INSEAD The Business School for the World®

Working Paper

2022/27/EFE

(Revised version of 2021/40/EFE)

When do Entrepreneurs Benefit from Acting Like Scientists? A Field Experiment in the UK

Elena Novelli
University of London, elena.novelli.1@city.ac.uk

Chiara Spina
INSEAD, chiara.spina@insead.edu

This paper investigates the role of a firm's degree of business development—defined as the extent to which the firm has already developed into a continuous and sustainable market participant—in moderating the benefits of a scientific approach to decision-making, which combines cognition and action-based components. We conducted a field experiment with 261 UK entrepreneurs with different degrees of business development. Our results show that all treated firms experience a positive effect on firm size, but not all on firm revenue. The treatment is associated with a positive effect on firm revenue for firms with a more advanced degree of business development, but a negative effect for firms with a lower degree of business development. We elaborate on the implications of these results for future re-search.

Keywords: Entrepreneurial Strategy; Experimentation; Field Experiment; Innovation; Value Creation

Acknowledgments: We gratefully acknowledge financial support from the Department for Business, Energy & Industrial Strategy (BEIS) of the UK Government through the BEIS Business Basics Grant, and support from the Innovation Growth Lab. This project was also made possible by support from The Business School (Formerly Cass), Bocconi University, and Oxford University. We thank scholars who provided comments on previous versions of this manuscript, particularly Alfonso Gambardella, Myriam Mariani, Teppo Felin, Arnaldo Camuffo, Ethan Mollick, and attendees from presentations at Warwick, ESADE, University of Bologna, EPFL, the Wharton Technology Conference, the Strategic Management Society Entrepreneurship and Strategy Seminar Series.

1. INTRODUCTION

A fundamental question for research in strategy concerns how decision-making approaches impact firm performance. This question is particularly relevant in entrepreneurial settings, where strategy makers face uncertainty in multiple domains, from technology (Folta, 1998; Gans & Stern, 2003; McGrath, 1997) to market preferences (Foss &Klein, 2012; Kirtley & O'Mahony, 2020; Sarasvathy, 2009), and the resolution of uncertainty is often endogenous to action (Agarwal et al., 2007; Moeen et al., 2020; Ott & Eisenhardt, 2020). Recent research highlights the relevance of a "scientific approach to decision making" (Camuffo et al., 2020). This approach combines the formulation of a theory of the problem under investigation (Felin et al., 2020a and b), the development of hypotheses that flow logically from the theory (Ehrig & Schmidt, 2022), the systematic collection of evidence that can support or refute those hypotheses (Bloom et al., 2012; Ries, 2011), and its disciplined assessment (Murray & Tripsas, 2004). It is called "scientific" as it resembles the approach followed by scientists as they develop new knowledge (Zellweger & Zenger, 2021). It addresses uncertainty by integrating a *cognitive*-based approach to decision-making (i.e., formulating a theory about the problem faced as in Csaszar & Laureiro-Martinez, 2018; Bingham & Eisenhardt, 2011) with an *action*-based component (Ghosh et al., 2020; Kohavi & Thomke, 2017; Ott, Eisenhardt & Bingham, 2017).

Prior studies suggest that a scientific approach to decision-making—and, more generally, approaches that combine cognition and action—can lead to superior learning and the definition of more effective strategies (Eisenhardt & Bingham, 2017; McDonald & Eisenhardt, 2020; Camuffo et al., 2020). However, work in this area is still in its infancy, and little is known about the conditions that determine the effectiveness of this type of approach. Despite recent theoretical and empirical work that points to its relevance (Camuffo et al., 2020; Zellweger &

Zenger, 2021), research in this stream has yet to examine, theoretically or empirically, its heterogeneous effects. This paper addresses this gap by advancing that a firm's *degree of business development* is fundamental in determining the impact of a scientific approach on its performance. A firm's degree of business development has been defined as the extent to which a firm has evolved from a nascent venture into a continuous and sustainable market participant (Churchill & Lewis, 1983; Davidsson, 2004; Reynolds, 2017). This evolution at the firm level is similar to what happens at the industry level where groups of firms develop through different phases, from preincubation to maturity (Agarwal et al., 2016; Moeen & Agarwal, 2017). Empirical evidence on approaches to decision-making is based mostly on studies involving firms at a relatively advanced stage of development (Bloom & VanReenen, 2010; Pillai et al., 2020; Yang et al, 2020). Far fewer studies have examined firms with a lower degree of business development and, as Kirtley & O'Mahony (2020) note, it is difficult to generalize results from more to less developed businesses. Thus, the performance implications of systematic decision-making approaches for the latter tend to be ambiguous (Bruhn et al., 2018; Camuffo et al. 2020; Karlan et al., 2015). Based on these premises, our study addresses the following research question: *What is the role of a firm's degree of business development in moderating the benefits of a scientific approach to decision-making?*

To answer this question, we build on strategy literature on firm evolution as well as literature on the entrepreneurial process, which suggest that firms' decision-making focuses on different problems as they develop over time: Firms with a lower degree of business development focus on broad questions such as the validation of the market (Churchill & Lewis, 1983). At these earlier stages the core elements of their business model still need to be defined (Siggelkow, 2002). In the later stages of development, firms typically face questions related to scaling up and generating revenue (Churchill & Lewis, 1983), and changing a core element is a rare event (Siggelkow, 2002, p. 127; Eisenhardt & Brown, 1999). We elaborate theoretically on how this evolution moderates the relationship between the use of a scientific approach and performance by affecting a) the degree of uncertainty resolution that the scientific approach is likely to support, (b) the quality of the value proposition to which the scientific approach is applied, and (c) the extent to which the use of the method directly results in changes in the value proposition (i.e., pivots). We predict that firm performance increases as a result of the implementation of a scientific approach to decision-making, but that this performance increase will be higher for firms with a higher degree of business development.

We test these predictions via a 9-month randomized control trial (RCT) with 261 UK entrepreneurial firms attending a strategy training program. Both treated and control firms underwent a training course and were taught to use elements of cognitive-based decision-making (reasoning through strategy frameworks and tools) and action-based decision-making (using data-gathering and testing techniques), for a total of 21 hours of training spread across 7 sessions. Entrepreneurs in the control group were encouraged to apply these components to their business but, typical for any business training course, they were not given specific instruction on how to do so. Those in the treatment group, instead, were taught to apply these concepts and tools using a scientific approach, i.e., combining cognitive- and action-based components. They were taught to use strategy frameworks to formally develop a theory of the problems faced and hypotheses consistent with that theory, test those predictions, and systematically evaluate the results. We collected observations on the performance of firms participating in our program before the intervention and for 9 months since the beginning of the intervention. We focused on two performance dimensions: firm size and firm revenue. Results show that, in line with our predictions, all treated firms achieved superior performance in terms of firm size, and firms with a higher degree of business development experienced a positive marginal effect of the treatment. However, if we look at firm revenue, an interesting picture emerges. While, on average, the treatment had a non-significant effect on firm revenue, once we distinguished between firms that started the training with a lower (vs. higher) degree of business development, we observed that the treatment had a negative impact for firms with a lower degree of business development but a positive impact for firms with a higher degree of business development.

Our study makes three main contributions. First, it contributes to strategy research on decision-making by providing causal evidence of the relationship between the use of a scientific approach to decision-making and firm performance. We show that this effect varies depending on the performance dimension analyzed and is contingent on the firm's degree of business of development. In doing so, we address an important empirical shortcoming of prior research by including a heterogeneous set of firms in terms of their degree of business development. Our results help understand earlier mixed findings on the relationship between the use of systematic managerial practices and firm performance (Bloom & Van Reenen, 2010; Bruhn et al., 2018; Camuffo et al. 2020; Karlan et al., 2015; Yang et al., 2020). Second, it contributes to research on entrepreneurial strategy. Both scholarly and practitioner-oriented work advocate for the value of experimentation to firms with a lower degree of business development (Kerr et al. 2014; Ries, 2011), whereas other studies (Agrawal et al., 2021; Gans et al., 2019; McDonald & Eisenhardt, 2020; Pillai et al. 2020) suggest that experimentation may be more fruitful for firms that are more developed and have already made key choices. Our results flesh out the nuances of these apparently conflicting insights by showing that firms with lower and higher degrees of business development all benefit from the use of this type of approach, but –because they address different types of problems—they benefit on different dimensions. Specifically, all firms benefit from approaches that use experimentation in terms of size, but the effect on revenue, at least within a relatively limited time window, is visible only for firms that already have a higher degree of business development.

Beyond its academic contribution, our study offers insight to policy makers and institutions looking to foster economic growth through programs that support innovation. Initiatives that offer training with a view to stimulate growth and productivity often yield limited results (Lerner, 2009; McKenzie, 2021). This study suggests a possible explanation: Firms benefit from training programs in terms of increased size, but not necessarily in terms of increased revenue, at least until they have reached a sufficient degree of business development. Awareness of how firms benefit from training programs could be a starting point for a more efficient admission process, and the provision of targeted forms of support for firms depending on their degree of business development.

2. THEORY

2.1 The effect of scientific decision-making on firm performance

Prior literature has emphasized two different types of structured processes that firms can employ when making decisions in the face of uncertainty. A first stream emphasizes the benefits of a *cognitive-based approach* to decision-making. This approach is centered on the development of a theory (Felin & Zenger, 2009; Felin & Zenger, 2017; Zenger, 2015), or mental representations (Csaszar & Laureiro-Martinez, 2018; Gary & Wood, 2011) as drivers of business innovation, performance heterogeneity, and superior strategy. A second stream emphasizes, instead, the importance of a more *action-based approach* to decision-making, relying on acting and then learning from the experience to guide subsequent action (Bingham & Eisenhardt, 2011; Leatherbee & Katila, 2020; McGrath, 2001). This approach ranges from trial and error to controlled variation of activities via experimentation to generate relevant feedback (Gans et al., 2019; McGrath, 1999; Murray & Tripsas, 2004; Ott et al., 2017; Pillai et al., 2020; Shepherd & Gruber, 2020). It also involves the systematic evaluation of the evidence that entrepreneurs gather (Bennett & Chatterji, 2019; Camuffo et al., 2020; Chatterji et al., 2018; Cohen et al., 2019)[1].

Cognitive- and action-based approaches can be complementary and mutually reinforcing (Gavetti & Levinthal, 2000; Ott, Eisenhardt & Bingham, 2017). Eisenhardt and Bingham (2017, p. 247) underline the importance of combining "*thinking and doing*" and of a holistic approach to decision-making that involves both a cognitive understanding of the "*playing field*" and action/learning via experimentation. McDonald and Eisenhardt (2020) emphasize the benefits of testing the assumptions underlying the cognitive templates used by firms, such as business models. They suggest that combining cognition with action reduces the uncertainty faced by entrepreneurs regarding the most appropriate model to use and helps them ground models in realistic and relevant information, leading to quicker and faster learning. In the same spirit, Camuffo et al. (2020) emphasize how scientists' rigor in the discovery process, which *simultaneously* involves a *cognitive* component (i.e., theory development and the formulation of hypotheses) and an *action-based* component (i.e., testing and evaluation of evidence), can be successfully applied to entrepreneurial decision-making. They call this the "scientific approach to decision-making" and we use the same terminology in this paper. Each individual component of the scientific approach leads to superior performance, but there are synergies when using these different components sequentially (Zellweger & Zenger, 2021).

A scientific approach to decision making involves four key steps: (1) the development of a theory; (2) its articulation into hypotheses that logically flow from it; (3) the collection of evidence that can either support or refute the hypotheses; (4) the disciplined assessment of the evidence collected. First, entrepreneurs who employ a scientific approach frame the problem

[1] Several useful toolkits aimed at supporting practitioners in experimenting have emerged in this area. They focus on different aspects of the experimentation process such as how to identify new business ideas (Gruber & Tal, 2017), how to strategize after identifying the initial business idea (Osterwalder, Pigneur & Clark, 2010), and how to experiment while searching for the right product-market fit (Ries, 2011).

they face using a theory—a cognitive representation of how their business generates value (Ehrig & Schmidt, 2022; Felin & Zenger, 2017). This helps them understand more clearly what the key dimensions of the problem are and what they should focus their attention on (Camuffo et al., 2020; Felin et al., 2020a). Second, the articulation of their theory into clear, falsifiable, predictions (Felin & Zenger, 2016) helps entrepreneurs modularize the problem into smaller, decomposable, and more addressable blocks, which reduces the level of causal ambiguity (Felin et al., 2020b; Leatherbee & Katila, 2020) and helps them generate more innovative ideas via recombination and modular addition (Ott & Eisenhardt, 2020). Third, gathering feedback through rigorous tests provides valuable feedback that can help entrepreneurs distinguish between businesses with good and bad outcomes (Thomke, 2003; Murray & Tripsas, 2004; Bingham & Eisenhardt, 2011; Ries, 2011; Gruber et al., 2013; Gans et al., 2019; Pillai et al., 2020; Shepherd & Gruber, 2020). Fourth, the systematic and critical assessment of the evidence gathered helps compare the signals collected against an ideal threshold to find support for key hypotheses (Boulding et al., 1997; Keil & Mahring, 2010).

The use of a scientific approach combines the four elements described above in a synergistic way to help resolve the uncertainty associated with each choice faced by entrepreneurs (Packard et al., 2017), guiding them toward more informed decisions based on logic reasoning and systematic testing (Zellweger & Zenger, 2021). Without a scientific approach, entrepreneurs lack focus in their thinking and action and tend to choose based on gut feelings (Hodgkinson & Sadler-Smith, 2018) or available evidence. Often, this process is not systematic (Forbes, 2005; Bennet & Chatterji, 2019), and thus is likely to produce untargeted variation with trial-and-error exploration (McBride & Wuebker, 2020). Resembling a process in which the decision maker combines cognitive and experiential search, the use of a theory both seeds and constrains the subsequent process of experimentation: It reduces the myopia of local search by seeding the search process in promising regions of the landscape and preventing it from taking root in less attractive regions (Gavetti & Levinthal, 2000; Gavetti & Rivkin, 2007; Levinthal, 2017). Since more informed decisions should ultimately result in superior performance (March, 1991; Gavetti & Menon, 2016), this suggests a positive effect on firm performance.

