clusters of varying shapes, a characteristic particularly beneficial for grouping topics in text data, where topic prevalence and cohesion can vary widely.

To interpret the clusters with representing topics, CountVectorizer and class-based Term Frequency-Inverse Document Frequency (c-TF-IDF) calculations were used. CountVectorizer counts how often each word appears in each cluster to determine the frequency of the words. c-TF-IDF calculates the importance scores for words within a cluster and extracts the most important words per cluster as the topic of that cluster. When using HDBSCAN, outliers that do not fall within any of the created topics might be created. Following the documentation , we employed a two-step strategy to reduce the outliers. First, the c-TF-IDF strategy was applied with a specific threshold. This was followed by a probabilities strategy to find the best matching topic for each outlier.

Table E.1 presents the results from BERTopic, with the counts and the representative keywords generated by the algorithm. Table E.2 presents the keywords used to categorize the responses to the survey question "Describe how you developed your alternatives."

Table E.1: BERTopic Modeling Results

Custom Name	**Count**	**Representative Keywords**
review restaurant	621	review, restaurant, reviewer, algorithm, rating
expand	448	expand, asia, africa, existing, europe
advertising	277	ad, advertising, pricing, change, click
user data	154	user, data, subscription, charge, provide
reservation booking	148	reservation, booking, add, expand, charge
food delivery	137	food, delivery, food delivery, chowhub, expand
exit	132	competitor, sell, exit, flex, acquire
national chain	122	chain, national, national chain, big, focus
cut operational	108	cut, operational, reduce, improvement, operational improvement
sales team	106	team, sale team, incentive, salesperson, change

Notes: This table shows the topic modeling results generated using the BERTopic package. The count represents the number of options that fall into the topic. The representation is generated by the algorithm.

Table E.2: Keywords Used to Categorize How Participants Developed Their Alternatives

Category	**Count**	**Keywords**
Problem-Solving	46	solve, solution, address, tackle, pain point, challenge, gap, problem, tradeoff, improv, optimis, reflection, case
Stakeholder	35	stakeholder, c levels, board, shareholder, customer, consumer, businesses, reviewer, sides of the platform, partner, ecosystem
External	34	market, industry, competitor, benchmark, trends, segment, demand, opportunities, external, positioning
Revenue-Cost	33	revenue, cost, profitability, profits, monetise, $, business model
Frameworks	28	framework, structure, blue ocean, swot, matrix, graph, map, issue tree
Expansion	16	expand, expansion, growing, growth, scale, organic
Analytical	15	data, pros and cons, analy
Internal	13	financials, strengths, internal, resource
Intuition	8	experience, intuition, whatever came to mind, intuitive, common sense, my own

Notes: This table shows the self-defined keywords used to categorize the responses to the survey question "Describe how you developed your alternatives." The count represents the number of participants whose responses contained the keyword. A response can cover multiple categories.

F Pre-registration differences

The experiments were pre-registered with the AEA RCT Registry. The key differences between the paper and the pre-registration are:

- A total sample size of 356 was pre-registered for the main experiments based on enrollment numbers. The final sample was reduced to 340 due to student absences and responses that lacked sufficient engagement with the task. Specifically, we excluded participants who generated only one option, as this suggested a lack of effortful engagement. Results remain robust when these participants are included in the analyses.
- A total sample size of 247 was pre-registered for follow-up experiments based on enrollment numbers. The final sample was reduced to 227 due to student absences and responses that lacked sufficient engagement with the task. Specifically, we excluded participants who generated only one option, as this suggested a lack of effortful engagement. Results remain directionally consistent when these participants are included, although the estimates are noisier.
- The pre-analysis plan describes secondary outcomes, some of which we noted at the time as subject to feasibility. We were indeed not able to obtain peer evaluations of option quality and a binary variable on how detailed the option is.
- The pre-analysis plan describes that alternatives generated by participants would be coded by two independent coders. In the follow-up experiments, a sub-sample of 462 alternatives generated by 89 participants was coded using a fine-tuned GPT-3.5 Turbo model trained on previously human-coded data due to budget constraints.

G Follow-up experiments

This appendix provides additional details on the follow-up experiments. Table G.1 shows the timeline of experimental interventions. Table G.3 shows the impact of frameworks on the alternatives participants considered. Table G.5 shows the impact of frameworks on the best alternative chosen.

Responses to the question "How did you approach coming up with strategic options for Rated? Please describe your thought process" were coded into different categories according to the keywords listed in Table G.6, and Figure G.1 shows the results.

Responses to the question "In your view, what is the key problem that Rated faces?" were coded into different categories according to the keywords listed in Table G.7, and Figure G.2 shows the results.

Responses to the question "Which specific data from the case did you rely on the most to complete this exercise?" were coded into different categories according to the keywords listed in Table G.8, and Figure G.3 shows the results. This question was only shown to a sub-sample of 89 participants in follow-up experiment 1, and to all the participants in follow-up experiment 2.

