

## RESEARCH ARTICLE



WILEY

# The lean startup method: Early-stage teams and hypothesis-based probing of business ideas

Michael Leatherbee<sup>1</sup> | Riitta Katila<sup>2</sup> <sup>1</sup>Departamento de Ingeniería Industrial y de Sistemas, Pontificia Universidad Católica de Chile, Santiago, Chile<sup>2</sup>Department of Management Science & Engineering, Stanford University, Stanford, California**Correspondence**

Michael Leatherbee, Departamento de Ingeniería Industrial y de Sistemas, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Macul Santiago, Chile.  
Email: mleatherbee@ing.puc.cl

**Funding information**

Comisión Nacional de Investigación Científica y Tecnológica, Grant/Award Number: NS130028; Fondo de Fomento al Desarrollo Científico y Tecnológico, Grant/Award Number: 11170537; National Science Foundation, Grant/Award Number: 1305078

**Abstract**

**Research Summary:** We examine a learning-by-doing methodology for iteration of early-stage business ideas known as the “lean startup.” The purpose of this article is to lay out and test the key assumptions of the method, examining one particularly relevant boundary condition: the composition of the startup team. Using unique and detailed longitudinal data on 152 NSF-supported lean-startup (I-Corps) teams, we find that the key components of the method—hypothesis formulation, probing, and business idea convergence—link up as expected. We also find that team composition is an important boundary condition: business-educated (MBA) members resist the use of the method, but appreciate its value ex post. Formal training in learning-by-thinking methods thus appears to limit the spread of learning-by-doing methods. In this way, business theory constrains business practice.

**Managerial Summary:** Lean startup methodology has rapidly become one of the most common and trusted innovation and entrepreneurship methods by corporations, startup accelerators, and policymakers. Unfortunately, it has largely been portrayed as a one-size-fits-all solution—its key assumptions subject to little rigorous empirical testing, and the possibility of critical boundary conditions ignored. Our empirical testing supports the key assumptions of the method, but points to business education of team members

as a critical boundary condition. Specifically, MBAs resist the use of the method despite being in a strong position to leverage it. Results from a post hoc analysis we conducted also suggest that more engagement with the method relates to higher performance of the firm in the 18-month period following the lean startup intervention.

#### KEYWORDS

business idea, hypothesis-based probing, lean startup method, learning-by-doing, learning-by-thinking, MBA education, young-firm teams

*We decided our company was a go! We have enough data to support our hypotheses and some good feedback from potential partners.*

*NSF I-Corps team, 2016.*

## 1 | INTRODUCTION

Bounded rationality—finite information, finite minds, and finite time—makes young firms imperfect decision-makers (Katila & Ahuja, 2002; Schmitt et al., 2018; March & Simon, 1958). Early-stage teams, especially under high levels of environmental uncertainty and in technology-based industries, are unlikely to decide on an optimal business idea<sup>1</sup> from the very beginning of their ventures (Contigiani & Levinthal, 2019). Too much relevant information is missing, and capturing it requires time and effort. To address this problem and reduce uncertainty about their businesses' viability, entrepreneurs are encouraged to iterate their business ideas using both “learning-by-thinking” and “learning-by-doing” methods that are outlined in research (Gans, Stern, & Wu, 2019; Ott, Eisenhardt, & Bingham, 2017). Understanding these methods is particularly important because getting the business idea “right” early on can have a high impact in the long run (McDonald & Gao, 2019; Zott & Amit, 2007).

In the past decade or so, a new learning-by-doing methodology known as the “lean startup” has emerged. The goal is simple: to help early-stage teams iterate business ideas until they are able to make a sound decision about them (Blank, 2003; Ries, 2011). The method's significant features are (a) formulation of hypotheses in nine pre-identified areas of the business idea (using a shorthand visualization called canvas) and (b) “getting out of the building” to probe each hypothesis by interviewing customers and other stakeholders. The expected outcome of this hypothesis-probing process is to converge on a business idea by “confirming” or “disconfirming” the hypotheses.<sup>2</sup>

Lean startup is a blend of previously identified learning-by-doing methods. It draws inspiration from blended approaches such as discovery-driven planning that similarly urge teams to articulate their underlying assumptions and to get data to iterate them (McGrath & MacMillan, 1995, 2009), and it builds on some existing principles of experimentation (i.e., hypotheses and their testing). It is novel, however, in its strong emphasis on interviewing customers and in its shorthand visualization of the core components of the business idea (Contigiani & Levinthal, 2019).

Lean startup is currently one of the most widely embraced entrepreneurship methods, and it is particularly valuable under environmental uncertainty (Contigiani & Levinthal, 2019; Kerr, Nanda, & Rhodes-Kropf, 2014). Nevertheless, the method's key assumptions and boundary conditions have not been subject to rigorous empirical analysis. Addressing this gap will thus be key to improving the method and to improving business idea processes in general.

The boundary conditions in particular require examination. The belief that early-stage teams would promptly revise their ideas in response to market information (e.g., customer interviews) is in sharp contrast with the description in recent research of some founders as reluctant to change their original ideas, even when external information prescribes it (Crilly, 2018; Grimes, 2017; Parker, 2006). In particular, teams deeply steeped in educational expertise often resist implementing changes in response to external feedback (Katila & Shane, 2005), suggesting educational background as a potentially relevant boundary condition.<sup>3</sup> A better understanding of team composition is also important because this factor can be adjusted by organizations directly, providing a clear point of intervention.

We examine 152 early-stage teams engaged in a standardized implementation of the lean startup method over an 8-week period. Our particular research setting is I-Corps—a National Science Foundation (NSF) program developed by Steve Blank using the lean startup method, which has been offered in a consistent manner since 2011. During the program, all teams iterate a business idea for a core technology originally created by a scientist (one of the team members) and financially supported by a grant from the National Science Foundation.

There are several findings. First, the results from our baseline hypotheses shed light on the *underlying assumptions* of the method. Our results confirm the central role of probing (i.e., customer interviews). Probing not only motivates convergence on a business idea, as the method predicts. We discover that it also motivates the formulation of new ideas, with their corresponding hypotheses-to-be-tested. Probing can thus help teams to dislodge an entrenched vision of the business idea, as needed, and find a new plan. We also find that if a team converges on an idea, it is likely to formulate fewer new hypotheses, which suggests that the method involves a natural stopping mechanism.<sup>4</sup> In contrast, we do not find confirmation for the central role of hypothesis formulation: formulating a large number of hypotheses does not drive or help teams' probing. Altogether, we find support for the main assumptions of the method, but we note that one criticism—that endless formulation of new hypotheses may tire out entrepreneurs and prevent scaling (Ladd, 2016)—is indeed correct, highlighting the dangers of hypothesis proliferation as an end in itself.

Second, we identify *boundary conditions* for the method. While it is beyond the scope of this study to examine all possible boundary conditions, our findings clearly show their importance for the method and its results. In particular, we find that certain teams embrace the method, while others tend to resist it. In particular, teams with formal *business education* (i.e., MBAs on the team) are the least likely to formulate hypotheses and to converge on a business idea. We suggest that this is because learning-by-thinking, which is the staple of MBA programs, can limit the appreciation of learning-by-doing methods like lean startup. However, compared to teams with no business education, MBA teams who *do* engage in hypothesis probing also formulate more new business ideas and their corresponding hypotheses *and* achieve idea convergence faster. MBAs thus may be particularly apt at some components (such as interpreting data from probing) if they engage with the method in the first place.

## 2 | RESEARCH BACKGROUND

Prior research outlines two main approaches—learning-by-thinking and learning-by-doing—that early-stage teams can use to iterate their business ideas (Gans et al., 2019; Ott et al., 2017). Iteration by learning-by-thinking, that is, thinking through potential consequences for alternative firm actions before choosing, received much early attention (e.g., Delmar & Shane, 2003; Gavetti, Levinthal, & Rivkin, 2005). Recently, scholarly attention has shifted to learning-by-doing, in other words, trying potential alternatives before choosing, most notably categorized as trial-and-error, bricolage, and experimentation approaches. Table 1 summarizes the differences between the two learning approaches and their relation to the lean startup method.

Learning-by-thinking methods use mental representations of the environment to plan how a firm's solutions address problems. A “cognitive structures” approach, for example, suggests that early teams form better strategies when they have a holistic understanding of opportunities, markets, and their own firms, and think through the potential consequences for alternative firm actions before choosing (Felin & Zenger, 2009; Ott et al., 2017). An exemplar

**TABLE 1** Learning-by-thinking, learning-by-doing, and the lean startup methods

KEY ELEMENTS	LEARNING-BY-THINKING			LEARNING-BY-DOING		
	COGNITIVE MODEL	ANALOGIES	LEAN STARTUP	TRIAL-AND-ERROR	EXPERIMENTATION	BRICOLAGE
Definition	Problem-solving by constructing mental representations to plan solutions that will address current problems	Problem-solving by using mental representations of past solutions to solve current problems	Problem-solving by hypothesis-probing using interview data	Problem-solving by trying many solutions in the hopes of stumbling upon a good one	Problem-solving by hypothesis-testing using controlled variation of activities and context	Problem-solving by combining resources that already exist in new ways to solve problems
Key references	Delmar & Shane, 2003	Gavetti et al., 2005		Callander, 2011	Camuffo et al. 2019	Baker & Nelson, 2005
Time horizon	Up-front, preplanned		<b>Real-time, experiential</b>	<b>Real-time, experiential</b>		
Method	Elaborate cognitive planning		<b>Rapid iterative experimentation in preselected target areas</b>	Iterative trial-and-error experimentation	<b>Spotlighting: Targeted, narrowly focused experiments</b>	Experimentation with recombinations of existing solutions to create new solutions
Hypotheses	N/A	N/A	<b>Formulation of hypotheses about solutions</b>	N/A	<b>Formulation of cause-effect relationships to test</b>	N/A
Seed for problem-solving	Some theory; prior packaged information about the environment	Past solutions	<b>Team's assumptions about solutions</b>	<b>Accumulated information from past solutions (local hill-climbing)</b>	Unclear: Theory, although not always	<b>Existing solutions and resources at hand</b>
Mental representations of the environment	<b>Environment is mapped and a vision for the future is envisioned</b>	Mapping of similarities between past and current environments	<b>Scaffolding: Visualization of preidentified components of the environment</b>	N/A	Unclear: Theory, although not always	N/A
Role of theory	Holistic framework may draw on theories	N/A	N/A	N/A	Causal theory may underlie hypotheses	N/A
Evaluation criteria	Intuition and thinking through solutions that the firm proposes	Intuition and thinking through the consequences of solutions that the firm proposes	<b>Parallel probing; hypothesis confirmation or disconfirmation by interviewees</b>	Comparison to past performance	<b>Hypothesis confirmation or disconfirmation</b>	Unclear: Subjective assessment relative to aspiration levels

<sup>a</sup>Source: Adapted from Ott et al., 2017 and Gans et al., 2019. Note: Bolded areas indicate overlap between the lean startup and the existing methods.



is Delmar and Shane's (2003) study of a random sample of 223 Swedish ventures, which finds that ventures with a holistic business plan are more likely to introduce products and less likely to fail. "Analogies," in turn, are mental representations of past solutions used to solve current problems. The case teaching approach common in business schools is an example. Although the lean startup features a blueprint of the environment (i.e., nine areas visualized in the canvas), in general it shares very few similarities with learning-by-thinking methods (see Table 1).