An example. To illustrate how a scientific approach works, consider the following hypothetical example related to an innovative firm called Palette that produces and sells vegetarian food. If Palette's founder, Felicia, were acting like a scientific entrepreneur, she would start by elaborating a theory that explained why her company creates value for customers. For instance, Felicia's theory might be that vegetarian food will rise in popularity because it represents a healthier, more sustainable choice than meat products and does not harm animals; however, negative perceptions of the taste of vegetarian food may limit its appeal. In line with this reasoning, Felicia would conclude that value can be generated by finding innovative ways of cooking vegetarian food to make it tastier. Her theory might also posit that younger consumers could be the ideal target market for her product as they care more about sustainability and a healthier lifestyle and are willing to pay a premium for tasty vegetarian food. As a scientific entrepreneur, Felicia would then combine the cognitive approach described so far with an action-based approach, collecting data to test and validate her theory. To do so, she would first derive the following testable hypotheses from the broad theory using a modular and sequential approach (i.e., starting from the most important hypothesis, one that allows her to understand if her business is viable): (1) her vegetarian food is as tasty as its non-vegetarian equivalent, and (2) conditional on the first hypothesis being supported, it is more likely to be purchased by younger customers. To put the first hypothesis to the test, she would conduct a blind taste test, inviting individuals to sample one of her vegetarian dishes alongside a non-vegetarian dish and rate them on their taste. She would pre-specify that she expects at least 75% of individuals to indicate that the two dishes are equally tasty, as she needs to be sure that the collected responses are different from random ones (where about 50% of customers would indicate a preference for either one of the two dishes). To test her second hypothesis, she could sell vegetarian food via a pop-up stall offering free samples prior to purchase and observe whether older or younger customers are more likely to purchase food after tasting it. Based on the results obtained, she could then use the findings to evaluate what to do next, e.g., whether to change the way the food is prepared.

In contrast, if Felicia were to act like a non-scientific entrepreneur, she would sell a product without a clear idea of the features valued by customers. In the absence of a theory, her decisions would be based on "what seems to work." She might sell different types of vegetarian food that vary in characteristics (i.e., produced with organic ingredients, gluten-free, etc.). Yet because these decisions would be neither driven by a clear theory nor formulated into distinct hypotheses to be tested individually, if she were, for example, to sell more food produced with organic ingredients, she would not know if it was because of the taste, the type of ingredients, or some other feature. Any evidence from sales data would be unlikely to help her understand what to do because there would be causal ambiguity in her test design. Overall, in the absence of a clear assessment of what customers value, new products or services would be introduced randomly rather than within a clear frame of reference. As this example clarifies, a non-scientific entrepreneur may serendipitously perform well, but the lack of clear theory, hypotheses, testing, and learning about how value is created for customers makes the venture less likely to consistently succeed. In line with prior research, we therefore suggest the following baseline hypothesis:

Hypothesis 1. A scientific approach to decision-making has a positive impact on firm performance.

2.2 A Scientific Approach for Firms with Different Degrees of Business Development

Prior research in strategy and entrepreneurship has emphasized that firms evolve by going through different stages (Cardon et al., 2017; Churchill & Lewis, 1983; Davidsson, 2004;

Siggelkow, 2002). This evolution can be thought of as a continuous journey from the start-up stage—in which the firm is merely "*a progressing nascent venture*"—to a more developed phase—in which the firm becomes "*a more sustainable and continued market participant*" (Davidsson, 2004, page 202). In this paper, we refer to this continuous construct as the *degree of business development*, defined as the extent to which the firm has already developed into a continuous and sustainable market participant (Churchill & Lewis 1983; Davidsson, 2004; Reynolds, 2017)[2]. Research on this topic notes that firms experience common problems and decisions about their business model when they are at similar stages in their development (Churchill and Lewis, 1983; Reynolds, 2017; Siggelkow, 2002) and that as firms progress from lower stages to higher stages of development the problems they face and decisions they have to make differ substantially (Alexy et al. 2021; Ambos & Birkinshaw, 2007; Santos and Eisenhardt, 2009; Siggelkow, 2002).

We conceptualize and measure the degree of business development as a continuous variable, but to better clarify the variation entailed by the continuum we focus our arguments on its two extremes, i.e., firms with lower degrees vs firms with higher degrees of business development. In the lower stages of development, the core elements of the business model still need to be defined (Eisenhardt & Brown, 1999). Firms focus on broad questions such as the development of the market proposition and the validation of the market: "*Can we get enough customers, deliver our products, and provide services well enough to become a viable business?*" (Churchill & Lewis, 1983, page 33). In later stages, changing a core element of the value proposition is a rare event, and firms are mostly focused on "*supporting or elaborating*

[2] Different authors refer to the construct of degree of business development with slightly different terms. Greiner (1972) talks about "stages of development"; Davidsson (2004) refers to it as to "emergence success." Other authors provide discrete classifications and use terms such as "existence" (Churchill & Lewis, 1983) or even "pre-existence," "conception," "gestation," "infancy and toddlerhood," and "childhood and adolescence" (Cardon et al. 2017) to refer to the earlier phases and terms such as "survival," "success," and "take off" (Churchill & Lewis, 1983) or "maturity" (Cardon et al., 2017) to refer to the later phases. We use the label "degree of business development" as we believe that it more clearly characterizes the continuous nature of the construct.

its core elements" to pursue fit and consistency (Siggelkow, 2002, p. 127). Firms typically face questions related to scaling up and generating revenue: "*The key problem shifts from mere existence to the relationship between revenues and expenses. (…) The main issues are as follows: In the short run, can we generate enough cash to break even and to cover the repair or replacement of our capital assets as they wear out? Can we, at a minimum, generate enough cash flow to stay in business and to finance growth to a size that is sufficiently large, given our industry and market niche, to earn an economic return on our assets and labor?*" (Churchill & Lewis, 1983, p. 34).

These differences in firms' primary concerns lead to differences in the use of decision-making approaches for firms at different degrees of development (Alexy et al. 2021). But the existing literature on the effect of systematic decision-making practices on performance has focused either on more developed or less developed businesses in isolation. In the context of firms with a *higher degree of development* Bloom and Van Reenen (2007), for instance, find that better management practices are associated with higher performance. Yang et al. (2020) find an association between the use of highly formalized, rigorous, deliberate processes by large firms and growth in employment. Other studies show that approaches to learning that rely on structure and codification are associated with superior performance in the context of acquisition integration for large and experienced acquirers (Heimeriks, et al., 2012; Zollo & Winter, 2004).

When it comes to the performance implications of systematic approaches for *less-developed businesses*, however, evidence is limited and mixed. Karlan et al. (2015) conducts an RCT in urban Ghana and finds that microenterprises exposed to systematic decision-making made changes in business practices, but they did not result in higher profits. In conducting an RCT with start-ups with a very low degree of development, Camuffo et al. (2020) find that treated entrepreneurs using a scientific approach are more likely to pivot to alternative ideas and terminate their projects early, but the effect on firm performance shows a high degree of variability. In related conceptual research, Zellweger and Zenger (2021) speculate that the use of a scientific approach to decision-making might have different implications for firms at different degrees of business development in terms of how they use the approach, learning potential, and performance. With regards to testing, Koning et al. (2022) find that A/B testing benefits performance slightly more for early-stage startups, but on specific metrics such as page views and new product features.

This suggests that investigating the extent to which the degree of business development moderates the relationship between the use of a scientific approach and performance is important. We propose that the use of a scientific approach to decision-making will have different performance implications when applied by firms at different degrees of business development. More developed firms are more likely to use the scientific approach to fine-tune an existing value proposition and improve their revenue generating model, supporting the core elements of the value proposition; less developed firms, instead, are more likely to apply the approach to the earlier questions of market validation and product development and to define the core elements of the value proposition (Churchill & Lewis, 1983; Siggelkow, 2002). This will moderate the relationship between the use of a scientific approach and performance by affecting (a) the degree of uncertainty resolution that the scientific approach is likely to support; (b) the quality of the value proposition to which the scientific approach is applied, and (c) the extent to which the use of the method will directly translate in pivoting, i.e., making strategic changes to the value proposition.

First, the application of the scientific approach will help firms reduce the uncertainty that they face, but this effect is likely stronger for firms with a higher degree of business development. Because more developed businesses have already operationalized the core elements of their business model (Davidsson, 2004; Siggelkow, 2002), their search space is more concentrated on the elaboration rather than on the definition of these core elements, and uncertainty reduction via a scientific approach is easier in this context. The identification of the theory components on which the entrepreneur is still unclear and their translation into more precise hypotheses is easier, as is defining and interpreting a full range of precise tests. Normally, each dimension of a value proposition creates a set of interdependencies (Ghosh, 2021) and will have to be evaluated holding other conditions constant when conducting tests. When the value proposition is already defined and key choices have already been made (for instance, about target customers, sales channels, etc.), gathering precise feedback will require fewer test permutations. Research notes that experiments conducted after many of the key choices have been made leads to "rich" information (Greenstein, 2012; Pillai et al. 2020; Rosenberg, 1982; Stern, 2005), reducing the level of uncertainty faced by the decision maker. This argument is also consistent with findings from Levinthal and Posen (2007), Ethiraj and Levinthal (2009), and Csaszar and Levinthal (2016) that show that reducing the number of dimensions of the business landscape a decision-maker faces is beneficial to simplifying decision-making and improving performance. In a similar vein, Gavetti and Levinthal (2001) note that actors faced with narrower problems obtain less ambiguous and faster feedback. The reduction of uncertainty provides entrepreneurs with a clear course of action regarding the changes that are required to improve the proposition, with a likely direct effect on performance.

Conversely, when a firm has a lower degree of business development and entrepreneurs are still making decisions on the core components of their business, a scientific approach might lead to further exploration and possibly uncertainty increase as opposed to uncertainty reduction. This is consistent with prior research that suggests that feedback obtained by entrepreneurs on "*early strategies*" (Gans et al. 2019, p.744) leads them to broaden their search rather than focus on and exploit one trajectory. To illustrate this logic, we provide the example of two companies providing fitness coaching for busy individuals. The first company, Coach Pod, focuses on delivering fitness services in small containers located in various residential locations so that individuals can exercise indoors in easy-to-access "portable gyms." As the business is already quite developed, the theory, hypotheses, testing, and evaluation can focus on specific aspects of the value proposition, such as how many people can access a container at the same time and which revenue model will be most successful. The remaining uncertainty gets resolved quickly. The second company, Coach Guru, is a less developed venture with a vision to provide fitness coaching for busy individuals. In the early stages, when the firm has a lower degree of development, the founders are unclear whether the service should be offered as a "gym van" driven to the customer's house or office, or via personal trainers visiting the customer's home, or through small fitness units in neighborhoods with no gym. As a result, they struggle to develop a theory to evaluate the pros and cons of each option to deliver value to their customers. This lack of theoretical clarity impacts the ability to formulate testable hypotheses and leads to tests producing "noisier" results because they are based on a prototyped version of the products as opposed to a real-life version. Coach Guru also struggles to test all the possible options on target customers because of limited time and resources, and the high cognitive load that such testing entails. Overall, uncertainty remains high in the short term.

Second, more developed businesses are likely to apply the scientific approach to "higher-quality" value propositions, resulting in a more positive performance for firms that apply the approach when they have a higher degree of business development. These are firms whose underlying core elements have survived, at least to some extent, a market test (Cusolito et al., 2020; Scott et al., 2015). Such value propositions might need fine-tuning rather than radical rethinking (Kirtley & O'Mahony, 2020), and the scientific approach will act as an effective tool in helping an entrepreneur systematically identify those incremental changes. In contrast, less developed firms use the approach to change a value proposition that has several untested elements (Ott et al., 2017); therefore, the quality and success potential of less developed businesses is likely more varied. In applying the scientific approach, some of the less developed firms will obtain positive signals and use the insights from the tests to further develop their value proposition. But others will discover that their initial logic is not supported by evidence (Kirtley & O'Mahoney, 2020). The scientific approach will still be beneficial as it will help them understand early on the flaws in their original idea (Camuffo et al. 2020), sending the entrepreneur back to the drawing board, possibly to come back later with a better trajectory.[3] The larger variation in the quality of the ideas among less developed firms will likely correspond to larger variation in their performance. Note, paradoxically, that with low degrees of development the scientific approach might even lead firms to perform worse in the short run than those that take a non-scientific approach. In the latter case, a firm might not realize the flaws in an idea and continue along an existing, though ultimately limited, trajectory. Consider two firms at lower degrees of development with the idea of offering a co-working space for women. A firm employing the scientific approach might have a theory about why women and not men would be interested in this type of value proposition, based on the assumption that women feel a higher need to meet other women during their work day than men. They might test this assumption through market research and find that it is not well supported. This would incentivize them to reconsider the positioning of their venture and re-market it for both men and women, (i.e., a broader market), leading to superior returns after the change but causing a delay in generating those returns. Conversely, a firm not employing the scientific approach would be more likely to continue business as usual. As the firm is generating some revenue, it might not realize the longer-term benefits of changing its value proposition.

[3] This situation is illustrated by Camuffo et al. (2020), who have described how Inkdome—a search-engine start-up for finding tattoo artists with a lower degree of development—used theory and tests to make radical changes and find a viable business model.

The two above arguments also translate to situations in which a scientific approach will directly result in a pivot, i.e., strategic changes to the value proposition. Prior research suggests that structured approaches to decision-making can support firms in making strategic changes or pivots to their value proposition (Camuffo et al. 2020; Kirtley & O'Mahony, 2020). Firms with a higher degree of business development, whose problems are related to fine-tuning their value propositions within a more defined search space, are likely to use the scientific approach to apply pivots to their proposition, and these are likely to translate in improvements to firm performance. At lower degrees of development, entrepreneurs are still understanding which elements of their value proposition matter for performance and what works in the context of their business (Kirtley & O'Mahony, 2020). Although they are also likely to significantly change fundamental aspects of their value proposition, these changes might require more time to be designed and might not happen in the short term. Also, due to the higher degree of uncertainty associated with these broader changes, they might not directly translate into performance improvements.