Table G.1: Timeline of experimental interventions

	Date of Experiment	Number of Sections	Number of Participants
Follow-up Experiment 1			
1	8 Feb 2024	2	138
2	26 Mar 2024	1	45
3	29 Mar 2024	1	44
	Sub-total	4	227
Follow-up Experiment 2			
4	25 Sep 2024	1	54
5	27 Sep 2024	1	62
6	1 Oct 2024	2	137
	Sub-total	4	253

Notes: This table shows the timeline and the corresponding details for each follow-up experiments. Follow-up Experiment 1 was pre-registered with AEA RCT Registry. Follow-up Experiment 2 was not pre-registered separately but follows the same pre-registration protocol as Follow-up Experiment 1, with the only difference being the treatment figure provided to participants. All other aspects of the experimental design remained identical to the pre-registered protocol.

Table G.2: The impact of frameworks on the alternatives generated

VARIABLES	(1) Number of Options	(2) Number of Strategic Options	(3) Mutually Exclusive	(4) Number of Continue Options	(5) Change Pricing	(6) Exit	(7) Proportion of Strategic Options	(8) Proportion of Continue Options
Treatment	-0.451*	-0.108	0.003	-0.313**	-0.063	-0.047*	0.051***	-0.028
	(0.240)	(0.200)	(0.018)	(0.151)	(0.045)	(0.025)	(0.020)	(0.020)
Constant	6.149***	4.269***	0.030	2.709***	0.623***	0.209***	0.691***	0.454***
	(0.361)	(0.318)	(0.035)	(0.254)	(0.092)	(0.072)	(0.034)	(0.042)
Observations	480	480	480	480	480	480	480	480
Gender-Consulting-Section FE	YES	YES	YES	YES	YES	YES	YES	YES
Control Variables	NO	NO	NO	NO	NO	NO	NO	NO
Adjusted R-squared	0.09	0.15	-0.04	0.00	0.06	0.10	0.09	0.05
Control Group Mean	6.17	3.89	0.04	2.82	0.51	0.12	0.62	0.47

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports the OLS linear regression results of the treatment binary indicator on the dependent variables. The results included a combined fixed effect for the gender-consulting-section strata. Robust standard errors are in parentheses.

Table G.3: The impact of frameworks on the alternatives generated

(a) Panel A: Follow-up 1

VARIABLES	(1) Number of Options	(2) Number of Strategic Options	(3) Mutually Exclusive	(4) Number of Continue Options	(5) Change Pricing	(6) Exit	(7) Proportion of Strategic Options	(8) Proportion of Continue Options
Treatment	0.145	0.208	-0.041	-0.005	-0.042	-0.073*	0.022	-0.008
	(0.258)	(0.214)	(0.025)	(0.195)	(0.065)	(0.040)	(0.028)	(0.029)
Constant	5.870***	4.121***	0.050	2.565***	0.613***	0.222***	0.705***	0.444***
	(0.370)	(0.319)	(0.037)	(0.262)	(0.094)	(0.074)	(0.036)	(0.042)
Observations	227	227	227	227	227	227	227	227
Gender-Consulting-Section FE	YES	YES	YES	YES	YES	YES	YES	YES
Control Variables	NO	NO	NO	NO	NO	NO	NO	NO
Adjusted R-squared	0.04	0.13	-0.00	-0.02	0.05	0.11	0.07	0.01
Control Group Mean	5.64	3.44	0.06	2.74	0.55	0.15	0.62	0.48

(b) Panel B: Follow-up 2

Treatment	-0.969**	-0.389	0.042*	-0.572**	-0.084	-0.030	0.077***	-0.043
	(0.386)	(0.325)	(0.024)	(0.224)	(0.061)	(0.030)	(0.028)	(0.028)
Constant	5.861***	4.011***	0.064	2.595***	0.455***	0.222**	0.691***	0.446***
	(0.496)	(0.351)	(0.059)	(0.319)	(0.108)	(0.089)	(0.045)	(0.051)
Observations	253	253	253	253	253	253	253	253
Gender-Consulting-Section FE	YES	YES	YES	YES	YES	YES	YES	YES
Control Variables	NO	NO	NO	NO	NO	NO	NO	NO
Adjusted R-squared	0.12	0.15	-0.03	0.02	0.05	0.03	0.11	0.08
Control Group Mean	6.65	4.30	0.02	2.89	0.47	0.09	0.62	0.45

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the OLS linear regression results of the treatment binary indicator on the dependent variables. The results included a combined fixed effect for the gender-consulting-section strata. Robust standard errors are in parentheses.

Table G.4: The impact of frameworks on best alternative chosen

VARIABLES	(1) Best Option is Strategic	(2) Best Option is Continue	(3) Best Option is Change Pricing	(4) Best Option is Exit
Treatment	0.076*	-0.002	-0.018	-0.002
	(0.043)	(0.045)	(0.033)	(0.011)
Constant	0.714***	0.376***	0.165**	0.032
	(0.082)	(0.091)	(0.069)	(0.034)
Observations	480	480	480	480
Gender-Consulting-Section FE	YES	YES	YES	YES
Control Variables	NO	NO	NO	NO
Adjusted R-squared	0.03	0.01	0.01	0.03
Control Group Mean	0.632	0.401	0.162	0.016

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports the OLS linear regression results of the treatment binary indicator on the dependent variables. The results included a combined fixed effect for the gender-consulting-section strata. Robust standard errors are in parentheses.