Learning-by-doing methods, in contrast, are based on taking action and learning from experience. There are three primary approaches: trial-and-error, bricolage, and experiment. In trial-and-error, many ideas are tried (typically in the neighborhood of existing good solutions) in the hopes of stumbling upon a promising one (Katila & Ahuja, 2002). An analogy may be made to evolutionary theory, where rounds of variation and selection produce adaptive results over time—although trial-and-error is less "blind" than natural selection is. Callander (2011) says, "Trial and error search begins where theory insight ends ... but is not blind as agents are using experience—accumulated information from successes and failures—to guide future choice." Lean startup resembles trial-and-error as it too accumulates information (i.e., interviewees' opinions) to gauge business ideas' potential success or failure. If feedback from the initial interviews is positive, focus is expanded in that direction; if not, focus is moved elsewhere and new hypotheses are formulated and probed.

Bricolage is another commonly studied learning-by-doing approach. In bricolage, solutions that already exist at hand are combined in new ways (Baker & Nelson, 2005). Effectuation is a close variant. Effectuation processes, like those in bricolage, imply that ventures draw on their knowledge and networks to select between possible effects that can be created using this known set of solutions (Sarasvathy, 2001). In lean startup, weekly business-idea iterations have some resemblance to bricolage (i.e., trying out different combinations until a solution is found).

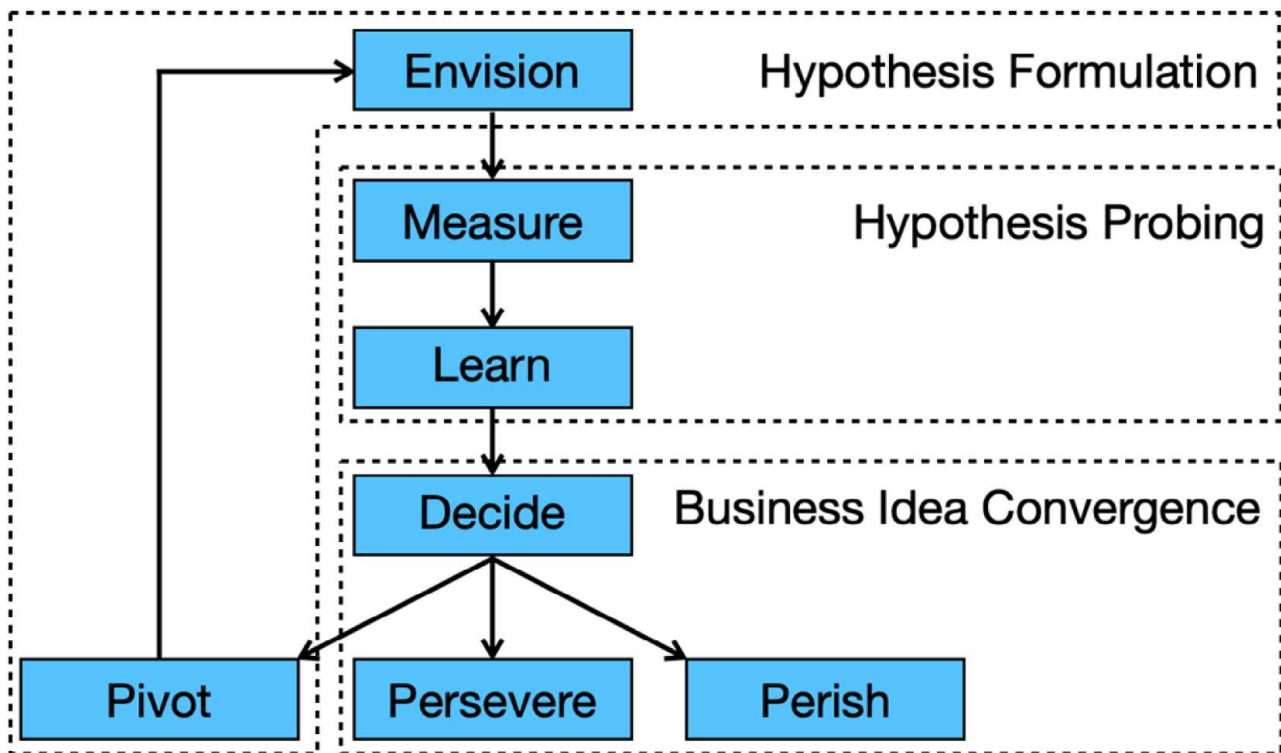
Experiment, the third major learning-by-doing approach, is based on testing hypotheses. Experimentation approaches include theory development; causal propositions; and controlled variation of activities with somewhat controlled circumstances for testing them (Aghion, Bolton, Harris, & Jullien, 1991; Camuffo et al., 2019)—or, alternatively, more light-weight experiments lacking controlled circumstances or clear theory (Miner, Bassoff, & Moorman, 2001). Again, the lean startup method shares several similarities with experiments, but also differs: the business-model canvas acts as a rough outline of areas where learning is needed (but no explicit theory is stated), and hypotheses are formulated and probed with customers (but not under controlled circumstances, nor systematically). Like many experiments (e.g., Andries, Debackere, & van Looy, 2013), it also embraces parallel probing of hypotheses.

Lean startup, as we have suggested and will now discuss in detail, is a blend of these three "learning-by-doing" approaches, particularly drawing from experimentation.

## 2.1 | The lean startup method

Popularized by practitioners Steve Blank and Eric Ries (Blank, 2003, 2013; Ries, 2011), the lean startup method aims to iterate business ideas, helping entrepreneurs make an early decision about their feasibility. To that end, it encourages entrepreneurs to make explicit their assumptions about a business idea (i.e., formulate hypotheses) and then probe them (i.e., interview customers).

One of the main tools of the lean startup is the business-model canvas introduced by Osterwalder and Pigneur (2010). Canvas provides structure for the teams to articulate their assumptions about the business idea in nine prespecified areas.<sup>5</sup> In particular, the teams use it to visually depict the goods and the services the firm plans to offer to customers, the parties participating, and the ways in which the ownership of goods and services is to be exchanged. These articulated assumptions then serve as hypotheses for the team to probe by "getting out of the building"—another significant feature of the method. Probing by interviewing customers and other stakeholders is done to confirm or disconfirm hypotheses. Through hypothesis-based probing, the lean startup method promises to



**FIGURE 1** Lean startup assumptions visualized [Color figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]  
Source: Adapted from Eisenmann, Ries, and Dillard (2013)

help early-stage teams reduce uncertainty about the viability of ideas. It may also encourage them to “pivot,” that is, to find a different idea when the initial idea is not confirmed. See Figure 1 for visualization of the method.

Despite some skepticism (Felin et al., 2020; Gans et al., 2019; Ladd, 2016), the lean startup method has been broadly embraced by entrepreneurs and corporate intrapreneurs, is taught widely in startup accelerators, and is embraced by policymakers in different countries. For example, Start-Up Chile, a renowned ecosystem startup accelerator created by the Chilean government (Clarysse, Wright, & Hove, 2015) with proven impact on the performance of participating startups (Gonzalez-Uribe & Leatherbee, 2018), encourages the use of the method by its members. Our online appendix summarizes the research on these initiatives.

Notwithstanding its popularity, the method's key assumptions have not been subject to rigorous empirical analysis. This is a surprising gap, as examining the key assumptions and components of the method is important to successfully implementing and improving it. Another key gap in our understanding is the boundary conditions. Because composition of the team influences its learning approach (Jung & Shin, 2019; Katila & Shane, 2005), the team's background is particularly relevant for identifying the boundaries for the use of the method. Prior research has tied educational backgrounds to different styles of learning, for example, including some teams deciding *not* to engage in learning, even when provided with evidence that supports change (Hallen & Pahnke, 2016). Analyzing how the lean startup method can be leveraged or be limited by team composition is important for identifying the boundaries of the method.

### 3 | HYPOTHESES

To better understand the steps of the lean startup method, we test the underlying assumptions of the method (H1–H4) followed by potential boundary conditions, particularly regarding the composition of the teams that use the method (H5–H8).

### 3.1 | Lean startup method and the underlying assumptions

*Formulating hypotheses.* In H1 we test the first underlying assumption of the lean startup method regarding hypothesis formulation. The assumption is that formulating several different hypotheses about the business idea serves as a nudge for the team to gather feedback on its idea from different perspectives. In particular, teams that formulate more hypotheses are proposed to probe the hypotheses more (i.e., interview more).

The practice of formulating several hypotheses—that is, breaking up problems and the proposed solutions into the nine areas of the business-model canvas—is believed to activate cognitive processing in teams. By explicitly stating a hypothesis, teams are expected to “turn on” a mindful, conscious activity rather than rely on habitual patterns of thought (Louis & Sutton, 1991). Sowing a seed of doubt in the team's mind about an assumption that was previously taken for granted (and thus ignored) is believed to help the team reconsider its assumptions. In particular, such a switch to active thinking is thought to “seed the search landscape” more broadly, highlighting areas in the team's opportunity landscape that are not fully understood. This process is then assumed to push the team to note the many elements in the environment which now warrant greater probing (from i to ii in Figure 2).

