Università Commerciale "Luigi Bocconi"
PhD School

PhD program in: Business Administration and Management

Cycle: XXXIII

Disciplinary Field (code): SECS P/08

Essays on predictive and non-predictive strategies: real and simulated experiments

Advisor: Alfonso Gambardella

Co-Advisor: Arnaldo Camuffo

PhD Thesis by

Danilo Messinese

ID number: 3054217

Year: 2022

Abstract

This work studies how adopting different decision-making approaches under uncertainty could impact differently the decision's outcomes. Humankind has been always obsessed by the future. Humans struggle to predict it to avoid negative outcomes. This is certainly true in the business context, where managers and, more broadly, decision-makers are continuously asked to make decisions with high impact on their firm's performance. While making decisions about the future might be an easy task under stable and clear conditions, it becomes a crucial aspect when decision-makers act under uncertainty. Even if the pathway is difficult and tiring, walking in the sunlight is always easier than walking in the dark. We are currently living in the era of uncertainty, being subject to a strong and fast technology evolution, to financial and economic crisis, to environmental crisis and ultimately to health crisis. At the same time, machines and algorithms are evolving at a fast pace, threatening and competing with humans in many tasks, included making decisions. With these motivations in mind, I believe that studying how decision's outcomes are impacted by the adoption of a certain decision-making approach is crucial. Since I started my reasoning from the concept of future, I focus my attention on the adoption of a predictive or a non-predictive approach to decision-making to explore innovative ideas. I also compare results from real contexts with results from a simulated reality. The main take-away is that how decision-makers acquire information is what triggers better decisions and, ultimately, better performance. In the first chapter I study the difference bwetween exploring ideas with a predictive and a non-predictive approach to decision-making, operationalized as the scientific and effectual approaches.

Managers and entrepreneurs do not have a structured approach and solid routines to make decisions under uncertainty, such as launch a new product or a new service, entering in a new market, or both. A survey from Harvard Business Review reports that 646 managers admit relying on their intuition instead of deciding with a systematic process. Research on cognitive biases has widely shown how managers and entrepreneurs tend to be overconfident and therefore more likely to incur in biased decisions and false positive outcomes (Astebro, 2003; Malmendier and Tate, 2008; Galasso and Simcoe, 2011; Astebro et al., 2014). Coherently with previous literature (Camuffo et al. 2020), I call this approach "scientific" since managers and entrepreneurs are asked to act as scientist would do in a business context. Indeed, this approach consists in developing theories and logic connections about the mechanism underlying future outcomes and test them with tailored experiments. "Scientific" managers and entrepreneurs are then called to analyse test results and make decisions accordingly. As mentioned, this approach helps decision-makers to improve their predictive power by probing the future with theory-based experiments but remains explorative. Managers and entrepreneurs explore other alternative ideas on which they can theorize on. On the non-predictive side, research on effectuation (Sarasvathy 2001; Dew et al. 2009; Chandler et Al. 2011) has shown how managers and entrepreneurs can deal with uncertainty by adopting a decision-making approach aimed to control the future instead of predicting it. Effectual decision-makers select alternative ideas based on loss affordability, experimentation, and flexibility. But, in this case, experiments are not guided by well-framed theories and are not part of a systematic process. Sarasvathy (2008) uses the metaphor of a patchwork quilt: managers and entrepreneurs see the business context as a table where all the pieces are there but must be assembled or even created as the future is unpredictable. With the aim to unfold the mechanism driving different termination rates of ideas from the adoption of these two approaches, I propose a stylized model with the aim to predict empirical results. The model proposes a Bayesian framework, where decision-makers acquire costly information to improve the precision of signals. Based on these informative signals, they act accordingly. I expect scientific decision-makers to react promptly to very informative bad signals. While I expect effectual decision-makers to react less to bad signals since they weight less predictive information. This translates into higher rates of termination for scientific decision-makers than effectual decision-makers. Moreover, scientific decision-makers terminate earlier than effectual decision-makers. I test the model with data from a Randomized Control Trial (RCT) with early-stage start-ups. The RCT was conducted online in Italy in 2020 and involved early-stage start-ups. Coherently with the model, I find that scientific start-ups terminate more and earlier because of more informative and negative signals. Effectual start-ups do act accordingly to more informative and negative signals. This is coherent with the nature of the two approaches, where the scientific one prescribes to use information to predict the future, while the effectual one prescribes to turn bad news into new opportunities and to exert control over the future. In the second chapter, I focus on the scientific approach solely. I provide evidence of the implications of a scientific approach to decision-making through four Randomized Control Trials, involving start-ups and small-medium firms (SMEs) across two countries, Italy and UK. The three main findings are that scientific decision-makers are more likely to terminate their idea in early stages, cofirming findings of the previous chapter. They pivot fewer times before committing to one or terminate the idea. They also perform better in terms of revenues. A model has been developed to explain empirical results. In the third chapter, I study a way to scale research findings by using a simulation game to replicate, to some extent, results of the previous two chapters about the scientific approach. I run a lab experiment using a simulation game I developed with the essential support of BUILT (Bocconi University Innovations in Learning and Technology) at Bocconi University (which hold all the rights on the game). The game simulates the launch of a start-up and allows players to adopt a more scientific approach to decision-making. The idea behind this chapter is that on the one hand running field experiments is extremely costly in many ways and, on the other hand, as other scientific fields, there is an increasing need to replicate research findings to a larger scale (Goldfarb and King, 2016, Astebro and Hoos, 2020). In this light, simulation games can help to scale the size of results and replicate findings.To validate the simulation game, I ran a lab experiment with master students where around half of them were trained to adopt a scientific approach to decision-making. I find that the simulation game replicates previous findings with real start-ups: trained students terminate more their start-up instead of launching it and get more conservative. Overall, this work shows robust findings about the adoption of a more scientific approach to decision-making under uncertainty. Moreover, it proposes a mechanism to explain what drives these results and compares them with findings from the adoption of a non-predictive approach to decision-making. It concludes with a first attempt to replicate these findings in a simulated reality that could help to scale research in the field at a faster pace. The work proceeds with the three chapters and the appendix of the first chapter has been inserted after the third chapter.

Acknowledgements

Unertainty is a lifestyle, not only matter of study. It is the dark that makes you looking for the light. I did love my PhD journey. I started from Physics years ago, learning something about Nature. After a few years, I landed at Bocconi to learn as much as possible about Society and Economics, soomething it was new for me. It has been tough and amazing. I have to thank my advisor Alfonso, for his non-stopping energy and inspirational attitude, and my co-advisor Arnaldo: two different sharp and smart points of view. I learned a lot and I will learn a lot from them. This journey has been possible only thanks to all the people who continously love me since years. And I do love them. They know. And I thank myself for having decided to get back to academia that crazy day, five years ago.

Let's change the world now.

Chapter 1

Exploring with predictive and control strategies under uncertainty

Prior research suggests that managers and entrepreneurs adopt explorative approaches to cope with uncertainty. This paper studies the implications of adopting (explorative) predictive and control-oriented strategies, empirically operationalized respectively as scientific and effectual approach (Camuffo et al. 2021; Sarasvathy 2001), and directly tests a potential mechanism driving the different decision outcomes. By analysing data from a Randomized Control Trial (RCT) involving 308 start-ups and 2,772 data points over time, I find that predictive decision-makers terminate ideas more frequently and earlier, while control-oriented decision-makers terminate less frequently and later, due different strategies of dynamic information acquisition during the exploration phase.

1.1 Introduction

When it comes to innovate, entrepreneurs and managers, hereafter Decision-Makers (*DMs*), face the crucial challenge of coping with uncertainty to decide whether to develop their business idea. They reduce the uncertainty by exploring different alternatives before committing to or terminating their business idea (Gans, Stern and Wu 2019; Jones and Pratap, 2020). Uncertainty theorists have defined different levels on uncertainty, which is commonly conceptualized to exist on a continuum, from weak to strong, and reasonate about to what extent uncertainty can be solved. Beside an interesting and profound debate around uncertainty mitigability (Alvarez and Barney, 2005; Wiltbank et al., 2006; Packard and Clark, 2020), most researchers seem to agree about the fact that *DMs* in possess of most and best information can, at least in principle, mitigate the perceived uncertainty (Townsend, Hunt, McMullen and Sarasvathy, 2018). While this debate is still open, it is clear how *DMs* face situations with different types and 'strengh' of uncertainty and struggle to adopt an effective approach to navigate these situations. In this light, prediction has been always playing a central role in strategy making under uncertainty because it allows to establish causal connections between actions and possible consequences to produce favourable outcomes. While pure predictive approaches, such as classic planning, might show limitations in dynamic situations where firms are called to adapt to changes (Mintzberg, 1990, Mosakowski, 1997), other predictive approaches, such as real options (McGrath, 1999) reavealed how prediction can be used also when fast adaptation and a more esplorative approach is required. Recent research (Camuffo et al. 2020a, 2021) studied how innovators can deal with high level of uncertainty by mirroring scientists to esplore an idea before making a decision. 'Scientific' *DMs* use prediction to test well-framed theories to learn and explore an idea in an uncertain environment. Other research works have highlighted how the power of prediction can be dependent on the environmental conditions (Mintzberg, 1994). In this light, *DMs* can opt for exert efforts to control the future instead of predict it. This means that they depict the future as difficult or even impossible to predict and, therefore, try to create and influence the evolution of events to induce favorable outcomes. Both exploring an idea with a predictive or a control-oriented approach involve beliefs about possible future events and require the information of acquisition. While the former seeks information to predict the future, the latter seeks information to understand how to create it, both struggling towards successful outcomes. The debate about what approach fits better different conditions of uncertainty is still ongoing, given an unclear explanation and empirical evidence of the mechanism underlying these two strategies and producing different outcomes. From an empirical point of view, for instance, Camuffo et al. (2020a, 2021) find that entrepreneurs exploring ideas with a predictive approach are more likely to terminate their idea and do so earlier than entrepreneurs making decisions based on their heuristics. They also perform better. On the control side, preminent scholars (Sarasvathy 2001, Dew et al. 2009) theorized that entrepreneurs that adopt a control-oriented strategy are less likely to abandon their idea and shape the future in their favour, setting an affordable level of risk. In this light, this paper has the goal to unfold a potential mechanism to explain *why DMs* adopting a predictive strategy or a control-oriented strategy 1) decide to commit to a business idea or terminate it and 2) *why* they do so ealier or later. I propose a stylized continuous-time sthocastic control problem to explain these two points, where *DMs* can control a precision parameter, namely information, the techonology to acquire information and the risk level. They then choose an irreversible action at an endogenous decision stopping time. The novelty with compared previous models is twofolds: a) I introduce the possibility to control the techonology of acquition of information and the level of risk b) I directly test the model in the empirical section. The main predictions of the model are that a) *DMs* that choose to acquire more precise informative signals with a more convex cost function and act coherently to signals, are more likely to terminate and do so earlier b) *DMs* that choose to acquire less precise informative signals with a less convex function and do not act coherently to signals, are less likley to terminate and do so later. The former *DMs* being *DMs* exploring ideas with a predictive strategy, the latter being *DMs* exploring ideas with a control-oriented strategy. I operationalize these two approaches respectively with a scientific approach (Camuffo et al. 2020a, 2021) and an effectual approach to decision-making (Sarasvathy 2001). I test the model, conceptualized before collecting any empirical evidence, by running a Randomized Control Trial (RCT) with early-stage start-ups. Key variables used in the model have been empirically measured with the scope of explicitly test this model. I compare the predictive and control-oriented strategies with classic search heuristics, like trial-and-error process (Nicholls-Nixon et al., 2000; Shepherd et al. 2012).

The RCT was conducted online in Italy in 2020 and involved 308 early-stage start-ups for a total of 2,772 data points. The RCT was structured in line with the four RCTs used by previous studies on the scientific approach conducted in different countries (Camuffo et al. 2021), but this is the first RCT that studies also a non-predictive, control-oriented, approach to entrepreneurial decision-making. This is also the only RCT held completely online due to the pandemic, allowing to involve start-ups from the whole country. The experimental design consisted of eight training sessions over a period of about four months and the data collection lasted for around one year. Start-ups have been randomly allocated to three different arms and received a similar training about how to deal with uncertainty and launch an innovative idea. 70 percentage of the training was the same for the treated (scientific and effectual) and control groups of start-ups. The three groups were taught how to use a business model canvas, how to conduct basic tests and collect data about customers such as interviews, surveys and A/B tests. The remaining 30 percentage was dedicated, for the scientific group, to teach start-ups how to develop theories, test hypotheses and make decision accordingly to the test results. On the other hand, the control-oriented group were taught how to select alternatives or terminate their idea by defining an affordable loss benchmark instead of focusing on the highest expected returns, how to get pre-commitments from potential stakeholders and leverage contingencies. Coherently with the model, I find that predictive *DMs* terminate more often and earlier because they acquire less but predictive information to improve signal precision and update their beliefs accordingly, coherently with Bayesian learning models (Cyert and DeGroot, 1987). Surprisingly, I find that also control-oriented *DMs* use their information to improve, to some extent, signal precision, but they acquire more information, as they learn by doing, and do not act coherently: they terminate less and later. This is coherent with the nature of the two strategies, where the predictive scientific strategy prescribes to use information to predict the future, while the control-oriented one prescribes to decide based on and affordable level of risk and get information to exert control over the future, weighting less predictive information. Accordingly with the model and previous research predictions (Camuffo et al. 2021), I also find that scientific *DMs* perform better in terms of revenues. This is in line with a logic based on the expected maximization of returns versus one that uses affordable loss as a benchmark: predictive *DMs* focus more on ideas that signal high returns, thus terminating less remunerative ideas and false positive, while control-oriented *DMs* keep exploring ideas until they fall below their affordable loss cutoff, hence showing being more likley to incur in false positive. This paper contributes to strategy research on decision-making by presenting and explicitly testing a mechanism to explain how explorating ideas with predictive and non-predictive strategies can lead to different outcomes in terms of survival time and rates of innovative firms. The paper continues as follows: in the next session I present a stylized model and the main predictions. Then I describe the RCT design and data collection. Finally, I discuss the empirical results and conclude.

1.2 Model

1.2.1 Premise

When Decision-Makers (hereafter *DMs*) make decisions under uncertainty, they often have imperfect and small information about the payoffs of the business ideas they want to pursue. Hence, *DMs* explore ideas by acquiring information about it and potential alternatives. In this model, I assume that *DMs* already have an idea and enter in an explorative phase to decide whether to commit to or terminate it. For example, when comparing technologies, a firm may not know the profitability of alternative technologies. The firm then spends money and time on R&D to identify the best technology to adopt. An important feature of this exploration is that the choice of “what and when to learn" implies a lot of aspects that can depend on the specific decision-making approach adopted. For instance, *DMs* may choose to conduct precise experiments to test the validity of hypotheses and improve their prediction power. Or they can choose to adopt a control-oriented strategy and learn from and engage with potential stakeholders and set a level of risk they can bear, until the uncertainty is reduced. To capture this richness, I consider *DMs* who can choose 'what to learn' as well as 'when to stop learning', and the payoff in the worst scenario, that is the loss they can afford, namely the 'affordable loss'. In this light, my model builds upon classic works on bayesian learning (Wald 1947; Arrow, Blackwell and Girschick, 1949) studying stopping learning problems with exogenous information, and works about optimal experimentation and information acquisition (Moscarini and Smith, 2001; Zhong, 2017) where *DMs* endogenize the learning process. The latter works, among other aspects, endogenize the "precision" feature of learning by allowing *DMs* to control the precision parameter of the signal collected during the exploration phase. So do I, adding the control of the static payoffs and the control of the techonology to acquire information. The model is a continuous-time sthocastic control problem where *DMs* can control these dimensions and choose an irreversible action at an endogenous decision time. The choice of information is designed by using a belief-based approach (Ely 2017; Zhong 2017) and a gaussian learning technology to describe signals. The model here presented has been conceptualized before collecting any empirical evidence and aims to predict and explain the mechanisms driving *DMs* decisions outcomes. Key variables used in the model have been empirically measured with the scope of explicitly test this model. I refer to the scientific approach (Camuffo et al. 2020a, 2021) and effectuation (Sarasvathy 2001) respectively as predictive, but explorative, and control-oriented strategies. I compare them with classic search heuristics, like trial-and-error processes (Nicholls-Nixon et al., 2000) or confimatory search (Shepherd et al. 2012). I will refer to them respectevely as 'explorative predictive' *DMs*, 'control-oriented' *DMs* and non explorative predictive/non control-oriented *DMs* (or simply 'others'). I flag variables referring to these three approaches with apex P, CO and N that stay respectively for *Predictive*, *Control-Oriented* and *Non predictive/Non control-oriented* .

1.2.2 Set-up

Decision problem. I consider a two-action, two-state world. *DMs* must choose between actions $a = (C, T)$, with C=*Commit to* and T=*Terminate* the idea, with payoffs π_a^ϑ in the states $\vartheta = \{S, F\} \in \theta$, where S and F stay for *Success* and *Failure*. The payoffs are such that no action is weakly dominant: $\pi_C^S = s > \pi_T^S = 0$, $\pi_C^F = f < \pi_T^F = 0$ and action C shows a strictly better payoff than action T in state S: $\pi_C^S = s > \pi_T^C = 0$:

	$\vartheta_1 = S$	$\vartheta_2 = F$
$a_1 = C$	$s > 0$	$f < 0$
$a_2 = T$	0	0

This is a general case where *DMs* are developing new business ideas that might not work ($f < 0$) and gain zero if they terminate the idea regardless the state. *DMs* do not know the true state and hold a prior belief $p \in \Delta(\theta)$ over the state S and $1-p$ over the state F, where $\Delta(\theta)$ is the set of all probability measures on θ: $\Delta(\theta) = \left\{p \in \mathbb{R}_+^\theta : \sum_{\vartheta\in\theta} p(\vartheta) = 1\right\}$. Therefore the expected payoff for an action a is $\pi_a(p) = p\pi_a^S + (1-p)\pi_a^F$ that, in my case becomes: $\pi(p) = \max\{0, ps + (1-p)f\}$. *DMs* are indifferent between committing to their business idea or terminate it for a certain belief $\hat{p} = \frac{-f}{s-f}$, such that $\pi(p)$ is strictly increasing in p $\forall p > \hat{p}$. *DMs* discount the expected payoff with rate $r > 0$. Before choosing an action, *DMs* enter in an explorative phase where they acquire informative signals about the state ϑ at each stage of exploration $t \in [0; \infty)$.