Table G.5: The impact of frameworks on best alternative chosen

(a) Panel A: Follow-up 1

VARIABLES	(1) Best Option is Strategic	(2) Best Option is Continue	(3) Best Option is Change Pricing	(4) Best Option is Exit
Treatment	0.078	0.012	0.013	-0.014
	(0.063)	(0.066)	(0.048)	(0.017)
Constant	0.714***	0.369***	0.150**	0.038
	(0.084)	(0.092)	(0.071)	(0.036)
Observations	227	227	227	227
Gender-Consulting-Section FE	YES	YES	YES	YES
Control Variables	NO	NO	NO	NO
Adjusted R-squared	0.02	-0.01	0.01	0.06
Control Group Mean	0.619	0.415	0.153	0.025

(b) Panel B: Follow-up 2

VARIABLES	(1) Best Option is Strategic	(2) Best Option is Continue	(3) Best Option is Change Pricing	(4) Best Option is Exit
Treatment	0.069	-0.008	-0.042	0.010
	(0.059)	(0.062)	(0.046)	(0.014)
Constant	0.844***	0.254***	0.019	0.079
	(0.076)	(0.095)	(0.021)	(0.058)
Observations	253	253	253	253
Gender-Consulting-Section FE	YES	YES	YES	YES
Control Variables	NO	NO	NO	NO
Adjusted R-squared	0.04	0.02	0.02	0.01
Control Group Mean	0.643	0.388	0.171	0.008

Robust standard errors in parentheses:
*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Notes: This table reports the OLS linear regression results of the treatment binary indicator on the dependent variables. The results included a combined fixed effect for the gender-consulting-section strata. Robust standard errors are in parentheses.

Table G.6: Keywords Used to Categorize Thought Process

Category	**Keywords**
Problem-Solving	solve, solution, address, tackle, pain point, challenge, gap, problem, tradeoff, improv, optimis, reflection, blind spot, case
Revenue-Cost	revenue, cost, profitability, profits, monetise, $, loss, business model
External	market, industry, competitor, benchmark, trends, segment, demand, opportunities, external, positioning, substitute
Stakeholder	stakeholder, c levels, board, shareholder, customer, consumer, businesses, reviewer, sides of the platform, partner, ecosystem, head of
Frameworks	framework, structure, blue ocean, swot, matrix, graph, map, issue tree, five forces
Expansion	expand, expansion, growing, growth, scale, organic
Analytical	data, pros and cons, analy, exhibit
Internal	financials, strengths, internal, resource, advantage, core
Intuition	experience, intuition, whatever came to mind, intuitive, common sense, my own, natural

Notes: This table shows the self-defined keywords used to categorize the responses to the survey question "How did you approach coming up with strategic options for Rated? Please describe your thought process." A response can cover multiple categories.

Table G.7: Keywords Used to Categorize Key Problem

Category	**Keywords**
Monetization	monetization, monetisation, monetizing, monetize, profit, profitability, margin, revenue, pricing, financial, value capture, make money
Product	product, ads, advertising, advertise, adtech, feature, technology, algorithm, review
User	user, engagement, customer, client, restaurant, retention, conversion
Cost	cost, expense, charge, spending, cashflow, cash flow, unit economics, expanding, bottom line
Positioning	positioning, position, market, segment, target, targeting, differentiation
Business Model	business model, model, strategy, plan
Competition	competition, competitor, competitive, compete
Growth	growth, expansion, scaling, new ideas, acquiring, changes, thinking big

Notes: This table shows the self-defined keywords used to categorize the responses to the survey question "In your view, what is the key problem that Rated faces?" A response can cover multiple categories.

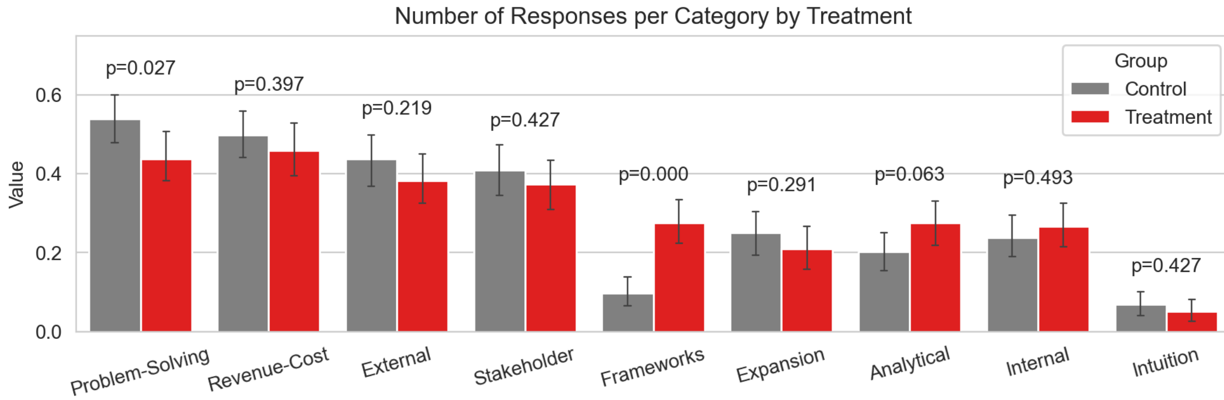
Table G.8: Keywords Used to Categorize Focus on Data