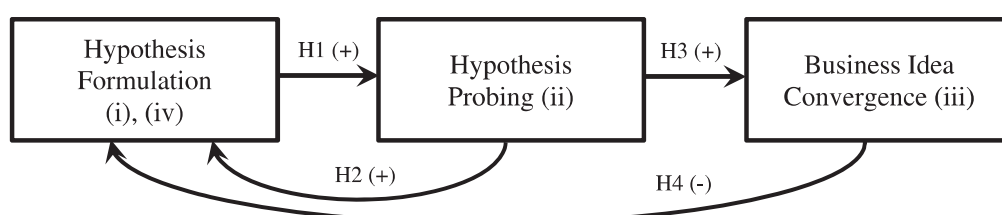
**Hypothesis (H1)** *Number of hypotheses formulated in a given period will relate positively to hypothesis probing in the following period.*

*Probing hypotheses.* In H2, we test another underlying (although rarely discussed) assumption of the lean startup method: namely, that probing is positively related to inspiration to formulate new hypotheses. Investigating one hypothesis helps the team think of new ones. This relationship is important to test as it contradicts much of the entrepreneurship literature that portrays listening to current customers and related stakeholders as counterproductive to forming new ideas (e.g., Felin et al., 2020). One of the main tenets of lean startup is that interviews help the team to learn about the feasibility of the hypothesized ideas and, if necessary, pivot to new ones. H2 tests this less discussed path to pivoting: that new information from interviews triggers new hypotheses that teams would not naturally think of in the absence of the method (from ii to iv in Figure 2).

First, because probes are tied to hypotheses, the new information probes generate often makes the team revisit the original cause-and-effect relationship. Without an explicit causal hypothesis, new information from an interview can seem obvious because it is easy to rationalize a novel phenomenon if no prior assumptions exist (Davis, 1971). In contrast, revisiting a stated hypothesis is likely to result in richer analysis of new information, questioning of prior beliefs, and formulation of new hypotheses.

Second, by nudging teams to probe hypotheses with external stakeholders they would not normally communicate with, probing is believed to encourage idea-networking (Dyer, Gregersen, & Christensen, 2008), which allows teams to access novel information embedded in social networks different from their own. Thus, teams may learn about areas in the opportunity landscape that they were not originally aware of, prompting new hypotheses.

Third, probing of hypotheses is expected to go hand-in-hand with formulating new hypotheses because it is through the concrete process of probing that members of the team become more acutely aware of the views held by their colleagues. Not all members of the early-stage team may fully comprehend other team members' assumptions



**FIGURE 2** Conceptual model of H1–H4

about the business idea. Probing thus has the potential to increase communication and idea-exchange within the team as interview data are unpacked on a weekly basis (cf. Edmondson, Bohmer, & Pisano, 2001 on extensive communication that enables executives to understand surprising results more effectively). Increases in idea exchanges between team members are then likely to inspire team members to think about new ideas not originally considered. And, as new ideas emerge that need to be tested, teams will likely formulate new hypotheses-to-test.

**Hypothesis (H2)** *Hypothesis probing in a given period will relate positively to new hypothesis formulation in the following period.*

*Business idea convergence.* Another assumption is that probing and convergence on a business idea—a team's belief that a hypothesis no longer requires testing—go hand in hand. When the team probes, gaining more understanding of the hypothesis, it is consequently more likely to converge on treating it as a quasi-fact. There are two reasons for the proposed relationship (from ii to iii in Figure 2). First, because probing is conducted with external stakeholders (such as customers), the lean startup method assumes that information gathered is “salient” and “vivid.” As such, the information is assumed to attract the team's attention and mobilize the team to converge. As Li, Maggitti, Smith, Tesluk, and Katila (2013) note, because salient and vivid information departs from expectations and norms of the team, it is more likely to enter team members' consciousness and affect subsequent action. Thus, this information will push the team to discuss alternatives, focus on possible tensions, and converge on a decision (cf. Edmondson et al., 2001; Li et al., 2013). In contrast, information that is perceived as similar to what the team already knows—such as information gathered from team members themselves—does not similarly call the team to action (see also Fiske & Taylor, 1991; Sullivan, 2010).

Second, because it involves outsiders, probing with external stakeholders is more likely to introduce novel data and information to the team. These novel data can then be used to resolve contradictory beliefs within the team and help dislodge barriers to convergence (Edmondson et al., 2001). Taken together, when the team is contrasting the cause-effect relationship outlined in a hypothesis against the information gained through probing (i.e., vivid and new information), the assumption of the lean startup method is that the process can lead to a reduction in the team's perceptions about the uncertainty of a given idea, encouraging convergence (from ii to iii in Figure 2). Moreover, as the team becomes ever more confident that they are reducing the uncertainty about their business idea, the assumption of the lean startup method, as proposed in H4, is that the need for new hypotheses should also diminish (from iii to iv in Figure 2).

**Hypothesis (H3)** *Hypothesis probing in a given period will relate positively to business idea convergence in the next period.*

**Hypothesis (H4)** *Business idea convergence in a given period will relate negatively to new hypothesis formulation in the next period.*

Although we have so far discussed the lean startup process as a “one-size-fits-all” solution, the arguments above point to much of the method's functioning being related to teams' level of engagement with the learning-by-doing aspects of the method. In H5–H8, we focus on the composition of the team as a potentially critical boundary condition.

### 3.2 | Boundary conditions: Lean startup method and the composition of early-stage teams

Research finds that the background and prior experiences of early-stage team members shape the decisions and strategic decision-making processes that ventures follow (Beckman, 2006; Eisenhardt & Schoonhoven, 1990; Lerner &

Malmendier, 2013). Early teams' educational background has been found to be particularly influential, at least partly because education is typically an intense, formative experience for individuals' thinking styles and worldview (Jung & Shin, 2019). Research finds, for example, that even long after graduation, MBA students closely follow academic theories taught in business schools (Priem & Rosenstein, 2000) and surgeon-CEOs' decisions about a venture's R&D strategies are strongly shaped by their medical education (Katila, Thatchenkery, Christensen, & Zenios, 2017).

In H5–H8, we argue that team members' educational background is particularly relevant for the use of the lean startup method (see Figure 3). Prior research has tied educational backgrounds to different styles of learning, including some teams deciding *not* to engage in learning, even when provided with evidence that supports change (Hallen & Pahnke, 2016). Understanding how such sources of inertia may be tied to the team's educational background is important for fully understanding the power and limits of the lean startup method. In the language of opportunity landscapes, team members with a particular educational background may have a fixed “x” landscape in mind that they are reluctant to reshape, either because they are averse to using the method to explore the landscape or because they are married to their original starting position within it.

*Business education.* Business education of team members is particularly relevant for lean startup because much of the required curriculum in U.S. business schools emphasizes learning-by-thinking rather than learning-by-doing methods (Navarro, 2017). Thus, in H5 we propose that formal business education, which we operationalize as team members with an MBA degree, is likely to particularly limit probing as a core learning-by-doing step in lean startup.

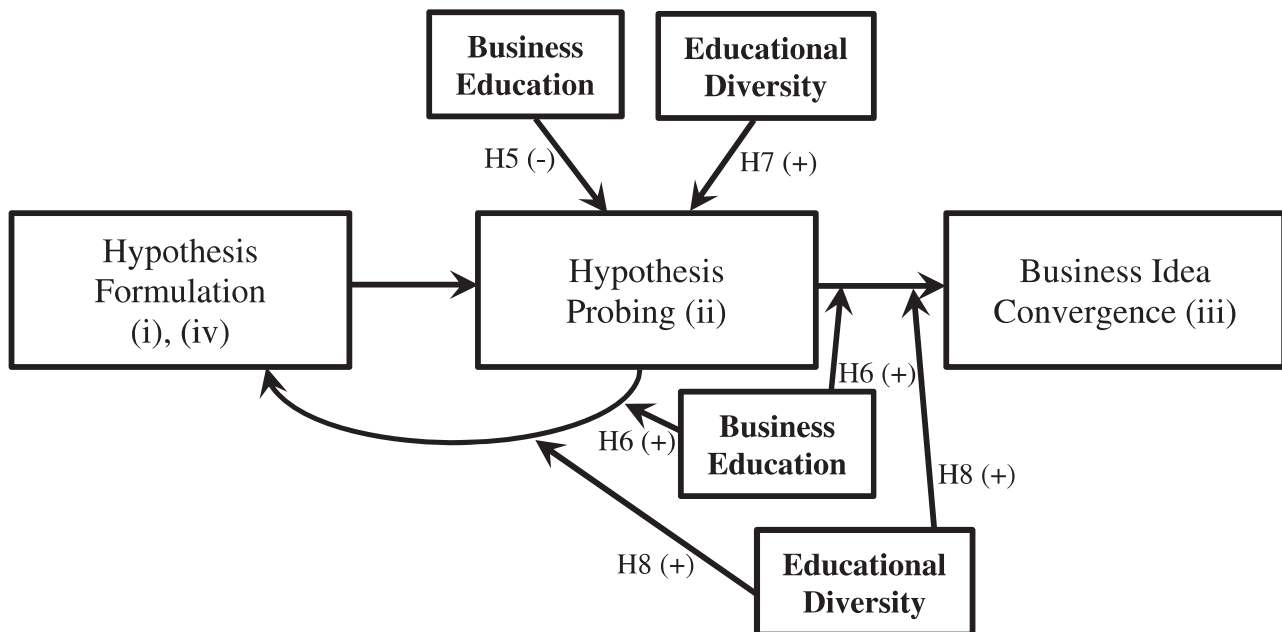
First, in their education, MBAs are trained to use learning-by-thinking methods, that is, the use of analogies (case studies), frameworks, and synthesized information such as market reports, to quickly and efficiently map out industry structures and strategies (Gary & Wood, 2011; Gavetti et al., 2005; Sarasvathy, 2001).<sup>6</sup> Corporate strategy, for example, a required course in 92% of the MBA programs in the United States, heavily emphasizes learning-by-thinking tools (Gavetti et al., 2005; Navarro, 2017). Corporate strategy staples such as the five forces (Porter, 1980), transaction costs (Williamson, 2019), and capabilities (Teece, Pisano, & Shuen, 1997) teach students to cognitively blueprint, synthesize, and think through the dimensions of firms and their environments before they act. This repeated activation and use of learning-by-thinking methods in an early stage of one's career is particularly likely to strengthen the individual's subsequent emphasis on problem-solving using learning-by-thinking methods, possibly making them skeptical of other methods. This point is consistent with a study by Gary and Wood (2011) where the authors showed, using a sample of 63 MBA students, that MBAs are comfortable using “learning-by-thinking” methods, in other words, holistic mental models of the business situation. All in all, it is likely that team members with formal business education would be less likely to embrace learning-by-doing methods. A close parallel is when prior industry experience with legacy technologies limits executives' ability to switch to new technologies, even in the face of discontinuous change (Furr, Cavarretta, & Garg, 2012).<sup>7</sup> We similarly propose that teams that are trained to use “legacy” methods may be skeptical of new learning-by-doing methods, and particularly of their distinct features such as probing.