Information. Let's describe the process of the signals as a diffusion process x_t with a drift μ^ϑ of state ϑ and a noise that *DMs* can reduce by acquiring information. x_t is then a Brownian Motion with an uncertain bivariate drift $\mu^\vartheta = \{\mu^S, \mu^F\}$. In other words, ideas are of two types: successful ($\vartheta = S$) with probability p and drift $\mu^\vartheta = \mu^S$ and unsuccessful ($\vartheta = F$) with probability $1-p$ and drift $\mu^\vartheta = \mu^F$. Hence, *DMs* do not know the true state but observe signals with the underlying rate μ^{ϑ} and a noise component:

$$dx_t = \mu^{\vartheta} dt + \frac{\sigma}{I_t} dW_t \tag{1.1}$$

where $W_t \sim N(0,t)$ is a Wiener process, with increments $W_s - W_t$ independent of ϑ (for $t < s$), and I_t is a measure of the informativeness (flow) of signals, or simply Information. *DMs* do not know μ^{ϑ} but observe x that describes μ^{ϑ} with a certain noise that they can control through I. In line with Zhong (2017), I define I_t as the speed at which the uncertainty is reduced:

$$I_t = -\mathbf{E}\left[\frac{dU(p_t)}{dt}\bigg|\mathcal{F}_t\right] \tag{1.2}$$

With p_t being the posterior process, assumed to be a martingale with natural filtration $\mathcal{F}_t$. $\mathcal{F}_t$ is the amount of past information available at t. $U : \Delta(\theta) \to \mathbb{R}$ is assumed concave and continuous and can be interpreted as a measure of uncertainty. This assumption will be useful to concavify the value function of the idea, as I will show in the next paragraphs. One classic example of U can be the entropy function. Hence, the faster the reduction of uncertainty after observing the signal, the higher the amount of information collected. By plugging (2) in (1), it is clear how increasing I_t reduces the variance of the signal as a consequence of the reduction of uncertainty. In a static environment, the information can be modelled as the distribution of posterior beliefs. Indeed, the distribution of posterior beliefs is induced by information *iff* the expectation of posterior beliefs is equal to the prior, according to Bayesian rule. Hence, in a dynamic environment, Bayesian rule should be satisfied at every exploration stage t, such that $\mathbf{E}[p_s] = p_t \ \forall s > t$.

Information technology. Flow quantities being equal, *DMs* can choose to purchase I with different technology. I define the *information technology* ω as given by the finite set of signals X_{ω} with probabilities $\omega_{\vartheta}(x) \in \Delta(\theta_{\omega})$. $\omega_{\vartheta}(x)$ is the probability that the signal observed is x when the true state is ϑ and induces the posterior probability on ϑ, according to Bayes rule. We can think at the information technology as a stochastic matrix of $\omega_{\vartheta}(x)$ with as many rows as states and as many columns as signals. Making the dependence of $I(p, \omega)$ on the prior p and ω explicit, an information technology ω is more valuable than another technology γ *iff* causes a larger reduction of the entropy U: $-\mathbf{E}\left[\frac{d}{dt}U(p_t, \omega)\right] > -\mathbf{E}\left[\frac{d}{dt}U(p_t, \gamma)\right]$. In other words, *DMs* can choose to purchase a similar amount of information but select a different technology of acquisition, where a more valuable technology is the one which reduces uncertainty faster and produces more informative signals such that the value of the information increases: $\nu(I(\omega)) > \nu(I(\gamma))$. We can think at the selection of ω as the choice of experimenting (at this level I do not make distinctions between experimenting to predict an exogenous environment of to control an endogenous environment). A direct logic consequence is that the cost of such technology is higher than less valuable technologies.

Cost of information. *DMs* purchase informative signals with a (flow) cost $c(I) \geq 0$. I assume $c(I)$ convex, twice differentiable and $\lim_{I \to \infty} c'(I) = \infty$. These assumptions are useful to avoid the possibility that *DMs* go through a lump-sum acquisition of information that would reduce at infinite speed the uncertainty. Indeed, this would imply an infinite marginal cost $c'(I)$. In other words, these assumption constraint *DMs* to acquire information smoothly across the exploration stages. By selecting a different technology of acquisition, *DMs* choose to pay a different of $c(I, \omega)$. I model this effect with a different cost convexity, with $c(I(p, \omega))$ being more convex than $c(I(p, \gamma))$: cost convexity is linear and strictly increasing in value of ω. Moreover, since $c(I)$ is a flow cost, a different convexity implies that the cost of acquiring new information depends on cost spent in the previous round of exploration. For instance, running a new experiment in the next round will cost more than the experiment in the previous round. It follows that adopting a more valuable technology could limit the amount of information that *DMs* can afford to purchase across rounds of exploration.

Figure 1.1: *DMs* can spend the same cost and acquire different amount of information with different values, by selecting a different information technology.

Beliefs. As mentioned before, the stochastic process p_t induced by the signal process is a martingale. Coherently, *DMs* update the prior beliefs in favour of μ^S *iff* signals rise faster than the expected drift: $\frac{dx_t}{dt} > p_t\mu + (1-p_t)(-\mu) = \mathbf{E}[\mu^\vartheta]$. Where $\mu^S = -\mu^F = -\mu$. To clarify the role of I_t, we may think to the case where prior belief p in the initial stage has Beta density

$$B(p|m_0, M_0) \propto p^{m_0}(1-p)^{M_0-m_0} \tag{1.3}$$

DMs form the prior by using M_0 observations from the past, of which $m_0 \in [0, M_0]$ are observations the idea is successful, i.e. $\vartheta = S$. We can interpret M_0 and m_0 as proxy of the amount of information available at $t = 0$. The prior is then $p_0 = \mathbf{E}\left[p|m_0, M_0\right] = \frac{m_0}{M_0}$. If *DMs* decide to go through a round of exploration, they can set the number of new observations N_0 and therefore control the variance of the (unknown) true number of successful observations n_0. $n_0 = N_0$ is like saying that $\mu^\vartheta = \mu^S = \mu$ and $n_0 = 0$ is like saying that $\mu^\vartheta = \mu^F = -\mu$. Controlling I is similar to control N_0. Beside the different conceptualization of signals, the intuition is the same as for the brownian motion: signals evolve as a distrubution, in this case the distribution of the share of successful observations, where *DMs* can control its variance.
After the round of exploration, the posterior has density

$$B(p|m_0, M_0, n_0, N_0) \propto p^{m_0+n_0}(1-p)^{N_0+M_0-n_0-m_0} \tag{1.4}$$

that generates the posterior in stage $t = 1$ equal to $p_1 = \frac{p_0 M_0 + x_1 N_0}{N_0 + M_0} = \alpha_0 p_0 + (1-\alpha_0)x_1$. Where $\alpha_0 = \frac{M_0}{N_0+M_0}$ being the weight that *DMs* put on the prior and $x_1 = \mathbf{E}\left[\frac{n_0}{N_0}\right]$ can be interpreted as the signal. At stage t, posterior becomes then $p_t = \alpha_{t-1}p_{t-1} + (1-\alpha_{t-1})x_t$, in line with Elfenbein and Knott (2015). Standard Bayesian updating requires α to increase over time as signals are collected. Summing up, beliefs are a martingale with traps at 0 and 1, evolve in a Bayesian way, with signals evolving with a Brownian motion whose variance is controlled by the means of I_t.

Stochastic control. *DMs* solve the following control problem:

$$V(p_0) = \sup_{\tau, I_t} \mathbf{E}\left[e^{-r\tau}\pi(p_\tau) - \int_0^\tau e^{-rt}c(I_t), dt\right] \tag{1.5}$$

where τ is a $< \mathcal{F}_t >$-measurable stopping time. According to equation (5), *DMs* acquire costly information that impacts p_t and choose the stopping time τ to maximize the discounted expected payoff $\mathbf{E}[e^{-r\tau}\pi(p_\tau)]$ less the discounted cost of information $\mathbf{E}\left[\int_0^\tau e^{-rt}c(I_t), dt\right]$. The supreme value V is assumed convex in p, a classic assumption for optimal learning. *DMs* stop acquiring information when V coincides with the static payoff π. I call beliefs p at which this condition is met, "threshold beliefs" $0 \leq \underline{p} \leq \bar{p} \leq 1$. *DMs* select action *Terminate* for $p \leq \underline{p}$, *Commit to* when $p \geq \bar{p}$ and esplore reducing uncertainty at speed $I(p)$ for $p \in \mathcal{E} = (\underline{p}, \bar{p})$, which I call *exploration zone.*

Equation (5) means that, at each round of exploration t, *DMs*, based on their prior belief p_0, decide whether to pay c to acquire new information I and discount the decision delay with r, in face of an expected return π. When the value is maximum, they stop learning and decide. According to the dynamic programming principle, (see Moscarini and Smith, 2001 for techincal proof), the supreme value V in (5) solves the Hamilton-Jacobi-Bellman (HJB) equation for the control problem, for different values of prior beliefs p. HJB decomposes the value function into two parts, the discounted future value plus the immediate return:

$$rv(p) = \sup_I \left\{ -c(I) + Iv''(p) \left(\frac{p(1-p)2\mu}{\sigma} \right) \right\} \tag{1.6}$$

Therefore $V = v$ is the recursive value for this stochastic optimal control problem with delay. The intuition here is that the problem of finding the value function v that optimizes the information acquitisition can be only solved backward and this is equal at each time point to the supreme value V. Figure 2 shows the value function $v = V$: the shaded zone is the exploration zone $\mathcal{E}$, while the non shaded zone is the stopping zone. The boundary conditions are: $v(\underline{p}) = \underline{p}\pi_T^S + (1-\underline{p})\pi_T^F = 0$ and $v(\bar{p}) = \bar{p}\pi_C^S + (1-\bar{p})\pi_C^F = \bar{p}s + (1-\bar{p})f$. The intuition here is that at each stage of exploration t, *DMs* choose whether it is worth to pay a cost $c(I)$ to gain a surplus given by I_t. By plugging the FOC of (6) into (6), it follows that $rv(p) = Ic'(I) - c(I)$: when the marginal surplus $MS = Ic'(I) - c(I)$ from acquiring information is greater then the discounted value rV, *DMs* keep exploring and $V(p) > \pi(p)$. When the marginal surplus $MS = rV$, they stop exploring. They do not want to pay a cost associated to the delay of the decision. $v(p)$ is tangent to the static payoff π at $\underline{p}$ and $\bar{p}$, such that $v'(\underline{p}) = \pi_T^S - \pi_T^F = 0$ and $v'(\bar{p}) = \pi_C^S - \pi_C^F = s - f$. *Property: $I(p)$ monotone in $v(p)$.* This property states that the informativeness measure I of the optimal signal is higher when Bellman value v is higher prior to stop exploring. Indeed, if $rv(p) = MS(I(p))$, it exists a strictly increasing inverse function $f = MS^{-1}$ such that $I(p) = f(rv(p))$. Hence, given that v, by convexity, follows the shape of π, so does $I(p) = f(rv(p))$. This property is independent on the formulation of the informativeness measure. Figure 2 summarizes the intuition behind the model. *DMs* acquire signals whose informativeness increases with the value and stop acquiring informative signals when the marginal surplus equates the cost of waiting rv for them. *DMs* choose to keep learning if the expected value of the speed of uncertainty reduction is worth the cost of exploring. They optimize the dynamic acquisition of information and, based on the direction of the update of beliefs, they choose whether to acquire more information or reduce the acquisition. Indeed, given the monotonicity of I in the value function v, *DMs* choose to increase the amount of information to acquire if the value increases and vice-versa. If the value decreases, they acquire less informative signals.

Figure 1.2: Value function, static payoff and Information (flow).

Payoff analysis. In this section I briefly study the implications of a change of payoffs from a geometric point of view. Figure 3 shows that if the static payoff $\pi_C^F = f$ rises (f_2 in the figure), so does the value function v and therefore I, according to its monotonicity property. A direct consequence of this analysis is that, if f rises, *DMs* acquire more information before stop exploring and are less likely to terminate as they see higher value v. Coherently, the exploration $\mathcal{E} = (\underline{p}, \bar{p})$ shifts down as they get less conservative.

Figure 1.3: When f increases, the value function v shifts up. So does the informativeness I. The grey regions are the so-called *Exploration zones*.

1.2.3 (Explorative) Predictive and Control-oriented Decision-Makers

Prediction plays a central role in strategy making, prescribing that a higher prediction power allows to control the future, to some extent. Wiltbank et al. (2006) highlight how, under different levels of uncertainty, classic prediction and control become distinct dimensions, with control strategies potentially driving successful outcomes under knightian uncertainty (Knight, 1921). In this light, a predictive but explorative, strategy as the scientific approach differentiates from other classic predictive approaches as it allows, with creative theories and tailored experiments, to navigate high levels of uncertainty. Indeed, on the one hand, it is based on the presumption that prediting the future could lead to more precise decisions and better performance (Camuffo et al., 2021; Novelli & Spina, 2021). On the other hand, scientific *DMs* use theories and experimentation to learn and explore new opportunities that may arise or be created. They develop models of the world and build precise experiments aimed to reduce the noise generated by wrong models of the world.

'Scientific' or, more in general, explorative-predictive strategies and control-oriented strategies have commonalities and differences. While explorative-predictive *DMs* embrace uncertainty by the means of theory and experimentation, control-oriented *DMs* face uncertainty by leveraging the means that are available to them and precommitting to how much they are willing to lose. In this light, I focus my attention on two key dimensions. The first one is the concept of *affordable loss*, defined as 'a predetermined level of affordable loss or acceptable risk' (Sarasvathy, 2001). For control-oriented *DMs*, I conceptualize the static payoff in the worst state $\pi_C^F = f$ with the affordable loss. The most important difference between $f \equiv f^{CO}$ for control-oriented *DMs* , (where CO stays for Control-Oriented) and f for predictive *DMs* is that affordable loss f^{CO} is based on endogenous aspects, such as *DMs*' financial constraints or preferences. Estimating the affordable loss f^{CO} does not depend on the idea, but on the *DMs* themselves and involve a trade-off between their resources availability and subjective risk attitude which they can control to some extent. On the other side, predictive *DMs* fix the static payoff $f \equiv f^P$ (where P stays for Predictive) coherently with the idea they are exploring and it can be interpreted as the development cost or investiment required to develop the idea and it is stritcly linked to the idea object of analysis. Since investments and development costs needed to launch a business can be high, I assume affordable loss being higher than that (less negative). Non-predictive and non-control-oriented *DMs* (which I flag with apex N) show $f \equiv f^N$ somewhat in the middle. This leads to my first assumption.

Assumption 1. *Control-oriented DMs have a higher f than others: $f^{CO} \geq f^N \geq f^P$*

The second dimension is choice of the information technology. Control-oriented *DMs* start from their available resources and social network and engage into a continuous interaction with people to learn and shape the future. They might form new markets and new means by getting pre-commitment from potential stakeholders. They use a 'logic of design' with the intention to design the future, to some exent, and collect information trying to converge to a viable idea, coherently with their affordable loss (Sarasvathy 2001; Wiltbank et al. 2006). They go through a countinuos experimentation aimed to shape and control an endogenous environment. On the other side, explorative predictive *DMs* use logical frameworks to build focused experiments aimed to predict an exogenous environment and increase the predictive value of information. They know "what to learn" and "where to search", as theories inform them about it. They fully embrace uncertainty guided by their theories and learn through experimentation. Non predictive and non control-oriented *DMs* do not follow any specific guideline or framework during their exploration phase. They acquire random information as a consequence of the fact that their exploration it is neither based on theories and precise experiments nor based on a continous attempt to expand cycle of resources to learn and adapt to future events. Indeed, on the explorative predictive side, framing theories and building precise experiments and, on the control-oriented side, negotiating and persuading potential stakeholders require more efforts than a random information acquisition process, with running predictive experiment being even more costly than a 'control experimentation' modality. In other words, explorative-predictive *DMs* choose the most valuable technology followed by control-oriented *DMs*: $\nu(\omega^P) > \nu(\omega^{CO}) > \nu(\omega^N)$. In geometric terms, both explorative-predictive and control-oriented *DMs* purchase more valuable information with higher the convexity of $c(I)$ than others. It follows that

Assumption 2. *Explorative-predictive DMs select the most more valuable information technology with highest convex costs of information, followed by control-oriented DMs*

From the payoff analysis and the monoticity property of I in v, it follows that control-oriented *DMs* see ideas with higher values v and are more likely to acquire higher amount of I than others. For the same reasons, explorative predictive *DMs* and non explorative-predictive/non control-oriented *DMs* behave in the opposite direction, seeing ideas with lower values and being more likely to acquire lower amount of I than others.

Moreover, from assumption 1 it follows that both control-oriented and explorative-predictive *DMs* acquire less information as the marginal surplus $MS = Ic'(I) - c(I)$ decreases faster as $c(I)$ increases faster. In sum, for control-oriented *DMs* these two effects balance each other to some extent as they are more likely to acquire more information as they see higher value v, but they also pay a higher cost for a more valuable information technology. Instead, for explorative-predictive *DMs* these two effects reinforce their attitude to acquire less information. They acquire less information as they see lower v and pay also more than others. Those who acquire more information and longer, are the non explorative-predictive/non control-oriented *DMs* as they see a value v which is somewhat in the middle between the explorative-predictive and control-oriented v and select a less valuable, and less costly, information technology acquisition.

Proposition 1. *Control-oriented DMs acquire higher amount of information than explorative-predictive DMs. Both acquire lower amount of information, but more valuable, than others:* $I^N > I^{CO} > I^P$

Classic literature on bayesian learning (Wald 1947; Arrow et al. 1949) studied the choice of "when to stop" learning, taking all the aspects of the learning process exogenous. As mentioned before, in the model I endogenize the *precision* aspect of learning, the technology of information and the choice of static payoffs, by letting *DMs* to control

I, ω (and therefore c) and choose f. While control-oriented *DMs* use the choice of f to drive their decision-making process, explorative predictive *DMs* put their efforts to control the precision of signals and choose the optimal strategy. The optimal strategy involves testing the idea in such a way that tests are diffucult to pass. Positive signals are difficult to collect and no news are bad news, since the explorative predictive *DMs*, as scientists would do, experiment to validate or falsify their theories. If the idea does not pass their tests, they become more conservative (Camuffo et al. 2021) about the idea and their future tests become increasingly more difficult to pass. This process is counterbalanced by the fact that positive signals come from highly informative tests. I is acquired with the aim to improve signal precision, as the technology is more valuable. On the other side, control-oriented *DMs* strive to control events to navigate an uncertain future. They deliberately exert efforts to make an exogenous future endogenous and set an affordable loss that takes away the need for prediction. The signals they construct come from negotiation and persuasion of other people from which they want to obtain inputs and a certain level commitment in their idea. Signals inform about states ϑ they think they can turn into success ($\vartheta = S$) through their relatively unique skills. They seek information that is more salient for their scope, where more salient means that grabs their attention, and allows them to converge towards a viable idea. Therefore, even if the nature of information and signals is not predictive as for explorative-predictive *DMs*, control-oriented *DMs* acquire information I to improve signal precision as well, but less than explorative-predictive *DMs* and more than others that instead go through a random acquisition process.Coherently with the selected information technology, we get:

Proposition 2. *Explorative-predictive DMs improve precision of x_t more than control-oriented DMs. Both improve precision more than others.*

While explorative-predictive *DMs* collect information I to reduce uncertainty and act accordingly to signals x_t, control-oriented *DMs* focus on things within their control and proceed coherently. As Dew, Sarasvathy, Read and Wiltbank (2009) stated, 'affordable loss reasoning is a biased mechanism' for committing to a business idea. Even if contingencies, failure rates, feedbacks from stakeholders are rowing against the odds of success, control-oriented *DMs* are less likely to terminate the idea. Moreover, affordable loss reasoning implies also that they set a constraint that could stop their entrepreneurial activity even in face of positive signals. Control-oriented *DMs* make a decision disconnected from signals, even if a bayesian decision-making framework would argue against their decision about terminating or committing to the idea. On the other side, explorative-predictive *DMs* focus on controlling the precision parameter I of signals driving their belief diffusion process and solving the stochastic control-stopping problem.