Category	**Keywords**
Financial	financial, 1, exhibit 01, revenue, expense, profitability, statement, p&l, balance sheet, consolidated, cost, exhibit one
Competitor	competitor, 3, competition, comparison
Stakeholder	customer, traffic, interview, 4, restaurant, visitor
Ad Market	advertising, 2, advertiser, ad product
Leadership	leadership, executive, ceo, head, execs, michael woods, c-suite, chief

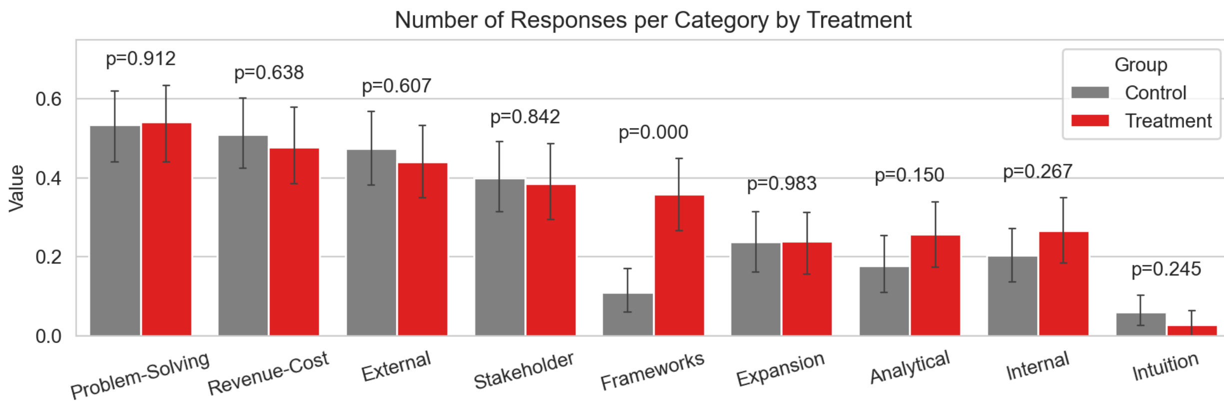
Notes: This table shows the self-defined keywords used to categorize the responses to the survey question "Which specific data from the case did you rely on the most to complete this exercise?" Numbers correspond to the exhibits shown in the case. A response can cover multiple categories.