These arguments imply that the effect of a scientific approach on performance will be greater for firms with more developed businesses, leading to the following hypothesis:

Hypothesis 2: A firm's degree of business development positively moderates the relationship between a scientific approach to decision-making and firm performance.

3. DATA AND METHODOLOGY

3.1 The RCT: Setting and Data Collection Process

To investigate the impact of a scientific approach to decision-making on firm performance, we conducted an RCT. Consistent with best practice, we pre-registered the field experiment before the intervention took place (Duflo et al., 2020)[4]. We embedded the field experiment in a business-support program designed and run by the authors in London, UK, from mid-February 2019 to November 2019. The treatment was administered through a training program, as similar interventions have been shown to affect outcomes for entrepreneurs (Anderson et al., 2018; Camuffo et al., 2020). We targeted entrepreneurial firms with less than 10 employees, as our empirical design required that the subjects receiving the treatment be key decision-makers, a condition more accurately met in the context of micro-businesses, where all employees tend to be involved in the management of the firm. We recruited firms with an ad-hoc marketing campaign using online media (such as social media, blogs, and online communities) and offline channels (flyers). Our final sample included 274 entrepreneurial firms. We did not impose any restrictions in terms of industry; firms admitted to the program operated in a wide range of sectors, from software to retail. Our setting enabled the recruitment of firms with different degrees of business development, a feature that set the program apart from other studies where only more developed (Campos et al., 2018; Chatterji et al., 2018; Guzman and Stern, 2016) or less developed businesses participated (Camuffo et al., 2020; Karlan et al., 2015).

The program involved an initial formal training period of 7 sessions (21 hours in total), which started in mid-February 2019 and finished in April 2019. Participants were divided into a treatment and a control group, and the sessions were used to administer the intervention. The training in both groups exposed participants to elements of both cognitive-based decision making, such as strategy frameworks and tools (for instance, the Business Model Canvas or Balance Scorecard), and action-based decision making (such as multiple data collection and testing techniques, including surveys, qualitative interviews and A/B testing to adapt to different entrepreneurial contexts). However, while the control group was not explicitly encouraged to

[4] We provided the registration number of this RCT in the cover letter of this manuscript, but we do not report it in the manuscript to preserve anonymity.

combine the two approaches, the treatment group was encouraged to do so, employing a scientific approach to decision-making. Specifically, the treatment group was encouraged to use the strategy frameworks presented in class to develop a theory of the problem they were facing and deriving hypotheses from it and was later encouraged to use the data gathering and analysis techniques learned in class to test those hypotheses.

For instance, one of the training sessions in both treatment and control groups was focused on the "Business Model Canvas." Both sets of entrepreneurs were taught to apply the tool to their business and discuss it with their peers, but only those in the treatment group were explicitly taught to reflect on how the different elements of the business model connected to each other in a cohesive theory and were subsequently asked to explicitly formulate that theory and break it down into separate hypotheses. Later in the program, entrepreneurs in both groups were taught about the importance of making decisions based on the evidence collected and were exposed to multiple evidence-gathering techniques (e.g., surveys, A/B testing, qualitative interviews). Entrepreneurs in the control group were free to apply those techniques based on their intuition, whereas entrepreneurs in the treatment group were explicitly encouraged to use these techniques to test the hypotheses developed in the previous sessions and reflect on how the evidence collected compared to their initial theory. We also designed a series of in-class activities and post-class assignments to encourage entrepreneurs to use the tools and techniques described in class. Consistent with our experimental design, entrepreneurs in both treatment and control groups were given the same number and types of in-class activities and post-class assignments. There were, however, differences in that the activities and assignments for the treatment group specifically focused on applying a scientific approach to decision-making, while the control group was not exposed to this approach. We provide an example of the differences in the training and in-class activities between treatment and control groups in Section 2 of the Appendix (Figure A1 and A2). This was an important feature of our training, which was highly engaging and experiential, involving hands-on activities and feedback from the instructors. To achieve this goal, we assigned entrepreneurs in both groups to smaller subgroups that were randomly matched with six experienced instructors who were recruited and trained for this study. The experiment was designed such that each instructor taught groups of entrepreneurs in both the treatment and control groups, allowing us to account for instructor-related differences in our regressions through fixed effects. All instructors received identical training material from the research team and underwent multiple "train-the-trainer" sessions to ensure they would deliver the content of the program in line with our research design.

Several measures were taken to ensure the internal validity of our results. We addressed contamination by teaching treated and control groups on different days of the week (Wednesday and Thursday) or different time slots of the same day (Saturday morning and afternoon), preventing them from meeting and discussing key elements of the treatment. We also kept communications about the program separate and discrete for the two groups.

We required all applicants to complete an extensive survey and participate in a 30-minute call with a member of the data collection team which aimed at collecting baseline information on their business and their approach to decision-making prior to the intervention. We then used this information to randomly assign firms to either the treatment or control group using statistical software (STATA); 139 firms were assigned to the treatment group and 135 firms to the control group.

Data collection and operationalization. The intervention ran between February and April 2019, but we monitored firms' performance and decision-making until the November 2019. Due to funding availability, we could only gather data over this relatively short time window, and we take this aspect into consideration when discussing our results. In addition to the pre-intervention survey and interview, we collected 8 data points through telephone interviews that focused on each firm's decision-making, key changes it had made in terms of value proposition, and performance. The first telephone interview post-intervention took place about 8 weeks after the training program had begun. We then collected data once a month until November 2019. In conducting these calls, we created a pre-defined protocol that included open- and close-ended questions, an approach in line with Bloom and Van Reenen's (2010) and Camuffo et al.'s (2020). We used open-ended questions to monitor entrepreneurs' decision-making process and let key themes emerge from narratives and closed-ended questions to elicit self-reported performance information. The open-ended questions substantially reduce concerns that respondents might answer in a way that complies with the research design, particularly since entrepreneurs were not aware their answers were scored against a pre-defined grid. The performance data provided were self-reported by the entrepreneurs, but we conducted cross-reference checks with external sources for 100 firms (for which we found correspondence between the information provided by the entrepreneurs and public records in 92.5% of the cases, with small discrepancies in other cases) and consistency checks across interview rounds. The final sample included 261 firms, as we excluded data provided by four participants that gave inconsistent information about their business and nine participants who were not willing to share data. Table 1 compares the baseline characteristics of the treated and control groups for the final sample of 261 firms. It shows that the two are not different in statistically meaningful ways.

Add Table 1 about here

To check that the treatment produced the intended result, we measured the level of adoption of the scientific approach based on the content of the telephone interviews. *Scientific Intensity* is a time-varying score (ranging from zero to five) that captures the level of adoption of the scientific approach. To calculate this score, we followed the data collection method used by Camuffo et al. (2020). A team of research assistants analyzed and coded each interview's content according to a pre-defined coding scheme and collected measures on the extent to which entrepreneurs used theory (measured with four variables), hypotheses (measured with four variables), tests (measured with four variables), and evaluations (measured with four variables). To adequately capture the multiple dimensions of each component, we identified some sub-components that measure the key aspects that define theory, hypotheses, tests, and evaluation. We provide details on each of the subcomponents of the approach in Table A1a in the Appendix. For each subcomponent, research assistants provided a score from 0 to 5, where a low score (say 0) indicates that the entrepreneur does not employ at all that specific aspect of his/her decision-making process; a high score (such as 5) reflects that the entrepreneur is adopting a specific aspect extensively. We then aggregated all variables by taking the average of the subcomponents to compute an overall scientific intensity score. The rationale for this choice is that the scientific approach is a holistic approach and entrepreneurs should use it in its entirety (Lazear, 2004; Zellweger and Zenger, 2021).

In Table A1b, we compare the level of scientific intensity of the treatment and control groups at the time of each interview. Results show that, while the difference between the two groups was not statistically significant at the baseline, the level of scientific intensity was significantly higher for treated firms in subsequent interviews, although it diminished in size and significance over time. Entrepreneurs completed their training between Interview 1 and Interview 2, indicating that the training effect was still visible months after completion[5].

[5] Only the treatment group was exposed to the scientific approach to decision-making. However, the scientific intensity of entrepreneurs in the control group is not zero. This is because, irrespective of our treatment, individuals might have a natural tendency to approach problems or make decisions in a scientific way, by developing a theory of the problem, elaborating explicit predictions, testing those predictions, and evaluating the results of those tests. We measured the natural tendency of entrepreneurs in both the treatment and the control groups toward the use of a scientific approach at the baseline (i.e., before any treatment). Our balance test confirmed that the randomization resulted in groups that were not significantly different. We then continued to measure scientific intensity throughout the program, and our results, reported in Table A1b, show that scientific intensity increased for the treated firms more than for the control firms, suggesting that our treatment was effective in improving the adoption of a scientific approach.

3.2 Methodology and Variable Operationalization

Methodology. To test Hypothesis 1 regarding the impact of the treatment on firm performance, we employed a classic *difference-in-difference* specification. We estimated it by fitting the following model:

$$\text{Performance}_i = \beta_0 + \beta_1 T_i + \beta_2 P_i + \beta_3 T_i P_i + \beta_4 Degree_i + \varepsilon_i$$

where T_i denotes the treatment and is equal to 1 for firms that were allocated to the treatment group and 0 for firms that were allocated to the control group and P_i denotes the time period post training, with $P_i = 0$ before the training program took place and $P_i = 1$ after the training program (at the end of the observation period) and $Degree_i$ is a variable that measures that degree of business development at the baseline. In this model the difference in difference estimand is the coefficient of the interaction term (β_3). It corresponds to the difference in performance before and after the training for treated firms vs. control firms.

To test Hypothesis 2 on the moderating effect of the level of business development, we employed a *triple difference specification* and analyzed how the difference in performance between treated and control firms is shaped by the Degree of Business Development, measured at the baseline. We fit the following model:

$$\text{Performance}_i = \beta_o + \beta_1 T_i + \beta_2 Degree_i + \beta_3 T_i Degree_i + \beta_4 P_i + \beta_5 T_i P_i + \beta_6 P_i Degree_i + \beta_7 T_i P_i Degree_i + \varepsilon_i$$

where $Degree_i$ is a variable that measures that degree of business development at the baseline. The coefficient of interest is β_7, the coefficient on the triple interaction term (Wooldridge, 2007). In all regressions, we clustered the standard errors at the firm level.

Variable operationalization.

Dependent variable: Performance. Our paper investigates the impact of a scientific approach to decision-making on performance. Prior research has extensively documented that firm performance is a multifaceted construct and that different performance measures capture different nuances of performance. Following prior research, we referred to two important measures of performance. First, we looked at *Firm size* as measured by the log of (1+) the number of employees of the firm in each period. This is an important performance measure and one that is often used in the literature (see Chatterji et al., 2018; Yang et al., 2020), particularly for less developed firms. Prior work emphasizes how these firms do not necessarily pursue revenue or profit but focus on growth instead, as this might put them on a better trajectory later (Fazio et al 2016; Tidhar & Eisenhardt, 2021). Our second performance measure is *Revenue*, measured as the log of (1+) the cumulative revenue generated up until the last period since the beginning of the program in thousand pounds sterling. These variables were calculated using data collected by our research assistants during each monthly interview.

Independent variables. Our first independent variable is *Treatment*, a dummy variable equal to 1 for firms in the treatment group and 0 for those in the control group. Our second independent variable is *Post,* a dummy variable equal to 0 at the baseline and 1 after the training program and until the end of our observation period. Finally, we included in the analysis the *Degree of Business Development.* Our theory suggests that firms with more developed businesses will benefit more from the intervention. As a proxy for the degree of business development we used the annual revenue of each firm in the year before it started the program (in thousands of GBP). This measure has been extensively used by prior research on firm evolution (Churchill & Lewis, 1983; Greiner, 1972). As emphasized by Greiner, *"a company's problems and solutions tend to change markedly as the number of employees and sales volume increase (...) in fact, organizations that do not grow in size can retain many of the same management issues and practices over lengthy periods"* (1972, p. 3). This is also in line with the measure used by Davidsson (2004) in the Panel Study of Entrepreneurial Dynamics (II version). We report the descriptive statistics and pairwise correlation in Table 2.

Add Table 2 about here

4 RESULTS

4.1 Firm performance: Firm Size

We start by examining the impact of our intervention on firm size. Column (1) in Table 3 reports the results of our difference in difference specification, estimated by OLS. As discussed in the previous section, the coefficient of interest is the interaction between Treatment and Post, which is positive and significant (B = 0.1160, p = 0.0197), suggesting that treated firms experienced an 11.6% increase in firm size after the treatment compared to control firms. This supports Hypothesis 1 and provides evidence that the application of a scientific approach to decision-making has a positive impact on firm performance as measured by size. To test Hypothesis 2, on the moderating effect of the degree of business development, we present the result of a triple difference specification in Column (2). Results show that when the degree of business development is equal to 0, treated firms grow about 10.5% more than control firms (B = 0.1053, p = 0.0511), but this effect is higher for firms with a higher degree of business development (B = 0.0002, p = 0.0834). Specifically, an increase in one standard deviation of the degree of business development (171.293) leads to a further 3.4% increase in firm size. This supports our second hypothesis. To control for unobserved heterogeneity, in Columns (3) and (4) we replicate the same set of analyses with the inclusion of firm fixed effect. In this specification, some of the terms that do not change over time (Treatment, Degree of Business Development at the baseline, and the interaction between the two) are completely absorbed by the fixed effects. Results are very similar in terms of magnitude and significance to those presented in Columns (1) and (2).