Proposition 3. *Control-oriented DMs terminate less (more) in face of negative (positive) signals than both explorative-predictive and other DMs. Explorative-predictive DMs terminate more (less) in face of negative (positive) signals than others.*

In sum, explorative-predictive *DMs* acquire more expansive predictive information I and then stop acquiring new information as learning incurs a high flow cost $c(I)$. They collect information to improve signal precision and decide accordingly, optimizing the information acquisition faster. They make their decision earlier than both control-oriented and non explorative predictive/non control-oriented *DMs* since the acquisition of information I slows down ealier, given assumption 1 and assumption 2 and proposition 1. Moreover, they use this information I to improve signal precision more than both control-oriented and non explorative-predictive/non control-oriented *DMs* by proposition 2. Explorative-predictive *DMs* act accordingly to the bayesian learning model proposed and are more likely to terminate in face of negative signals, and viceversa, than control-oriented and non explorative predictive/non control-oriented *DMs*. Control-oriented *DMs*, by adopting an affordable loss reasoning, are less likely to act as bayesian decision-makers when it comes to act in face of signals, even if they accumulate higher levels of expansive information and improve to some extent signal precision. Thefore, for explorative-predictive *DMs*, the reasoning goes as: *a)* Precise signals arrive early as they control I and the technology ω to improve their predictive power. *b)* They act coherently with signals. *c)* As they see ideas with lower value, by assumption 1, and falsify theories with tailored experiments, they are more likely to terminate ideas. For control-oriented *DMs*, the reasoning goes as: *a)* They set an affordable loss f and get signals from potential stakeholders, controlling the techonolgy ω. *b)* Even if I improves signal precision to some extent, they are less likely to act coherently with signals. *c)* As they see ideas with higher value and are guided by an endogenous level of risk, they are less likely to terminate and do so later.

Proposition 4. *Explorative-predictive DMs are more likely to terminate projects, and do so earlier, than both control-oriented DMs and others. Control-oriented DMs terminate less and and do so later.*

In the next sections I will test directly the propositions provided by the model.

1.3 Examples

I provide two examples of start-ups exploring an idea from the RCT coherent with the assumptions about explorative-predictive and control-oriented DMs, respectevely operationalized with the scientific (Camuffo et al 2020a, 2021) and effectual (Sarasvathy 2001) approaches which are in line with the two strategies discussed.

1.3.1 Example 1: scientific approach

Talia is a start-up that wanted to launch an online marketplace where customers could buy sustainable fashion items from small local sustainable brands made in Italy. The founder had a theory about their idea. First, the climate change crisis would push people to be more aware about the pollution produced by the fashion industry. Linked to the first point, people would search for sustainable fashion items online as, even if many existing e-commerce platforms have already flagged some of their items as sustainable, customers do not trust incumbents. Third, local brands are more likely to be perceived as sustainable from customers than large brands. Fourth, small local sustainable brands would need a channel to increase their sales volume. Fifth, the target market are young (18-24yo) women since they should be more likely to buy clothes online. Starting from this theory, the founder started to compute the investment required to build the platform by benchmarking with similar platforms and to get evidence about the theory. She spent time refining the theory and interviewing few small local brands. She launched A/B tests to measure the extent to which target customers were likely to prefer sustainable local brands to more famous brands. She launched surveys on the main social networks to test the extent to which customers would search online for sustanaible fashion items. She also developed a page on Instagram to engage with local brands and build a customer base, before launching the e-commerce, to run tailored and unbiased tests. She chose to spent time and money to collect highly informative signals with the precise goal to predict the probability of success. By analyzing test results, she soon realized that 1) despite awareness of the climate change crisis, young customers were less likely to search online for sustainable fashion items than expected 2) existing e-commerce were perceived more reliable than expected about sustainability 3) many small local sustainable brands have invested to develop their own online website and were not willing to sell on another platform. At the end, she abandoned the idea as she got precise negative informative signals about the low probability of success in face of high development costs and it was not worth to acquire more information.

1.3.2 Example 2: effectual approach

Kora is a start-up that wanted to sell tailor-made swimwear for young women. The founder stated that she started from her passion for fitness and was pushed by her insicurities with her body when wearing swimwear. When she was a teenager, she "embarked on a fitness journey to change" her body, but still this did not solved her unpleasant sensations she was experiencing when wearing swimwear. Over the years, since she was interested in figuring out what type of swimwear would fit best her, she become knowledgeable about design, fabric and quality of swimwear. Therefore, she decided to put things together and to explore the idea of starting her own business of tailor-made swimwear for young female customers. She clearly set the amount of money and time that she could lose in case of failure. Instead of developing a theory about why this idea would work and run precise tests with the aim to predict the future, she told family and friends about her idea. Despite an initial skepticism, she managed to persuade her dressmaker aunt to help sewing and drawing the first pieces of her collection. Instead of investing in a marketing campaign, she contacted from her Instangram profile an influencer and, after a tough negotiation, she convinced her to get on board, promote the collection and become part of the venture. However, despite her efforts, the founder was not able to sell her collection and received harsh feedbacks about her idea from potential customers she was constantly in contact with. Signals were informative and precise, to some extent, but she ignored them and decided to keep trying until she could afford the loss. As she felt to have reduced the uncertainty about the possibility to launch her business, she finally launched the business, which is currently operative. With compared to these two examples, non scientific/non effectual entrepreneurs in my RCT did not frame any theory and did not acquire information by running precise tests. Also, they did not attempt to persuade potential stakeholders trying to keep them on board and did not set any clear affordable loss. They did not exert strong efforts in collecting information, but sometimes ran some small tests or engaged into brainstorming. They mainly kept trying different options without a clear mental framework and without leveraging their existing skills or social networks. They kept collecting information from basic market research, by asking friends and family or by developing small prototypes with the aim to guess a successful solution, without a clear strategy to predict or control future events.

In sum, scientific *DMs* develop general theories that they try to falsify or validate by running precise highly informative experiments and make coherent decisions. Effectual *DMs* start from their skill and knowledge, exert efforts to engage with and learn from potential stakeholders and decide coherently with their affordable loss.

1.4 Empirical context

1.4.1 The RCT: Experimental Design

I leverage data from an RCT (pre-registered before the field experiment took place) I conducted in Italy with 308 early-stage start-ups. Start-ups have been randomly allocated to three groups where two of them have been treated respectively with a scientific and an effectual training to decision-making and the third with a control training, that is without any training about the scientific or effectual approach ($\approx 33\%$ scientific group, $\approx 33\%$ effectual group, $\approx 33\%$ control group). Entrepreneurs provided data about their decision-making process and decisions outcomes from before the program and for around one year after the program. All start-ups undergoing the 3 different arms received 8 sessions of online training for a total of 24 hours, from mid-October 2020 until February 2021. Figure 2 provides more detail on the timeline. These sessions included interactive lectures and coaching by qualified instructors, each working with a subset of the start-up sample. All start-ups in the 3 experimental arms received the same amount of training on entrepreneurial decision making (on topics like business model canvas, customers' interviews, minimum viable products/services, concierge/prototype, etc.). However, while the control group was not encouraged to adopt any decision-making process, the other two groups have been explicitly trained to combine those tools with a specific decision-making process, namely a scientific and an effectual approach. A large team of research assistants has been recruited to advertise the program keep in touch with the entrepreneurs and collect data. Four qualified instructors have been hired and deeply trained about the scientific and effectual approach to deliver the training. The training material has been complitely designed by the research team in chief of all the operations.

Figure 1.4: Timeline of the RCT: experiment and data collection.

The training program has been conducted entirely online due to the pandemic affecting Italy and the entire world. On the one hand, this might have discouraged some entrepreneurs and make the development of their ideas slower, even if there is some evidence about the opposite direction as well ("Startup Surge: Pandemic Causes New Businesses To Double", Forbes, Jan 2021). On the other hand, this has allowed to train start-ups from almost each region of Italy, without the constraint of physically travelling and incurring in potential additional costs. The program has been widely advertised online on the main social media and leveraging the brands of the academic institutions involved, which are the top business and tech universities in Italy. The marketing campaign promoted the training program as a pre-accelerator program for early-stage start-ups, covering basic business topic and free of charge. The call for application was open to start-ups operating in any sector. To manage all the operations online, I leveraged the help from a team of developers to build a website (www.innoventurelab.org) where entrepreneurs could apply and create a profile for their start-up and their contact information. They were asked to fill an online survey and to take a phone interview and provide data about them, their business idea and their decision-making process. Only entrepreneurs that completed all these steps have been admitted to the program. Each training session was designed to be highly interactive: the three hours were split into 3 moments of frontal lecture and three moments were entrepreneurs, within the same class, were randomly allocated to breakout rooms to directly apply what they learnt in the previous moment of frontal lecture. Instructors could enter in these breakout rooms to provide advice coherently with the treatments. Indeed, entrepreneurs were randomly allocated to subgroups, within treatments, matched with four experienced instructors. Each instructor was teaching to the three groups, such that they could be included as fixed effect in the analysis, controlling for different teaching styles that could impact the absorption and efficacy of the treatments. Conducting the experiment online allowed to better ensure the internal validity or the experimental results. Indeed, entrepreneurs did not meet in person avoiding any possible contamination. Moreover, the training sessions were conducted in different time slots (Saturday morning and afternoon and Sunday morning). Any communication was kept separated across the three groups and across subgroups as well. As mentioned, the recruiting process required to complete an extensive survey and a deep interview with a member of the data collection team, with the goal of collecting baseline data to allocate randomly entrepreneurs to the three groups. By using a statistical software (STATA), 101 start-ups were assigned to the control group, 102 start-ups to the scientific group and 105 start-ups to the effectual group, for a total of 308 start-ups. The Appendix reports the results of the balance tests across groups, comparing each of the two treatment groups with the control group and the two treatment group with the other one (tables A1, A2, A3).

1.4.2 The training

All start-ups in the 3 experimental arms received the same amount of training on entrepreneurial decision making (on topics like business model canvas, customers' interviews, minimum viable products/services, concierge/prototype, etc.). However, while the control group was not encouraged to adopt any decision-making process, the other two groups have been explicitly trained to combine those tools with specific decision-making approaches, that is a more scientific or a more effectual approach. To make it clear, all the groups were exposed to the use of the Business Model Canvas as a tool to represent their idea. However, only the scientific group has been trained to fill the Business Model Canvas coherently with a well-framed theory, linking all the components with a consistent logic and formulate precise hypotheses starting from them. On the other side, the effectual group has been trained to use the Business Model Canvas as a blurred guide to track changes from the initial idea and, for instance, focus more on key partners and resources instead of on revenues and costs, coherently with the idea to get pre-commitment from potential stakeholders and start from available means. In later training sessions, entrepreneurs have been taught how to get feedbacks and some evidence about their idea, by interviewing and surveying potential stakeholders or conducting classic tests, such as A/B test. While scientific entrepreneurs were trained to build precise and powerful experiments to test theory-based hypotheses, the effectual group was trained to get feedbacks and build networks to open new markets and eventually get new means to develop their idea. Entrepreneurs in the control group were instead left free to use their intuition and heuristics when using these tools to collect data. While the scientific group has been trained to evaluate test results to decide accordingly to the signals collected and discard not promising ideas, the effectual group was encouraged to set an affordable loss and keep exploring ideas even in face of bad news, staying coherent with the risk they endogenously chose to take. Again, the control group was free to make decisions using the, more or less, informative signals by using intuition and natural heuristics. To further clarify differences across groups, I provide a brief example. The fifth session was about analysing feedbacks and results from interviews, surveys or other kind of tests. Entrepreneurs in the scientific group were exposed to a profound session about how to set a threshold upfront to be unbiased when making decisions after looking at tests results. They were taught to score the test in terms of how informative and precise it was and set the condition for which the idea passed the tests or not. While entrepreneurs in the effectual group were trained about how to evaluate feedbacks and interactions in terms of to what extent they could involve the actors providing feedbacks in their venture. They were trained to evaluate negative feedbacks and results against their affordable loss. The control group, instead, was free to evaluate test results and feedbacks in the way they found more appropriate, coherently with their heuristics.

1.4.3 Data Collection

A team of research assistants has been recruited and deeply trained on how to conduct interviews following a well-structured protocol. They were selected among top students interested in entrepreneurial topics and their skills were tested to ensure they would have performed in line with expectations. Research assistants periodically submitted a survey before conducting a phone call that included both open and closed questions. Entrepreneurs had to fill an online survey before starting the periodic phone call with research assistants. The main variables used in the anlaysis refer to the informativeness, signals collected, revenues and whether they terminated the idea were collected through closed-ended questions. In addition to the baseline data point, 8 datapoints have been collected where the first data point after the baseline took place in November 2020, about 5 weeks after the start of the training program. After that, data have been collected every 6 weeks, on average, until December 2021.

1.5 Methodogy and Results

As mentioned in the premise of the model, I conceptualized the model before the program started. This gave me the chance to ask questions and collect data with the precise scope to test my model and unfold a potential mechanism triggering different outcomes from the adoption of an explorative-predictive or a control-oriented approach to decision-making.

Mechanism (independent) variables.

My main independent variable is Intervention, a variable taking value of 1 for treated firms, 0 for the control group. When I compare effectuation with the scientific approach, the variable takes value 1 for effectual firms and 0 for scientific firms.

My second independent variable is Information. Gans and Stern (2019) stated that exploring alternative ideas requires some level of information that "can only be gained through experimentation", while Choi et al (2008) show how entrepreneurs have thresholds about the amount of information accumulated to decide whether to stop exploring and exploit a certain business idea, to gain a better picture of production processes, market landscape and customer preferences. In this light, I measure the amount of information (flow), that entrepreneurs perceive during the exploration phase as the average of 4 items, ranging from 0 to 1, asking at each data point how much the amount of information they got about the sector the start-up operates in, the competition, the resources needed to commit with the business idea and the customers. This allowed me to get a measure *perceived* amount of information acquired, as the model prescribes a full control over the dimension, since it is defined as the speed at which the *perceived* uncertainty is reduced. The form of the question explicits the assumption that *DMs* search in a bounded space, where, if they acquire all the available information, they would achieve the upper bound equal to 1. This is the main variable the guides the mechanism that explains the outcomes, as shown in the model. Entrepreneurs have been allocated randomly to the three groups with respect their initial level of Information. In line with the model, the next variable is Signal. This is a variable, ranging from 0 to 1, measuring the direction and the intensity of signals received or collected at each data point, with 0 measuring extremely negative signals, 0.5 neutral signals and 1 extremely positive signals. The variable it is equal to the prior beliefs of success at the baseline datapoint, which has been randomized across groups, since the model prescribes signals impacting from $t = 1$ on beliefs evolution.

At the baseline data point, I also randomized entrepreneurs for the weight entrepreneurs put on prior beliefs α, operationalized as *overprecision* (Moore and Healy, 2008) i.e. the excessive confidence about the accuracy of prior beliefs (see Table A1-A3 and A8 of the Appendix for more details). Assuming that α is not impacted by the treatment, this allows me to test my model by looking at the only signals instead of beliefs as well. Indeed, if the weight on the prior α decreases with signal precision, I should find an increasing correlation between signals and beliefs. If the weight on the prior α increases, I should find a blurred effect by only looking at signals, meaning that, effects on the signals should provide conservative results as I am interested in reaction to signals. In any case, looking at signals provides either realiable or conservative results.

Dependent variables.

Termination. In line with other research about whether and when abandoning not promising ideas (Meyer and Zucker, 1989; Lowe and Ziedonis, 2006; Elfeinbein, Knott and Croson 2017; Camuffo et al. 2020a, 2021; Dew, Sarasvathy, Read and Wiltbank, 2009), I focus on a crucial outcome for entrepreneurs and, more in general, for decision-makers that explore innovative ideas, that is whether they terminate or not their idea. Again this is in line with the model, where *DMs* could choose between terminate or commit to it. This is a dummy equal to 1 if entrepreneurs terminated their idea within the data collection window, 0 otherwise.

Week of termination. This variable measures the week at which the entrepreneurs terminated their business idea, that is when they stop exploring and abandoned their entrepreneurial project.

Overall, these two dependent variables are sufficient to test what and when *DMs* decided to terminate or not their idea, while the independent variables are aimed to test the mechanism, as prescribed by the model. Table 1 and 2 provide descriptive statistics of these variables. Longitudinal observations about the 308 firms across nine data points, refer to observations until the moment on termination, if entrepreneurs terminated their idea. For this reason, time varying variables display 2154 observations instead of a total of 2772 observations. Twenty nine four percent of the start-ups, 89 start-ups in absolute terms, terminated their ideas within the data collection window. On average, they gained EUR

2,892.8 of revenues, with a right skewed distribution as the majority of start-ups shows no revenues within the data collection period. Entrepreneurs acquire on average sixty three percent out of all the available information, increasing the amount of information before entering in the training program, as shown by Figure 5.

Table 1.1: Variables and Descriptive Statistics - Cross Section Sample

Variable	Description	Obs	Mean	SD	Min	Max
Termination	Dummy equal to 1 if the firm terminated the project within the observation window; 0 otherwise	308	0.29	0.5	0	1
Revenue	Firm's cumulative revenue in EURO	308	2,892.80	14,059.15	0	124,000
Intervention	Dummy equal to 1 if the firm was not treated; 2 if treated with the scientific approach, 3 if treated with the effectual approach	308	2.01	0.82	1	3
Information I	Firm's cumulative amount of information about Sector, Resources, Customers, Competitors (range, 0 to 1)	308	0.63	0.18	0.04	1
Signal x	Variable ranging from 0 to 1, measuring the overall direction and the intensity of signals received or collected. It is equal to the prior belief p_0 at the baseline time point	308	0.70	0.17	0	1

Table 1.2: Variables and Descriptive Statistics - Longitudinal Sample

Variable	Description	Obs	Mean	SD	Min	Max
Termination	Dummy equal to 1 if the firm terminated the project within the observation window; 0 otherwise	2154	0.04	0.20	0	1
Week of Termination	Week at which the firm terminated the project	308	50.77	14.22	11	58
Revenue (Flow)	Firm's revenue (flow) in EURO	2154	384.74	2819.11	0	60000
Intervention	Dummy equal to 1 if the firm was treated; 0 otherwise	2772	2.01	0.82	1	3
Information I (flow)	Firm's amount of information (flow) about Sector, Resources, Customers, Competitors (range, 0 to 1)	2154	0.09	0.25	-0.6	1
Signal x	Variable ranging from 0 to 1, measuring the overall direction and the intensity of signals received or collected. It is equal to the prior belief p_0 at the baseline time point	2154	0.72	0.20	0	1

To test the Propositions of the model and unfold the mechanism driving the termination of ideas, I use a three stage least squares (3SLS) strategy where I instrument the Information (flow) I with the Intervention, then I instrument Signal x with the instrumented I, and I determine the impact of the instrumented Signal on Termination. Indeed, as beliefs p evolve with signals x by equation (2), and since I is convex in p and follows the shape of v, an increase in I should imply an increase in x as well. Therefore, increasing the precision of x by the means of I translates in a positive correlation between these two dimensions.

1.6 Mechanism

In the section I will gradually show how Proposition 4 about termination is triggered by the mechanism, as prescribed by Propositions 1-3. Proposition 4 is in line with previous findings about the scientific approach and previous predictions about effectuation. Indeed Camuffo et al. (2020a, 2021), by analysing data from four different RCTs in different countries, show how scientific entrepreneurs are systematically more likely to terminate and do so earlier. On the effectual side, Dew et al. (2009) state that entrepreneurs adopting an affordable loss reasoning are more likely to fully commit to the idea and, therefore, less likely to terminate it. Results support the propositions of the model. Scientific *DMs* dynamically optimize information acquisition and acquire less information to make their decisions and terminate more and earlier. Effectual *DMs* acquire more information then scientific *DMs*, but improve precision of signals less and do not act accordingly in face of bad signals. They terminate less and later than everyone. Entrepreneurs in the control group acquire more information than both scientific and effectual entrepreneurs, but improve less then them the precision of signals and do not act accordingly. They terminate less and later than scientific entrepreneurs, but more and later than effectual entrepreneurs. The entire analysis is mainly using longitudinal data where I included time fixed effects, along with mentor dummies, and standard error clustered at the intervention-mentor level.