Second, we propose that because MBA training is about business, the team is more likely to treat members with an MBA as domain experts in forming business ideas. Thus, they would likely give outsize emphasis to such members' preferences in setting and iterating the business idea. This tendency may reflect the authority principle (Cialdini, 2001), by which team members become less likely to challenge a member with business education, believing that the expert is more likely to identify the best solution. Either way, with an MBA on the team, we propose that the team will feel less compelled to apply probing, the core step of the lean method.

**Hypothesis (H5)** *Business education (i.e., MBA) on the team will relate negatively to hypothesis probing.*

In H6 we particularly focus on how business education (MBA) of team members interacts with the use of the lean startup method in terms of probing (see Figure 3). Research on learning-by-doing methods has singled out learning from data and experiments (i.e., probing) as particularly significant yet challenging for some executives. In particular, executives that lack relevant expertise to interpret data benefit less from experiments, because experimentation by an inexperienced team member is likely to be costly and error-prone (Thomke, 2003); and conversely, because





**FIGURE 3** Conceptual model of H5–H8

there is some evidence that teams with more experience derive more benefits through better interpretation (cf. Koning, Hasan, & Chatterji, 2019 on A/B testing).

Although these prior studies were about experiments, not the lean startup, they indicate that team composition is likely related to how the results of probing are utilized. Formal business training can potentially interact with probing in several ways, depending on how team members use their formal training to process the data that are gained. One possibility is that the learning-by-thinking training makes MBAs skeptical of any interview data and thus could make them less likely to even start processing the data. In this scenario we would see a nonsignificant or even a negative interaction. Another possibility, which we propose in H6, is that learning-by-thinking training acts as a complement to learning-by-doing. In this scenario, MBAs who embrace the learning-by-doing method could use appropriate analogies and frameworks from their formal training (including how to interpret traditional focus group data) to interpret what they have learned through the interviews and from their interviewees, suggesting a positive interaction. Thus, it may be more difficult for the teams that lack business education to interpret the results from probing or become inspired with new ideas (cf. Katila et al., 2017; Sorenson, 2003). In contrast, team members with MBAs are more likely to have relevant experience, helping them to interpret the results of probing more effectively:

**Hypothesis (H6)** *Business education (i.e., MBA) in the team will amplify the positive relation between (a) hypothesis probing and business idea convergence and (b) hypothesis probing and hypothesis formulation.*

*Diversity in education.* In contrast to the previous two hypotheses that focused on MBA-teams, in H7 and H8 we move to consider the education of all team members more broadly. Although we have prior understanding of how diversity (e.g., in industry experience; Furr et al., 2012) makes teams more flexible users of learning-by-thinking methods, we do not understand well how diversity is related to the use of learning-by-doing methods such as lean startup. In H7 and H8, we focus on the diversity regarding a team's education (i.e., range of degrees in disciplines including law, medicine, engineering, and business), as each degree is tied to different styles of learning and knowledge bases. In H7, we propose that educational diversity in the team is likely to prompt the use of the lean startup method, and probing in particular.

First, overconfidence in the team's existing approach and overoptimism regarding initial business ideas is more likely when the team is homogeneous. This is because homogeneous teams share a similar domain knowledge and

are likely to consolidate their preexisting beliefs when observing that their teammates have the same beliefs. In contrast, more diverse teams are likely to be collectively less confident in their methods and in their starting position, because of the lack of consensus about a given domain of knowledge and how it relates to the opportunity landscape. Consistent with this argument, Almandoz and Tilcsik (2016) found that more diverse teams were more “skeptical,” whereas in more homogeneous teams (where nonexperts in the domain were a minority), it was “difficult ... to challenge experts and consider alternatives.” Therefore, we propose that the presence of educational diversity in the team would go hand-in-hand with a predisposition to engage in probing.

Second, educationally diverse teams can more easily envision access to a broader set of potential stakeholders, making it mechanically easier for such teams to engage in hypothesis probing (such as arranging meetings for interviews and absorbing insights). Ability to reach out to—and meet with—a greater number of stakeholders is thus likely to make these teams more receptive to the method and to probing in particular.

Third, we propose that, by having a broader knowledge-base and stakeholder arena to build on, educationally diverse teams are likely to have a broader view of the opportunity landscape. This broader view could help diverse teams see both needs and solutions from different perspectives (Gruber, MacMillan, & Thompson, 2013), making them more prone to probe. A more expansive view leads to more expansive questioning.

**Hypothesis (H7)** *Educational diversity in the team will relate positively to hypothesis probing.*

We further propose that educationally diverse teams that probe more will be able to interpret the data from probing better. This is because they are likely to have broader absorptive capacity—that is, their collective knowledge base is more heterogeneous. As a consequence, they are quicker to converge on an assessment about the soundness of their ideas. At the same time, we expect that probing across a broader and more diverse group of stakeholders will help diverse teams generate new hypotheses. Conversely, it is more difficult for teams that lack educational diversity to decide how to interpret results or become inspired by new hypotheses, as they are likely to have a narrower base to draw on. Thus, we propose:

**Hypothesis (H8)** *Educational diversity in the team will amplify the positive relation between (a) hypothesis probing and business idea convergence and (b) hypothesis probing and hypothesis formulation.*

## 4 | METHODS

### 4.1 | Research setting

*Lean Startup at I-Corps.* To test our hypotheses about the lean startup method, we use the U.S. National Science Foundation's Innovation-Corps (I-Corps) program as our setting. This Congress-approved program was launched in 2011 to offer the lean startup training to prior NSF grantees. The program is aimed at helping NSF-funded scientists and engineers to identify business ideas for their technology-based inventions. While implementation of the lean startup method varies widely across the world, NSF has strived to train in a highly standardized manner, strictly adhering to method's founding principles (Arkilic, 2019).<sup>8</sup> “We imposed a standard curriculum,” an early program officer at NSF's I-Corps said in an interview. “We wanted uniformity ... to get the innovation process to scale.”

Standardized features of I-Corps include a standard curriculum, a lead instructor who trains all the trainers, standardized length (2–3 day in-person kickoff, 8-week online modules, 2 day in-person wrap-up), and uniform requirements for team formation. All I-Corps teams are also required to use the common LaunchPad Central platform (which incorporates the business-model canvas). Other uniform features include transparent reporting on the platform including details of hypotheses and interviews, enabling rapid feedback from instructors and

subsequent iteration by teams; weekly reporting deadlines; parallel probing of multiple hypotheses per interview; and weekly performance goals. These features were imposed from the first training and have not changed (Arkilic, 2019).

The original motivation for NSF to use the lean startup method was the lack of knowledge among scientists (i.e., NSF grantees) about business and the related perennial challenge in academic entrepreneurship, namely, low rates of commercialization of technologies by scientists and engineers (Wright, Piva, Mosey, & Lockett, 2009). The learning-by-doing orientation of the I-Corps program was also deemed a particularly good fit for the educational background of NSF scientists.

*Program elements.* During the I-Corps, each team participates in an intensive 8-week program with a cohort of about 12–15 other teams. The program includes both on-site (in a local I-Corps node) and off-site weeks, and is frequently described by participants as “intense” and a “24–7 bootcamp” (postparticipant surveys).

Before the start of the 8-week intervention, all teams are required to submit the first version of their business idea using the program's online platform, LaunchPad Central. At this point, the teams' business ideas are likely to be ill-defined and imprecise. Nevertheless, they must be explicitly outlined for the team to start the program. In particular, all teams are required to articulate the underlying assumptions of their business idea in nine prespecified areas of the business-model canvas, that is, to formulate the first hypotheses.

Program delivery is standardized. Once a week, teams attend a “flipped classroom session,” meaning that material on lean startup is mostly studied beforehand, with classroom time devoted to promoting deeper understanding. During the classroom sessions, the teams present weekly progress, and receive critique from other teams and the instructors.

The teams' core tasks during the program are to formulate and log “hypotheses to probe,” the “number of weekly interviews”, the people interviewed (role and typically the name of the interviewee), the key insights from the interview, and changes in the team's business-model canvas (e.g., confirmed or disconfirmed hypotheses, new hypotheses). It is mandatory to log each of these changes on the program's online platform on a weekly basis. Instructors monitor the weekly progress and provide written feedback to teams through the platform.

Our detailed longitudinal data on the week-by-week evolution of each team's business-model canvas (using LaunchPad Central) records the changes that teams undergo. For example, a team may start with an assumption of specific market segments that would benefit from their technology, only to remove or add market segments in the following weeks as a result of the probing conducted with potential business stakeholders.

*Probing.* Consistent with the lean startup method, the program strongly emphasizes the idea of “getting out of the building” to conduct customer and stakeholder interviews. A weekly target is set for the teams to conduct about 10 interviews with real potential business stakeholders (mostly customers) to probe the hypotheses, including parallel probing so that multiple hypotheses can be probed in a single interview. Teams are required to use more than one interview to “confirm” or “disconfirm” each hypothesis.

To qualitatively validate teams' probing efforts, we asked all 14 teams in one cohort to answer, on a weekly basis, the questions “What were the 2 to 3 most important decisions you made this week related to your project? Why did you make each decision?” We content-analyzed these responses to isolate the instances where teams referred to probing that they had conducted. We also counted the number of statements in each week that made a direct reference to the effort of gathering real information from the market (see Tables 9a and 9b in the online appendix). Despite our prompt making no reference to probing, participants described it frequently. This analysis strongly validates probing as the central component in lean startup.

## 4.2 | Sample

Our dataset consists of 1,061 early-stage team members (divided into 388 teams from 23 cohorts) that participated in the National Science Foundation's I-Corps program across the United States over 8 weeks. While our sample

spans several cohorts (we include cohort fixed effects), the teaching materials, supervision and feedback methods, length of the program, and evaluation protocols are standardized across all cohorts.