Figure G.1: Number of Responses per Category by Treatment – Thought Process

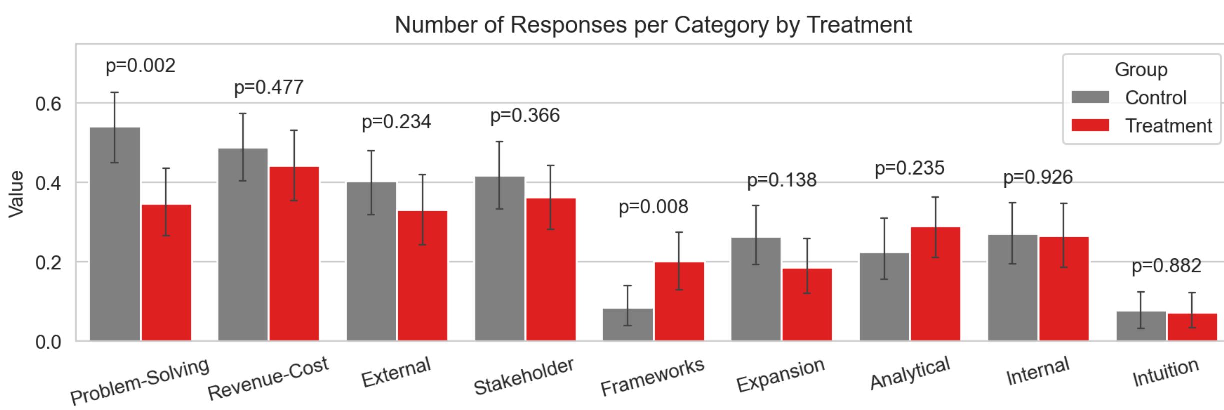
(a) Combined Results



(b) Follow-Up Experiment 1



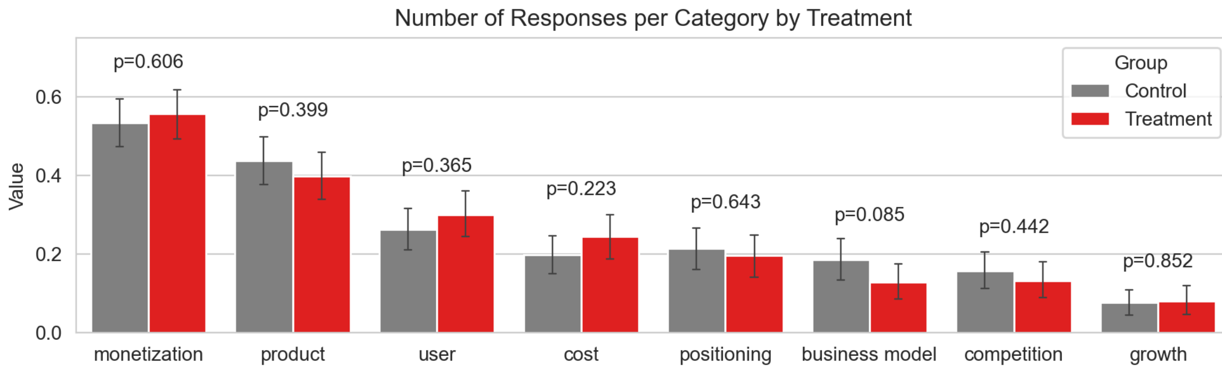
(c) Follow-Up Experiment 2



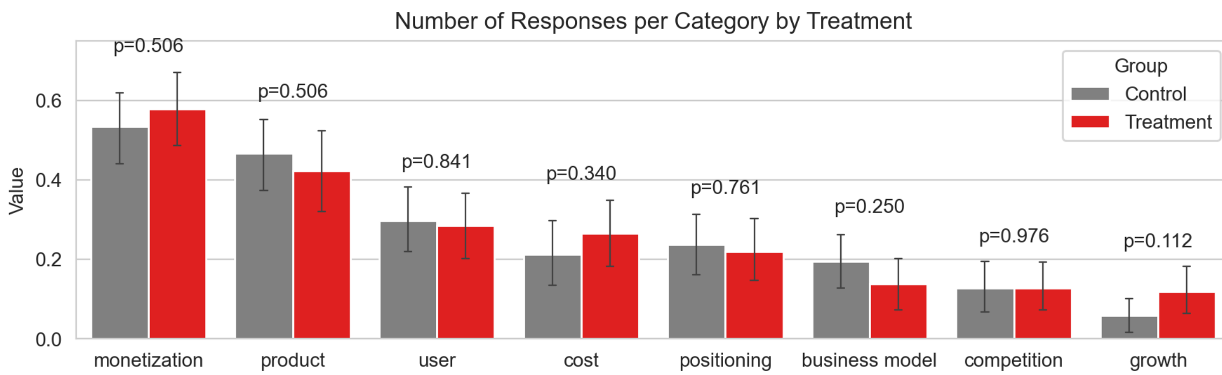
Notes: These figures plot the categorization of responses to the survey question, "How did you approach coming up with strategic options for Rated? Please describe your thought process." The x-axis shows the different categories. The y-axis shows the likelihood that a participant covered the category. The p-value indicates the t-test results comparing the average likelihood between the control and the treatment groups. A response can cover multiple categories.