1.6.1 Information

Figure 3 shows that I_t (cumulative) is increasing and concave in time. *DMs* acquire information in the early stage and, as uncertainty is reduced, the accumulation of information slows down and becomes flat. This suggests that *DMs* are saturating their bounded search space and when acquiring new information gets costly, they stop this acquisition process as the boundary condition $MS = rv$ is approaching or met. The graph also supports Proposition 1, with the control group accumulating more information than effectual and scientific *DMs* and scientific *DMs* acquiring less information. Table 3 shows the impact of the two treatments on I (flow), running a panel linear probability model. Column (1) shows that the effectual treatment increases by 0.5 percentage points (p=0.023) the flow of information with respect to the scientific treatment. Column (2) and (3) show respectevely that the entrepreneurs in the control group acquire more information than scientific and effectual entrepreneurs and the impact is statistically significant. Proposition 1 is fully supported.

Figure 1.5: The acquisition of I_t slows down over time as *DMs* reduce uncertainty. Overall, scientific *DMs* aquire less I_t than effectual *DMs* and others.

Table 1.3: Information (flow)

	(1) Information (flow) OLS Panel Effectuation vs Scientific	(2) Information (flow) OLS Panel Scientific vs Control	(3) Information (flow) OLS Panel Effectuation vs Control
Intervention	0.005*	-0.008***	-0.004***
	(0.023)	(0.002)	(0.000)
Constant	0.579***	0.619***	0.606***
	(0.000)	(0.000)	(0.000)
Observations	1,470	1,385	1,453
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentorr
Number of id	207	203	206

Robust pval in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In column (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Column (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

1.6.2 Informative signals

The second step of the model prescribes that *DMs* acquire information to improve the precision of signals. Table 4 shows the impact of I on signals using two-stage least squares with a panel linear model, instrumenting I with the intervention. In this way, I can test whether signals get more intense, i.e. more precise, as entrepreneurs acquire information, establishing a causal relationship between the treatment and the intensity of signals. Column (1) in Table 4 shows that increasing I (flow) by one percentage point decreases the intensity of signals if Information (flow) is instrumented with the effectual treatment as opposed to the scientific treatment. Instead column (2) and (3) show that, when instrumented respectevely with the scientific and effectual treatment against the control group, an increase in I increases signals. In sum, as prescribed by the model, scientific entrepreneurs acquire information with the specific aim to improve signal precision and reduce the noise induced by the brownian motion. Effectual entrepreneurs improve to some extent the signal precision, but statistically significantly less then scientific entrepreneurs. Entrepreneurs in the control group, even if they acquire more information, improve less signal precision. This is also in line with assumption 2: entrepreneurs in the control group acquire select a less valuable information technology and acquire cheap information that does not help to reduce the uncertainty, while scientific entrepreneurs acquire less but expansive valuable information and reduce uncertainty more. Effectual entrepreneurs acquire as well expansive valuable information, but they *choose* to not use their information to improve their predictive power. Beside a weak statistical significance, Proposition 2 is supported.

Table 1.4: Sub-mechanism 1: Instrumenting Information (flow) - 2SLS

	(1) Signals OLS Panel Effectuation vs Scientific	(2) Signals OLS Panel Scientific vs Control	(3) Signals OLS Panel Effectuation vs Control
Information (flow)	-2.374*	2.192	9.545***
	(0.091)	(0.109)	(0.008)
Observations	1,470	1,385	1,453
R-squared	-2.486	-1.334	-33.223
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in column (1). In column (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Column (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

1.6.3 Reaction to signals

Since the the Bellman value v is convex in p and therefore in x, an increase of signals would imply higher value of the idea and viceversa. Scientific entrepreneurs are trained to optimize signal precision such that they make a decision accordingly. While effectual entrepreneurs are trained to act coherently with their affordable loss, instead of in face of positive and negative signals. Table 5 shows that results support Proposition 3, using two-stage least squares with a panel linear model, instrumenting signals x with the intervention. Indeed, Column (1) shows that an increase (decrease) of signals is positively (negatively) correlated to the choice to terminate the idea, when signals are instrumented with the effectual treatment against the scientific treatment. Column (2) shows that scientific entrepreneurs are less (more) likely to terminate in face of positive (negative) signals than entrepreneurs in the control group: an increase of signals, instrumented with the scientific treatment, impacts negatively the choice to terminate the idea. Column (3) shows that signals instrumented with the effectual treatment negatively correlates with the choice to terminate, when compared with the control group: effectual entrepreneurs do not act coherently with an optimization framework, but are less likely to terminate even when facing bad signals. This is in line with their willingness to control future events instead of rely on predictive signals.

Table 1.5: Sub-mechanism 2: Instrumenting Signals - 2SLS

	(1) Termination OLS Panel Effectuation vs Scientific	(2) Termination OLS Panel Scientific vs Control	(3) Termination OLS Panel Effectuation vs Control
Signals	1.426*	-0.232*	0.258***
	(0.087)	(0.071)	(0.000)
Observations	1,470	1,385	1,453
R-squared	-2.550	0.017	-0.144
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in column (1). In column (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Column (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

1.6.4 Full mechanism

We are now ready to put all things together and unfold the whole mechanism. Table 3 supports Proposition 1, Table 4 supports Proposition 2 and Table 5 supports Proposition 3. Table 6 here below shows again support Proposition 3 but with the whole mechanism in action, by using a three-stage least squares linear panel model that I run using the *cmp* command in STATA that allows to estimate multi-equation systems with linear and probit models, and cluster the standard error. Intervention instruments Information (flow) I which in turn instruments signals x that impact on termination. This is the full mechanism. *DMs* optimally acquire information to improve signal precision and inform their decision about terminate or commit to the idea. Column (1) shows that effectual entrepreneurs are less (more) likely to terminate (commit to) than scientific entrepreneurs in face of bad informative signals and vice-versa. Column (2) instead shows that scientic entrepreneurs are more likely to terminate in face of bad informative signals. This is a crucial result. Scientific *DMs* acquire less information but with a more valuable technology as they need to impreve their predictive power. They get more precise information and trust signals the they think are informative, optimizing the dynamic acquisition of information. On the other side, column (3) shows that effectual entrepreneurs, with compared to the control group, are again less likely to act coherently with the signals, even if these are more informative then signals collected by the control group. Effectual entrepreneurs do not make decisions by optimizing their information acquisition process, but keep exploring in the belief to shape the future. All these findings are statistically significant.

Table 1.6: Full Mechanism: Instrumenting Signals with Information (flow) and Information (flow) with Intervention - 3SLS OLS

	(1) Termination OLS Panel Effectuation vs Scientific	(2) Termination OLS Panel Scientific vs Control	(3) Termination OLS Panel Effectuation vs Control
Signals	1.427*	-0.231*	0.258***
	(0.073)	(0.083)	(0.001)
Constant	-0.842*	0.140*	-0.158***
	(0.074)	(0.099)	(0.001)
Observations	1,470	1,385	1,453
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in column (1). In column (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Column (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

1.7 Termination

1.7.1 Termination rates

I close my analysis by looking at the only impact of the intervention on termination, as prescribed by Proposition 4. Table 7 I report results from a linear probability longitudinal analysis. The effectual intervention decreases by 2.3 percentage points (p=0.000) the probability of terminating the idea at any point in time, with respect to the scientific treatment, and by 1.9 percentage points (p=0.000) with compared to the control group, as shown respectevely in columns (1) and (3). Column (2) reports a positive but not statistically significant impact of the scientific treatment on the probability of termination, when compared to the control group. This effect gets significant when I run a probit panel regression as in table 8. Marginal effect at values reported in column (2) shows that the scientific approach increases by 4.3 percentage points (p=0.079) the probability of terminating the idea, while the effectual treatment decreases this probability by 1.1 percentage points (p=0.000), as shown in column (3), both compared with the control group. Marginal effect, based on the results reported in column (1), shows that the effectual treatment decreases by 1.9 percentage points (p=0.000) the probability of termination also when compared with the scientific treatment. As a robusteness check, in line with previous research (Camuffo et al. 2021), table 9 reports the results of a cross section linear probability model. Column (1) shows that effectual entrepreneurs are 9.2 percentage points (p=0.005) less likely to terminate than scientific and 4.8 percentage points (p=0.001) less likely to terminate than entrepreneurs in the control group (column (3)), in line with previous predictions about the adoption of a more effectual approach (Dew et al. 2009). Column (2) confirms previous findings about the scientific approach (Camuffo et al. 2020a, Camuffo et al. 2021): scientific entrepreneurs are more likely to terminate than entrepreneurs in the control group and the effect is significant. Table 10 shows the same findings, but by running a cross section probit regression. Marginal effect at values reported in column (1) and (2) show that entrepreneurs treated with an effectual versus a scientific approach are 8.9 percentage points (p=0.000) less likely to terminate, while scientific entrepreneurs are 3.3 percentage points (p=0.033) more likely to terminate than entrepreneurs in the control group. The marginal effect for the effectual treatment in column (3) shows that the effectual treatment decreases the probability of termination by 4.8 percentage points (p=0.000) with compared to entrepreneurs in the control group.

Table 1.7: Termination OLS Panel

	(1) Termination OLS Effectuation vs Scientific	(2) Termination OLS Scientific vs Control	(3) Termination OLS Effectuation vs Control
Intervention	-0.023***	0.004	-0.019***
	(0.000)	(0.182)	(0.000)
Constant	0.034***	-0.008	0.010
	(0.007)	(0.105)	(0.261)
Observations	1,470	1,385	1,453
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

Table 1.8: Termination Probit Panel

	(1) Termination Probit Effectuation vs Scientific	(2) Termination Probit Scientific vs Control	(3) Termination Probit Effectuation vs Control
Intervention	-0.234***	0.047*	-0.136***
	(0.000)	(0.079)	(0.000)
Constant	-5.287***	-5.718***	-5.524***
	(0.000)	(0.000)	(0.000)
Observations	1,470	1,385	1,453
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

Table 1.9: Termination OLS Cross-Section

	(1) Termination OLS Effectuation vs Scientific	(2) Termination OLS Scientific vs Control	(3) Termination OLS Effectuation vs Control
Intervention	-0.092***	0.033*	-0.048***
	(0.005)	(0.079)	(0.001)
Constant	0.440***	0.297***	0.289***
	(0.000)	(0.000)	(0.000)
Observations	207	203	206
R-squared	0.031	0.014	0.009
Dummies for mentors	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

Table 1.10: Termination Probit Cross-Section

	(1) Termination Probit Effectuation vs Scientific	(2) Termination Probit Scientific vs Control	(3) Termination Probit Effectuation vs Control
Intervention	-0.271***	0.094**	-0.146***
	(0.000)	(0.033)	(0.000)
Constant	-0.131	-0.534***	-0.557***
	(0.448)	(0.000)	(0.000)
Observations	207	203	206
Dummies for mentors	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

1.7.2 Time to termination

Table 11 reports the results of a Cox proportional hazard model to fully support Proposition 4, in line with the dynamic optimization presented in the model. Hazard rate of termination in column (1) is lower for effectual entrepreneurs than for scientific ones, while column (2) and (3) compare these respectevely the scientific and effectual approach with the control training. The scientific group show a higher hazard rate ratio than the control group, while start-up treated with the effectual approach show a lower ratio. As a robusteness check, in table 12 I report an OLS regression that predict the week of termination. Findings are in line with the survival analysis. Effectual entrepreneurs terminate later than both entrepreneurs in the scientific and control group. Scientific entrepreneurs, again in line with previous findings about the scientific approach (Camuffo et al. 2020, 2021) terminate earlier than the control group. They terminate earlier than effectual entrepreneurs as well.

Table 1.11: Termination Time - Survival

	(1) Hazard of termination Survival Effectuation vs Scientific	(2) Hazard of termination Survival Scientific vs Control	(3) Hazard of termination Survival Effectuation vs Control
Intervention	-0.452***	0.112**	-0.272***
	(0.000)	(0.035)	(0.000)
Observations	207	203	206
Dummies for mentors	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

Table 1.12: Termination Time - OLS

	(1) Week of termination OLS Effectuation vs Scientific	(2) Week of termination OLS Scientific vs Control	(3) Week of termination OLS Effectuation vs Control
Intervention	4.974***	-0.389	4.466***
	(0.004)	(0.357)	(0.008)
Constant	45.447***	50.894***	48.963***
	(0.000)	(0.000)	(0.000)
Observations	207	203	206
R-squared	0.050	0.013	0.029
Dummies for mentors	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

Overall, I find a strong support for Proposition 4: scientific *DMs* are more likely to terminate their idea and do so earlier, effectual *DMs* are less likey to terminate their idea and do so later, *DMs* in the control group stay in the middle.

1.8 Additional results: Performance

My model proposes a mechanism to explain termination rates and when termination happens, but does not provide explicit predictions on performance. Figure 6 shows that, within the observation window, firms treated with a scientific approach perform better in terms of revenues over time, with firms in the control and effectual group performing at similar levels. Table 13 provides the impact of the treatments on the cumulative revenues (in EUR) in the last data point. Column (1) shows what seen in Figure 6: firms treated with a scientific approach earn EUR 3,544.39 more than firms in the effectual group (p=0.020) and EUR 3,672.65 more than firms in the control group (p=0.043), as reported in column (2). Firms treated with the effectual intervention show very small difference with compared to firms in the control group (column (3)). As a robusteness check, in table 14 I report a longitudinal analysis where the dependent variable is the revenue flow of the firm in each period. Findings are in line with the cross section analysis: firms in the scientific group perform better at any point in time than both effectual and control firms. Beside the small effect, results can be interpreted coherently with the previous findings, but with *cum grano salis*. While, the results about the comparison between entrepreneurs in the scientific and control group are in line with previous research (Camuffo et al. 2020a, 2021), results about entrepreneurs in the effectual group deserve more attention. Indeed, scientific *DMs* might be more likely to avoid false positives given their focus on controlling the precision parameter I to improve their prediction power. On the other side, effectual *DMs*, being less conservative, are less likely to avoid false positives as they keep exploring and trying to turn bad ideas into performing ideas. This might imply that they can start performing in the medium-long term. I leave this point open, since the model does not provide a compelling explanation.

Figure 1.6: Performance (euros).

Table 1.13: Performance OLS Cross-Section

	(1) Revenue OLS Effectuation vs Scientific	(2) Revenue OLS Scientific vs Control	(3) Revenue OLS Effectuation vs Control
Intervention	-3,544.395**	3,672.655**	-159.116
	(0.020)	(0.012)	(0.570)
Constant	2,926.237	1,347.278	696.710**
	(0.197)	(0.184)	(0.043)
Observations	207	203	206
R-squared	0.023	0.014	0.028
Dummies for mentors	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

Table 1.14: Performance OLS Panel

	(1) Revenue (flow) OLS Effectuation vs Scientific	(2) Revenue (flow) OLS Scientific vs Control	(3) Revenue (flow) OLS Effectuation vs Control
Intervention	-498.74**	551.282***	9.676 *
	(0.002)	(0.000)	(0.520)
Constant	-92.774	-379.898**	-107.631
	(0.765)	(0.013)	(0.206)
Observations	1,470	1,385	1,453
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in model (1). In model (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Model (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization (see Appendix for more details).

Overall, I believe that these results shed a light on what mechanism drives survival rates and timing in innovative contexts and what are the implications of exploring an idea by adopting an explorative-predictive approach or a control-oriented approach to decision-making under uncertainty. Results show a clear path. *DMs* adopting a explorative-predictive approach to explore their idea are more likely to acquire a low amount of information but such that it is sufficient to improve more and faster the precision of signals and act accordingly in face of precise bad signals and terminate more do so ealier. They optimize the dynamic acquisition of information more than others. Control-oriented *DMs* acquire more information than explorative-predictiv *DMs* since their focus is on controlling the risk parameter f, which I interpret what they can afford to lose in the worst scenario $\vartheta = F$, namely the affordable loss. They improve to some extent the precision of signals as well, but their decisions are guided by the affordable loss and not by signals, producing different termination timing and rates than explorative-predictive *DMs*.

1.9 Conclusion

In the economics of entrepreneurship and innovation management, a key problem is represented by the excessive entry rate of start-ups and, more in general, the launch of potentially innovative ideas in face of low returns (Geroski, 1995; Caves, 1998, Astebro, Jeffrey, and Adomdza, 2007, Elfeinbein and Knott, 2017). This is due to high levels of uncertainty that entrepreneurs and managers must face and often the outcome of their decisions turns out to be suboptimal. Different strategies can be used to deal with uncertainty. Assuming the beneficial aspects of exploration before deciding, in this paper I study the timing and the rate of termination of ideas of decision-makers adopting a more predictive, but still explorative, approach and a control-oriented strategy. I develop a stylized optimization model of dynamic acquition of information that I empirically test by running a Randomized Control Trial engaging 308 early-stage start-ups in Italy in 2020. I operationalize the explorative-predictive strategy with the scientific approach (Camuffo et al. 2020a, 2021) and the control-oriented approach with the effectual startegy (Sarasvathy 2001). In this light, start-ups in the RCT were randomly allocated to 3 groups and trained respectively with 1) a scientific approach: build precise experiments to falsify theories aimed to predict the future 2) an effectuatual approach: set an affordable loss and get pre-commitment from potential stakeholders to control and shape the future and 3) a 'control' approach: to base decisions on intuition and natural heuristics. I found that start-ups in the scientific group terminate more and earlier than the control group, while start-ups in the effectual group terminate less and later than the control group. The underlying mechanism, which I test, is that the scientific approach is a more valuable technology to acquire information: scientific *DMs* acquire less but more expansive and more precise information which produce more informative signals. They dynamicaly optimize thsi process and terminate in face of bad and reliable signals. The effectual treatment generates a similar but less intense effect as the effectual approach provides a valuable technology as well which produces less precise signals. Coherent with their attitude towards control, effectual *DMs* tend to ignore signals if they can bear the endogenous level of risk. They terminate less and later. I find also that scientific start-ups perform better than the control group, while there is no significative effect for effectual start-ups with compared to the control group. The result for scientific firms is in line with previous research. The different results between scientific and effectual start-ups can be explained in two ways. First, higher rates of termination imply that scientific entrepreneurs are more likely to avoid false positive ideas that would have implied no revenues if they did not terminate them. Second, effectual entrepreneurs are less likely to avoid false positive ideas, and, in face of bad news, they try to turn them into new opportunities. This may imply that they might play in the long term and show lower performance in the short term, but better performance in the long term. These results are subject to several limitations. They take into consideration a limited observation window, and this might impact the average rate of termination. But more than one year is a reasonable time window for early-stage start-ups when looking at termination rates. Moreover, the pandemic, and the consequent lockdown during the training, might have impacted the possibility to interact with other stakeholders or build offline experiments. To limit this issue, we also trained entrepreneurs how to interact efficiently online with potential clients, suppliers and allies and how to test theories online. I also acknowledge that the measure of information might be incomplete, since legal issues, technology constraints or other fields of knowledge might impact how informative the signal is. I see this as an opportunity for future research, since it can be interesting disentangling the concept of information in a business context and understand what kind of the information is mostly impacting outcomes and under what conditions. Moreover, I do not measure the cost of acquiring information. Measuring the cost of information and analysing its impact on decisions, coherently with the proposed model, could be another opportunity for future research. This study has also managerial and practical implications. Indeed, we could apply this model to several contexts, such as R&D in innovative firms. These findings suggest that firms might control the efforts to acquire new information and set thresholds to guide their decision-making process. R&D managers might develop theories and tailor precise experiments to predict the probability of success of a certain project in their portfolio, or they might set an affordable loss for their R&D unit and explore different innovative projects if they believe in an unpredictable future. The model and the findings suggest that in the first case, it is more likely that the R&D manager will terminate more projects in the portfolio and earlier, focusing on the most promising to perform eventually better in the short term. In the second case, the R&D manager will terminate less projects, eventually open new projects, and perform eventually better in the long term. I hope this study helps to uncover a potential mechanism driving termination rates and timing of innovative ideas using predictive and non-predictive strategies. Given the importance of these themes, I think these findings might encourage further research to deeply understand the underlying mechanisms and implications of predictive and control-oriented decision-making approaches.