To construct our team-level characteristic variables, we handpicked data from team members' LinkedIn profiles and cross-validated with data from Crunchbase and preprogram surveys (Gonzalez-Uribe & Leatherbee, 2018). We gathered biographical and demographic information, including each member's gender, age, level of education, and prior work and entrepreneurial background. Teams with incomplete individual-level information were dropped. Our final sample consists of 152 early-stage teams (381 team members from 16 different cohorts) and 1,216 team-week observations.<sup>9</sup>

Our sample design is particularly appropriate for our research objectives. All teams are early-stage but have a preconceived business idea (NSF-funded technology to be commercialized). Eligibility for I-Corps excludes teams with significant revenues and private financing, consistent with our focus on early-stage teams. NSF further mandates that each team have the following members: entrepreneurial lead, technical lead (PI), and an industry lead (mentor) reducing team heterogeneity. Further, we have complete data on all sample teams' decision-making regarding business ideas with *high levels of granularity* and in *real time*.

*Effort by teams.* Due to the competitive nature of the I-Corps program (participants must go through an application process and commit to an intense program-related workload), teams are highly motivated to spend considerable effort on iterating their business ideas during the 8 weeks of the program. Altogether, over the course of the program, teams typically solicit and perform close to 100 interviews (estimated interview workload is 20–30 hr/week), give 10 presentations, and participate in individual meetings with instructors.

In addition to the documented intensity of teams' engagement with the program, both the pre- and follow-on surveys<sup>10</sup> conducted with the teams indicate that they are indeed building a company, and not merely engaging in an executive education course. On average, 55% of the teams in our sample had incorporated their company within 6 months after the I-Corps intervention; and about 80% of these firms were still operational 18 months after I-Corps and had sought additional funding.

*Entrepreneurial team background.* The I-Corps team members in our sample closely represent the broader population of *professional entrepreneurs* in science-driven businesses. Based on our data, the average age of team members is 32 years old, and 23% are female. Slightly under 12% of the members have an MBA, 38% hold a Master of Science degree, and 50% have a Ph.D. (as is typical in science-intensive sectors; Belz, Giga, & Zapatero, 2018). At the time of the study, 24% are university professors, typically in sciences and engineering. On average, participants have held five previous job positions in the past. Twenty-four percent have previously held a corporate C-level position, and 30% of the participants had founded a company prior to joining the program.

As noted above, teams are required to have three members, with an occasional team having a co-entrepreneurial or a co-industry lead. Fourteen percent of the I-Corps teams have a member with an MBA, while 80% of the teams have at least one member with a Ph.D. Only 7% of the teams are entirely female, while 58% of the teams are entirely male.

### 4.3 | Measures

*Core variables of interest.* There are three core variables of interest: hypothesis formulation, hypothesis probing, and business idea convergence. We measured *hypothesis formulation* as the number of new assumptions that each team articulated about their business idea on the program's platform in a given week.

We measured *hypothesis probing* by the number of stakeholder interviews a team conducted each week. Because we focus on how often the teams check the feasibility of their ideas with external stakeholders, the number of interviews by a given team is an appropriate measure. Detailed minutes of each meeting were uploaded on the program's platform and the quality was monitored by the program instructors, further bolstering the measure. In alternate tests we included a squared term of the variable with no changes in our hypothesized findings.

We measured *business idea convergence* as the number of items in a given week that were either removed from a team's business-model canvas (i.e., they were deemed invalid assumptions) or added to it without being associated with a hypothesis (i.e., they were considered a fact, not an assumption). Because teams register changes in their business-model canvas on a weekly basis, we are able to track items that are “removed” or “disconfirmed”. By removing an item, teams are explicitly declaring that their perception of the validity of the item was incorrect or undesirable. In contrast, if teams *add* new hypotheses (hypothesis formulation) in a given week, we consider this a sign that they recognize they are still dealing with relatively uncertain elements. These new hypotheses are not included in *business idea convergence*, but rather are directly measured as part of the *hypothesis formulation* variable.

We measured each team's *business education* by a binary variable labeled *MBA on team* that takes a value of one if the team has at least one member with an MBA degree and zero otherwise. An MBA, or master's degree in business administration, focuses on providing expertise in management analysis and is accredited to ensure consistency and quality of education. We do not consider undergraduate business degrees or certificates of executive education programs because their curricula are less standardized than the MBA. During the time period of our study, as noted above, the “learning-by-thinking” methods were the staple of MBA education making *MBA on team* a particularly relevant measure.<sup>11</sup>

Like Harrison and Klein (2007), we measured diversity in expertise using a Blau index of *educational diversity*, based on the classification of each team member's highest degree (Doctor of Law, Doctor of Medicine, MBA, Ph.D., MS, BS). The index takes the value of zero if all members have the same degree and higher values indicate higher diversity.

*Controls.* We included several controls. Because bigger teams may be able to codify more of their assumptions, conduct more interviews, and have a greater capacity to change (Barker & Duhaime, 1997), we controlled for team size, measured as the number of members in each team. Moreover, we controlled for gender diversity, as diverse teams may be more likely to embrace new tools such as lean startup. Gender diversity was measured using a Blau index based on the number of same-gender team members.

We also controlled for the training cohorts using *cohort fixed effects*. Although the goal of the NSF is to provide a standardized training experience across all cohorts, we included cohort fixed effects to control for any unobserved differences not controlled for otherwise. In addition, we included *week fixed effects* to account for systematic program-level differences during the course of each 8-week program, and because teams may have become more skilled at hypothesis-based probing or idea convergence as they matured in the program.

#### 4.4 | Statistical analysis

The data consist of a panel of observations on team-weeks. We used negative binomial models as our primary models and Poisson models for robustness because our dependent variable is a count variable with overdispersion (variance exceeds the mean; Hilbe, 2007). To account for team heterogeneity (repeat observations of the same team across multiple weeks), we used Generalized Estimating Equations (GEE) regressions (Liang & Zeger, 1986).

We repeated the analysis with *fixed-effects Poisson regression* that attempts to account for time-invariant team heterogeneity with team fixed effects; we chose this method as it is generally regarded as a “truer” fixed effects model than fixed-effects negative binomial models. We also included a lagged dependent variable (e.g., *hypothesis formulation*) to control for time-variant unobserved team heterogeneity and to facilitate causal inference (Heckman & Borjas, 1980). We report all results with robust standard errors.

*Causal inference.* While few research designs can completely rule out reverse causality or other possible alternative explanations (see Rathje & Katila, 2021), we have tried to account for these issues in several ways. In addition to lagging the dependent variable, we also used random-effects analysis, which, unlike fixed effects, allows estimation of variables that do not change over time. Findings were highly consistent (available from the authors). We also used an *instrumental variables* analysis, as described in more detail in the online appendix, although we recognize its



limitations. To further facilitate causal inference, we included a rich-control set and lagged our independent variables by 1 week. Finally, we conducted several post hoc analyses to better understand the results.

## 5 | RESULTS

Table 2 shows the descriptive statistics and correlations for all variables. Each week, teams formulated 3 to 4 hypotheses, conducted around 10 interviews to probe hypotheses, and converged on 3 business ideas, on average. Overall, the independent variables show considerable variance, and the correlation matrix indicates low to moderate correlations among them (team size and MBA on team were entered separately and together, with no changes in the results).

The results for the GEE negative binomial regressions are shown in Tables 3 through 6.<sup>12</sup> All models include week and cohort fixed effects. Model 1 in all tables includes the control variables only. Of the control variables, path-dependencies are interesting: probing in the past is not related to probing in the future, and formulating more hypotheses is related to fewer new hypotheses in the future. However, teams that converged in the past are likely to do so again.

Model 2 introduces hypothesis formulation, hypothesis probing, and business idea convergence variables, corresponding to H1-H4. Model 3 introduces MBA on the team (business education) and its interaction with the main independent variables (H5, H6). Model 4 repeats with educational diversity (H7, H8). We use the full model (Model 5 in all tables) to interpret the results because omitting any key explanatory variables can lead to bias in the estimation of the remaining parameters (Kennedy, 1998: 103). Model 6 replicates the full model but excludes teams with extensive preprogram joint experience (i.e., the roughly 20% of the teams that had formed before the start of the I-Corps program<sup>13</sup>). After excluding teams with prior joint experience, the results (Models 6 in all main tables) were consistent and strongly supported our original findings, as described in detail below.

Table 3 focuses on the relationship between hypothesis formulation and hypothesis probing (see Figure 2). H1 argued that formulation would go hand in hand with probing. Our results did not support this expectation. The only exception is educationally diverse teams, for whom there is some indication that more hypotheses result in more probes compared to less diverse teams. We return to this unexpected finding in the Discussion.

Table 4 focuses on the relation between hypothesis probing and hypothesis formulation. In H2, we argued that probing would inspire teams to come up with new ideas not previously considered, priming new hypotheses. Our results confirm the hypothesis.

Tables 5 and 6 focus on convergence. We argued that probing hypotheses would help teams gather information that would allow them to converge (H3), and converging on a business idea would result in fewer new hypotheses (H4). Our results support both H3 and H4.

**TABLE 2** Descriptive statistics and correlations

		Mean	SD	1	2	3	4	5	6
1	Hypothesis formulation	3.63	6.22						
2	Hypothesis probing	10.45	7.40	.07					
3	Business idea convergence	2.97	5.01	.00	.28				
4	MBA on team	0.14	0.35	-.01	-.07	-.01			
5	Educational diversity	0.39	0.31	-.01	-.03	.07	.19		
6	Team size	2.51	1.06	.01	.06	.08	.44	.29	
7	Gender diversity	0.16	0.22	-.01	-.01	.05	.22	.09	.30

Note: N = 152 teams.

**TABLE 3** Negative binomial GEE regressions on Hypothesis probing<sub>t</sub>

	(1)	(2)	(3)	(4)	(5)	(6)
Hypothesis formulation <sub>t-1</sub>		-0.000 (.520)	-0.000 (.828)	-0.001 (.084)	-0.001 (.094)	-0.000 (.280)
MBA on team			0.005 (.518)		0.006 (.485)	0.008 (.313)
MBA on team × Hypothesis formulation <sub>t-1</sub>			-0.001 (.193)		-0.001 (.110)	-0.002 (.031)
Educational diversity				-0.002 (.736)	-0.003 (.646)	0.000 (.940)
Educational diversity × Hypothesis formulation <sub>t-1</sub>				0.001 (.066)	0.002 (.036)	0.001 (.059)
Gender diversity	0.001 (.830)	0.001 (.838)	0.001 (.842)	0.001 (.824)	0.001 (.818)	0.006 (.406)
Team size	0.004 (.285)	0.004 (.277)	0.004 (.291)	0.004 (.311)	0.004 (.330)	0.003 (.389)
Lagged DV	0.000 (.630)	0.000 (.596)	0.000 (.531)	0.000 (.610)	0.000 (.525)	0.000 (.979)
Constant	-0.176 (.000)	-0.175 (.000)	-0.175 (.000)	-0.173 (.000)	-0.173 (.000)	-0.140 (.000)
Observations	1,064	1,064	1,064	1,064	1,064	833
Deviance	806.8	806.7	806.0	805.4	804.4	677.9

Note: p-Values in parentheses, two-tailed tests. All models include week and cohort fixed effects. Robust standard errors. Model 6 excludes teams with prior joint experience.