Figure G.2: Number of Responses per Category by Treatment – Problem Formulation

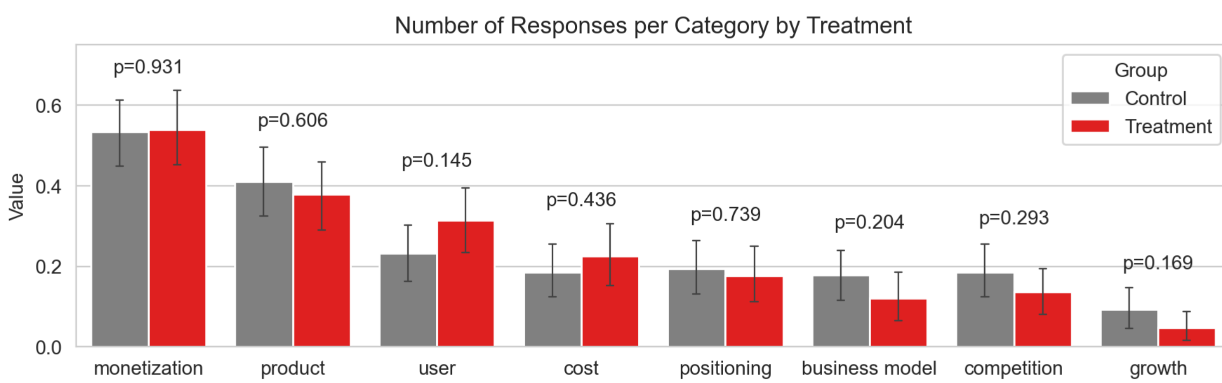
(a) Combined Results



(b) Follow-Up Experiment 1



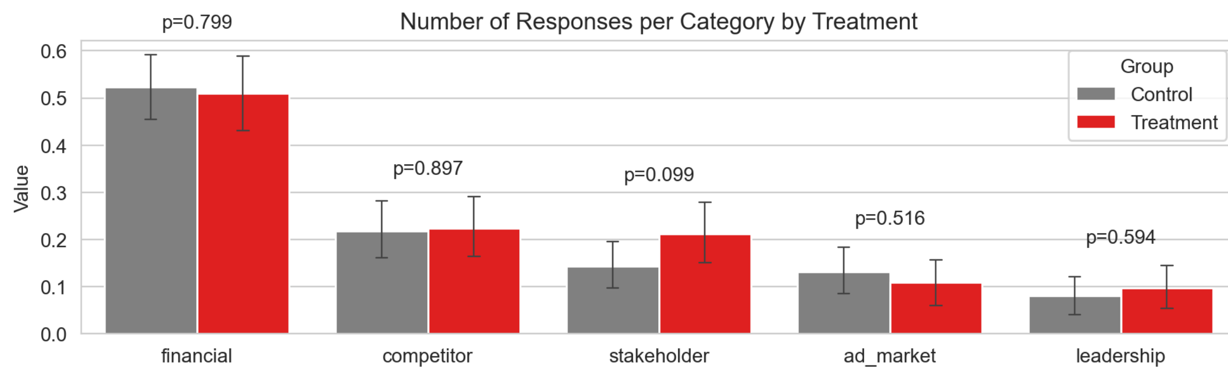
(c) Follow-Up Experiment 2



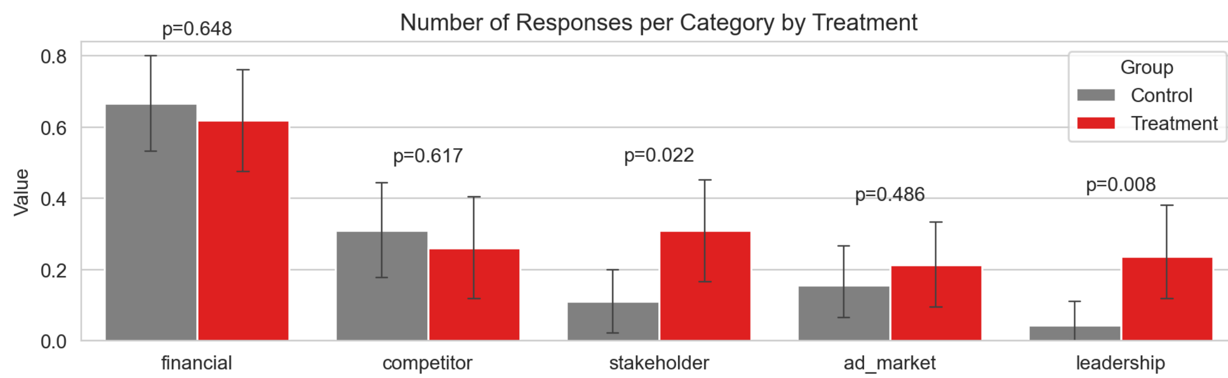
Notes: These figures plot the categorization of responses to the survey question, "In your view, what is the key problem that Rated faces?" The x-axis shows the different categories. The y-axis shows the likelihood that a participant covered the category. The p-value indicates the t-test results comparing the average likelihood between the control and the treatment groups. A response can cover multiple categories.

Figure G.3: Number of Responses per Category by Treatment – Data Attended

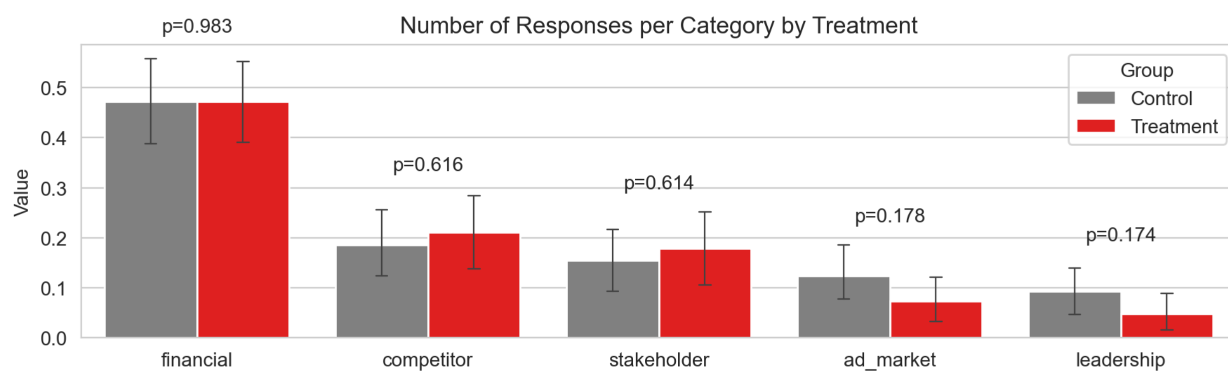
(a) Combined Results



(b) Follow-Up Experiment 1



(c) Follow-Up Experiment 2



Notes: These figures plot the categorization of responses to the survey question, "Which specific data from the case did you rely on the most to complete this exercise?" The x-axis shows the different categories. The y-axis shows the likelihood that a participant covered the category. The p-value indicates the t-test results comparing the average likelihood between the control and the treatment groups. A response can cover multiple categories.

H LLM-based simulation

This section provides more details on large language models (LLMs) based simulation. Traditionally, decision-makers have relied on human mental representations to process information. With the development of advanced algorithms like large language models (LLMs), there is growing interest in understanding how these algorithms could augment how decision-makers process information, represent problems, and make strategic decisions Csaszar et al. (2024). Furthermore, LLMs serve as a channel for distributing and facilitating access to certain external representations. They aid in visualizing and processing information, much like traditional external representations, but with added dynamic and interactive capabilities, driven by more complex representations, to provide tailored insights and clarifications on specific aspects of a scenario or decision.