References

Aghion, P. and Howit P. (1992) "A model of growth through creative destruction". *Econometrica* 60, 323-351

Agrawal, R., and 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., and Stern, S. (2021). Enabling entrepreneurial choice. *Management Science.*

Alvarez, S. A., and Barney, J. B. 2005. How do entrepreneurs organize firms under conditions of uncertainty? *Journal of Management*, 31(5): 776-793.

Åstebro T, Herz H, Nanda R, Weber RA. 2014. Seeking the roots of entrepreneurship: insights from behavioral economics. *Journal of Economic Perspectives* 29(3): 49–69.

Åstebro T, Jeffrey SA, Adomdza GK. 2007. Inventor perseverance after being told to quit: the role of cognitive biases. *Journal of Behavioral Decision Making* 20(3): 253–272.

Arrow K. J., Blackwell D. and Girshick M. A. (1949). Bayes and Minimax Solutions of Sequential Decision Problems. *Econometrica* Vol. 17, No. 3/4 pp. 213-244

Blank SG (2006) The Four Steps to the Epiphany: Successful Strategies for Products That Win (KS Ranch, Pescadero, CA).

Csaszar, F. A., Ostler J. (2019). A Contingency Theory of Representational Complexity in Organizations. *Organization Science* (Forthcoming)

Camerer, Colin, and Dan Lovallo. 1999 "Overconfidence and excess entry: An experimental approach." *American Economic Review* 89 (1): 306-318.

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

Camuffo, A., Gambardella, A., and Spina, C. (2020b). Small changes with impact: Experimental evidence of a scientific approach to the decision-making of entrepreneurial firms. Working Paper.

Camuffo, A., Gambardella, A., Messinese, D., Novelli, E., Paolucci, E., and Spina, C. (2021). A scientific approach to innovation management. evidence from four field experiments. NBER

Caves R. E., (1998). Industrial Organization and New Findings on the Turnover and Mobility of Firms. *Journal of Economic Literature* Vol. 36, No. 4 (Dec., 1998), pp. 1947-1982

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

Choi Y. R., Lévesque M., Shepherd D. A. (2008). When should entrepreneurs expedite or delay opportunity exploitation? *Journal of Business Venturing* 23 (2008) 333–355

Cyert, R.M., DeGroot M.H., (1987). Bayesian analysis and uncertainty in economic theory. Springer

Dew N, Read S, Sarasvathy SD, Wiltbank R (2009). Effectual versus predictive logics in entrepreneurial decision-making: differences between experts and novices. *Journal of Business Venturing* 24 (4), 287–309.

Eisenmann, T., Ries, E. and Dillard, S. (2013). Hypothesis-driven Entrepreneurship: the Lean Startup. *Harvard Business School Note* 9-812-095.

Elfeinbein D. W., Knott A. M., and Croson R., (2017) Equity stakes and exit: An experimental approach to decomposing exit delay. Strategic Management Journal, Volume 38, Pages 278-299

Ely J. C. (2017) Beeps. *The American Economic Review* Vol. 107, No. 1, pp. 31-53

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

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

Felin, T., and Zenger, T. (2017). The theory-based view: Economic actors as theorists. *SSRN Electronic Journal*.

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

Geroski, P.A. (1995). What do we know about entry? *International Journal of Industrial Organization*, 13(4)—421-440.

Goldfarb, B., and King, A. A. (2016). Scientific apophenia in strategic management research: Significance tests mistaken inference. *Strategic Management Journal*, 37(1), 167-176.

Jones, John Bailey, and Sangeeta Pratap. 2020. "An Estimated Structural Model of Entrepreneurial Behavior." *American Economic Review*, 110 (9): 2859-98.

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

Kirzner, Israel M. (1973).Competition and Entrepreneurship. Chicago: *University of Chicago Press*.

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

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

Koning, R., Hasan, S., and Chatterji, A. (2019). Experimentation and startup performance: Evidence from A/B testing (No. w26278). *National Bureau of Economic Research*.

Knight F. H. (1921). Risk, Uncertainty and Profit . Houghton Mifflin: New York.

Krabel S., Mueller P. (2009). What drives scientists to start their own company?: An empirical investigation of Max Planck Society scientists. *Research Policy*, 38 (6), pp. 947-956

Lowe R, Ziedonis A. 2006. Overoptimism and the performance of entrepreneurial firms. *Management Science* 52(2): 173–186.

McGrath R. G. (1999). Falling forward: real options reasoning and entrepreneurial failure. *Academy of Management Review* 24(1): 13-30.

Meyer M.W., Zucker L.G. (1989). Permanently Failing Organizations. Sage: Newbury Park, CA.

Mintzberg H. (1990). The design school: reconsidering the basic premises of strategic management. *Strategic Management Journal* 11(3): 171-195.

Mintzberg H. (1994). The Rise and Fall of Strategic Planning: Reconceiving Roles for Planning, Plans, Planners. Free Press: New York.

Mosakowski E. (1997). Strategy making under causal ambiguity: conceptual issues and empirical evidence. *Organization Science* 8(4): 414-44

Mises, L. v. (1998). Human action: A treatise on economics(4threv.ed.). Auburn, AL: Ludwig von Mises Institute.

Moore, D. A., and Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2), 502–517.

Moscarini G., Smith L., (2001). The Optimal Level of Experimentation. *Econometrica*, Nov., 2001, Vol. 69, pp. 1629-1644

Murray, F. (2004). The role of academic inventors in entrepreneurial firms: sharing the laboratory life. *Research policy* 33 (4), 643-659

Nanda, R., and Rhodes-Kropf, M. (2017). Cost of Experimentation and the Evolution of Venture Capital (No. 15-070). Harvard Business School.

Nelson T. (2003). The persistence of founder influence: management, ownership, and performance effects at initial public offering. *Strategic Management Journal*, 24: 707-72

Nicholls-Nixon, C. L., Cooper, A. C., and Woo, C. Y. (2000). Strategic experimentation: Understanding change and performance in new ventures. *Journal of Business Venturing*, 15(5), 493-521.

Ott T. E., Eisenhardt K. M., Bingham C. B. (2017). Strategy Formation in Entrepreneurial Settings: Past Insights and Future Directions. *Strategic Entrepreneurship Journal*, 11(4), pp. 306-325.

Packard, M. D. 2017. Where did interpretivism go in the theory of entrepreneurship? *Journal of Business Venturing*, 32: 536–549

Packard, M. D. and Clark B. B., 2020. On the Mitigability of Uncertainty and the Choice between Predictive and Nonpredictive Strategy. *Academy of Management Review* 2020, Vol. 45, No. 4, 766–786.

Pfeffer, J., and Sutton, R. I. (2006). Hard facts, dangerous half-truths, and total nonsense: Profiting from evidence-based management. *Harvard Business Press.*

Reymen I.M., Andries P., Berends H., Mauer R., Stephan U., Burg E. (2015). Understanding dynamics of strategic decision making in venture creation: A process study of effectuation and causation. *Strategic Entrepreneurship Journal*, 9 (4), pp. 351-379

Ries E (2011) The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses (Crown Business, New York).

Sarasvathy, S. D. (2001). Causation and effectuation: Toward a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of management Review*, 26(2), 243-263.

Sarasvathy S. D. (2008). Effectuation: Elements of Entrepreneurial Expertise. *New Horizons in Entrepreneurship*

Scott, E. L., Shu, P., and Lubynsky, R. M. (2015). Are better ideas more likely to succeed?: An empirical analysis of startup evaluation (No. 16-013). Harvard Business School.

Shepherd, D. A., Haynie, J. M., and McMullen, J. S. (2012). Confirmatory search as a useful heuristic? Testing the veracity of entrepreneurial conjectures. *Journal of Business Venturing*, 27(6), 637-651.

Toole A. A., Czarnitzki D. (2009). Exploring the Relationship Between Scientist Human Capital and Firm Performance: The Case of Biomedical Academic Entrepreneurs in the SBIR Program. *Management Science*, 55(1), pp. 101-114

Townsend D. W., Hunt R., McMullen J. S., Sarasvathy S. D. (2018). Uncertainty, Knowledge Problems, and Entrepreneurial Action *Academy of Management Annals* 2018, Vol. 12, No. 2, 659–687

Tversky, A., Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131.

Wald, A. (1947). Foundations of a General Theory of Sequential Decision Functions. *Econometrica*, 15, 279-313.

Wiltbank, R., Dew, N., Read, S., Sarasvathy, S. D. (2006). What to do next? The case for non-predictive strategy. *Strategic management journal*, 27(10), 981-998.

Zhong W. (2017). Optimal dynamic information acquisition. Available at SSRN 3024275

APPENDIX

1.10 Balance checks

Table A1: Balance Checks Scientific vs Control

Variable Name	Description	Scientific		Control		Difference	
		Mean	SD	Mean	SD	b	p
Gender (male)	Gender of the main founder	0.84	0.37	0.82	0.38	-0.02	(0.68)
Age	Age of the main founder	30.23	8.79	30.09	8.01	-0.14	(0.91)
Phase	Phase of development of the start-up (1 Problem analysis; 2 Prototype; 3 Prototype with customers; 4 On the market but no revenues; 5 On the market with revenues)	1.67	1.02	1.65	1.09	-0.01	(0.93)
Location	Dummy variable equal to 1 if the start-up is location in the North of Italy, 0 otherwise	0.58	0.50	0.55	0.50	-0.03	(0.63)
Sector	Sector the start-up operates in (1 if Software, 0 otherwise)	0.09	0.29	0.11	0.31	0.02	(0.62)
Work experience	Number of years of work experience	6.88	7.29	6.91	7.65	0.03	(0.98)
Managerial experience	Number of years of managerial experience	1.36	3.52	1.52	4.10	0.16	(0.76)
Eduction level	Highest educational level attained by team members (4= PhD, 3=MSc, 2=BA, 1=high school or no degree; main founder)	2.03	1.02	2.10	0.98	0.07	(0.62)
Team size	Number of team members	2.20	1.52	2.18	1.70	-0.02	(0.94)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	17.97	17.47	15.20	12.67	-2.77	(0.20)
Overprecision	Number between 0 and 1 counting the average number of correct answers falling in a 90% confidence interval	0.57	0.24	0.60	0.25	0.03	(0.36)
Information I_0	Amount of information about Sector, Customers, Competitors, Resource (range, 0 to 1)	0.61	0.19	0.62	0.20	0.01	(0.70)
Prior belief p_0	Self-assessment that the idea will be successful (range, 0 to 1)	0.65	0.20	0.64	0.24	-0.01	(0.64)
Probability to terminate	Probability of terminating the project	0.15	0.19	0.17	0.19	0.02	(0.34)
Probability to pivot	Probability of changing the business idea	0.26	0.21	0.29	0.24	0.03	(0.30)
Scientific intensity	Score reflecting the extent to which the entrepreneur's decision-making process follows a scientific approach (range, 0 to 5)	1.80	1.16	1.63	1.05	-0.17	(0.29)
Bird in hand	Score reflecting the extent to which the entrepreneur leverages her/his existing knowledge, skills and social network (range, 0 to 5)	3.26	1.13	3.20	1.14	-0.06	(0.44)
Affordable loss	Score reflecting the extent to which the entrepreneur chooses waht they can afford to lose and make decisions accordingly (range, 0 to 5)	2.56	1.48	2.49	1.61	-0.07	(0.75)
Crazy quilt	Score reflecting the extent to which the entrepreneur interact with potential stakeholders (range, 0 to 5)	1.11	1.17	1.18	1.19	0.06	(0.72)
Observations		102		101		203	

Table A2: Balance Checks Effectuation vs Control

Variable Name	Description	Effectuation		Control		Difference	
		Mean	SD	Mean	SD	b	p
Gender	Gender of the main founder	0.78	0.42	0.82	0.38	0.041	(0.47)
Age	Age of the main founder	29.4	8.16	30.09	8.01	0.69	(0.54)
Phase	Phase of development of the start-up (1 Problem analysis; 2 Prototype; 3 Prototype with customers; 4 On the market but no revenues; 5 On the market with revenues)	1.50	1.01	1.65	1.09	0.16	(0.28)
Location	Dummy variable equal to 1 if the start-up is location in the North of Italy, 0 otherwise	0.58	0.50	0.55	0.50	-0.04	(0.60)
Sector	Sector the start-up operates in (1 if Software, 0 otherwise)	0.06	0.23	0.11	0.31	0.05	(0.18)
Work experience	Number of years of work experience	5.85	6.61	6.91	7.65	1.06	(0.29)
Managerial experience	Number of years of managerial experience	1.14	2.81	1.52	0.38	0.16	(0.43)
Eduction level	Highest educational level attained by team members (4= PhD, 3=MSc, 2=BA, 1=high school or no degree; main founder)	2.01	0.98	2.10	0.98	0.07	(0.61)
Team size	Number of team members	2.08	1.22	2.18	1.70	0.10	(0.63)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	14.59	15.94	15.20	12.67	0.61	(0.76)
Overprecision	Number between 0 and 1 counting the average number of correct answers falling in a 90% confidence interval	0.58	0.22	0.60	0.25	0.02	(0.62)
Information I_0	Amount of information about Sector, Customers, Competitors, Resource (range, 0 to 1)	0.60	0.19	0.62	0.20	0.03	(0.36)
Prior belief p_0	Self-assessment that the idea will be successful (range, 0 to 1)	0.63	0.22	0.64	0.24	0.01	(0.83)
Probability to terminate	Probability of terminating the project	0.17	0.19	0.17	0.19	0.00	(0.87)
Probability to pivot	Probability of changing the business idea	0.26	0.20	0.29	0.24	0.03	(0.30)
Scientific intensity	Score reflecting the extent to which the entrepreneur leverages her/his existing knowledge, skills and social network (range, 0 to 5)	1.67	1.14	1.63	1.05	-0.04	(0.81)
Bird in hand	Score reflecting the extent to which the entrepreneur leverages her/his existing knowledge, skills and social network (range, 0 to 5)	3.08	1.01	3.20	1.14	0.12	(0.44)
Affordable loss	Score reflecting the extent to which the entrepreneur chooses waht they can afford to lose and make decisions accordingly (range, 0 to 5)	2.19	1.50	2.49	1.61	0.3	(0.17)
Crazy quilt	Score reflecting the extent to which the entrepreneur interact with potential stakeholders (range, 0 to 5)	1.09	1.11	1.18	1.19	0.09	(0.58)
Observations		105		101		206	

Table A3: Balance Checks Scientific vs Effectuation

Variable Name	Description	Scientific		Effectuation		Difference	
		Mean	SD	Mean	SD	b	p
Gender	Gender of the main founder	0.84	0.37	0.78	0.42	0.06	(0.26)
Age	Age of the main founder	30.23	8.79	29.4	8.16	0.83	(0.48)
Phase	Phase of development of the start-up (1 Problem analysis; 2 Prototype; 3 Prototype with customers; 4 On the market but no revenues; 5 On the market with revenues)	1.67	1.02	1.50	1.01	0.17	(0.23)
Location	Dummy variable equal to 1 if the start-up is location in the North of Italy, 0 otherwise	0.58	0.50	0.58	0.50	-0.00	(0.97)
Sector	Sector the start-up operates in (1 if Software, 0 otherwise)	0.09	0.29	0.06	0.23	0.03	(0.39)
Work experience	Number of years of work experience	6.88	7.29	5.85	6.61	1.035	(0.29)
Managerial experience	Number of years of managerial experience	1.36	3.52	1.14	2.81	0.22	(0.62)
Eduction level	Highest educational level attained by team members (4= PhD, 3=MSc, 2=BA, 1=high school or no degree; main founder)	2.03	1.02	2.01	0.98	0.00	(0.99)
Team size	Number of team members	2.20	1.52	2.08	1.22	0.12	(0.55)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	17.97	17.47	14.59	15.94	3.38	(0.15)
Overprecision	Number between 0 and 1 counting the average number of correct answers falling in a 90% confidence interval	0.57	0.24	0.58	0.22	0.02	(0.64)
Information I_0	Amount of information about Sector, Customers, Competitors, Resources (range, 0 to 1)	0.61	0.19	0.60	0.19	0.02	(0.58)
Prior belief p_0	Self-assessment that the idea will be successful (range, 0 to 1)	0.65	0.20	0.63	0.22	0.02	(0.54)
Probability to terminate	Probability of terminating the project	0.15	0.19	0.17	0.19	-0.02	(0.43)
Probability to pivot	Probability of changing the business idea	0.26	0.21	0.26	0.20	-0.00	(0.97)
Scientific intensity	Score reflecting the extent to which the entrepreneur interact with potential stakeholders (range, 0 to 5)	1.80	1.16	1.67	1.14	0.13	(0.42)
Bird in hand	Score reflecting the extent to which the entrepreneur leverages her/his existing knowledge, skills and social network (range, 0 to 5)	3.26	1.13	3.08	1.01	0.18	(0.25)
Affordable loss	Score reflecting the extent to which the entrepreneur chooses waht they can afford to lose and make decisions accordingly (range, 0 to 5)	2.56	1.48	2.19	1.50	0.37	(0.08)
Crazy quilt	Score reflecting the extent to which the entrepreneur interact with potential stakeholders (range, 0 to 5)	1.11	1.17	1.09	1.11	0.03	(0.86)
Observations		102		105		207	