**TABLE 4** Negative binomial GEE regressions on Hypothesis formulation<sub>t</sub>

	(1)	(2)	(3)	(4)	(5)	(6)
Hypothesis probing <sub>t-1</sub>		0.004 (.001)	0.004 (.000)	0.002 (.038)	0.007 (.000)	0.015 (.000)
MBA on team			-0.206 (.059)		-0.214 (.065)	-0.490 (.051)
MBA on team × Hypothesis probing <sub>t-1</sub>			0.011 (.026)		0.009 (.071)	0.030 (.047)
Educational diversity				0.001 (.990)	0.029 (.639)	-0.053 (.379)
Educational diversity × Hypothesis probing <sub>t-1</sub>				0.001 (.707)	0.000 (.917)	0.004 (.081)
Gender diversity	-0.073 (.124)	-0.007 (.881)	0.084 (.003)	0.071 (.002)	-0.005 (.896)	0.003 (.944)
Team size	0.002 (.899)	0.034 (.134)	-0.004 (.734)	-0.007 (.616)	0.040 (.035)	0.112 (.000)
Lagged DV	-0.011 (.000)	-0.006 (.001)	-0.010 (.000)	-0.011 (.000)	-0.007 (.000)	-0.010 (.000)
Constant	0.001 (.972)	-0.544 (.000)	-0.242 (.023)	0.006 (.897)	-0.427 (.001)	0.310 (.176)
Observations	1,064	1,064	1,064	1,064	1,064	833
Deviance	1,441	1,371	1,378	1,431	1,361	1,139

Note: p-Values in parentheses, two-tailed tests. All models include week and cohort fixed effects. Robust standard errors. Model 6 excludes teams with prior joint experience.

**TABLE 5** Negative binomial GEE regressions on Business idea convergence<sub>t</sub>

	(1)	(2)	(3)	(4)	(5)	(6)
Hypothesis probing <sub>t-1</sub>		0.012 (.000)	0.010 (.000)	0.006 (.000)	0.006 (.000)	0.005 (.000)
MBA on team			-0.203 (.000)		-0.166 (.001)	-0.192 (.000)
MBA on team × Hypothesis probing <sub>t-1</sub>			0.019 (.000)		0.018 (.000)	0.015 (.000)
Educational diversity				-0.080 (.362)	-0.048 (.475)	-0.020 (.747)
Educational diversity × Hypothesis probing <sub>t-1</sub>				0.010 (.030)	0.006 (.052)	0.006 (.052)
Gender diversity	-0.006 (.925)	0.036 (.012)	-0.029 (.535)	0.070 (.187)	-0.033 (.414)	0.007 (.883)
Team size	0.033 (.036)	-0.011 (.017)	0.062 (.000)	0.034 (.013)	0.033 (.033)	0.051 (.005)
Lagged DV	0.004 (.000)	0.007 (.000)	0.009 (.000)	0.007 (.000)	0.007 (.000)	0.008 (.000)
Constant	-0.822 (.000)	-0.024 (.046)	-0.897 (.000)	-0.907 (.000)	-0.814 (.000)	-0.644 (.000)
Observations	1,064	1,064	1,064	1,064	1,064	833
Deviance	1,495	1,607	1,441	1,454	1,450	1,140

Note: p-Values in parentheses, two-tailed tests. All models include week and cohort fixed effects. Robust standard errors. Model 6 excludes teams with prior joint experience.

H5 through H8 focused on team composition as a boundary condition. Models 3 and 5 in Tables 3–6 test H5. Teams with MBAs formulate fewer hypotheses and are slower to converge on their business ideas, but the results on probing are not significant. However, the positive coefficient on the interaction between MBAs on the team and probing (Tables 4 and 5) indicates that probing results in more new hypotheses formulated and more convergence in teams with an MBA member, confirming H6. These results are consistent with the argument that MBAs as business experts can possibly evaluate the results of probing more effectively.

In H7 we proposed that educational diversity is related with more probing, while H8 focused on the interactions. We do not find the expected effect on probing (H7) in our results. However, post hoc analyses show that educational diversity is positively related with *major* business idea convergence (results available from the authors). We also find that probing results in more convergence and more new hypotheses in educationally diverse teams, as we expected (H8), although the results are only moderately significant.

Because standard errors can be inaccurately reduced in models with a lagged dependent variable, resulting in overstated significance, we also ran the models by *excluding the lagged dependent variable*. Our original findings stayed. Because standard errors were almost identical across the models that included or excluded the lagged dependent variable, artificially small standard errors are less likely to have influenced the significance of our main findings.

Our post hoc analyses also provide evidence on long-term performance outcomes of lean startup. As reported in the online appendix, teams that engage more with the method (more hypotheses, more interviews) have more positive outcomes in the 18-month period following the lean startup intervention (venture foundings, employment), further lending credence to our findings. Other sensitivity analyses (reported in the online appendix) use instruments, and examine several alternative explanations for our findings including idea quality and teams' startup experiences, again confirming the original findings.

**TABLE 6** Negative binomial GEE regressions on Hypothesis formulation<sub>t</sub>

	(1)	(2)	(3)	(4)	(5)	(6)
Business idea convergence <sub>t-1</sub>		-0.006 (.012)	-0.006 (.013)	-0.009 (.000)	-0.006 (.011)	-0.003 (.166)
MBA on team			-0.147 (.085)		-0.135 (.131)	-0.329 (.050)
MBA on team × Hypothesis probing <sub>t-1</sub>			0.006 (.121)		0.004 (.258)	0.020 (.002)
Educational diversity				-0.082 (.105)	-0.049 (.389)	-0.169 (.093)
Educational diversity × Hypothesis probing <sub>t-1</sub>				0.012 (.000)	0.005 (.045)	0.018 (.000)
Gender diversity	-0.073 (.124)	0.017 (.635)	0.023 (.520)	0.018 (.588)	0.011 (.800)	0.009 (.848)
Team size	0.002 (.899)	0.043 (.066)	0.047 (.048)	0.024 (.330)	0.042 (.084)	0.071 (.003)
Lagged DV	-0.011 (.000)	-0.004 (.032)	-0.004 (.038)	-0.005 (.010)	-0.003 (.052)	-0.008 (.003)
Constant	0.001 (.972)	-0.525 (.000)	-0.460 (.002)	-0.429 (.004)	-0.443 (.003)	-0.362 (.049)
Observations	1,064	1,064	1,064	1,064	1,064	833
Deviance	1,441	1,370	1,366	1,373	1,361	1,057

Note: p-Values in parentheses, two-tailed tests. All models include week and cohort fixed effects. Robust standard errors. Model 6 excludes teams with prior joint experience.

## 6 | DISCUSSION

Our research started with the observation that, although the lean startup methodology is increasingly popular, its key assumptions and boundary conditions have been subject to little rigorous testing. Ours is one of the first empirical studies—if not the first—to pin down the assumptions underlying the method, as implemented by NSF's I-Corps, which arguably most faithfully represents its principles. Our results also highlight that the method is not a one-size-fits-all solution; rather, critical boundary conditions such as team composition may influence its functioning.

### 6.1 | Key Findings

There are several key findings. First, our results challenge some common assumptions about the lean startup method while confirming others. They pinpoint the central role of *probing* (i.e., customer interviews), confirming the significance of “getting out of the building” in the method. It is probing that motivates business idea convergence—the key tenet of the method. Probing also motivates new business ideas (with corresponding hypotheses), a channel that is not often discussed by the method's adherents or anticipated by its critics (Felin et al., 2020). Intriguingly, probing can be helpful for teams to dislodge their original vision of the business idea, as needed, and find a new plan. Thus, we confirm that probing is a central component in the lean startup method, for reasons that are both more and less obvious.

In contrast, hypothesis formulation is not as central as expected. In fact, teams that formulate more hypotheses subsequently probe *fewer* of them. One possible reason is that one crisp, well-formulated hypothesis may lend itself

much better to probing than many vague and poorly formulated ones. If this reasoning is correct, it suggests a need to emphasize quality over quantity in hypothesis formulation. Emphasis on quantity seems particularly problematic given the benefits of probing outlined above. Another possible reason for the unexpected result hinges on differences in how different types of teams approach hypothesis formulation: perhaps MBAs think that if they have thought through a hypothesis, they no longer need to probe it, an interpretation that is consistent with MBAs' affinity for learning-by-thinking methods. In contrast, educationally diverse teams seem to embrace hypotheses and their probing more. Either way, the goal should be to formulate a meaningful, but not too large, number of hypotheses.

Our results also address some of the recent criticisms of the method. Lean startup is often criticized as a self-reinforcing loop from hypothesis formulation to probing and further formulation of new hypotheses, which may tire out and prevent scaling (Ladd, 2016). A closer examination of the results suggests a more nuanced interpretation however. The criticism may indeed be correct to highlight the need to formulate hypotheses only in moderation, but we also find that convergence leads to *fewer* new hypotheses; that is, there is a natural stopping mechanism. At least in the context of the NSF's I-Corps program, the concerns related to endless iteration with no clear end result in sight do not seem as alarming as previously assumed.

Our second key area of examination relates to boundary conditions. Our findings extend prior literature by suggesting that expertise using the “legacy” learning approaches can be a handicap in the team's use of newer approaches. Specifically, we find that MBA teams comfortable with learning-by-thinking methods are resistant to use the lean startup, a learning-by-doing method. In contrast, teams with no business education (no MBAs) are more likely to formulate new hypotheses and converge on new business ideas than business-educated (MBA) teams, suggesting more affinity for learning-by-doing methods. However, it is noteworthy that the lean startup can be an effective tool if business-educated MBAs try it in the first place, because probes by MBA-teams are more likely to lead to more new ideas (with the corresponding hypotheses) and more business idea convergence, suggesting that MBA training helps interpret data from probes. It is also intriguing that, while teams with an MBA are significantly *less* likely to conduct a *major* change in their business idea in general, they are *more* likely to make major changes if they engage in hypotheses probing (these additional results are available from the authors). This is particularly encouraging for corporate entrepreneurship, as many established corporations are likely to have MBA-led teams. So, to assuage Felin et al.'s (2020) fears, the use of the method appears to actually help MBA-teams make changes.