Add Table 3 about here

4.2 Firm performance: Firm Revenue

Table 4 reports the results of our analyses that investigate the impact of the treatment on firm revenue. In this case, too, Column (1) reports the results of the difference-in-difference specification. The coefficient of the interaction between Treatment and Post is negative and not significant (B = -0.1197, p = 0.4596). This suggests that, when we examine firm performance as measured by revenue, the treatment did not have a significant effect. Hypothesis 1 is therefore supported when performance is measured as firm size, but not when measured as firm revenue. Results from the triple difference specification, reported in Column (2), paint a very interesting picture. They show that the effect of the treatment is negative and significant when the degree of business development is equal to 0 (B = -0.2998, p = 0.0690), with the treatment resulting in 30% lower revenue than the control group. However, for firms with a higher degree of business development, the treatment has a positive effect on firm revenue (B = 0.0035, p=0.0090). An increase in one standard deviation of the degree of business development (171.293) leads to a very substantial 60% increase in firm revenue. The overall impact of the treatment is positive for firms that have annual revenue equal to at least £86,000 at the baseline, which corresponds to about the top 15% firms of the sample. This supports our second hypothesis. To control for unobserved heterogeneity, we report the results of the specifications that include firm fixed effects in Columns (3) and (4); these results are consistent with those presented in Columns (1) and (2) both in terms of signs and significance.

Add Table 4 about here

4.3 Robustness checks

Alternative measure of Degree of Business Development

To ensure that our results are robust to the operationalization of the variable Degree of Business Development, we provide some robustness check with an alternative measure. We asked entrepreneurs to provide us with an estimate of the probability of making changes to their business (on a scale from 0% to 100%). For clarity of interpretation, we constructed a variable that we call *Probability of <u>not</u> making changes to the business proposition* that we calculate as the difference between 100 and the probability of making changes (provided by the entrepreneur).

We expect that entrepreneurs with a *lower* Degree of Business Development will report a *lower* probability of *not* changing their business, whereas entrepreneurs with more developed businesses will indicate that the probability of not making changes to their business is higher. Results from these analyses are reported in Table A2 in the Appendix. When considering firm size as the dependent variable, the results confirm the positive impact of the treatment for all firms (Column 1). The small positive effect associated with firms with a higher degree of business development is not confirmed (Column 2). We then examine revenue as a dependent variable. Results are consistent with those presented in Table 4. They show that when the Degree of Business Development is equal to 0 the effect of the treatment on Firm Revenue is negative. However, the treatment has a positive effect on the revenue of firms with a higher Degree of Business Development.

Outliers

We checked if the results might have been driven by the presence of outliers in our sample by replicating the analysis after 99% winsorization. We report these results in Table A3 in the Appendix; they are consistent with those reported in Tables 3 and 4.

Attrition

Not all the firms in our sample continued to participate in the interviews until the end of the study (see Table A4a for the distribution of attrition over time). Notoriously, attrition is more the norm than the exception in field experiments (Gerber & Green, 2012). To address this issue, we designed the program so that the training was followed by a series of monthly events focused on relevant themes for entrepreneurs delivered in the same way for treated and control firms but on separate days. Participation in these events was conditional on firms' continued engagement with the program and data collection. Nevertheless, some firms dropped out before the last interview round. Entrepreneurs that were not available for interviews indicated that their incentive to participate in interviews was lower after the training was over. To verify that attrition did not affect our results, we followed the best practice outlined by Gerber and Green (2012). First, we checked there was no significant difference between treated and control groups in terms of early withdrawal from the program. In Table A4b in the Appendix, we estimate early withdrawal from the program as a function of the intervention, which we show has no significant impact. Second, we addressed attrition by inputting the missing values of those who left the study. We followed Gerber and Green (2012) and used different case scenarios. As a starting point, the main analyses presented in Tables 3 and 4 made the conservative assumption that the performance of firms that left the program remained at the same degree as when they left the program. This assumption is consistent with previous studies that have used similar data (Camuffo et al. 2020). We then replicated the analyses by assuming that the performance of firms who left the program grew at the average rate of growth for firms in the sample. We present these analyses in Table A5 in the Appendix, Columns (1) and (3). Finally, we replicated these analyses using an unbalanced panel and retaining firms in the sample only up until the time at which they left the program. We included interview dummies to control for the fact that different firms left the program at different points in time. Results are reported in Columns (2) and (4) and are overall consistent with those presented in the main analyses, supporting the idea that our results are robust despite attrition.

4.4 Excluding alternative explanations

Experience and confidence

In our main analyses, we used firms' annual revenue before entering the program (in thousands of GBP) as a proxy for their degree of business development. We then asked: Is the positive effect for treated firms with higher revenue driven by other factors that are also associated with higher revenue? We consider two possible alternative explanations. First, research has extensively emphasized the importance of prior experience for firm survival (Agarwal & Shah, 2014; Klepper & Sleeper, 2005) and performance (Agarwal et al., 2016; Azoulay et al., 2020;

Gruber et al., 2013; Shah et al., 2019), so one might expect that prior experience affects these results. Second, one could argue that the results might be driven by the confidence of the entrepreneur in the project (Bennett & Chatterji, 2019; Chen et al., 2020; Hayward et al., 2010). To address these alternative explanations we report in Table A6 in the Appendix the results obtained replicating the analyses and introduce the interactions between the intervention and (a) prior work experience of the team at baseline (measured as the average number of years of work experience in any role of all individuals working in the firm) (Columns 1 and 3) and (b) the level of confidence of the entrepreneur at the baseline (measured as their agreement on a 1–5 scale with statements related to confidence) (Columns 2 and 4). Results show that these alternative explanations do not account for the results from our main analyses: The interaction terms are not significant, except for the interaction of the treatment with the degree of work experience,[6] which has a negative and significant impact on firm size, suggesting that the treatment had a lower impact on performance for more experienced firms. This supports the theoretical intuition that the effect observed in our main regressions was driven by the degree of development of the business instead of other factors such as experience or confidence.

Higher investment in the business

We also examined if entrepreneurs with more developed firms invest more resources in their business, thus incurring higher costs. Table A8 in the Appendix reports the results of a triple difference specification where the dependent variable is the logarithm of (1+) the cumulative cost incurred by the firm up until the last period since the beginning of the program in thousand pounds sterling. Results from Model 1 show that, overall, the intervention does not increase costs. However, results in Model 2 show that the effect of the intervention is not statistically significant when the degree of business development is 0 (B = -0.0881, p = 0.5943), but it is

---

[6] As alternative measures of prior experience we used the average number of years of industry, managerial, and entrepreneurial experience of the individuals working in the firm. Results, reported in Table A7 in the Appendix, are consistent with those presented in Table A6.

positive and significant (B = 0.0041, p = 0.0026) for higher degrees of business development suggesting that more developed, treated firms spend significantly more. If the increase in cost were the only channel through which more developed treated firms obtained higher revenue, it would imply that these firms are not appropriating value but merely transferring value to customers (Brea-Solis et al. 2015). To explore whether the superior performance in terms of revenue for treated firms with a higher degree of development is associated with superior value creation we focus on value added, defined as the logarithm of the sum of (i) the difference between revenue and costs (in thousands of GBP) and (ii) the absolute value of the minimum of the difference between revenue and costs (in thousands of GBP) in the sample. Results from Column (3) show that the effect of the intervention is negative but not statistically significant (B = -0.0730, p = 0.1315), while the interaction is positive and significant (B = 0.0005, p = 0.0157), suggesting that treated firms with a higher degree of development significantly increased their productivity, while there is high variation for firms with a lower degree of development.

4.4 Exploring the mechanisms

Differences in problems and goals

Our theory builds on prior research in strategy (e.g., Churchill and Lewis, 1983) that suggests that firms with a higher vs. lower degree of business development apply and benefit from the scientific approach differently because they face different problems and have different objectives. We provide evidence of these differences by analyzing data collected for all entrepreneurs at the baseline, when we asked them what their main problems and goals were. We use multiple variables. First, we employ two dummy variables, which we labelled *Multiple problems* and *Multiple goals*, respectively, to indicate whether a firm had indicated more than one problem or goal. We then grouped firms' problems into two main categories: (1) *market validation problems* (including, for instance "identifying the product market fit," "developing the product," and "setting up the business plan") and (2) *scale-up problems* (including, for instance, "increasing growth" and "improving operational efficiency"). Consistently, we also grouped their goals into two main categories: (1) market validation goals (including, for instance, "entering the market") and (2) scale-up goals (including, for instance, "increasing revenue" and "increasing the customer base"). We then assigned firms to two categories: lower degree of business development (if their annual revenue at baseline was lower than the median) and higher degree of business development (if their annual revenue at baseline was higher than or equal to the median). Results, reported in Table 5, show that firms with a lower (vs. higher) degree of business development do not differ in terms of facing multiple problems or goals, but they do differ with regard to the types of problems and goals that they face. Consistently with literature on this subject and our theory, we find that firms with a lower degree of development are significantly more likely to face problems and goals related to market validation and firms with a higher degree of development are significantly more likely to face problems and goals associated with scaling up.

Add Table 5 about here

Differences in the level of (un) certainty

An important aspect of our theory is that when the scientific approach is applied to a more defined value proposition (as in the case of more developed firms), it helps to reduce uncertainty more than when the value proposition is less defined (as it often is for firms with a lower degree of business development). To find evidence of this mechanism, we analyzed the text of entrepreneurs' interviews. Due to funding constraints, we could only transcribe a random sample of all baseline and final interviews for 173 firms, corresponding to 66% of firms in our sample. We analyzed the text of these interviews using Linguistic Inquiry and Word Count (LIWC), a dictionary-based software developed by psychologists to measure validated constructs (such as emotions, cognitive thought processes, grammatical features, etc.)

in written text (Tausczik & Pennebaker, 2010). LIWC has been extensively used to analyze business-related communications and has the advantage of providing systematic and scalable measures of large text (Hannigan et al., 2019).

We used LIWC and its dictionary for the concept of certitude to calculate a score associated with each interview that reflected the degree of certainty expressed by the entrepreneur during the baseline and the final interview they conducted. Our variable *Certainty* was measured as the logarithm of (1+) this score. First, we performed a t-test to check that there were no significant differences in the level of certainty in the language used by treated and control entrepreneurs at the baseline interviews ($\beta = 0.00$, $p = 0.991$) nor between entrepreneurs characterized by a degree of business development above versus below the median ($\beta = 0.02$, $p = 0.733$) at the baseline interview. We then performed a triple difference analysis to check if the certainty expressed by entrepreneurs in their interviews varied for treated and control entrepreneurs at the baseline versus at the last interview conducted, particularly depending on their degree of development. Results are reported in Table 6. These results show that the treatment has no effect on the degree of certainty expressed when entrepreneurs are talking about their venture for firms with a degree of development equal to 0. However, the treatment has a positive and significant effect on the degree of certainty used by more developed firms in their interviews. This effect is not large, which is understandable since it is unlikely that our intervention would drastically change the way in which entrepreneurs talk about their idea. But the fact that the effect is significant supports the idea that the application of the scientific approach reduces the level of uncertainty faced by entrepreneurs with more developed businesses, while it has a varied effect on entrepreneurs with less developed businesses.

Add Table 6 about here

Difference in quality

Our theory also suggested that entrepreneurs with a higher degree of business development have value propositions of higher quality, having already passed the first phases of development (Cusolito et al., 2020; Scott et al., 2015). We therefore checked if the quality of the value propositions developed by entrepreneurs explains these results. We measured this construct in two ways. First, we used the value of the entrepreneurial idea at the baseline (measured as the entrepreneur's own estimation of the value of their business idea, ranging from 0 to 100[7]). Second, we collected an external evaluation of the business idea by having two research assistants score the quality of the proposals submitted by the entrepreneurs at the baseline based on three criteria: Feasibility, Originality, and Target Market Size. Both research assistants were blind to the treatment and unaware of the goals of the study. Results are reported in Table 7 (Columns 1 and 2) and show that the entrepreneurs' own evaluations do not explain our results, in line with the idea that entrepreneurs are often biased regarding the value of their own ideas (Cain et al., 2015; Chen et al., 2018; Cooper et al., 1988). Instead, a higher external evaluation of the entrepreneurial idea is positively associated with revenue (see Table 7, Column 3 and 4). This supports the intuition that entrepreneurs with more developed businesses have higher quality business ideas, and this increases their revenue. This suggests the intriguing possibility that a scientific approach might be more effective for entrepreneurs with better ideas: The approach enables them to fine-tune and reinforce those ideas.

Add Table 7 about here

Differences in pivoting

Our theory suggested that the scientific approach would make firms with a higher degree of business development more likely to directly make strategic changes to their value proposition

[7] The question that we asked entrepreneurs at the baseline was: "Considering 0 = 'This business is not going to be making revenues' and 100 = 'This business is going to be extremely successful in making revenues,' please report the minimum and maximum amount of revenues that you expect your business to make."

(i.e., pivot). We therefore examined if our intervention led firms with different degrees of development to pivot. During the regular interviews our research assistants asked each entrepreneur to indicate whether they had made any changes to their business model in any of the nine areas identified in the Business Model Canvas (customer segment, distribution, revenue stream, value proposition, activities, resources, cost structure, partners or customer relationships). We classified each of these changes as a pivot. Our measure *Number of Pivots* was calculated as the logarithm of (1+) the total number of pivots made by each firm at the end of the observation period. [8]

Results in Table 8 (Model 1) show that the treatment does not have a significant impact on the number of pivots in general. However, results in Model 2 show that the interaction between *Treatment* and the *Degree of Business Development* has a positive and statistically significant impact on the total number of pivots (B = 0.0013, p = 0.0302). When the degree of business development is equal to 0, however, the treatment does not have a significant impact on the number of pivots. This supports our theoretical intuition that treated firms with more developed businesses are more likely to directly employ the scientific approach to fine-tune their business proposition by changing components of the business to enhance their key value proposition to customers.