1.11 Variables details

1.11.1 Scientific intensity

Table A4: Scientific Intensity Components

Component	Sub-component	Definition	Score
Theory	Clarity of theory	The extent to which the theory is understandable	0 (not clear at all) to 5 (extremely clear)
Theory	Articulation of theory	The extent to which the theory is detailed	0 (not detailed at all) to 5 (extremely detailed)
Theory	Consideration of alternatives	The extent to which the theory includes alternative possible options	0 (no consideration of alternatives at all) to 5 (careful consideration of many alternatives)
Theory	Theory based on evidence	The extent to which the theory is based on objective evidence	0 (not based on objective evidence at all) to 5 (extremely based on objective evidence)
Theory	Hierarchy of theory	The extent to which the theory helps to prioritize the problems to be solved	0 (problems are not prioritized at all) to 5 (extremely prioritized)
Theory	Theory modularity	The extent to which the theory decomposes a problem into sub-problems	0 (problems are decomposed at all) to 5 (problems are totally decomposed in subproblems
Hypotheses	Explicitness of hypotheses	The extent to which the respondent can articulate the fundamental assumptions that make his/her business viable	0 (not explicit at all) to 5 (extremely explicit)
Hypotheses	Coherence of hypotheses	The extent to which hypotheses are coherent with the theory	0 (not coherent at all) 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 (not detailed at all) 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 (not falsifiable at all) to 5 (extremely falsifiable)
Hypotheses	Testability of hypotheses	The extent to which it is possible to decide whether the hypothesis is true or false based on tests	0 (not testable at all) to 5 (extremely testable)
Hypotheses	Alternative hypotheses	The extent to which the hypotheses are aimed at falsifying one thing and supporting another as a direct (alternative) consequence	0 (not aimed to discover alternatives at all) to 5 (extremely aimed to discover alternatives)
Tests	Coherence of tests	The extent to which the test is coherent with the hypotheses	0 (not coherent at all) 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 (not valid at all) 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 (not representative at all) 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 (not rigorous at all) to 5 (extremely rigorous)
Tests	Causality of tests	The extent to which the test measures a causal link between two variables tested	0 (not causal links at all) to 5 (extremely high causality)
Tests	Test sampling	The extent to which the test is carried out on a sample with reduced selection and self-selection bias	0 (no bias reduction at all) to 5 (extremely unbiased)
Evaluation	Data-based assessment	The extent to which the evaluation is based on data	0 (not based on data at all) 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 (not coherent at all) to 5 (extremely coherent)
Evaluation	Systematic evaluation	The extent to which the evaluation is based on systematically collected and analysed data	0 (not systematic at all) 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 (not explanatory at all) to 5 (extremely explanatory)
Evaluation	Estimate of measure of performance	The extent to which the evaluation is based on the estimate of a measure of performance that is then used to make a decision	0 (not estimated at all) to 5 (extremely estimated)
Evaluation	Evaluation of alternatives	The extent to which the data collected help to estimate the value of the alternative component to the one tested	0 (not useful to evaluate aternatives at all) to 5 (extremely useful to evaluate aternatives)
Evaluation	Evaluation of negative results	The extent to which the evaluation of negative test results allow to learn new exploration possibilities	0 (not explorative at all) to 5 (extremely explorative)
Decision	Decision based on thresholds	The extent to which the decision to commit to, terminate or pivot is based on thresholds of test results value of the alternative component to the one tested	0 (not based on thresholds at all) to 5 (extremely based on thresholds)
Decision	Decision based on calibrated thresholds	The extent to which the thresholds take into account the quality of the tests and the type of data collected	0 (not based on calibrated thresholds at all) to 5 (extremely based on calibrated thresholds)

1.11.2 Effectual intensity components

Table A5: Bird in hand Components

Component	Definition	Score
Leverage who they are	The extent to which entrepreneurs develop the idea starting from who they are, that is, from their skills and abilities	0 (not based on who they are at all) to 5 (extremely based on who they are)
Leverage who they know	The extent to which entrepreneurs develop the idea starting from who they know, that is, from their family, friends, work network	0 (not based on who they know at all) to 5 (extremely based on who they know)
Leverage what they know	The extent to which entrepreneurs develop the idea starting from what they know, that is, from their background and experience	0 (not based on what they know at all) to 5 (extremely based on what they know)

Table A6: Affordable loss Components

Component	Definition	Score
Maximum affordable loss	The extent to which entrepreneurs used the maximum resources they can afford to lose	0 (not used resources at all) to 5 (extremely used resources)
Affordable loss risk level	The extent to which entrepreneurs did not add resources (even money) to those initially arranged, excluded those coming from external sources	0 (not added at all) to 5 (extremely added)
Focus on affordable loss	The extent to which entrepreneurs focused their attention on not losing more than they can afford, instead of focusing on expected return	0 (not focused at all) to 5 (extremely focused)

Table A7: Crazy quilt Components

Component	Definition	Score
Crazy quilt competitors	The extent to which entrepreneurs entered into agreements or collaboration with possible competitors	0 (not entered at all) to 5 (extremely entered)
Crazy quilt suppliers	The extent to which entrepreneurs entered into agreements or collaboration with suppliers who have shown interest before commercialization	0 (not entered at all) to 5 (extremely entered)
Crazy quilt customers	The extent to which entrepreneurs entered into agreements or collaboration with customers who have shown interest before commercialization	0 (not entered at all) to 5 (extremely focuenteredsed)

1.11.3 Information

Table A8: Information Components

Component	Definition	Score
Sector	how much information the entrepreneur has about the sector the start-up operates in	0 (not information at all) to 1 (all information available)
Customers	how much information the entrepreneur has about the potential customers	0 (not information at all) to 1 (all information available)
Competitors	how much information the entrepreneur has about the potential competitors	0 (not information at all) to 1 (all information available)
Resources	how much information the entrepreneur has about the resources needed to develop the idea	0 (not information at all) to 1 (all information available)

1.11.4 Weight on prior beliefs: α_0

1.12 Overprecision - α_0

Overprecision is measured coherently with previous research (Moore and Healy, 2008). I asked entrepreneurs to answer seven general questions by setting ranges which the right answers might be in, with 90% of confidence. In case they had absolutely no idea where the answer lied, they would indicate the maximum range possible for the question (i.e., if the unit is a percentage, fill in with 0% to 100%). Then I scored 1 for each correct range and 0 whether the right answer lied outside the proposed range and averaged the seven scores. Overall, I obtained a variable ranging from 0 to 1, where 1 means that they do not react to signals and the posterior probability equals the prior and 0 means that they completely rely on signals.

Table A9: Overprecision Question

(a) Please answer the following questions by indicating the range in which you think the correct answer is included. The range should contain the correct value in 90% of cases. If you have no idea what the answer is, enter the widest possible range (e.g. from 0 to 1000).

Question	Lower limit	Upper limit
What is the total GDP in Italy in 2019?	(bln EUR)	(bln EUR)
What is the average age of the Italian population in 2019?	(years)	(years)
What is the percentage of the urban population in Italy in 2019?	(%)	(%)
How many employees per company are there on average in Italy in 2019?	(number)	(number)
What is the percentage of companies with a website in Italy in 2019?	(%)	(%)
What percentage of Italian companies sell via the web in 2019?	(%)	(%)
What is the population of Piedmont in 2019?	(mln)	(mln)

1.13 Reaction to signals - Probit

Table A10: Sub-mechanism 2: Instrumenting Signals - 2SLS Probit

	(1) Termination Probit Panel Effectuation vs Scientific	(2) Termination Probit Panel Scientific vs Control	(3) Termination Probit Panel Effectuation vs Control
Signals	4.475***	-0.946	2.759****
	(0.000)	(0.532)	(0.000)
Constant	-4.269***	-5.765***	-5.977***
	(0.000)	(0.000)	(0.000)
Observations	1470	1385	1453
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentorr
Number of id	207	203	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a scientific approach, in column (1). In column (2) and (3) it is equal to 1 if start-ups have been treated respectevely with a scientific and an effectual approach, 0 otherwise. Column (1) controls for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization.

1.14 Full mechanism - Probit

Table A11: Full Mechanism Effectuation vs Scientific - 3SLS Probit

	(1) Stage 3 Termination Probit Panel Effectuation vs Scientific	(2) Stage 2 Signals Probit Panel Effectuation vs Scientific	(3) Stage 1 Informativeness (flow) Probit Panel Effectuation vs Scientific
Signals	4.478***		
	(0.000)		
Information (flow)		-2.367***	
		(0.001)	
Intervention			0.005***
			(0.001)
Constant	-3.192***	2.008***	0.590***
	(0.000)	(0.000)	(0.000)
Observations	1470	1470	1470
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	207	207	207

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if treated with the scientific approach. Regressions control for the variable *affordable loss* that were unbalanced between the scientific and effectual group despite randomization.

Table A12: Full Mechanism Scientific vs Control - 3SLS Probit

	(1) Stage 3 Termination OLS Probit Scientific vs Control	(2) Stage 2 Signals OLS Probit Scientific vs Control	(3) Stage 1 Information (flow) OLS Probit Scientific vs Control
Signals	-0.896		
	(0.549)		
Information (flow)		2.192	
		(0.133)	
Intervention			-0.008***
			(0.001)
Constant	-1.224	-0.727	0.619***
	(0.249)	(0.414)	(0.000)
Observations	1385	1385	1385
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	203	203	203

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an scientific approach, 0 if with a control training.

Table A13: Full Mechanism Effectuation vs Control - 3SLS Probit

	(1) Stage 3 Termination OLS Probit Effectuation vs Control	(2) Stage 2 Signals OLS Probit Effectuation vs Control	(3) Stage 1 Information (flow) OLS Probit Effectuation vs Control
Signals	2.768***		
	(0.000)		
Information (flow)		9.918***	
		(0.000)	
Intervention			-0.004***
			(0.002)
Constant	-2.904***	-5.135***	0.606***
	(0.000)	(0.000)	(0.000)
Observations	1453	1453	1453
Dummies for mentors	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor
Number of id	206	206	206

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. Intervention is a dummy variable equal to 1 if start-ups have been treated with an effectual approach, 0 if with a control training.

Chapter 2

A Scientific Approach to Innovation Management: Theory and Evidence from Four Field Experiments

This paper studies the implications of an approach in which managers and entrepreneurs make decisions under uncertainty by formulating and testing theories such as scientists do. By combining the results of four Randomized Control Trials (RCTs) involving 754 start-ups and small-medium enterprises and 10,730 data points over time, we find that managers and entrepreneurs who adopt this approach terminate more projects, do not experiment with many new ideas, and perform better. We develop a model that explains these results.

2.1 Introduction

Managers and entrepreneurs do not have solid routines or methods to make decisions under uncertainty, such as decisions regarding the launch of new products, services or new businesses. This observation is supported by ample evidence from the world of practice and academic research. For instance, a Harvard Business Review report surveyed 646 managers and showed that many managers rely on gut feeling rather than systematic and well-organized judgments (Harvard Business Review Analytic Services, 2016). McKinsey analyzed the decisions of 2207 executives, and found that in 28% of the cases they make good decisions, in 60% they make bad decisions, and in 12% they make infrequent good decisions (Lovallo and Sibony, 2010). Moreover, CEOs and entrepreneurs tend to be overconfident (Camerer and Lovallo, 1999; Astebro, 2003; Malmendier and Tate, 2008; Galasso and Simcoe, 2011; Astebro et al., 2014) and 84.8% of the US start-up do not report successful results at least within their first 7 years (Fairlie and Miranda, 2016).

Against this background, this paper studies an approach to managerial decisions under uncertainty that starts with a rigorous framing of the problem, develops theories that predict the outcomes of actions from logical connections between antecedents and consequences, and tests these theories using existing data or data drawn from well-defined experimental designs. We call this approach "scientific" because it overlaps to a good extent with the approach that scientists use to develop new knowledge.

Extant studies have documented the poor logic of managerial decisions under uncertainty, and some scholars and practitioners advocate greater rigor and logic in defining frameworks, formulating models, and testing them (e.g. Martin 2009; Felin and Zenger, 2009; Csazar and Levinthal, 2016; Eisenhardt and Bingham, 2017). Our study is also part of a broader effort to highlight the importance of managerial practices (Bertrand and Schoar, 2003; Bloom and Van Reenen, 2007; Bloom et al., 2013; Gosnell, List and Metcalfe, 2020).

We provide evidence of the implications of a scientific approach to decision-making through four randomized control trials (RCTs) involving start-ups and small-medium enterprises (SMEs). We obtain three main results. First, scientific decision-makers are more likely to terminate projects in early stages. Second, they change ideas fewer times before they commit to one or terminate the project. Third, they perform better.

We develop a model that provides a guide to interpret the empirical results. Recent models of entrepreneurial behaviour (Akcigit and Kerr, 2018; Jones and Pratap, 2020) focus on exploration for innovation decisions, but neglect the heterogeneity in how entrepreneurs decide to explore the environment and act based on feedback gathered through exploration. Our model focuses on some key features of the scientific approach and their implications for decision-making.

Our research follows an innovative research design based on four randomized control trials (RCTs) in which we conduct an evaluation of the same intervention across contexts (with the exception of minor changes due to operational constraints), following recent studies that have implemented a similar design (see, for instance, Allcott, 2015; Banerjee et al. 2015, Bowers et al. 2017). Along with a more accurate assessment of the impact of the intervention (Angrist and Pischke, 2010; List, Maniadis and Tufano, 2017), we address one important shortcoming of research based on field experiments, whose results may be sensitive to differences in populations, implementation, and economic environments (Levitt and List, 2009; Allcott, 2015; Banerjee et al. 2015, Bowers et al. 2017).

We conducted our four RCTs in Milan (in 2016 and 2017), Turin (2018), and London (2019) and involved start ups and small and medium enterprises (SMEs). Taken together, the four RCTs involve 754 firms and 10,730 data points. We decided, deliberately, to develop three new RCTs, using the same intervention and research design, after observing the results of our first RCT that uses data on 116 firms (Camuffo et al., 2020). This study finds encouraging results, which however depend too much on the empirical specifications or the methods of estimation. We then thought that seeking scale and statistical power was a better allocation of our research time than turning to new research designs and questions. In this way, we can produce more robust and reliable results about one important phenomenon rather than less robust knowledge about two or more different phenomena. In the conclusions of this paper we report where and how the larger scale of the four RCTs produce new or more robust results than Camuffo et al. (2020). In the Appendix we also show the results of the individual RCTs.

In all four RCTs, the experimental design involved the same training sessions over a period of about sixteen weeks, and the collection of data for about one year. All firms received training on how to manage innovation decisions using standard content on innovation management and entrepreneurship. About 80% of the content was the same for both treatment and control groups. Specifically, both treated and control groups were exposed to frameworks that they could use to support decision making, such as for instance the business model canvas and the balance scorecard; both groups were also exposed to the importance of using data and evidence to make decisions and where therefore exposed to multiple techniques for collecting data from customers, suppliers and key partners such as qualitative interviews, surveys and A/B testing. However, the treatment group was exposed to a scientific approach to decision-making: they were taught to use the frameworks learnt to develop a theory of the problem and draw hypotheses that flew logically from it; they were then taught to use the evidence-gathering techniques learnt in class to test those hypotheses and to evaluate the results in a disciplined way.

Our results highlight the importance of systematic decision-making in innovation and entrepreneurial contexts characterized by uncertainty. Prior work has emphasized the importance, in these contexts, of flexibility and experimentation (Kerr, Nanda and Rhodes-Kropf, 2014, Ewens, Nanda and Rhodes-Kropf, 2018), but this paper suggests that when entrepreneurs conduct experiments like scientists — which is, complementing experiments with theories and hypotheses — they benefit greatly. In other words, a theoretical framework helps entrepreneurs to envision scenarios more precisely, design more informative tests, and more successfully update priors. As noted by Kerr Nanda and Rhodes-Kropf, (2014, page.38) “Successful experimentation requires being able to capitalize on experiments that reveal positive outcomes, and these upside scenarios can be as tricky as termination decisions". The scientific approach helps to capitalize on experiments with positive or not so positive outcomes (when decision-makers need to pivot). As the results show, this has powerful consequences for the number and type of innovative projects that are commercialized, and for the performance of these projects. This explains an important mechanism related to the 'up or out' pattern (i.e. start-ups that fail to grow will abandon an economy) documented by Haltiwanger, Jarmin and Miranda, (2013) and Decker et al. (2014), among others.

Section 2 presents our model and its main predictions. Section 3 describes the experimental design, sample, and data collection. Section 4 presents our results. Section 5 concludes.

2.2 Model

2.2.1 Set-up

Decision-makers explore ideas in stages. The goal of exploration is to identify a strategy or an activity that they can pursue and that produces an outcome. For short, we call this strategy or activity an "idea." The outside option of decision makers is $\pi_0 > 0$. We assume for simplicity that in each stage of exploration decision-makers run an experiment that tests one idea. The focus on one idea per experiment helps to streamline our discussion. The relevant twist of this assumption is that we rule out that decision-makers have the resources to run all their ideas once and in parallel.

We assume that decision-makers enter the exploration process with a new idea different from the outside option. Thus, in our model the entry into the process is given, and the question is how the process unfolds after entry. At the beginning of the first stage of exploration, decision-makers receive a signal on whether it is worth running a costly experiment on this new idea to obtain more information on whether its value is higher than the outside option. Based on this signal they decide whether to run this experiment. If they do, at the end of the experiment they observe the value of the new idea and compare it to the outside option. During the experiment, they also obtain information about new ideas that they could test with a new experiment. If they decide to run the new experiment, they enter a new stage of exploration. Thus, at the end of the experiment in each stage, decision-makers decide whether to *terminate* the project and earn the outside option, *pivot* to a new idea, which means running a new experiment to test a new idea, or *commit* to the idea they have just experimented with.

They terminate if the idea they have just experimented with is less valuable than the outside option and they do not see any signal about new ideas such that the expected value of a new experiment is higher than the outside option. They pivot if the value of this new experiment is higher than the current idea and the outside option. They commit to the current idea if it is more valuable than both the outside option and the expected value of a new experiment.

We assume for simplicity that all the ideas that decision-makers can experiment with in the different stages of exploration can take values $\pi > \pi_0$ or $0 < \pi_0$. Before the experiment they do not observe these realizations. However, the probability that they infer π, the value of the idea, is $p \in (0, 1)$ and depends on the characteristics of decision-makers, such as the fact that they are cautious in evaluating their idea, or other characteristics such as their education, experience or knowledge of the phenomenon they deal with. We assume that $p \sim F(p)$, where F is a cumulative probability distribution of p across decision-makers. The experiment reveals whether the value of the idea is π or 0. Our simplification implies that if decision-makers observe π they commit to the idea they have just experimented with and do not pivot because in a new costly experiment the best they can do is to earn π again. If decision-makers observe 0, either they pivot, if the expected value of the new experiment exceeds the outside option, or they terminate. In the Appendix we develop an extended model in which ideas can take a continuum of values, which allows for pivoting even when the experiment is successful. This extended model does not change our main results.

At the beginning of each stage, decision-makers know that if at the end of the experiment they observe 0, they earn the outside option. Thus, before the experiment, the expected value of running the experiment is $\pi p + \pi_0(1-p) - k > \pi_0$, or

$$(\pi - \pi_0)\, p - k > 0 \tag{2.1}$$

where k is the cost of running the experiment. Solving for p, it is easy to see that the share of decision-makers who satisfy this inequality is $1 - F(x)$ where $x \equiv \left(\frac{k}{\pi - \pi_0}\right)$. Again heterogeneity across decision-makers, such as in terms of education or experience, implies that some decision-makers have lower k or F, for given p. They run experiments more efficiently, and are more likely to satisfy inequality (2.1).

2.2.2 Decision process

The first step of the exploration process is determined by whether decision-makers run the first experiment. The share of decision-makers who run this experiment is $1-F(x)$. Since p and all the parameters of equation (2.1) depend on given characteristics of decision-makers, if this inequality is satisfied in stage 1, it will be satisfied in all future stages of exploration. Moreover, as noted, if at the end of the experiment in stage 1 decision-makers observe π they commit to the current idea they have just experimented with because they do better than the outside option and it is not worth paying k again for an experiment that cannot produce better outcomes than π.