Fourth, our findings indicate that lean startup may be a particularly suitable method for technology-driven academics as a first step to start building their ventures. A core challenge is that academic scientists often “do not leave the building” (Franklin, Wright, & Lockett, 2001; Wright et al., 2009), meaning that they err on the side of too little iteration and too little customer information. Prior work has suggested (Wright et al., 2009) that connecting MBA-degree holding alumni with scientists could be a solution. Our results suggest that a more nuanced interpretation is in order and that future work should consider important boundary conditions for involvement of MBAs, particularly if learning-by-doing methods are used.

Finally, and more broadly, we provide both qualitative and quantitative data about *early-stage teams*. In particular, we help address a gap in quantitative work focused on these teams and their business-idea probing (cf. Demil, Lecocq, Ricart, & Zott, 2015). Our rich empirical data (hand-collected longitudinal data, week-by-week granularity during the intervention, and a lagged dependent variable and fixed effects to address team heterogeneity) are key to a more fine-grained understanding of how early-stage teams go about deciding what business idea to pursue.

## 6.2 | Limitations

As in any study, there are limitations. Because we studied real teams pursuing real business ideas, we were not given the freedom to assign a control group or to randomize, as one would have in a laboratory experiment. Thus, the causal interpretation of our results is limited. However, because all teams went through the same standardized intervention, as described above, we were able to isolate the effects of the method particularly well, shielding from



potential contamination by the “social context” (e.g., interactions with the rest of the organization that might blur the results in larger established firms that use the method).

Despite the advantages of the setting, future work should examine the use of the lean startup method outside of a standardized intervention such as I-Corps. Future research could also identify the many other possible boundary conditions on which the key relationships in the method might hinge (e.g., team composition, feedback from instructors and stakeholders, team formation process, program length)—that is, precisely the elements that were relatively standardized in I-Corps but are likely to differ in other implementations of the method.

It is also possible that lean startup as a method is particularly suitable in situations of greater uncertainty (e.g., Schmitt et al., 2018; Zhang et al., 2020), or for engineers and scientists, for whom technological development is relatively well thought out but market needs are not (Franklin et al., 2001), as is the case in I-Corps. If the conditions were reversed (e.g., teams lack technology expertise but understand the market well), the value of the method could be different. It is also unclear whether we can observe similar results in contexts where the solution is more evident for the team, such as tastier meals in a fast food restaurant, or where uncertainty about the opportunity landscape is low. Again, this is a path for future work.

Another possible limitation is that our sample teams are relatively small (about three). Although this is close to a prototypical team size in new firms (Eesley, Hsu, & Roberts, 2014), it is not clear how our findings extend to larger teams in more established organizations. Perhaps, in these larger teams, diversity plays a more prevailing role than what we found in our setting.

Moreover, we do not measure venture performance beyond the preliminary post hoc analyses reported in the online appendix that indicate that more intense interviewing is related to more ventures formed. Performance effects of the method are thus ripe for more exploration.

Finally, our measure of business idea convergence is from the perspective of the teams. Therefore, we are not able to identify whether what teams are learning is actually valuable. These questions offer intriguing ideas for future research.

### 6.3 | Conclusion

Although the scholarly community is still in its early stages of embracing lean startup as a learning-by-doing method for entrepreneurship, our research anchors the concept theoretically, and shows that it has potential to be an engine for innovation and change in organizations. Increasing consensus about theoretically grounded assumptions and boundary conditions of lean startup should help elevate the theory-practice understanding of the method and help improve it and similar methods in the future. Altogether, empirical research on entrepreneurial learning methods holds great promise in helping us understand better how new business ideas are formed and shaped.

#### ACKNOWLEDGEMENTS

We are grateful to Steve Blank, Tom Byers, Kathy Eisenhardt, Todd Morrill, Alexander Osterwalder, Yves Pigneur, and Michelle Rogan for discussing the research design and interpretation of the findings with us. Participants at the Strategic Management Society Special Conference in Santiago, the Frontiers of Entrepreneurship Conference at UNC, STVP Research Seminar at Stanford, and the Academy of Management Conference in Vancouver provided thoughtful feedback. The authors gratefully acknowledge the financial support from the National Science Foundation (Grant # 1305078), the Stanford Technology Ventures Program, Stanford Institute for Innovation in Developing Economies, Chile's National Commission for Scientific and Technological Research (CONICYT) Núcleo Milenio Research Center in Entrepreneurial Strategy Under Uncertainty (NS130028), ANID (The Chilean Agency for Research and Development) for its support through FONDECYT Iniciación Project number 11170537, CORFO, and the Aubrey Chernick Foundation. We thank Delara Fadavi, Macarena Gonzalez, Cassi Janakos, Nicholas Morrill, and the Bay Area I-Corps Node personnel for their logistical and research assistance. Much appreciation goes to Co-Editor Christoph Zott and two anonymous reviewers for their constructive feedback.

## ORCID

Michael Leatherbee  <https://orcid.org/0000-0002-9125-8271>

Riitta Katila  <https://orcid.org/0000-0001-8125-515X>

## ENDNOTES

- <sup>1</sup> We define “business idea” as startup's business concept including all the elements that would typically be included in a business plan (Delmar & Shane, 2003).
- <sup>2</sup> The lean startup method uses the terms “hypothesis testing,” “validation,” and “experimentation” in reference to the process of checking the assumptions underlying the business idea. Throughout this paper we use the term “probing,” as it more accurately reflects what the teams are actually doing, that is, interviewing customers, but not necessarily testing or running experiments in a scientifically rigorous way. We also use the term “convergence” rather than validation to clearly differentiate the method from scientific experiments that are not the method's focus. Convergence is defined as a team's belief that a particular assumption in the business-model canvas now has reduced level of uncertainty and is not considered an ongoing hypothesis to test.
- <sup>3</sup> Parker (2006), for example, found that older entrepreneurs were less willing to change business ideas even when external feedback suggests otherwise, but there is very little evidence about team expertise and “learning-by-doing” methods in general.
- <sup>4</sup> These findings are particularly interesting given prior literature's critique of the method that probing leads entrepreneurs to change ideas too frequently, become disheartened and give up (Ladd, 2016); or converge toward business ideas that are bounded by customers' explicit feedback, and thus obvious in nature (Felin, Gambardella, & Zenger, 2020).
- <sup>5</sup> The components of the canvas include value proposition, customer segments, exploitation activities, strategic partners, resources, customer relationships, distribution channels, revenue streams, and cost structure.
- <sup>6</sup> Other examples include studies that have shown that management training (i.e., finance or operations) improved business exploitation and efficiency (Bloom, Eifert, Mahajan, Mckenzie, & Roberts, 2013; Bruhn, Karlan, & Schoar, 2010).
- <sup>7</sup> Prior industry experience is typically cited as the major barrier. For instance, Furr et al.'s (2012) analysis of photovoltaic companies over a 15-year period showed that if the CEO's prior experience was in the solar industry, the firm was more resistant to switch to new photovoltaic technologies. In contrast, CEOs with no industry experience were quick to embrace technological changes.
- <sup>8</sup> Steve Blank advised the NSF on the program's design and implementation.
- <sup>9</sup> Observations drop to 1,064 when the lagged dependent variable is used in regression models.
- <sup>10</sup> We had access to two sets of data: our own surveys and those of NSF. Response rates per team in our follow-on surveys range from 75 to 95% for the 6-month post survey. The NSF data were anonymized, so we could not link respondents to their teams or to individual characteristics.
- <sup>11</sup> Note that team members with Ph.D. degrees have overwhelmingly received their doctorate training in science and engineering fields (i.e., typical NSF grantees) so they have *not* been taught the learning-by-thinking skills in business that are the staple of MBA education.
- <sup>12</sup> Because Poisson fixed-effects models drop time-invariant covariates, they allow for testing Hypotheses H4 through H7 only. Results are consistent with the reported GEE negative binomial ones (results available from the authors).
- <sup>13</sup> For this analysis, we collected objective data for our teams about whether a venture had been formed before the I-Corps program. For robustness, we triangulated our data with anonymized I-Corps participant surveys that NSF shared with us. In these self-reported data, 23% of the respondents report that a venture was formed before the start of the program, increasing confidence in our team-level data that reports a corresponding number of 20%.

## REFERENCES

- Aghion, P., Bolton, P., Harris, C., & Jullien, B. (1991). Optimal learning by experimentation. *The Review of Economic Studies*, 58(4), 621.
- Almandoz, J., & Tilcsik, A. (2016). When experts become liabilities: Domain experts on boards and organizational failure. *Academy of Management Journal*, 59(4), 1124–1149.
- Andries, P., Debackere, K., & van Looy, B. (2013). Simultaneous experimentation as a learning strategy: Business model development under uncertainty. *Strategic Entrepreneurship Journal*, 7(4), 288–310.
- Arkilic, E. (2019). Raising the NSF innovation corps. In M. Wisnioski, E. Hintz, & M. Stettler Kleine (Eds.), *Does America need more innovators?* (pp. 69–82). Cambridge, MA: The MIT Press.