Add Table 8 about here

5 DISCUSSION AND CONCLUSIONS

Is it beneficial for entrepreneurs to use a scientific approach to decision-making? And, if so, when is it beneficial to use this approach? Recent research indicates that a scientific approach leads to improved learning (Camuffo et al., 2020; Zellweger & Zenger, 2021), but there is limited and mixed empirical evidence on the effect of this approach or its sub-components (Koning et al., 2022) on performance. We conducted an RCT with 261 entrepreneurial firms

[8] We employed a cross-sectional analysis as the variable pivot cannot be defined at the baseline, making this approach more appropriate than a difference in difference approach.

in the UK to evaluate the impact of a scientific approach on performance and its differential effect on different types of firms. In the RCT, we created a business support program and taught treated firms a scientific approach to decision-making, while the control firms received comparable training without a scientific approach. We collected detailed observations on performance and key business choices at baseline and after the training. Our sample is unique as it includes different types of firms at various degrees of business development. Results show that a scientific approach is beneficial for all treated firms as they all increased their number of employees, but only the more developed businesses increased their revenue. We examined the mechanisms behind these results and found evidence that more developed businesses face different types of problems and goals, that firms with a higher degree of business development apply the scientific approach to higher quality value propositions, and that the scientific approach results in higher degrees of uncertainty resolution and higher number of pivots for firms with a higher degree of business development. We interpret this result to suggest that as less developed firms apply the approach, it helps them in "figuring out the right value proposition," which directly translates to firm growth, but will require more time to translate into revenues.

Our study makes several contributions to research in strategy and entrepreneurship. First, our results provide insight on the performance implications of the use of a scientific approach to decision-making, and, more broadly, on the use of decision-making approaches that combine cognitive-based and action-based components. The growing literature on entrepreneurial strategy (Gans et al., 2019; Zellweger & Zenger, 2021) has advanced (mostly through theoretical work) that experimentation—and particularly scientific-like experimentation—improves performance. Preliminary evidence from qualitative studies (McDonald & Eisenhardt, 2020) and observational data (Koning et al, 2019) supports this idea. Our results provide an important empirical test of these recent concepts through an RCT that allows us to identify cause-effect relationships between a scientific approach and performance. We show that all treated firms grow in terms of employee number. To the extent that employee growth can be interpreted as an early measure of performance, these results suggest the possibility that firms at all degrees of business development benefit from the approach although differently.

With regards to revenue, we find that only firms with a higher degree of businesses development benefit from the scientific approach, at least within the time window we consider. This finding represents our second contribution, as it adds a novel perspective to literature on strategy and firm development. Existing work on decision-making has focused on either firms with a higher (Bloom & Van Reenen, 2007; Yang et al, 2020; Zollo &Winter, 2004) or lower degree of business development (Camuffo et al., 2020) in isolation. What remains unclear is if systematic approaches to decision-making unfold differently for firms with different degrees of development, prompting a shift in the academic conversation from "are these approaches beneficial?" to "what types of firms are they beneficial for?" These results also suggest the existence of a possible generalizability bias in previous studies focusing on more developed firms and suggest that their findings should be applied with caution to less developed businesses.

Our paper also contributes to research on strategic entrepreneurship that advocates for the importance of testing and purposeful experimentation for firm performance (Bingham & Eisenhardt, 2011; Gruber & Tal, 2017; Murray & Tripsas, 2004; Shepherd & Gruber, 2020; Thomke, 2003). In line with these studies, our results support the view that testing and experimentation, particularly combined with theory and hypothesis development, can be useful. However, the finding that more developed businesses benefit more from the use of a scientific approach to decision-making complements the literature that emphasizes that firms with a lower degree of business development can successfully gather feedback through experimentation. The Lean Start-up movement advances the idea that experimentation, customer feedback, and iterative design are superior choices compared to planning, top-down innovation, and up-front design investments (Blank, 2013; Ries, 2011). The underlying assumption of these studies is that being as nimble and as flexible as possible will help entrepreneurs adjust more easily in a context characterized by high uncertainty, delaying important choices and substantial investments until they reach a stage where they have enough evidence to commit to a course of action. Indeed, a key tenet of this philosophy is "Build fast and fail fast," using minimum viable products to obtain feedback on ideas that are nor very developed. Our results suggest, instead, that firms with different degrees of business development all benefit from the use of an approach that use experimentation (particularly when combined with theory), but –because they address different types of problems—they benefit on different dimensions, with firms with a more developed value proposition benefiting both in terms of size and revenue growth.

Finally, our results contribute to the strategy literature on the synergies between cognitive and experiential search (Gavetti & Levinthal, 2000; Gavetti & Rivkin, 2007; Levinthal, 2017). Research in this area has emphasized the existence of synergies between these two approaches, whose combination enables decision makers to find a "middle ground" between the myopia of local search and the omniscience assumed in economic analysis (Gavetti & Levinthal, 2000). By exploring theoretically and testing empirically the impact on performance of a decision-making approach that combines elements of cognitive-based search (theory and hypotheses development) with elements of experiential search (gathering evidence and learning from it), this paper represents a relevant empirical test in the context of entrepreneurship.

In making these considerations, we also acknowledge the limitations of this study, which point to opportunities for future research. First, our study is focused on firms with less than ten employees. This is an advantage in that it allowed us to ensure that the treatment was administered to the individuals directly involved in the firm's decision-making. However, an open question is whether the treatment would produce the same effect with larger firms limits the generalizability of our findings. We see this as an opportunity for future research. Second, it would be important for future research to replicate our analyses in a longer time window to support the intuition that firms with a lower degree of business development (not only those with a higher degree) can benefit from the approach in terms of revenues as well in the long run. Third, future research could examine additional factors that might contribute—in conjunction with a scientific approach to decision-making—to better performance. For instance, resources might play an important role in shaping performance outcomes and might afford entrepreneurs with more opportunities to experiment scientifically.

A final contribution is to offer insights to policymakers. Encouraging entrepreneurship has been a major means to spur economic growth (Bennett & Chatterji, 2019; Decker et al., 2014; Lerner, 2009; McKenzie, 2021). Bennett and Chatterji's (2019) nationally representative survey on the pre-entry activities conducted by potential entrepreneurs in the US found that fewer than half of those who consider starting a business take the lowest-cost steps, such as searching the Internet for potential competitors or speaking with a friend, a phenomenon they attribute to the psychological costs associated with learning the true promise of an idea. They conclude that one way to increase the quality and quantity of entrepreneurial ventures would be to lower the cost of experimentation at the very *beginning* of the entrepreneurial process. Our results show that an intervention intended to encourage systematic experimentation to support decision-making was helpful for both more and less developed firms in terms of leading to size increases, but it led to a revenue increase only for *firms with a higher degree of business development*, having a negative impact on the revenue of less developed firms, at least within the observed time window. These results, therefore, underline the importance of understanding the ideal time window as well as the most effective performance dimensions on which programs targeted at less developed firms should be evaluated to appreciate their effectiveness.

Given the importance of this topic for the economy, we consider this a promising path for future research.

REFERENCES

Agarwal, R., Audretsch, D., & Sarkar, M. B. (2007). The process of creative construction: knowledge spillovers, entrepreneurship, and economic growth. *Strategic Entrepreneurship Journal*, *1*(3-4), 263—286.

Agarwal, R., Campbell, B. A., Franco, A. M., & Ganco, M. (2016). What do I take with me? The mediating effect of spin-out team size and tenure on the founder–firm performance relationship. *Academy of Management Journal*, *59*(3), 1060—1087.

Agarwal, R., & Shah, S. K. (2014). Knowledge sources of entrepreneurship: Firm formation by academic, user and employee innovators. *Research policy*, *43*(7), 1109—1133.

Agrawal, A., Gans, J. S., & Stern, S. (2021). Enabling entrepreneurial choice. *Management Science*.

Alexy, O., Poetz, K., Puranam, P., & Reitzig, M. (2021). Adaptation or persistence? Emergence and revision of organization designs in new ventures. *Organization Science*, *32*(6), 1439—1472.

Ambos, T. C., & Birkinshaw, J. (2007). How do new ventures evolve? The process of charter change in technology ventures. *Academy of Management Proceedings* 2007(1), 1—6.

Anderson, S. J., Chandy, R., & Zia, B. (2018). Pathways to profits: The impact of marketing vs. finance skills on business performance. *Management Science*, *64*(12), 5559—5583.

Azoulay, P., Jones, B. F., Kim, J. D., & Miranda, J. (2020). Age and high-growth entrepreneurship. *American Economic Review: Insights*, *2*(1), 65—82.

Bennett, V., & Chatterji, A. (2019). The entrepreneurial process: Evidence from a nationally representative survey. *Strategic Management Journal*.

Bingham, C., & Eisenhardt, K. (2011). Rational heuristics: The ‘simple rules’ that strategists learn from process experience. *Strategic Management Journal*, *32*(13), 1437—1464.

Blank, S. (2013). Why the lean startup changes everything, *Harvard Business Review*. https://hbr.org/2013/05/why-the-lean-start-upchanges-everything

Bloom, N., Sadun, R., & Van Reenen, J. (2012). The organization of firms across countries. *The Quarterly Journal of Economics*, *127*(4), 1663—1705.

Bloom, N., & Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The Quarterly Journal of Economics*, *122*(4), 1351—1408.

Bloom, N., & Van Reenen, J. (2010). Why do management practices differ across firms and countries? *Journal of Economic Perspectives*, *24*(1), 203—24.

Boulding, W., Morgan, R., & Staelin, R. (1997). Pulling the plug to stop the new product drain. *Journal of Marketing research*, *34*(1), 164—176.

Brea-Solís, H., Casadesus-Masanell, R., & Grifell-Tatjé, E. (2015). Business model evaluation: Quantifying Walmart's sources of advantage. *Strategic Entrepreneurship Journal*, *9*(1), 12—33.

Bruhn, M., Karlan, D., & Schoar, A. (2018). The impact of consulting services on small and medium enterprises: Evidence from a randomized trial in Mexico. *Journal of Political Economy*, *126*(2), 635—687.

Cain, D. M., Moore, D. A., & Haran, U. (2015). Making sense of overconfidence in market entry. *Strategic Management Journal*, *36*(1), 1—18.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., & Mensmann, M. (2018). Is personal initiative training a substitute or complement to the existing human capital of women? Results from a randomized trial in Togo. *AEA Papers and Proceedings*, 108, 256—61.

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.

Cardon, M. S., Post, C., & Forster, W. R. (2017). Team entrepreneurial passion: Its emergence and influence in new venture teams. *Academy of Management Review*, *42*(2), 283—305.

Chatterji, A., Delecourt, S., Hasan, S., & Koning, R. (2018). When does advice impact startup performance? *Strategic Management Journal*, *40*(3), 331—356.

Chen, J. S., Croson, D. C., Elfenbein, D. W., & Posen, H. E. (2018). The impact of learning and overconfidence on entrepreneurial entry and exit. *Organization Science*, 29(6), 989—1009.

Chen, J. S., Elfenbein, D. W., & Posen, H. E. (2020). The problems and promise of entrepreneurial partnerships: Decision making, overconfidence, and learning in founding teams. *Academy of Management Review, (Forthcoming)*.

Churchill, N., & Lewis, V. (1983). The Five Stages of Small Business Growth*, Harvard Business Review*. Retrived from https://hbr.org/1983/05/the-five-stages-of-small-business-growth.

Cohen, S. L., Bingham, C. B., & Hallen, B. L. (2019). The role of accelerator designs in mitigating bounded rationality in new ventures. *Administrative Science Quarterly*, *64*(4), 810—854.

Cooper, A. C., Woo, C. Y., & Dunkelberg, W. C. (1988). Entrepreneurs' perceived chances for success. *Journal of business venturing*, *3*(2), 97—108.

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

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

Cusolito, A. P., Dautović, E., & McKenzie, D. J. (2020). Can government intervention make firms more investment-ready? A randomized experiment in the Western Balkans. *The Review of Economics and Statistics*, 103(3), 428—442.

Davidsson, P. (2004). *Researching entrepreneurship* (Vol. 5). New York: Springer.

Decker, R., Haltiwanger, J., Jarmin, R., & Miranda, J. (2014). The role of entrepreneurship in US job creation and economic dynamism. *Journal of Economic Perspectives*, *28*(3), 3—24.

Duflo, E., Banerjee, A., Finkelstein, A., Katz, L. F., Olken, B. A., & Sautmann, A. (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics. *National Bureau of Economic Research* (No. w26993).

Ehrig, T., & Schmidt, J. (2022). Theory-based learning and experimentation: How strategists can systematically generate knowledge at the edge between the known and the unknown. *Strategic Management Journal*. https://doi.org/10.1002/smj.3381

Eisenhardt, K.M., & Bingham, C.B. (2017). Superior strategy in entrepreneurial settings: Thinking, doing, and the logic of opportunity. *Strategy Science*, *2*(4), 246—257.

Eisenhardt, K. M., & Brown, S. L. (1999). Patching. *Harvard business review*, *77*(3), 72—73.

Ethiraj, S. K., & Levinthal, D. (2009). Hoping for A to Z while rewarding only A: Complex organizations and multiple goals. *Organization Science*, *20*(1), 4—21.

Fazio, C., Guzman, J., Murray, F., & Stern, S. (2016). A new view of the skew: Quantitative assessment of the quality of American entrepreneurship. *Kauffman Foundation New Entrepreneurial Growth, Kansas City, MO: February*.

Felin, T., Gambardella, A., Stern, S., & Zenger, T. (2020a). Lean startup and the business model: Experimentation revisited. *Long Range Planning*, *53*(4) pagine?.

Felin, T., Kauffman, S., & Zenger, T. (2020b). Microfoundations of resources: a theory. *Available at SSRN.*

Felin, T., & Zenger, T. (2009). Entrepreneurs as theorists: on the origins of collective beliefs and novel strategies. *Strategic Entrepreneurship Journal*, 3(2), 127—146.

Felin, T., & Zenger, T. (2016). CROSSROADS—Strategy, problems, and a theory for the Firm. *Organization Science*, 27(1), pagine?.

Felin, T., & Zenger, T. (2017). The theory-based view: Economic actors as theorists. *SSRN Electronic Journal*, 2(4), 211—287.