Let Φ be the probability that decision-makers observe that the value of the idea is 0. Note that Φ is different from $1 - p$. The latter is the probability inferred by decision-makers that the idea is worth 0 before running the experiment, while Φ is the actual share of decision-makers who observe 0 after the experiment. In this respect, Φ is the "true" probability of observing 0. Decision-makers pivot, instead of terminating, if they observe 0 and a valuable potential new idea. We make two assumptions that describe the discovery process in our model.

Assumption 1. *Decision-makers search for new ideas in a closed space made of A ideas*

Assumption 2. *If at the end of stage of exploration $s \geq 1$ decision-makers observe* 0*, they:*

- *discover with probability $\theta_s \in [0, 1]$ a new idea that they can test in a new experiment*
- *learn that the value of other $\lambda_s A$ unexplored ideas is also* 0*, with $0 < \lambda_s \leq 1$, where $\lambda > 0$ because at least the idea actually tested in the experiment will be discarded.*

Assumption 3 says that decision-makers search in a space bounded by their experience, knowledge or by their preferences to focus on one activity (e.g. they do not want to operate in some industries or markets).

The first part of Assumption 2 says that at the end of a failed experiment, decision-makers have the chance to see new ideas. The second part of Assumption 2 says that the decision-makers apply logical frameworks that enable them to link the outcome of the experiment focused on a specific idea to other ideas. Specifically, if the experiment indicates that the value of the focal idea is 0, they realize, prior or without making a new experiment, that $\lambda_s A$ ideas not yet assessed in previous experiments will also have value 0. Therefore, they do not need to test them. We call θ_s the discovery rate of ideas and λ_s the exclusion rate.

Assumption 2 is crucial in our analysis. It says that while decision-makers see new ideas, they can use their logic to compare them to ideas they have already tested with negative results, and thus realize that it is not worth paying the cost of testing them in a new experiment. As we will see in the next section, and illustrate with examples, this ability to logically identify ideas that would also be associated with a negative outcome without testing them is an important characteristic of scientific decision-makers.

Thus, at the end of stage 1, if the experiment yields a negative outcome, decision-makers see with probability θ_1 a new idea they can experiment with. However, the failure of the idea they have just tested implies that they rule out $\lambda_1 A$ ideas from their search space of A ideas. Thus, the probability that the new idea is a genuinely new idea worth testing is $\theta_1(1-\lambda_1)$, where $(1-\lambda_1)$ is the probability that the new idea is not one of the ideas that decision-makers exclude *a priori.*

To summarize, at the end of stage 1, $F(x)$ decision-makers have terminated before running the first experiment; $[1-F(x)]\,\Phi\theta_1(1-\lambda_1)$ run the first experiment, observe 0 and pivot to a new idea they have identified when running the first experiment; $[1-F(x)]\,\Phi\,[1-\theta_1(1-\lambda_1)]$ run the first experiment, observe 0 and terminate; and $[1-F(x)]\,(1-\Phi)$ run the first experiment, observe π and therefore commit.

The process continues across stages in the same fashion. Thus, if decision-makers run the experiment in stage 2, at the end of it, either they observe π and commit, or they observe 0 with probability Φ. In this case, they see a new idea they can pivot to with probability θ_2. However, the failed experiment in stage 2 suggests that a share λ_2 of the $(1-\lambda_1)A$ ideas available at the end of stage 1 are also unfeasible. Thus, the probability that the decision-maker pivots is $\theta_2\,[1-\lambda_1-\lambda_2(1-\lambda_1)] = \theta_2(1-\lambda_1)(1-\lambda_2)$, where you now rule out both the ideas discarded at the end of stage 1 and those discarded at the end of stage 2.

If at the end of stage 3 the experiment fails again, decision-makers see a new idea with probability θ_3, and rule out $1-(1-\lambda_1)(1-\lambda_2)-\lambda_3(1-\lambda_1)(1-\lambda_2) = (1-\lambda_1)(1-\lambda_2)(1-\lambda_3)$ ideas. More generally, at the end of stage s the probability of pivoting conditional on reaching this stage and observing a negative outcome in the experiment in stage s is $\theta_s\prod_{j=1}^{s}(1-\lambda_j)$. Overall, the share of decision-makers who terminate at the end of this stage is $[1-F(x)]\,\Phi\left[1-\theta_s\prod_{j=1}^{s}(1-\lambda_j)\right]$, the share of decision-makers who pivot is $[1-F(x)]\,\Phi\theta_s\prod_{j=1}^{s}(1-\lambda_j)$, and those who commit to the idea experimented in stage s is $[1-F(x)]\,(1-\Phi)$

2.2.3 Scientific Decision-Makers

We focus on two characteristics of scientific decision-makers.

First, scientific decision-makers analyze problems using logical frameworks. They develop models of the problem identifying its components and the relationship between them. This helps them to identify logical gaps in their ideas, and more generally in their reasoning about the potential success of their ideas. Also, high quality data and rigorous tests are more likely to deliver signals on the value of the ideas that are more closely associated with their actual value. Both factors push decision-makers to assess problems more objectively, reducing the natural tendency of managers and entrepreneurs to overconfidence (Astebro, 2003; Malmendier and Tate, 2008; Galasso and Simcoe, 2011; Astebro et al., 2014). Second, they better understand problems because of better frameworks that help them to better understand the space of solutions. This has two implications. The first one is that a better assessment of the solutions makes them more efficient in discovering new ideas. The second implications is that the logical links in their frameworks help them to associate, through logic, analogies, or both, the outcomes of their experiments to other potential ideas. Our examples in the next section illustrate these two characteristics of scientific decision-makers. Here we summarize them in the following two assumptions.

Assumption 3. *Given x and p, scientific decision-makers are more likely to show a higher value of the distribution function F*

Assumption 4. *Scientific decision-makers are more likely to exhibit higher θ_s and λ_s, $s \geq 1$*

The higher F implies that, other things being equal, scientific decision-makers are more likely to exhibit a lower probability p because they are more cautious in evaluating the idea. This makes it less likely that they satisfy equation (2.1). The higher θ_s implies that they are more likely to discover new ideas, while the higher λ_s implies that, if they see a negative outcome from an experiment, they identify a higher number of untested ideas with negative outcomes.

The following two propositions capture the predictions that we test with our RCT. First, in early stages, a scientific approach is more likely to lead to terminate projects. Second, scientific decision-makers are less likely to pivot many times.

The intuition of the first proposition is that the reduction of overconfidence makes scientific decision-makers more cautious, leading to a distribution F with greater weight on lower values of p. Assumption ass: overconfidence then implies that they are less likely to satisfy condition equation (2.1), and therefore they are more likely to terminate before running the experiment in stage 1.

However, for ideas that have not been terminated in the early stages, scientific decision makers are likely to show a lower level of termination due to the fact that, following Assumption 6, scientific decision makers are also characterized by a higher probability of finding new ideas θ_1, which may offset the effect of F. It may take a few stages to outweigh the initial higher share of termination of scientific decision-makers. Thus, in early stages scientific decision-makers are more likely to terminate.

The intuition of the second proposition is that, because the explorable space of search is bounded, the higher discovery rate together with the higher exclusion rate implies that scientific decision-makers cover the search space more quickly. Thus, scientific decision-makers commit or terminate earlier. In earlier stages, we cannot say if the probability of pivoting increases or decreases. The initial effect produced by the higher discovery rate θ_1 may increase the probability of pivoting. However, the combined effect of covering the space more quickly, and the higher exclusion rate of ideas, make it unambiguously more likely that, eventually, scientific decision-makers are less likely to pivot.

Proposition 1. *If scientific decision-makers exhibit a higher F, in early stages they are more likely to terminate projects*

Proof. The share of decision-makers who terminate up to the end of stage s is

$$F(x) + [1 - F(x)]\,\Phi \sum_{i=1}^{s} \left[1 - \theta_i \prod_{j=1}^{i} (1 - \lambda_j) \right]$$

Assumption 3 implies that scientific decision-makers exhibit higher F. Therefore, the first term of this expression indicates that they are more likely to terminate before launching the first experiment. Assumption 2 implies that we cannot sign the subsequent terms unambiguously. However, the probability of termination after the first experiment is $[1-F(x)]\,\Phi\cdot\,[1-\theta_1(1-\lambda_1)]$. Using $dF>0$ and $d\,[\theta_1(1-\lambda_1)]$ to denote differences between scientific and non-scientific decision-makers, the overall difference in the probability of termination at the end of the first stage is $\{1-\Phi\,[1-\theta_1(1-\lambda_1)]\}\,dF-[1-F(x)]\,\Phi d\,[\theta_1(1-\lambda_1)]$. The first term of this expression is positive, suggesting that, other things being equal, the overall effect of the higher F still implies that scientific decision-makers exhibit a higher probability of termination at the end of the first stage. We cannot sign unambiguously $d\,[\theta_1(1-\lambda_1)]$, or the other terms of the probability of termination at the end of stage s. However, we cannot rule out that the initial boost will be absorbed gradually, and at least up to some initial stages of exploration scientific decision-makers are more likely to terminate. ■

Proposition 2. *Scientific decision-makers do not pivot many times*

Proof. In the generic stage s, the probability that decision-makers pivot more than s times is $[1-F(x)]\,\Phi\theta_s\prod_{j=1}^{s}(1-\lambda_j)$. Approximate $\theta_s\prod_{j=1}^{s}(1-\lambda_j)$ with a geometric Brownian motion with expected value $\theta e^{-\lambda s}$, where $-\lambda<0$ is the negative drift. Assumption 4 implies that scientific decision-makers exhibit higher θ and λ. Using $d\theta$, $d\lambda>0$ to denote differences for scientific decision-makers this implies $d\theta e^{-\lambda s}=e^{-\lambda s}d\theta-\theta s e^{-\lambda s}d\lambda>0$ for $s<\frac{d\theta}{\theta d\lambda}$. Since $F(x)$ is higher for scientific decision-makers, for s sufficiently large scientific decision-makers are unambiguously less likely to pivot. ■

2.2.4 Examples

We provide two examples of companies from our RCTs consistent with the assumptions about scientific decision-making in the previous section.

Inkdome is a start-up that decided to launch an online search engine to find the right artist for getting tattoos. They developed a clear theory. They expect the service to be viable if four hypotheses are corroborated: 1) consumers use different artists for new tattoos; 2) they search online; 3) the search takes time; 4) they can find online all the information they need to find the ideal artist for the service they seek. Before testing these hypotheses with data from a survey interview, they set the following rule for launching the project: a) all hypotheses have to be corroborated; b) one hypothesis is corroborated if more than 60% of the interviewees respond positively. Using this 60% rule, the experiment corroborated the first three hypotheses, but not the fourth one. Inkdome abandoned the project.

Under similar conditions, many counterfactual entrepreneurs in our RCT did not set such clear hypotheses and tests. They based their assessments on generic discussions and interviews, and unclear rules. The lack of good frameworks implied that they did not set clear rules or consider that their assessment might be affected by over- (or under-) confidence. More generally, at the end of the first RCT in Milan in 2015 we asked the 88 surviving start-ups to score between 1-7 how likely they would be to close a potential new start-up that they found. The median score of the treatment group was 4.4 vs 3.2 of the control group, with a p-value smaller than 1%. Of course, we cannot exclude that the additional level of caution triggered by the treatment might lead entrepreneurs to even become underconfident. While we cannot object to this statement, we rely on the evidence from past literature that suggests that managers and particularly entrepreneurs are typically overconfident.

A good example of our assumption about the implications of good frameworks is represented by Mimoto. Mimoto is a start-up that planned to offer an electric moto-sharing service in Milan. The founders started with a theory that focused on target customers. The logic is that analogous services such sa car- or bike-sharing are likely to meet a wide demand by different types of customers. Motos, instead, are likely to be used by special groups of people. They then theorized that the ideal target customers are young people, with mobility needs, and ability to pay. This prompted them to focus on college students.

However, when they tested their theory in practice, the use of the service was disappointing, both by college students and others. They went then back to their theory. They still thought that the focus was young people with ability to pay, but they had to look for customers with *unpredictable* mobility needs. College students have predictable schedules dictated by the regularity of the timing of their classes. This makes public or private transportation competitive because they can plan mobility needs in advance either by following public transportation schedules or by planning when and where to park depending on the schedule. A moto service is most useful when you suddenly have to grab some means of transportation to reach quickly a relatively far location in the city. Mimoto then saw clearly where to turn for a new target customer. It focused on young professionals, who are also young and need mobility in large cities, but have unpredictable mobility needs.

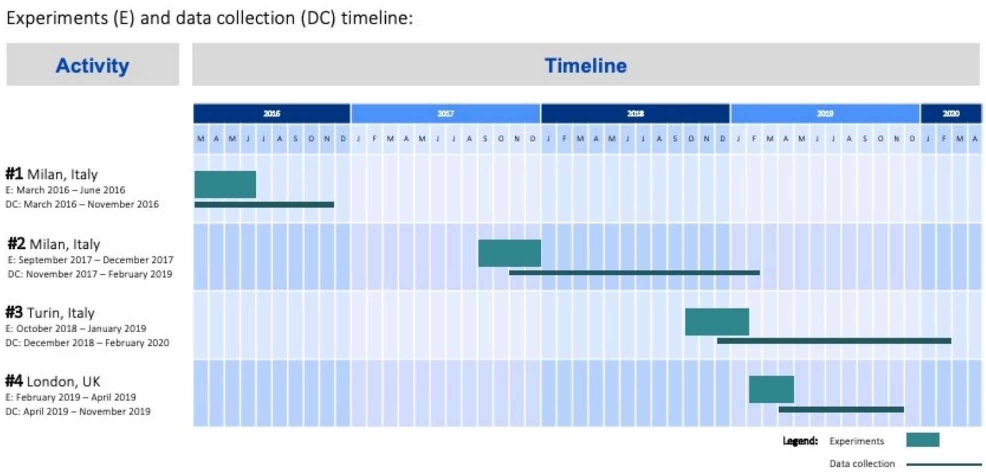
The general counterfactual non-scientific entrepreneurs in our RCTs did not have a clear theory and did not run rigorous experiments. As a result, many of them did not have clear directions about the actions to take after negative signals. On some occasions, they run experiments one after the other, and on other occasions they did not run any experiment but just used the evidence as a basis for internal discussion. Also, during the experiment Mimoto noted that women had a hard time using the sturdy motorcycles that they employed in the experiment. They realized that while this was a problem for women, it was in fact a general problem that could also be faced by other people, while they observed that in Milan many different people employed scooters. Thus, they switched from motos to scooters, ruling out all combinations of target customers and motos. To summarize, thanks to general frameworks, scientific decision-makers know more promptly where to go when they receive feedback information from experiments. This helps them not to wander by trying several ideas without a solid logic.

2.3 Context and Research Design

2.3.1 Experimental Design

We offered four training programs free of charge to entrepreneurs in Milan (Italy), Turin (Italy), and London (UK). We randomly assigned 50% of the participants to a training condition (training using the scientific approach) and the other 50% to a control condition (training without the scientific approach). Entrepreneurs provided data on their decision-making processes and performance before and for several months after the end of the training programs. The structure, type of intervention, and the data collection process of the four experiments were the same, with a view to conduct an internal replication of the same study across samples and regions. Figure 2.1 provides more detail on the timeline of each RCT and of the data collection. For each study, the research

Figure 2.1: Timeline of the Four RCTS



team created new training programs for entrepreneurs. Training programs are effective ways to treat entrepreneurs (Field, Jayachandran and Pande, 2010; Campos et al., 2017; Anderson et al., 2018). The four programs were asynchronous with a gap of at least a few months between each other, given the heavy set-up required for each experiment. In every location, the research team recruited large teams of research assistants that were trained on how to advertise the program, interact with entrepreneurs, and collect data. In addition, we recruited qualified instructors and taught them how to deliver the training material, which was designed by the research team.

2.3.2 Step 1: Recruitment of Participants

In each study, we advertised the program at a national level over multiple online and offline channels, including social media posts, newsletters, magazines and events. One of the reasons why the program was appealing to entrepreneurs is that it was advertised and delivered under the brand of the countries' top business schools. We conducted the advertisement campaign over several weeks to ensure the recruitment of at least 100 entrepreneurs. Regardless of the media used, the campaign promoted the training program generically as a business support program (to avoid self-selection based on interest in a specific topic) offered free of charge to entrepreneurs operating in any industry. To apply to the program, we required entrepreneurs to provide information about their business, team, and decision-making practices via an online survey and a brief telephone interview. We did not admit to the program entrepreneurs who failed to complete the survey or the interview.

Since our study investigated the decisions that entrepreneurs make about their business, our empirical design required that the subjects receiving the treatment were the key decision-makers in the firm. We were more likely to meet this condition in micro-enterprises (with less then 10 employees), where owners are highly involved in the management of the firm. In the field experiment in London we admitted in the program only firms that met this condition. In the three field experiments conducted in Italy entrepreneurs met this condition naturally because all firms were at a very initial stage and had yet to start offering their service or product to the market.

The recruitment campaigns attracted entrepreneurs from different Italian regions (in the first three experiments) and from different parts of England (in the fourth experiment). The final sample included 754 entrepreneurial firms.

2.3.3 Step 2: Intervention Details

We assigned entrepreneurs in each experiment to either a treatment or a control group through simple randomization. We also broke down the treatment and control groups into smaller groups, and randomly assigned each subgroup to an experienced instructor. We administered a baseline survey to all entrepreneurs prior to the intervention, and we used the information in the survey to check that the observable characteristics were balanced across the treatment and the control groups using balance tests. The Online Appendix reports the results of the balance tests across groups for each of the four RCTs and for the variables that were common across all RCTs. As the four RCTs were conducted asynchronously, the research team had the opportunity over time to introduce additional relevant dimensions to the baseline survey. As a result, the list of observables is larger for later RCTs.

Treatment and control groups attended the same number of sessions, covering the same topics related to strategy and entrepreneurship. The sessions were highly experiential and smaller groups ensured that instructors provided feedback to each participant. About 80% of the content in the two classes was the same in terms of topics delivered and teaching material. Specifically, both treated and control groups were taught frameworks that they could use to support decision making, such as the business model canvas or the balance scorecard; both groups were also exposed to evidence gathering techniques such as qualitative interviews, surveys and A/B testing. Both groups were taught to apply these frameworks and techniques to their specific contexts and were given feedback from their peers and instructor. However, the treatment group was taught to apply the frameworks and techniques used to make decisions using a scientific approach, which is by developing a theory of the problem and hypotheses that flew logically from it, by testing those hypotheses and eventually by evaluating the results of the test in comparison with the theory originally developed. The control group, instead, was free to apply these frameworks and techniques in the way they found more appropriate.

An example can help clarify the difference between the two groups. One of the first sessions of the training program focused on the Business Model Canvas (henceforth, BMC), a tool widely used in entrepreneurial education that concisely and visually represents a company's business model. It is composed of nine elements that describe a firm's customer value proposition, customer segments, channels, customer relationships, revenue streams, key resources, key partners, key activities, and cost structure. The control group was exposed to the basic content of the BMC and was taught to use this tool to provide a general overview of their business and discuss its implications. Entrepreneurs in the control group were then encouraged to apply the framework to their own business and given a dedicated time slot to discuss about it with their peers. Entrepreneurs were then encouraged to present the application in front of the classroom and received general feedback from the instructor and peers. This is the typical way in which BMC is taught in MBAs and Executive programs. Similarly, the treatment group was exposed to the basic content of the BMC, asked to apply it to their business and discuss it with their peers. But differently from the control group, the treatment group was also explicitly invited to develop a theory that emerged from the application of the BMC to their business and develop explicit hypotheses. For example, imagine an entrepreneur that, when filling in their BMC, indicated that they were running an electronics retail business using an online distribution channel. If this entrepreneur was part of the control group, they would be invited to generally discuss about the motivation behind this choice, its alignment with other choices made by the company, and would be given feedback on its suitability. The same entrepreneur in the treatment group, instead, would be explicitly asked to formulate the hypotheses underlying this choice, which, if supported, would make this choice a valuable one. For example, one such hypothesis might be "the majority of my target customers in the city where I am located buys electronics online".