Building on Csaszar et al. (2024), we use LLMs as simulated agents to explore the extent to which large language models (LLMs) can resemble the human decision-making process, as well as how they might be influenced by specific frameworks. We used GPT-4 to simulate the main experiment and the follow-up experiment in-silico, generating a comparable sample size of virtual participants, which we randomly assigned to consider the same problem as our human participants, either with or without the framework. Notably, the framework used in the main experiment was developed recently and is not in the public domain, ensuring it is not part of the LLMs' training data.

We used GPT-4 to simulate virtual agents for the same experimental task: generating strategic options for the review platform Rated. We used the "gpt-4-vision-preview" model, which can analyze the framework from the figure that was provided to the model. The "temperature" parameter was set to 1 to introduce variability into the responses. To provide more context for the task, the main text of the case was used as a system prompt. We asked the agents to act as MBA students from a leading business school, without specifying additional demographic details. We ran two sets of simulations: the control group was given only the general prompts (see Table H.1 for more details); the treatment group received the strategic framework(s) in addition to the prompts. For the simulation of the main experiment, we provided Figure 2 and collected 2,240 alternatives from 333 full responses. For the simulation of the follow-up experiment, we provided Figure 3b and collected 3,595 alternatives from 596 full responses. All options were coded using a fine-tuned GPT-3.5 Turbo model trained on previously human-coded data.

Our results show that the strategic framework prompted LLM agents to generate a larger number of strategic options, consistent with findings from our main experiment. On average, prompting LLM agents with the framework increased the number of strategic options generated by 1 ($p < 0.001$), translating into a 37% (=1.02/2.74) increase (see Table H.2 for more details). This result is consistent with evidence from marketing and political science suggesting that LLMs can be used to simulate human responses to surveys and stimuli in ways consistent with economic theory and well-documented consumer behavior patterns Argyle et al., 2023; Brand et al., 2023; Horton, 2023; Li et al., 2024, and provide additional evidence that LLM agents display strategic decision-making behavior that is similar to MBA and executive education students enrolled in a leading business school. However, it is worth noting that prompting LLM agents with the framework also increased the total number of options generated by 0.9 ($p < 0.001$), translating into a 15% (=0.92/6.27) increase, highlighting how frameworks may also be a complement for ideation when using LLMs.

In addition, LLMs demonstrated the ability to process and integrate multiple frameworks to simultaneously to uncover additional strategic options. In the simulation using a set of frameworks, prompting LLM agents with the framework increased the number of strategic options generated by 0.12 ($p = 0.04$), translating into a 4% (=0.12/2.7) increase (see Table H.2 for more details), although the magnitude is smaller than providing a single framework.

These findings suggest that LLM agents may be able to aid in strategic decision-making by generating more strategic alternatives. Moreover, they raise the possibility that LLMs can help with the challenge of fixed mental models by considering and integrating different representations and new information. Furthermore, since frameworks shift problem formulation, LLMs have the potential to influence not just the ideation stage, but also the problem formulation stage in strategic decision-making.

Table H.1: Prompt used in GPT-based simulations

#	**Prompt**
1	Identify the primary challenge that Rated is currently facing and outline the contributing factors. First, conduct a comprehensive review of Rated's position in the industry. Then, analyze the main difficulty confronting Rated and the various elements or circumstances that are adding to this challenge. Your response should be clear and concise. Please think through the steps needed to answer the question, but only tell me the primary challenge and the contributing factors.
2	Describe your approach to identifying the primary challenge for Rated. First, recall the process and the factors you considered when identifying the primary challenge and the contributing factors. Then, describe how you synthesized the information to identify the primary challenge and the contributing factors. Your response should provide a clear and logical explanation of the steps you took. Please think through the steps needed to answer the question, but only tell me your approach to identifying the primary challenge.
3	(Treatment) Identify and outline a comprehensive list of strategic options available for Rated **based on the framework provided in the image**. Please provide as many strategic options as possible, ensuring that each option is distinct and well-defined. First, analyze the core businesses, main competitors, and the competitive advantage of Rated. Then, consider which areas should be focused on for strategic development. Building on the analyses, develop a comprehensive list of strategic options that Rated could consider. Your response should encompass a wide range of potential strategic choices that Rated could consider. Your strategic options should be creative, precise, specific, and relevant. Please think through the steps needed to answer the question, but only tell me the end result. (Control) Identify and outline a comprehensive list of strategic options available for Rated. Please provide as many strategic options as possible, ensuring that each option is distinct and well-defined. First, analyze the core businesses, main competitors, and the competitive advantage of Rated. Then, consider which areas should be focused on for strategic development. Building on the analyses, develop a comprehensive list of strategic options that Rated could consider. Your response should encompass a wide range of potential strategic choices that Rated could consider. Your strategic options should be creative, concise, specific, and relevant. Please think through the steps needed to answer the question, but only tell me the end result.