Folta, T. (1998). Governance and uncertainty: the trade-off between administrative control and commitment. *Strategic Management Journal*, *19*(11), 1007—1028.

Forbes, D. P. (2005). Managerial determinants of decision speed in new ventures. *Strategic Management Journal*, *26*(4), 355—366.

Foss, N., & Klein, P. (2012). *Organizing Entrepreneurial Judgment*. Cambridge: Cambridge University Press.

Gans, J., & Stern, S. (2003). The product market and the market for "ideas": commercialization strategies for technology entrepreneurs. *Research Policy*, *32*(2), 333—350.

Gans, J., 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. (2001). Bringing cognition back in and moving forward. *Journal of Management and Governance*, *5*(3), 213-216.

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

Gavetti, G., & Menon, A. (2016). Evolution cum agency: Toward a model of strategic foresight. *Strategy Science*, *1*(3), 207—233.

Gavetti, G., & Rivkin, J. W. (2007). On the origin of strategy: Action and cognition over time. *Organization Science*, *18*(3), 420—439.

Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. New York, NY: WW Norton.

Ghosh, S. (2021). Entrepreneurship Development.

Ghosh, S., Thomke, S., & Pourkhalkhali, H. (2020). The effects of hierarchy on learning and performance in business experimentation. *Academy of Management Proceedings*, 2020(1), 20500).

Greenstein, S. (2012). Economic experiments and the development of Wi-Fi. *History and strategy*, 29, 3—33.

Greiner, L.E. (1972). Evolution and revolution as organizations grow. *Harvard Business Review*, *76*(3), 55—64.

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.

Gruber, M., & Tal, S. (2017). *Where to play*. Upper Saddle River, New Jersey: FT Publishing International.

Guzman, J., & Stern, S. (2016). The state of American entrepreneurship: New estimates of the quantity and quality of entrepreneurship for 15 US states, 1988-2014. *National Bureau of Economic Research*, (w22095).

Hannigan, T. R., Haans, R. F., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., ... & Jennings, P. D. (2019). Topic modeling in management research: Rendering new theory from textual data. *Academy of Management Annals*, *13*(2), 586—632.

Hayward, M. L., Forster, W. R., Sarasvathy, S. D., & Fredrickson, B. L. (2010). Beyond hubris: How highly confident entrepreneurs rebound to venture again. *Journal of Business Venturing*, *25*(6), 569—578.

Heimeriks, K. H., Schijven, M., & Gates, S. (2012). Manifestations of higher-order routines: The underlying mechanisms of deliberate learning in the context of postacquisition integration. *Academy of Management Journal*, *55*(3), 703—726.

Hodgkinson, G. P., & Sadler-Smith, E. (2018). The dynamics of intuition and analysis in managerial and organizational decision making. *Academy of Management Perspectives*, *32*(4), 473—492.

Karlan, D., Knight, R., & Udry, C. (2015). Consulting and capital experiments with microenterprise tailors in Ghana. *Journal of Economic Behavior & Organization*, *118*, 281—302.

Keil, M., & Mähring, M. (2010). Is your project turning into a black hole?. *California Management Review*, *53*(1), 6—31.

Kerr, W., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives*, *28*(3), 25—48.

Kirtley, J., & O'Mahony, S. (2020). What is a pivot? Explaining when and how entrepreneurial firms decide to make strategic change and pivot. *Strategic Management Journal*.

Klepper, S., & Sleeper, S. (2005). Entry by spinoffs. *Management Science*, *51*(8), 1291—1306.

Kohavi, R., & Thomke, S. (2017). The surprising power of online experiments. *Harvard Business Review*, *95*(5), 74—82.

Koning, R., Hasan, S. and Chatterji, A. (2022). Experimentation and Start-up Performance: Evidence from A/B Testing. *Management Science*.

Lazear, E. P. (2004). Balanced skills and entrepreneurship. *American Economic Review*, *94*(2), 208—211.

Leatherbee, M., & Katila, R. (2020). The lean startup method: Early-stage teams and hypothesis-based probing of business ideas. *Strategic Entrepreneurship Journal*, *14*(4), 570—593.

Lerner, J. (2009). *Boulevard of broken dreams: why public efforts to boost entrepreneurship and venture capital have failed--and what to do about it*. Princeton, New Jersey: Princeton University Press.

Levinthal, D.A. (2017). Mendel in the C-Suite: Design and the Evolution of Strategies. *Strategy Science*, *2*(4), 282—287.

Levinthal, D., & Posen, H. E. (2007). Myopia of selection: Does organizational adaptation limit the efficacy of population selection?. *Administrative science quarterly*, *52*(4), 586—620.

March, J.G. (1991). Exploration and exploitation in organizational learning. *Organization science*, *2*(1), 71—87.

McBride, R., & Wuebker, R. (2020). Social objectivity and entrepreneurial opportunities: Implications for entrepreneurship and management. *Available at SSRN 2427142*.

McDonald, R.M., & Eisenhardt, K.M. (2020). Parallel play: Startups, nascent markets, and effective business-model design. *Administrative Science Quarterly*, 65(2), 483—523.

McGrath, R. (1997). A real options logic for initiating technology positioning investments. *Academy of Management Review*, *22*(4), 974—996.

McGrath, R. (1999). Falling forward: Real options reasoning and entrepreneurial failure. *The Academy of Management Review*, *24*(1), 13.

McGrath, R. (2001). Exploratory learning, innovative capacity, and managerial oversight. *Academy of Management Journal*, *44*(1), 118—131.

McKenzie, D. (2021). Small business training to improve management practices in developing countries: re-assessing the evidence for ‘training doesn’t work’. *Oxford Review of Economic Policy*, *37*(2), 276—301.

Moeen, M., & Agarwal, R. (2017). Incubation of an industry: Heterogeneous knowledge bases and modes of value capture. *Strategic Management Journal*, *38*(3), 566—587.

Moeen, M., Agarwal, R., & Shah, S. K. (2020). Building industries by building knowledge: Uncertainty reduction over industry milestones. *Strategy Science, 5*(3), 218—244.

Murray, F., & Tripsas, M. (2004). The exploratory processes of entrepreneurial firms: The role of purposeful experimentation. *Advances in Strategic Management*, 21, 45—75.

Osterwalder, A., Pigneur, Y., & Clark, T. (2010). *Business Model Generation*. Hoboken, New Jersey: John Wiley & Sons.

Ott, T., & Eisenhardt, K. (2020). Decision weaving: Forming novel, complex strategy in entrepreneurial settings. *Strategic Management Journal*, *41*(12), 2275—2314.

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.

Packard, M. D., Clark, B. B., & Klein, P. G. (2017). Uncertainty types and transitions in the entrepreneurial process. Organization Science, 28(5), 840—856

Pillai, S. D., Goldfarb, B., & Kirsch, D. A. (2020). The origins of firm strategy: Learning by economic experimentation and strategic pivots in the early automobile industry. *Strategic Management Journal*, *41*(3), 369—399.

Reynolds, P. D. (2017). When is a firm born? Alternative criteria and consequences. *Business Economics*, 52(1), 41—56.

Ries, E. (2011). *The Lean Startup: How today's entrepreneurs use continuous innovation to create radically successful businesses*. New York, NY: Crown Publishing Group.

Rosenberg, N. (1982). *Inside The Black Box*. Cambridge: Cambridge University Press.

Santos, F. M., & Eisenhardt, K. M. (2009). Constructing markets and shaping boundaries: Entrepreneurial power in nascent fields. *Academy of Management Journal*, *52*(4), 643—671.

Sarasvathy, S. D. (2009). *Effectuation: Elements of entrepreneurial expertise*. Cheltenham: Edward Elgar Publishing.

Scott, E. L., Shu, P., & Lubynsky, R. M. (2015). Are better ideas more likely to succeed?: An empirical analysis of startup evaluation. *Harvard Business School.* (Working Paper 16-013)

Shah, S. K., Agarwal, R., & Echambadi, R. (2019). Jewels in the crown: Exploring the motivations and team building processes of employee entrepreneurs. *Strategic Management Journal*, 40(9), 1417—1452.

Shepherd, D., & Gruber, M. (2020). The lean startup framework: closing the academic–practitioner divide. *Entrepreneurship Theory and Practice*, 45(5), 967-998.

Siggelkow, N. (2002). Evolution toward fit. *Administrative science quarterly*, *47*(1), 125—159.

Stern, S. (2005). *Economic experiments: The role of entrepreneurship in economic prosperity. In understanding entrepreneurship: A research and policy report*. Ewing Marion Kauffman Foundation, 16–20. Available at SSRN: https://ssrn.com/abstract=1262370

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology*, *29*(1), 24—54.

Tidhar, R., & Eisenhardt, K. (2021). Measure twice, cut once: Scaling novel business models in the nascent online fashion industry. *Academy of Management Proceedings* ,2021(1), 12264.

Thomke, S. (2003). *Experimentation Matters*. Boston, Mass.: Harvard Business School Press.

Wooldridge, J. (2007, April). What's new in econometrics? Lecture 10 difference-in-differences estimation. *NBER Summer Institute*, *9*(2011), 85. www. nber. org/WNE/Slides7–31–07/slides_10_diffindiffs. pdf.

Yang, M. J., Christensen, M., Bloom, N., Sadun, R., & Rivkin, J. (2020). How do CEOs make strategy?. *National Bureau of Economic Research*, (w27952).

Zellweger, T. M., & Zenger, T. R. (2021). Entrepreneurs as scientists: A pragmatist approach to producing value out of uncertainty. *Academy of Management Review*, (ja).

Zenger, T. (2015). New ways to compete: Build a theory, then a strategy. *Harvard Business Review.*
Zollo, M. and Winter, S.G. (2004). Deliberate learning and the evolution of dynamic capabilities. *Organization science*, *13*(3), 339—351.
.

TABLES

Table 1. Balance checks

		Treatment		Control		Difference	
Variable	Definition	Mean	SD	Mean	SD	b	p
Business Age	Age of the business (years)	2.48	3.22	3.28	5.17	0.8	(0.14)
Team size	Number of team members	2.14	1.95	2.31	2.14	0.18	(0.49)
Gender (Female)	Proportion of women in the team	0.42	0.42	0.5	0.44	0.08	(0.15)
Age	Age (team average)	35.77	8.56	36.37	9.2	0.6	(0.59)
Hours - Total Weekly	Weekly hours dedicated to the company (team average)	31.55	18.57	29.61	17.18	-1.94	(0.39)
Background-Economics	Team members with Economics backgrounds (%)	0.15	0.29	0.15	0.29	0	(0.94)
Background - STEM	Team members with a STEM (Science Technology Engineering Mathematics) backgrounds (%)	0.3	0.39	0.36	0.43	0.06	(0.26)
Education	Highest educational level attained by team members (5= PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise; team average)	2.67	0.81	2.58	0.79	-0.1	(0.34)
Confidence	Agreement on a 1-5 scale with the following statements (team average): "We are confident in our entrepreneurial skills", "We are sure we are deploying the best strategy for our business", "We are confident in our ability to manage our business", "We master the competences necessary for our venture", "We are sure there is no better business model for our idea"	3.41	0.7	3.34	0.76	-0.07	(0.44)
Probability Pivot Idea	Probability of making a radical change to the business	45.85	28.18	42.12	26.99	-3.72	(0.28)
Probability Pivot Problem	Probability of changing the problem and customer segment	38.18	26.16	40.55	26.26	2.38	(0.47)
Probability Expansion	Probability of expanding the business outside of the current industry or market	68.25	27.4	66.59	28.12	-1.67	(0.63)
Turnover Annual	Annual turnover (2018) £	50616.11	145448.79	71977.35	195899.81	21361.24	(0.32)
Turnover Monthly	Monthly turnover (January 2019) £	5113.83	17734.76	6099.5	24490.47	985.67	(0.71)
Hours - % Innovation yearly	Working hours dedicated to the design of new products or services in the last year (2018, %)	46.05	33.35	40.02	32.68	-6.04	(0.14)
Hours - % Innovation monthly	Working hours dedicated to the design of new products or services in the last month (January 2019, %)	39.46	34.16	36.84	34.59	-2.62	(0.54)
Idea Value - Mean	Estimated value of the project (mean, 0 to 100)	66.73	17.05	66.62	20.22	-0.11	(0.96)
Idea Value - Range	Estimated value of the project (range, 0 to 100)	39.26	22.03	38	21.94	-1.26	(0.65)
Experience - Industry	Number of years of experience in industry (Team Average)	6.75	6.47	7.7	7.56	0.95	(0.28)

Experience - Work	Number of years of work experience (Team Average)	13.02	7.98	13.53	8.59	0.51	(0.62)
Experience - Entrepreneurial	Number of years of entrepreneurial experience (team average)	3.85	3.49	4.64	5.95	0.79	(0.20)
Experience - Managerial	Number of years of managerial experience (team average)	5.96	5.29	6.22	6.16	0.26	(0.73)
		133		128		261	