- Baker, T., & Nelson, R. E. (2005). Creating something from nothing: Resource construction through entrepreneurial bricolage. *Administrative Science Quarterly*, 50, 329–366.
- Barker, V. L., & Duhaime, I. M. (1997). Strategic change in the turnaround process: Theory and empirical evidence. *Strategic Management Journal*, 18(1), 13–38.
- Beckman, C. M. (2006). The influence of founding team company affiliations on firm behavior. *Academy of Management Journal*, 49(4), 741–758.
- Belz, A. P., Giga A., & Zapatero F. (2018). Early-stage financing for academic entrepreneurs (In review). Retrieved from <https://ssrn.com/abstract=3221288>
- Blank, S. (2003). *The four steps to the epiphany: Successful strategies for products that win*.
- Blank, S. (2013). Why the lean start-up changes everything. *Harvard Business Review*, 91(5), 63–72.
- Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., & Roberts, J. (2013). Does management matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1), 1–51.
- Bruhn, B. M., Karlan, D., & Schoar, A. (2010). What capital is missing in developing countries? *The American Economic Review*, 100(2), 629–633.
- Callander, S. (2011). Searching and learning by trial and error. *American Economic Review*, 101(6), 2277–2308.
- 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), 503–1004.
- Cialdini, R. B. (2001). The science of persuasion. *Scientific American*, 284(2), 76–81.
- Clarysse, B., Wright, M., & Hove, J. V. (2015). *A look inside accelerators: Building businesses*. London, England: Nesta.
- Contigiani, A., & Levinthal, D. A. (2019). Situating the construct of lean startup: Adjacent ‘conversations’ and possible future directions. *Industrial and Corporate Change*, 28(3), 551–564.
- Crilly, N. (2018). ‘Fixation’ and ‘the pivot’: Balancing persistence with flexibility in design and entrepreneurship. *International Journal of Design Creativity and Innovation*, 6(1–2), 52–65.
- Davis, M. S. (1971). That’s interesting: Towards a phenomenology of sociology and a sociology of phenomenology. *Philosophy and Social Science*, 4, 309–344.
- Delmar, F., & Shane, S. A. (2003). Does business planning facilitate the development of new ventures? *Strategic Management Journal*, 24(April), 1165–1185.
- Demil, B., Lecocq, X., Ricart, J. E., & Zott, C. (2015). Introduction to the SEJ special issue: Business models within the domain of strategic entrepreneurship. *Strategic Entrepreneurship Journal*, 9, 1–11.
- Dyer, J. H., Gregersen, H. B., Christensen, C. (2008). Entrepreneur behaviors, opportunity recognition, and the origins of innovative ventures. *Strategic Entrepreneurship Journal*, 2(4), 317–338.
- Edmondson, A., Bohmer, R., & Pisano, G. (2001). Disrupted routines: Team learning and new technology implementation in hospitals. *Administrative Science Quarterly*, 46(4), 685–716.
- Eesley, C. E., Hsu, D. H., & Roberts, E. B. (2014). The contingent effects of top management teams on venture performance. *Strategic Management Journal*, 35(12), 1798–1871.
- Eisenmann, T., Ries, E., & Dillard, S. (2013). Hypothesis-driven entrepreneurship: The lean startup. *Harvard Business School (July)*, 1–26.
- Eisenhardt, K. M., & Schoonhoven, C. B. (1990). Organizational growth: Linking founding team, strategy, environment, and growth among U.S. semiconductor ventures, 1978–1988. *Administrative Science Quarterly*, 35(3), 504–529.
- Felin, T., Gambardella, A., Stern, S., & Zenger, T. (2020). Lean startup and the business model: Experimentation revisited. *Long Range Planning*, 53(4), 101889.
- Felin, T., & Zenger, T. R. (2009). Entrepreneurs as theorists: On the origins of collective beliefs and novel strategies. *Strategic Entrepreneurship Journal*, 14(6), 127–146.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York, NY: McGraw-Hill.
- Franklin, S. J., Wright, M., & Lockett, A. (2001). Academic and surrogate entrepreneurs in university spin-out companies. *Journal of Technology Transfer*, 26(1–2), 127–141.
- Furr, N. R., Cavarretta, F., & Garg, S. (2012). Who changes course? The role of domain knowledge and novel framing in making technology changes. *Strategic Entrepreneurship Journal*, 6(3), 236–256.
- Gans, J. S., Stern, S., & Wu, J. (2019). Foundations of entrepreneurial strategy. *Strategic Management Journal*, 40(5), 736–756.
- Gary, M. S., & Wood, R. E. (2011). Mental models, decision rules, and performance heterogeneity. *Strategic Management Journal*, 32(6), 569–594.
- Gavetti, G., Levinthal, D. A., & Rivkin, J. W. (2005). Strategy making in novel and complex worlds: The power of analogy. *Strategic Management Journal*, 26(8), 691–712.
- Gonzalez-Urbe, J., & Leatherbee, M. (2018). The effects of business accelerators on venture performance: Evidence from start-up Chile. *Review of Financial Studies*, 31(4), 1566–1603.

- Grimes, M. G. (2017). The pivot: How founders respond to feedback through idea and identity work. *Academy of Management Journal*, 61(5), 1692–1717.
- Gruber, M., MacMillan, I. C., & Thompson, J. D. (2013). Escaping the prior knowledge corridor: What shapes the number and variety of market opportunities identified before market entry of technology start-ups? *Organization Science*, 24(1), 280–300.
- Hallen, B. L., & Pahnke, E. C. (2016). When do entrepreneurs accurately evaluate venture capital firms' track records? *Academy of Management Journal*, 59(5), 1535–1560.
- Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Journal*, 32(4), 1199–1228.
- Heckman, J. J., & Borjas, G. J. (1980). Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Economica*, 47(187), 247–283.
- Hilbe, J. (2007). *Negative binomial regression*. Cambridge, England: Cambridge University Press.
- Jung, J., & Shin, T. (2019). Learning not to diversify: The transformation of graduate business education and the decline of diversifying acquisitions. *Administrative Science Quarterly*, 64, 337–369.
- Katila, R., & Ahuja, G. (2002). Something old, something new: A longitudinal study of search behavior and new product introduction. *The Academy of Management Journal*, 45(6), 1183–1194.
- Katila, R., & Shane, S. A. (2005). When does lack of resources make new firms innovative? *Academy of Management Journal*, 48(5), 814–829.
- Katila, R., Thatchenkery, S., Christensen, M. Q., & Zenios, S. (2017). Is there a doctor in the house? New ventures, expert end-users in organizational roles, and innovation. *Academy of Management Journal*, 60(6), 2415–2437.
- Kennedy, P. (1998). *A guide to econometrics*. Malden, MA: Blackwell Publishers.
- Kerr, W. R., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives*, 28(3), 25–48.
- Koning, R., Hasan, S., & Chatterji, A. (2019). Experimentation and startup performance: Evidence from A/B testing. *NBER Working Paper w26278*.
- Ladd, T. (2016). The limits of the lean startup method. *Harvard Business Review*. <https://hbr.org/2016/03/the-limits-of-the-lean-startup-method> [23 November 2018].
- Lerner, J., & Malmendier, U. (2013). With a little help from my (random) friends: Success and failure in post-business school entrepreneurship. *The Review of Financial Studies*, 26(10), 2411–2452.
- Li, Q., Maggitti, P. G., Smith, K. G., Tesluk, P. E., & Katila, R. (2013). Top management attention to innovation: The role of search selection and intensity in new product teams. *Academy of Management Journal*, 56(3), 893–916.
- Liang, K.-Y., & Zeger, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1), 13–22.
- Louis, M. R., & Sutton, R. I. (1991). Switching cognitive gears: From habits to mind to active thinking.pdf. *Human Relations*, 44(1), 55–76.
- March, J. G., & Simon, H. A. (1958). *Cognitive limits on rationality*. Organizations: McGraw-Hill.
- McDonald, R., & Gao, C. (2019). Pivoting Isn't enough? Managing strategic reorientation in new ventures. *Organization Science*, 30(6), 1125–1393.
- McGrath, R. G., & MacMillan, I. C. (1995). Discovery driven planning. *Harvard Business Review*, 73(4), 44–54.
- Miner, A. S., Bassoff, P., & Moorman, C. (2001). Organizational and improvisation learning: A field study. *Administrative Science Quarterly*, 46(2), 304–337.
- Navarro. (2017). The MBA core curricula of top-ranked U.S. business schools: A study of failure? *The Academy of Management Learning and Education*, 7(1), 108–123.
- Osterwalder, A., & Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers*. Hoboken, NJ: John Wiley & Sons.
- Ott, T. E., Eisenhardt, K. M., & Bingham, C. B. (2017). Strategy formation in entrepreneurial settings: Past insights and future directions. *Strategic Entrepreneurship Journal*, 11(3), 306–325.
- Parker, S. C. (2006). Learning about the unknown: How fast do entrepreneurs adjust their beliefs? *Journal of Business Venturing*, 21(1), 1–26.
- Porter, M. E. (1980). Industry structure and competitive strategy: Keys to profitability. *Financial Analysts Journal*, 36(4), 30–41.
- Priem, R. L., & Rosenstein, J. (2000). Is Organization theory obvious to practitioners? A test of one established theory. *Organization Science*, 11(5), 509–524.
- Rathje, J., & Katila, R. (2021). Enabling technologies and the role of private firms. *Strategy Science*.
- Ries, E. (2011). *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. New York, NY: Crown Business.
- Sarasvathy, S. D. (2001). Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review*, 26(2), 243–263.

- Schmitt, A., Rosing, K., Zhang, S. X., & Leatherbee, M. (2018). A dynamic model of entrepreneurial uncertainty and business opportunity identification: Exploration as a mediator and entrepreneurial self-efficacy as a moderator. *Entrepreneurship Theory and Practice*, 42(6), 835–859.
- Sorenson, O. (2003). Strategy as quasi-experimentation. *Strategic Organization*, 1(3), 337–343.
- Sullivan, B. N. (2010). Competition and beyond: Problems and attention allocation in the organizational rulemaking process. *Organization Science*, 21, 432–450.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509–533.
- Thomke, S. (2003). *Experimentation matters: Unlocking the potential of new technologies for innovation*. Boston, MA: Harvard Business Press.
- Williamson, O. E. (2019). Transaction-cost economics: The governance of contractual relations. *Corporate Governance: Values, Ethics and Leadership*, 22(2), 133–162.
- Wright, M., Piva, E., Mosey, S., & Lockett, A. (2009). Academic entrepreneurship and business schools. *Journal of Technology Transfer*, 34(6), 560–587.
- Zhang, S. X., Gao, R., Odeh, N., & Leatherbee, M. (2020). A microfoundational model of real options reasoning: The roles of individual search propensity and perceived uncertainty. *Strategic Entrepreneurship Journal*.
- Zott, C., & Amit, R. (2007). Business model design and the performance of entrepreneurial firms. *Organization Science*, 18(2), 181–199.

## SUPPORTING INFORMATION

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

**How to cite this article:** Leatherbee M, Katila R. The lean startup method: Early-stage teams and hypothesis-based probing of business ideas. *Strategic Entrepreneurship Journal*. 2020;14:570–593. <https://doi.org/10.1002/sej.1373>