4	Describe your approach to developing strategic options for Rated. First, recall the primary challenge you identified and the factors you considered when generating these options. Then, describe how the primary challenge influenced the range of options you developed and how you synthesized the information to develop a comprehensive list of strategic options. Your response should provide a clear explanation of the steps you took to develop strategic options for Rated, including any analytical frameworks, data sources, or external inputs you utilized in the process. Your response should also provide a clear and logical explanation of the relationship between the primary challenge and the strategic options you formulated. Your response should be clear and concise. Please think through the steps needed to answer the question, but only tell me your approach to developing strategic options.
5	Please select the best strategic option from the list of options you have developed for Rated and format your response to include the index of the option and its content. Explain why you consider this option to be the best choice and provide supporting details or reasoning. Please ensure that your explanation highlights the specific strengths and advantages of the selected option and demonstrates a thorough analysis of its potential impact. First, evaluate the effectiveness of each strategic option in addressing Rated's challenges and opportunities. Then, think about the long-term implications of each option. If possible, conduct a cost-benefit analysis for each option. Your response should be clear and concise, and provide a compelling argument for the chosen option. Please think through the steps needed to answer the question, but only tell me the end result.
6	Please articulate your mental representation of this task in detail, using descriptive language to convey your underlying mental model. First, reflect on the task and the thought process you engaged in to complete it. Then, identify the primary elements of your mental model. Finally, construct a description of your mental model using clear and precise language. Your response should be clear and concise. Please think through the steps needed to answer the question, but only tell me your mental representation of this task.

Notes: This table shows the prompts used in the GPT-based simulations. Both groups received the same prompt except for Question 3, where the difference is bolded and underscored. The framework provided in the image is the same as in the main experiment. The main text of the case and additional instructions regarding the format of the output are also provided.

Table H.2: Regression results of GPT-based simulations with a single framework

VARIABLES	(1) Number of Options	(2) Number of Strategic Options	(3) Mutually Exclusive	(4) Number of Continue Options	(5) Change Pricing	(6) Exit	(7) Best Option is Strategic	(8) Best Option is Continue	(9) Best Option is Change Pricing	(10) Best Option is Exit
Treatment	0.917***	1.023***	0.006	-0.128	0.030	0.422***	-0.047	0.041	0.065	0.000
	(0.219)	(0.149)	(0.006)	(0.128)	(0.029)	(0.041)	(0.038)	(0.036)	(0.041)	(0.000)
Constant	6.269***	2.743***	0.000*	3.760***	0.910***	0.030**	0.162***	0.856***	0.796***	0.000
	(0.167)	(0.104)	(0.000)	(0.102)	(0.022)	(0.013)	(0.029)	(0.027)	(0.031)	(0.000)
Observations	333	333	333	333	333	333	333	333	333	333
Adjusted R-squared	0.05	0.12	0.00	0.00	0.00	0.24	0.00	0.00	0.00	.
Control Group Mean	6.27	2.74	0.00	3.76	0.91	0.03	0.16	0.86	0.80	0.00

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports the OLS linear regression results of the treatment binary indicator on the dependent variables. Robust standard errors are in parentheses. The variable “best option is exit” shows no variance because none of the observations identified exit as the best option.

Table H.3: Regression results of GPT-based simulations with a set of frameworks

VARIABLES	(1) Number of Options	(2) Number of Strategic Options	(3) Mutually Exclusive	(4) Number of Continue Options	(5) Change Pricing	(6) Exit	(7) Best Option is Strategic	(8) Best Option is Continue	(9) Best Option is Change Pricing	(10) Best Option is Exit
Treatment	0.00	0.12**	0.00	-0.15***	0.01	0.00	-0.03	0.03*	0.01	0.00
	(0.06)	(0.06)	(0.00)	(0.05)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.00)
Constant	6.03***	2.70***	0.00	3.45***	0.99***	0.00	0.05***	0.96***	0.97***	0.00
	(0.05)	(0.04)	(0.00)	(0.04)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.00)
Observations	596	596	596	596	596	596	596	596	596	596
Adjusted R-squared	-0.00	0.01	.	0.01	-0.00	-0.00	0.00	0.00	-0.00	.
Control Group Mean	6.03	2.70	0.00	3.45	0.99	0.00	0.05	0.96	0.97	0.00

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Robust standard errors in parentheses:
*** p<0.01, ** p<0.05, * p<0.1
Notes: This table reports the OLS linear regression results of the treatment binary indicator on the dependent variables. Robust standard errors are in parentheses. The variables “mutually exclusive” and “best option is exit” show no variance because none of the observations were coded as mutually exclusive, or identified exit as the best option.

References

Argyle, L., et al. (2023). Out of one, many: Using language models to simulate human samples. *Political Analysis*, *31*(3), 337–351.

Brand, J., Israeli, A., & Ngwe, D. (2023). Using gpt for market research [Available at: https://papers.ssrn.com/abstract=4395751]. *SSRN Scholarly Paper*.

Csaszar, F., Ketkar, H., & Kim, H. (2024). Artificial intelligence and strategic decision-making: Evidence from entrepreneurs and investors. *SSRN*.

Horton, J. (2023). Large language models as simulated economic agents: What can we learn from homo silicus? *Working Paper*.

Hugging Face. (2022, November). Sentence-transformers/all-mpnet-base-v2 · hugging face [Available at: https://huggingface.co/sentence-transformers/all-mpnet-base-v2].

Li, P., Castelo, N., Katona, Z., & Sarvary, M. (2024). Frontiers: Determining the validity of large language models for automated perceptual analysis. *Marketing Science*, *43*(2), 254–266.