Table 2 Descriptive statistics and pairwise correlations

		Obs	Mean	SD	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	Intervention	522	0.510	0.500	0	1.000	1.000																						
2	Revenue (Log 1+,£ 000)	522	1.274	1.673	0	7.290	-0.048	1.000																					
3	Employees (Log 1+)	522	0.794	0.722	0	2.773[9]	0.022	0.354	1.000																				
4	Degree of Business Development	522	60.510	171.293	0	1500.000	-0.066	0.478	0.319	1.000																			
5	Probability of *not* changing the business	522	56.000	27.418	0	100.000	-0.067	0.089	0.119	0.102	1.000																		
6	Prior experience (work)	522	13.244	8.174	0	40.000	-0.032	0.086	-0.013	0.010	-0.091	1.000																	
7	Prior experience (industry)	522	7.152	6.919	0	35.000	-0.072	0.069	-0.026	0.049	-0.125	0.605	1.000																
8	Prior experience (managerial)	522	6.011	5.566	0	30.000	-0.025	-0.028	-0.023	-0.020	-0.150	0.745	0.517	1.000															
9	Prior experience (entrepreneurial)	522	4.190	4.774	0	30.000	-0.081	-0.009	-0.043	-0.040	-0.149	0.556	0.463	0.608	1.000														
10	Confidence	522	3.378	0.723	1.2	5.000	0.048	0.042	0.122	0.117	0.252	-0.115	-0.048	-0.033	0.054	1.000													
11	Cost (Log 1+, in £ 000)	522	1.225	1.468	0	6.998	-0.022	0.679	0.297	0.328	0.031	0.116	0.081	0.040	0.026	0.052	1.000												
12	Value Added Log (Abs(Min Value Added) +) (in £ 000)	522	5.102	0.291	0	6.315	-0.034	0.352	0.038	0.263	0.017	0.044	0.027	-0.001	0.004	-0.033	-0.023	1.000											
13	Multiple Issues	522	0.134	0.341	0	1.000	0.094	-0.004	-0.098	-0.015	0.028	-0.047	0.062	-0.043	-0.044	-0.004	0.007	0.035	1.000										
14	Market Validation Issue	522	0.142	0.349	0	1.000	0.025	-0.071	-0.052	-0.079	-0.058	-0.031	0.064	-0.017	-0.111	-0.024	-0.070	0.017	0.227	1.000									
15	Scale Up Issue	522	0.318	0.466	0	1.000	0.045	0.181	0.021	0.099	-0.097	0.013	0.064	-0.025	0.102	0.064	0.126	0.070	0.118	-0.278	1.000								
16	Multiple Goals	522	0.115	0.319	0	1.000	-0.007	-0.016	-0.053	0.029	-0.015	-0.074	-0.048	-0.065	-0.068	0.028	-0.036	0.005	0.211	0.060	0.063	1.000							
17	Market Validation Goal	522	0.123	0.328	0	1.000	0.063	-0.124	-0.100	-0.111	-0.013	-0.040	-0.063	0.060	0.015	-0.024	-0.068	-0.024	0.059	0.150	0.021	0.232	1.000						
18	Scale Up Goal	522	0.674	0.469	0	1.000	-0.028	0.176	0.026	0.009	0.018	0.123	0.068	0.003	0.040	-0.042	0.108	0.035	-0.062	-0.140	0.088	-0.237	-0.538	1.000					
19	Certainty (Log 1+)	346	0.652	0.227	0	1.230	0.019	-0.045	-0.139	-0.070	0.119	0.059	0.066	0.091	0.125	-0.006	-0.057	-0.025	0.054	-0.014	0.010	0.012	-0.010	-0.048	1.000				
20	Idea quality (self-assessed)	522	66.730	18.502	1	100.000	0.006	0.128	0.170	0.143	0.174	-0.066	-0.043	-0.055	0.083	0.300	0.153	0.025	0.023	-0.132	0.033	0.063	-0.028	-0.024	0.020	1.000			
21	Idea quality (self-assessed)	522	3.047	0.416	2	4.167	-0.047	0.007	0.121	0.062	0.037	-0.157	-0.122	-0.082	-0.026	0.136	0.042	0.011	-0.094	-0.063	-0.024	0.022	0.000	0.019	-0.046	0.162	1.000		
22	Number of pivots (Log 1+)	522	0.580	0.896	0	3.258	-0.008	0.296	-0.052	-0.050	-0.074	0.035	0.001	0.053	0.067	-0.020	0.454	0.035	0.037	-0.013	0.028	0.008	0.052	0.069	0.058	0.000	0.045	1.000	
23	Scientific intensity	522	2.680	1.147	0	5.000	0.075	0.051	0.086	-0.043	0.054	-0.193	-0.175	-0.132	-0.141	0.022	0.190	0.013	-0.064	0.023	-0.040	-0.015	0.044	0.002	-0.045	0.098	0.276	0.268	1.000

.

[9] The requirement that entrepreneurs have fewer than 10 employees was a criterion we imposed for firms at the time of our baseline observation. As some of these firms grew during the program, some of them reached a higher number of employees. The maximum value of the number of employees in the table equals 15 and reflects this increase in size. The maximum number of employees at the baseline was 10 (instead of what we would expect, i.e., 9). This is because four firms had 10 employees at the baseline. Even if we targeted firms with fewer than 10 employees and admitted into the program only firms that met this threshold, four firms told us that their baseline answer was incorrect subsequently to their admission into the program. We therefore kept these firms in our sample but reported them with the correct number of employees. We have replicated our analyses without including these firms and results were consistent with those reported in the paper.

Table 3. Impact of the treatment on performance: Firm size

VARIABLES	(1) Log (1+) Employ- ees OLS Panel	(2) Log (1+) Employees OLS Panel	(3) Log (1+) Employ- ees OLS Panel	(4) Log (1+) Employ- ees OLS Panel
Treatment X Post	0.1160	0.1053	0.1160	0.1053
	(0.0197)	(0.0511)	(0.0202)	(0.0515)
Treatment X Post X Degree of Business Development		0.0002		0.0002
		(0.0834)		(0.0837)
Treatment	0.0043	-0.0312		
	(0.9573)	(0.7174)		
Post	-0.0763	-0.0766	-0.0763	-0.0766
	(0.0247)	(0.0390)	(0.0253)	(0.0394)
Degree of Business Development	0.0014	0.0011		
	(0.0002)	(0.0104)		
Treatment X Degree of Business Development		0.0006		
		(0.4369)		
Post X Degree of Business Development		0.0000		0.0000
		(0.9588)		(0.9587)
Constant	0.7183	0.7368	0.8024	0.8024
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	522	522	522	522
R-squared	0.112	0.119	0.022	0.026
Number of id	261	261	261	261
Firm FE	-	-	Yes	Yes
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table 4. Impact of the treatment on performance: Firm Revenue

VARIABLES	(1) Log (1+) Revenue (in £ 000) OLS Panel	(2) Log (1+) Revenue (in £ 000) OLS Panel	(3) Log (1+) Revenue (in £ 000) OLS Panel	(4) Log (1+) Revenue (in £ 000) OLS Panel
Treatment X Post	-0.1197	-0.2998	-0.1197	-0.2998
	(0.4596)	(0.0690)	(0.4594)	(0.0694)
Treatment X Post X Degree of Business Development		0.0035		0.0035
		(0.0090)		(0.0093)
Treatment	0.0043	-0.0832		
	(0.9696)	(0.4289)		
Post	1.0572	1.0869	1.0572	1.0869
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Degree of Business Development	0.0047	0.0038		
	(0.0000)	(0.0000)		
Treatment X Degree of Business Development		0.0014		
		(0.2361)		
Post X Degree of Business Development		-0.0004		-0.0004
		(0.2763)		(0.2759)
Constant	0.4915	0.5565	0.7757	0.7757
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	522	522	522	522
R-squared	0.301	0.331	0.372	0.409
Number of firms	261	261	261	261
Firm FE	-	-	Yes	Yes
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table 5. Problems and goals

	Lower Degree of Business Development		Higher Degree of Business Development		Difference	
	Mean	SD	Mean	SD	b	p
Multiple Problems	0.16	0.37	0.1	0.3	0.06	(0.142)
Market Validation Problems	0.19	0.4	0.09	0.28	0.11	(0.012)
Scale Up Problems	0.26	0.44	0.38	0.49	-0.12	(0.044)
Multiple Goals	0.13	0.33	0.1	0.3	0.02	(0.536)
Market Validation Goal	0.19	0.4	0.05	0.21	0.15	(0.000)
Scale Up Goal	0.57	0.5	0.79	0.41	0.22	(0.000)
Observations	134		127		261	

Table 6 Uncertainty resolution: Degree of certainty

	(1) Certainty (Log 1+)	(2) Certainty (Log 1+)
VARIABLES	OLS Panel	OLS Panel
Treatment X Post	0.0335	0.0092
	(0.4266)	(0.8366)
Treatment X Post X Degree of Business Development		0.0006
		(0.0403)
Treatment	-0.0011	-0.0097
	(0.9727)	(0.7807)
Post	0.0093	0.0307
	(0.7623)	(0.3379)
Degree of Business Development	-0.0001	-0.0003
	(0.1496)	(0.1858)
Treatment X Degree of Business Development		0.0002
		(0.3365)
Post X Degree of Business Development		-0.0005
		(0.0435)
Constant	0.6488	0.6550
	(0.0000)	(0.0000)
Observations	346	346
Number of id	173	173
Firm FE	-	-
Clustered Errors	Firm	Firm

Robust pval in parentheses

Table 7 Quality of the value proposition

	(1) Log (1+) Employees Own eval	(2) Log (1+) Revenue (in £ 000) Own eval	(3) Log (1+) Employees Ext eval	(4) Log (1+) Revenue (in £ 000) Ext eval
Treatment X Post	0.184	-0.503	0.4396	-2.3783
	(0.3303)	(0.3254)	(0.2086)	(0.0595)
Treatment X Post X Idea Quality (self-evaluation)	-0.001	0.0057		
	(0.7212)	(0.4692)		
Treatment X Post X Idea Quality (external evaluation)			-0.1052	0.7418
			(0.3514)	(0.0735)
Treatment	-0.2863	-0.2503	-0.5718	0.3088
	(0.357)	(0.6372)	(0.3363)	(0.7961)
Idea Quality (self- evaluation)	0.0039	0.0074		
	(0.2166)	(0.1471)		
Treatment X Idea Quality (self-evaluation)	0.0039	0.0022		
	(0.4043)	(0.7897)		
Post	-0.2560	0.7840	-0.4869	2.0632
	(0.0663)	(0.0431)	(0.0705)	(0.0269)
Post X Idea Quality (self- evaluation)	0.0027	0.0041		
	(0.1826)	(0.4922)		
Idea Quality (external evaluation)			0.0829	0.0669
			(0.5517)	(0.8205)
Treatment X Idea Quality (external evaluation)			0.1813	-0.1344
			(0.3535)	(0.7323)
Post X Idea Quality (external evaluation)			0.1339	-0.3281
			(0.1156)	(0.2667)
Constant	0.5581	0.3349	0.5616	0.6218
	(0.0088)	(0.3015)	(0.1947)	(0.4849)
Observations	522	522	522	522
Number of id	261	261	261	261
Firm FE	-	-	-	-
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table 8 Pivot

	(1)	(2)
	#Pivot (Log 1+) OLS	#Pivot (Log 1+) OLS
	Cross Section	Cross Section
Treatment	-0.0424	-0.1186
	(0.7246)	(0.3411)
Treatment X Degree of Business Development		0.0013
		(0.0302)
Degree of Business Development	-0.0006	-0.0011
	(0.1147)	(0.0002)
Constant	0.9366	0.9799
	(0.0000)	(0.0000)
Observations	261	261
R-squared	0.038	0.050
Mentors dummies	Yes	Yes
Clustered Errors	Firm	Firm

Robust pval in parentheses

APPENDIX

Section 1

Table A1a Scientific intensity components

Component	Sub-component	Definition	Score
Theory	Clarity of theory	The extent to which the theory is understandable	0 (no theory) or from 1 (not clear) to 5 (extremely clear)
Theory	Articulation of theory	The extent to which the theory is detailed	0 (no theory) or from (not detailed) to 5 (extremely detailed)
Theory	Consideration of alternatives	The extent to which the theory includes alternative possible options	0 (no theory) or from 1 (no consideration of alternatives) to 5 (careful consideration of many alternatives)
Theory	Theory based on evidence	The extent to which the theory is based on objective evidence	0 (no theory) or from 1 (theory not based on objective evidence) to 5 (extremely based on objective evidence)
Hypotheses	Explicitness of hypotheses	The extent to which the respondent can articulate the fundamental assumptions that make his/her business viable	0 (no hypotheses) or from 1 (not explicit hypotheses) to 5 (extremely explicit)
Hypotheses	Coherence of hypotheses	The extent to which hypotheses are coherent with the theory	0 (no hypotheses) or from 1 (not coherent) to 5 (extremely coherent)
Hypotheses	Level of details of hypotheses	The extent to which hypotheses clearly indicate the details of what the entrepreneur wishes to learn and how to measure it	0 (no hypotheses) of from 1 (not detailed) to 5 (extremely detailed)
Hypotheses	Falsifiability of hypotheses	The extent to which it is possible to clearly determine (after tests) whether the hypotheses are supported or not	0 (no hypotheses) or from 1 (not falsifiable) to 5 (extremely falsifiable)
Tests	Coherence of tests	The extent to which the test is coherent with the hypotheses	0 (no tests) or from 1 (not coherent) to 5 (extremely coherent)
Tests	Validity of tests	The extent to which the test has been conducted in a context similar to which the business operates	0 (no hypotheses) or from 1 (not valid) to 5 (extremely valid)
Tests	Representativeness of tests	The extent to which the test has been conducted with a sample that is representative of the broad group the firm targets	0 (no hypotheses) or from 1 (not representative) to 5 (extremely representative)
Tests	Rigorousness of tests	The extent to which the appropriate test and procedure for that type of test have been chosen for hypotheses-testing	0 (no hypotheses) or from 1 (not rigorous) to 5 (extremely rigorous)
Evaluation	Data-based assessment	The extent to which the evaluation is based on data	0 (no hypotheses) or from 1 (not based on data) to 5 (extremely based on data)
Evaluation	Coherence of measures	The extent to which the measure used are consistent with the learning objective the entrepreneur has in mind	0 (no hypotheses) or from 1 (not coherent) to 5 (extremely coherent)
Evaluation	Systematic evaluation	The extent to which the evaluation is based on systematically collected and analyzed data	0 (no hypotheses) or from 1 (not systematic) to 5 (extremely systematic)
Evaluation	Explanatory power of evaluation	The extent to which the evaluation results in clarity on the main findings from the test and their implications for the business	0 (no hypotheses) or from 1 (not explanatory) to 5 (extremely explanatory)

Table A1b Scientific intensity

	Treatment		Control		Difference	
Scientific intensity	Mean	SD	Mean	SD	b	p
Interview 0	2.56	1.23	2.35	1.29	-0.2	(0.20)
Interview 1	2.93	1.07	2.69	1.18	-0.25	(0.08)
Interview 2	2.98	1.01	2.73	1.04	-0.25	(0.05)
Interview 3	3.01	0.98	2.76	1.01	-0.24	(0.05)
Interview 4	2.95	0.93	2.73	1.02	-0.22	(0.06)
Interview 5	2.94	0.95	2.75	1.02	-0.19	(0.12)
Interview 6	2.95	0.93	2.76	0.99	-0.19	(0.12)
Interview 7	2.97	0.94	2.78	0.99	-0.18	(0.13)
Interview 8	2.97	0.95	2.83	0.99	-0.14	(0.24)
Observations	133		128		261	

Table A2 Alternative measures of degree of business development: Probability of *not* making changes

VARIABLES	(1) Log (1+) Employees OLS Panel	(2) Log (1+) Employees OLS Panel	(3) Log (1+) Revenue (in £ 000) OLS Panel	(4) Log (1+) Revenue (in £ 000) OLS Panel
Treatment X Post	0.1160	0.1890	-0.1197	-1.1561
	(0.0197)	(0.1139)	(0.4596)	(0.0029)
Treatment X Post X Degree of Business Development (Probability of *not* making changes)		-0.0013		0.0184
		(0.4844)		(0.0020)
Treatment	-0.0144	-0.0458	-0.0813	-0.1432
	(0.8657)	(0.8086)	(0.5786)	(0.6315)
Post	-0.0763	-0.1421	1.0572	1.6783
	(0.0247)	(0.1266)	(0.0000)	(0.0000)
Degree of Business Development (Probability of *not* making changes)	0.0032	0.0027	0.0052	0.0052
	(0.0432)	(0.2526)	(0.1141)	(0.1799)
Treatment X Degree of Business Development (Probability of *not* making changes)		0.0005		0.0011
		(0.8651)		(0.8328)
Post X Degree of Business Development (Probability of *not* making changes)		0.0011		-0.0107
		(0.3963)		(0.0221)
Constant	0.6310	0.6612	0.5234	0.5272
	(0.0000)	(0.0000)	(0.0097)	(0.0168)
Observations	522	522	522	522
Number of id	261	261	261	261
Firm FE	-	-	-	-
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table A3. Winsorization

	(1) Log (1+) Employees OLS Panel	(2) Log (1+) Employees OLS Panel	(3) Log (1+) Revenue (in £ 000) OLS Panel	(4) Log (1+) Revenue (in £ 000) OLS Panel
Treatment X Post	0.1132	0.1026	-0.1360	-0.2767
	(0.0207)	(0.0542)	(0.3984)	(0.0956)
Treatment X Post X Degree of Business Development		0.0002		0.0027
		(0.0856)		(0.0714)
Treatment	0.0043	-0.0312	0.0011	-0.0832
	(0.9574)	(0.7174)	(0.9924)	(0.4289)
Post	-0.0763	-0.0766	1.0566	1.0861
	(0.0247)	(0.0390)	(0.0000)	(0.0000)
Degree of Business Development	0.0014	0.0011	0.0045	0.0038
	(0.0002)	(0.0104)	(0.0000)	(0.0000)
Treatment X Degree of Business Development		0.0006		0.0014
		(0.4369)		(0.2361)
Post X Degree of Business Development		0.0000		-0.0004
		(0.9588)		(0.2767)
Constant	0.7183	0.7368	0.5018	0.5565
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	522	522	522	522
Number of id	261	261	261	261
Firm FE	Yes	Yes	Yes	Yes
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table A4a. Attritrion

Interview Number	In	Withdrawn	With-drawn %
0	261	0	
1	223	38	15%
2	212	11	4%
3	207	5	2%
4	193	14	5%
5	185	8	3%
6	173	12	5%
7	163	10	4%
8	147	16	6%

Table A4b. Attrition: Probability of withdrawing from the program

VARIABLES	(1) Early Withdraw OLS Cross-section
Treatment	-0.0167
	(0.7862)
Constant	0.4453
	(0.0000)
Observations	261
R-squared	0.000
Clustered Errors	Firm

Table A5. Attrition: Alternative specifications

VARIABLES	(1) Log (1+) Employees OLS Panel	(2) Log (1+) Employees OLS Panel	(3) Log (1+) Revenue (in £ 000) OLS Panel	(4) Log (1+) Revenue (in £ 000) OLS Panel
Treatment X Post	0.0970	0.0709	-0.4508	-0.1344
	(0.1379)	(0.0343)	(0.5062)	(0.0803)
Treatment X Post X Degree of Business Development	0.0003	0.0000	0.0041	0.0014
	(0.0642)	(0.7719)	(0.0213)	(0.0061)
Treatment	-0.0312	0.0032	-0.0832	-0.2486
	(0.7174)	(0.9697)	(0.4289)	(0.1354)
Post	-0.0564	-0.0558	1.9572	1.0442
	(0.2471)	(0.0698)	(0.0011)	(0.0000)
Degree of Business Development	0.0011	0.0011	0.0038	0.0036
	(0.0104)	(0.0084)	(0.0000)	(0.0000)
Treatment X Degree of Business Development	0.0006	0.0008	0.0014	0.0034
	(0.4369)	(0.3088)	(0.2361)	(0.0390)
Post X Degree of Business Development	-0.0000	0.0000	-0.0015	-0.0003
	(0.7854)	(0.5227)	(0.1455)	(0.0657)
Constant	0.7368	0.7160	0.5565	0.5991
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	522	2,349	522	2,349
Number of id	261	261	261	261
Interview effect	-	Included	-	Included
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table A6 Excluding alternative explanations

VARIABLES	(1) Log (1+) Employees OLS Panel	(2) Log (1+) Employees OLS Panel	(3) Log (1+) Revenue (in £ 000) OLS Panel	(4) Log (1+) Revenue (in £ 000) OLS Panel
Treatment X Post	0.2373	0.1864	0.1090	-0.4123
	(0.0148)	(0.3884)	(0.7152)	(0.5777)
Treatment X Post X Work Experience	-0.0093		-0.0166	
	(0.0951)		(0.4032)	
Treatment X Post X Confidence		-0.0211		0.0865
		(0.7422)		(0.6889)
Treatment	-0.0962	-0.4612	-0.2634	0.1016
	(0.5572)	(0.2575)	(0.3072)	(0.8853)
Work Experience	-0.0021		0.0023	
	(0.7521)		(0.8319)	
Treatment X Work Experience	0.0053		0.0126	
	(0.5778)		(0.4383)	
Post	-0.0950	-0.1500	0.7079	1.1809
	(0.1504)	(0.3168)	(0.0016)	(0.0328)
Post X Work Experience	0.0014		0.0259	
	(0.7310)		(0.0930)	
Confidence		0.0558		0.1290
		(0.5120)		(0.3908)
Treatment X Confidence		0.1263		-0.0619
		(0.2979)		(0.7684)
Post X Confidence		0.0221		-0.0370
		(0.5994)		(0.8151)
Constant	0.8441	0.6293	0.7954	0.3959
	(0.0000)	(0.0269)	(0.0000)	(0.4329)
Observations	522	522	522	522
Number of id	261	261	261	261
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table A7 Alternative measures of experience

VARIABLES	(1) Log (1+) Employees OLS Panel	(2) Log (1+) Employees OLS Panel	(3) Log (1+) Employees OLS Panel	(4) Log (1+) Revenue (in £ 000) OLS Panel	(5) Log (1+) Revenue (in £ 000) OLS Panel	(6) Log (1+) Revenue (in £ 000) OLS Panel
Treatment X Post	0.1496	0.1853	0.2051	-0.1343	-0.0751	-0.0054
	(0.0406)	(0.0139)	(0.0034)	(0.5494)	(0.7496)	(0.9809)
Treatment X Post X Industry Experience	-0.0052			0.0040		
	(0.4011)			(0.8699)		
Treatment X Post X Managerial Experience		-0.0118			-0.0073	
		(0.0909)			(0.8084)	
Treatment X Post X Entrepreneurial Experience			-0.0238			-0.0245
			(0.0081)			(0.5329)
Treatment	-0.1209	-0.1085	-0.0544	-0.1797	-0.0985	-0.1543
	(0.3056)	(0.3780)	(0.6416)	(0.3480)	(0.6375)	(0.4395)
Industry Experience	-0.0067			0.0039		
	(0.3270)			(0.7298)		
Managerial Experience		-0.0065			-0.0099	
		(0.5083)			(0.5009)	
Entrepreneurial Experience			-0.0038			-0.0176
			(0.6348)			(0.1751)
Treatment X Work Experience	0.0132			0.0125		
	(0.2043)			(0.4948)		
Treatment X Managerial Experience		0.0137			-0.0008	
		(0.3265)			(0.9692)	
Treatment X Entrepreneurial Experience			0.0066			0.0105
			(0.7053)			(0.7156)
Post	-0.0699	-0.0790	-0.0660	0.9648	1.0213	0.9341
	(0.1448)	(0.0845)	(0.1307)	(0.0000)	(0.0000)	(0.0000)
Post X Industry Experience	-0.0008			0.0121		
	(0.8404)			(0.4996)		
Post X Managerial Experience		0.0005			0.0058	
		(0.9243)			(0.7865)	
Post X Entrepreneurial Experience			-0.0022			0.0269
			(0.6662)			(0.2701)
Constant	0.8669	0.8556	0.8330	0.7973	0.8877	0.9076
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	522	522	522	522	522	522
Number of id	261	261	261	261	261	261
Clustered Errors	Firm	Firm	Firm	Firm	Firm	Firm

Robust pval in parentheses

Table A8 Cost and value added

VARIABLES	(1) Log (1+) Cost (in £ 000) OLS Panel	(2) Log (1+) Cost (in £ 000) OLS Panel	(3) Log (Abs(Min Value Added) +) Value Added (in £ 000) OLS Panel	(4) Log (Abs(Min Value Added) +) Value Added (in £ 000) OLS Panel
Treatment X Post	0.1251	-0.0881	-0.0502	-0.0730
	(0.4500)	(0.5943)	(0.2718)	(0.1315)
Treatment X Post X Degree of Business Development		0.0041		0.0005
		(0.0026)		(0.0157)
Treatment	-0.0645	-0.0830	0.0150	-0.0002
	(0.4362)	(0.3687)	(0.0593)	(0.9820)
Degree of Business Development	0.0028	0.0023	0.0004	0.0002
	(0.0029)	(0.0359)	(0.0071)	(0.3807)
Treatment X Degree of Business Development		0.0001		0.0002
		(0.9092)		(0.4192)
Post	1.3588	1.3974	0.0215	0.0126
	(0.0000)	(0.0000)	(0.2154)	(0.4820)
Post X Degree of Business Development		-0.0005		0.0001
		(0.3260)		(0.2121)
Constant	0.3767	0.4137	5.0694	5.0858
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	522	522	522	522
Number of id	261	261	261	261
Clustered Errors	Firm	Firm	Firm	Firm

Robust pval in parentheses

Section 2

Examples of differences in slides and in-class activities between treatment and control

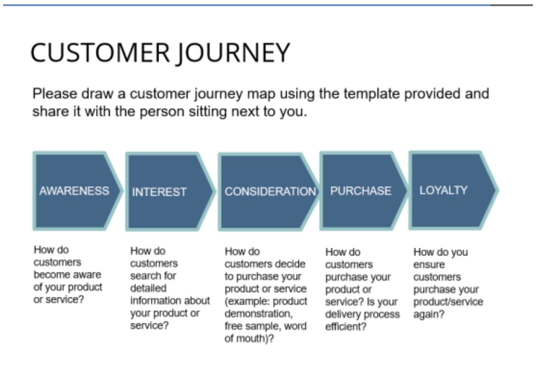
In Training Session 2, both groups learned about the customer journey map as a tool to systematically examine what their potential customers were currently doing and where they might experience bottlenecks. As shown in Figure A1 below, where we reproduce some of the slides used during the training program, both groups learned about the key building blocks of the customer journey and were exposed to the same example (IKEA). Entrepreneurs in the treatment group were explicitly shown how the customer journey map could be used by IKEA to develop theory and hypotheses, and to test and evaluate them. The control group, instead, devoted more time to learning about the content that was common to both groups. We highlight in green the part that was different for the treatment group.

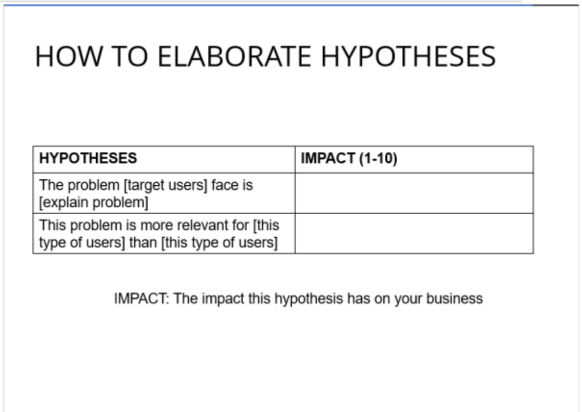
Figure A1. Extract from Session 2 slides: Treatment vs. control group

After this portion of the lecture, both groups conducted an in-class activity, during which they were invited to complete a customer journey map for their business. The in-class activity was followed by a debrief guided by the instructor, where entrepreneurs received feedback on their customer journey maps. These activities were helpful in getting entrepreneurs to apply the content of the class to their business right away, as well as to clarify any doubts entrepreneurs might have with the instructors. Instructions provided to the two groups are reproduced below in Figure A2. For the treatment group, part of the exercise involved the application of the scientific approach, while an alternative question was asked to the control group. Since Training Session 2 was one of the earlier sessions, the focus of the lecture for the treatment group was on developing theory and hypotheses, but in later sessions the focus broadened to include how to test the hypotheses and evaluate results.

Figure A2. **In-class activities from Session 2 slides: Treatment vs. control group**

TREATMENT GROUP





CONTROL GROUP