In subsequent sessions, entrepreneurs in both groups were taught techniques to collect data in support of their decisions. For instance, they were taught about qualitative interviews, surveys, and experiments, and the strengths and weaknesses of each of these methodologies. Entrepreneurs in both groups were then invited to think about which techniques they could use in their businesses and discuss with their peers and instructor about one specific implementation in their context. The control group was let free to choose the context or problem to which they applied those techniques and was given general feedback on the way in which the technique was applied. The treatment group, instead, was explicitly invited to use these techniques to test the hypotheses formulated in the previous sessions and was given specific feedback on whether the proposed design was consistent with the hypothesis that they wanted to test. Of course, entrepreneurs in both groups were given genuine and valuable feedback. For example, if an entrepreneur proposed to administer a survey to a very small sample of target customers, the instructor would recommend them to increase their sample size irrespective of whether they belonged to the treatment or the control group; or if an entrepreneur formulated a survey question in a way that could be improved in terms of clarity, they would be offered suggestions regarding how to improve it, irrespective of whether they were in the treatment or in the control group.

It is important to note that for every RCT, each instructor was teaching both a treatment and a control group at different times of the day or different days of the week. This choice allowed to include instructor-fixed effects in our analyses and control for the different teaching styles of instructors and ultimately affect the absorption of the content taught to participants. Although instructors were not blind to the treatment, we directly supervised the delivery of each session to ensure high teaching standards and that the content was in line with the experimental design described above.

To prevent participants from meeting and potentially discussing key elements of the treatment, we offered training sessions on different days of the week or on the same day of the week to both groups, but at different times of the day. To further prevent contamination, the research team kept all communication to the two groups of entrepreneurs attending the program discrete and separate. The research team also checked if applicants to the program knew other applicants and made sure to allocate all of those that knew each other to the same experimental group.

2.3.4 Step 3: Data Collection

We systematically collected data on all participants through telephone interviews conducted by a team of research assistants over the span of several months. We hired research assistants for the purpose of these experiments and the research team trained them extensively. Research assistants were undergraduate or graduate students that were selected on the basis of their academic performance, basic knowledge of the entrepreneurial process, communication and analytical skills. The research team interviewed research assistants, and tested their communication and analytical skills through various activities (analysis of a business case, interviewing an entrepreneur and coding responses according to a simple, predefined coding scheme), to ensure they would be able to perform the tasks required by the project.

Research assistants performed regular phone calls that followed a predefined script that included open and closed-ended questions focusing on changes in the business model, decision-making, and performance outcomes. In all RCTs but the first one, we recorded telephone interviews and subjected them to random checks to ensure that research assistants were conducting calls in accordance with the guidelines provided by the research team. The main variables used in this study refer to outcomes such as termination, pivot and amount of revenue and were therefore collected through closed-ended questions. Following an approach similar to the one used by Bloom et al. (2012), we also included a number of open-ended questions that elicited — without asking leading questions — what type of approach to decision-making entrepreneurs were using. Specifically, we instructed research assistants to code for the occurrence and the extent to which entrepreneurs employed themes related to theory, hypotheses, tests and evaluation. We use these data in supplementary analyses and provide more detail about this in Section 7 of the Online Appendix.

The data collection process continued for up to 14 months after the training program ended. In one of the RCTs (London), we could only collect observations for 7 months after the training program due to funding constraints. We take into account the duration of the data collection process in discussing our results.

2.4 Results

2.4.1 A Glimpse at the Data

Tables 2.1 and 2.2 report the descriptive statistics for the variables collected after the intervention through telephone interviews and refer, respectively, to the cross-sectional and longitudinal samples.

Firms in our sample pivot at most six times during the observation window. Fifty nine percent of the sample never pivots, and six percent of the firms pivot more than two times. To measure pivoting, we referred to the BMC taught to entrepreneurs during the training program. During each of the interviews, we asked entrepreneurs to report and describe any changes made to any of the nine dimensions of the BMC (value proposition, customers, channels, customer relationships, key activities, key partners, key resources, revenue streams, cost structure). We classified a firm as having pivoted at time t if they reported a major change to their value proposition or customer segment, two key dimensions of their business. Thirty four percent of the firms terminate their projects within the observation window. The average amount of revenue is EUR 15,538, with large variation in the sample since a substantial number of firms has zero revenue within the observation window. The number of entrepreneurs participating to each RCT changes mostly in light of budget constraints and venue capacity of each study. In all RCTs, half of the sample is assigned to the treatment condition. In all of our analyses we will report the results obtained on the full sample of 754 firms. However, in the Online Appendix, we report the results obtained for each RCT separately, for each of the models we present in this paper.

Figure 2.2 (a) and 2.2 (b) provide a visual representation of our data. The evidence is consistent with our predictions. In Figure 2.2 (a) the number of treated firms that terminate the project within the observation window is higher than the number of control firms. Moreover, as we will show with our regressions, they are more likely to terminate earlier, as predicted by Proposition 1. Also, treated firms are more likely to pivot once, while control firms are more likely to pivot more times. This result, which we confirm below with our regressions, is interesting in light of our model. Our model predicts that a higher share of pivoting in early stages depends on the fact that the ability of scientific decision-makers to learn how to run better experiments from initial observations outweigh the more conservative effect of precision that encourages termination. Figure 2.2 (b)

Table A1: Variables and Descriptive Statistics - Cross Section Sample

Variable	Description	Obs	Mean	SD	Min	Max
Number of Pivots	Number of times the firm pivoted within the observation period	754	0.68	1.05	0	6
Pivot=0	Dummy equal to 1 if the firm did not pivot within the observation window; 0 otherwise	754	0.59	0.49	0	1
Pivot=1	Dummy equal to 1 if the firm pivoted once within the observation window; 0 otherwise	754	0.24	0.43	0	1
Pivot=1-2	Dummy equal to 1 if the firm pivoted once or twice within the observation window; 0 otherwise	754	0.35	0.48	0	1
Pivot=2	Dummy equal to 1 if the firm pivoted twice within the observation window; 0 otherwise	754	0.11	0.31	0	1
Pivot=2+	Dummy equal to 1 if the firm pivoted more than twice within the observation window; 0 otherwise	754	0.06	0.24	0	1
Termination	Dummy equal to 1 if the firm terminated the project within the observation window; 0 otherwise	754	0.34	0.48	0	1
Revenue	Firm's cumulative revenue in EURO	754	15538	83240.95	0	1489026
RCT1	Milan 1	754	0.15	0.36	0	1
RCT2	Milan 2	754	0.33	0.47	0	1
RCT3	Turin	754	0.17	0.37	0	1
RCT4	London	754	0.35	0.48	0	1
Intervention	Dummy equal to 1 if the firm was treated; 0 otherwise	754	0.5	0.5	0	1
Average Scientific Intensity	Score reflecting the extent to which the firm's decision making process follows a scientific approach	754	2.23	1.21	0	5
Postgraduate	Dummy equal to 1 if team average on a score that reflected the highest level of education of each member of the team reported by the entrepreneur was higher than 3, where the educational level attained by team members is coded as follow: 5= PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise	754	0.17	0.38	0	1
Experience: Industry	Dummy equal to 1 if the team average experience in the focal industry is higher than 5 years; 0 otherwise	754	0.28	0.45	0	1
Experience: Managerial	Dummy equal to 1 if the team average managerial experience is higher than 5 years	754	0.23	0.42	0	1
Mature	Dummy equal to 1 if the team average age is higher than 30 years. NA for RCT1	638	0.56	0.5	0	1

Table A2: Variables and Descriptive Statistics - Longitudinal Sample

Variable	Description	Obs	Mean	SD	Min	Max
Termination	Dummy equal to 1 if the firm terminated the project within the observation window; 0 otherwise	8508	0.03	0.17	0	1
Week of Termination	Week at which the firm terminated the project	754	41.72	16.89	6	66
Revenue (Flow)	Firm's revenue (flow) in EURO	10730	1098.48	8581.3	0	231000
Intervention	Dummy equal to 1 if the firm was treated; 0 otherwise	10730	0.5	0.5	0	1
Scientific Intensity	Score reflecting the extent to which the firm's decision making process follows a scientific approach	10730	2.24	1.26	0	5

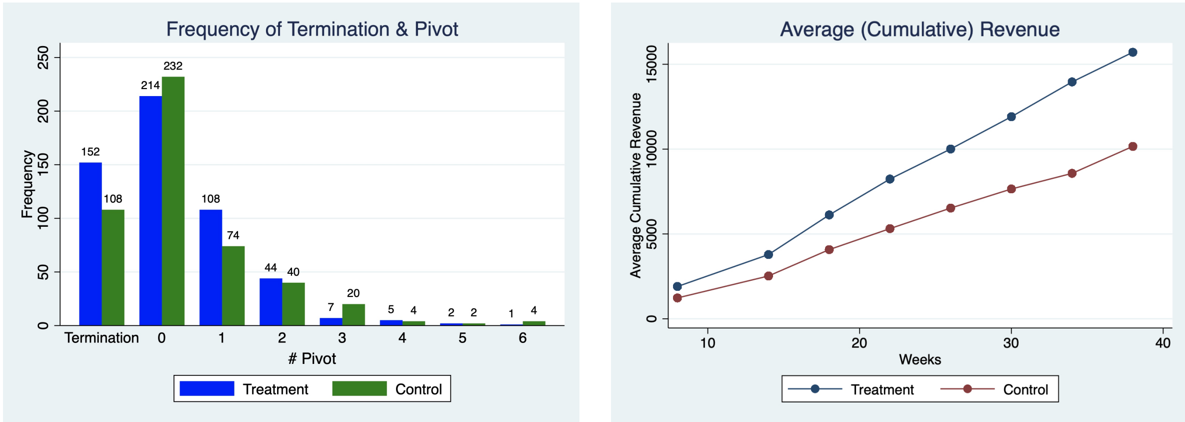
shows that, on average, the revenue of treated firms grows faster than that of control firms.

2.4.2 Termination

We start our analysis by examining the impact of the intervention on termination. Table 2.3 reports the results of our analyses. Column (1) reports the results of a cross sec-

Figure 2.2: Termination, Pivot and Performance

(a) Termination and Pivot (b) Performance (euros)



...

tion linear probability model that shows that the intervention raises the probability of termination by 10.4 percentage points (p=0.001). As a robustness check we also run a probit regression and we report the results in Column (2). The marginal effect calculated at the observed values suggests that treated firms report a probability to terminate that is 10.4 percentage points higher than that of control firms(p=0.000). As a further robustness check, Column (3) reports the results of a longitudinal analysis that includes time fixed effects, along with mentor and RCT dummies, and standard error clustered at the intervention-mentor-RCT level. Results show that that the intervention increases the probability of terminating the project by 2.4 percentage points at any moment in time (p=0.000). Finally, Column (4) reports the results of a longitudinal analysis from a probit model. Based on these results, the marginal effect of intervention calculated at the observed values is of 1 percentage point at any moment in time and is statistically significant (p=0.001).

Table 2.4 reports the results of a Cox proportional hazard model in Column (1). We corroborated the proportionality assumption using the Schoenfeld residuals. We find that the hazard rate of termination is higher for treated firms than for control firms. In Column (2) we replicate this analysis using a regression estimated with OLS to predict the week of termination. We find that, on average, treated firms terminate their project about 2.3 weeks earlier than control firms (p=0.012). Overall, we find that scientific decision-makers are more likely to terminate their projects earlier.

In the Online Appendix, we compare these results (Column (1) in each of the Tables from A6 to A11) with the results of obtained for each individual RCT (Columns 2-5).

Table A3: Termination

VARIABLES	(1) Termination OLS Cross-Section Full Sample	(2) Termination Probit Cross-Section Full Sample	(3) Termination OLS Panel Full Sample	(4) Termination Probit Panel Full Sample
Intervention	0.104***	0.299***	0.024***	0.164***
	(0.001)	(0.000)	(0.000)	(0.000)
Constant	0.283***	-5.038***	0.027	-5.639***
	(0.000)	(0.000)	(0.211)	(0.000)
Observations	754	754	8,508	8,508
R-squared	0.078			
Dummies for mentors and RCT	Yes	Yes	Yes	Yes
Time FE	-	-	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT
Number of id			754	

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization. Specifically, Model (3) controls for Percentage STEM and Percentage Economics, Model (4) controls for Self regulation, Model (1) controls for the interaction between each RCT dummies and variable that was unbalanced in that specific RCT.

Table A4: Termination Time

VARIABLES	(1) Hazard of termination Survival Full Sample	(2) Week of termination OLS Full Sample
Intervention	0.375***	-2.322**
	(0.000)	(0.012)
Constant		32.446***
		(0.000)
Observations	754	754
R-squared		0.242
Dummies for mentors and RCT	Yes	Yes
Time FE	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.4.3 Pivot

Table 2.5 reports the results of our analysis on the number of pivots. Column (1) reports the results of a cross section regression, estimated with OLS where the dependent variable is the number of pivots made by the firms within the observation window. The regression includes dummies for mentors and RCTs and we clustered standard errors at the intervention-mentor-RCT level and reports a not significant effect of the intervention. In Column (2) we report the results of a linear probability regression in which the dependent variable is a dummy equal to 1 if the firm has pivoted only once within the observation window and equal to 0 otherwise (i.e., if it has not pivoted or has pivoted more than once). By distinguishing between the choice to pivot once or many times, the results in Column (2) show that the intervention raises the probability of pivoting only once by 8.7 percentage points (p-value = 0.001) vis-à-vis no pivot or more than one pivot. As noted earlier, this is an interesting result. Our model suggests that this effect depends on the fact that scientific decision-makers' learning from observing the results of the experiments enable them to outweigh the conservative effect of their greater precision.

Table A5: Number of Pivots

VARIABLES	(1) # Pivots OLS Cross-Section Full Sample	(2) Pivoting once OLS Cross-Section Full Sample
Intervention	-0.032	0.087***
	(0.654)	(0.001)
Constant	0.432***	0.083***
	(0.000)	(0.007)
Observations	754	754
R-squared	0.120	0.082
Dummies for mentors and RCT	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Results from a multinomial probit specification, reported in Table 2.6, confirm this finding. Columns (1), (2) and (3) refer, respectively, to the probability that a firm pivots once, twice, or more than twice vis-à-vis the no-pivot baseline. In Figure 2.3 we show the marginal effects of intervention calculated at the observed values for the entire sample. The intervention raises the probability of pivoting once or twice and lowers the probability of not pivoting or pivoting more than twice. Specifically, when we look at the full sample, the intervention decreases the probability of not pivoting by 5.7 percentage points (although this result is not significant at the conventional level, with p=0.120), increases the probability of pivoting once by 8.6 percentage points (p=0.002), increases the probability of pivoting twice by 0.9 percentage points (although this result is not significant at the conventional level, with p=0.672), and decreases the probability of pivoting more than twice by 3.7 percentage points (p=0.005).

Also in this case, we report in the Online Appendix, these results in comparison with the results of obtained for each individual RCT (Tables A12-A17).

Table A6: Pivot Multinomial Probit

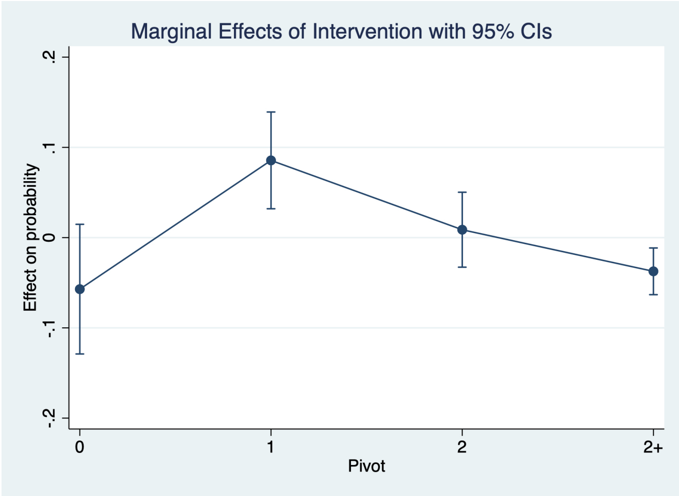
VARIABLES	(1) Pivoting only once Multinomial Probit Cross-Section Full Sample	(2) Pivoting twice Multinomial Probit Cross-Section Full Sample	(3) Pivoting more than twice Multinomial Probit Cross-Section Full Sample
Intervention	0.370***	0.148	-0.287
	(0.010)	(0.397)	(0.117)
Constant	-1.374***	-2.104***	-2.438***
	(0.000)	(0.000)	(0.000)
Observations	754	754	754
Dummies for mentors and RCT	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.4.4 Performance

Our model predicts that in early stages treated firms are more likely to terminate their projects and they do not pivot many times. However, it does not predict whether this yields higher performance because performance depends on whether the conjectures they make are correct. In particular, it may be that control firms linger on projects that

Figure 2.3: Marginal Effects of Intervention on Pivot



should not be terminated, or it might be that by not pivoting, or pivoting many times, they enjoy better outcomes. We therefore explore this question empirically.

In Table 2.7, Column (1) reports the results of our analysis of the impact of the intervention on the cumulative revenue of firms in our sample (in EUR) at the date of our last observation in each trial, estimated by OLS. This represents a different time period for each one of our trials. However, our RCT dummies control for these differences, on average. We also employ mentor dummies and cluster the standard errors at the intervention-mentor-RCT level. Results show that on average treated firms earn EUR 6,504.108 more than control firms (p=0.046). The small effect size reflects the fact that many firms in our first three RCTs earn no revenue as they are start-ups that started with our training program. Some of the firms started earning revenues of the order of dozen thousand EUR, very much in line with the share and extent of revenues earned by start-ups in Italy in their first few months of operation. The increase in revenue between the time of the first interview and the last ranges from 0 to EUR 1,320,396, with revenues increasing of EUR 23,100 at the 90th percentile. This result is further confirmed in Column (2), where we report the results of a longitudinal analysis where the dependent variable is the revenue flow of the firm in each period. On average the revenue of treated firms is EUR 677.342 higher than control firms (p=0.075).

In Tables A18-A19 of the Online Appendix, we report these results in comparison with the results of obtained for each individual RCT.

Table A7: Performance OLS

VARIABLES	(1) Revenue OLS Cross-section Full Sample	(2) Revenue (Flow) OLS Panel Full Sample
Intervention	6,504.108**	677.342*
	(0.046)	(0.075)
Constant	9,039.968***	820.151*
	(0.006)	(0.086)
Observations	754	10,730
R-squared	0.086	
Dummies for mentors and RCT	Yes	Yes
Time FE	-	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT
Number of id		754

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.4.5 Instrumenting Scientific Intensity

Our analyses so far provided estimates of the intent-to-treat effect. This is interesting from a policy perspective as it provides an estimate of the effect of the treatment that takes into account that not all individuals targeted by an intervention are necessarily compliers (Gelman, Hill and Vehtari, 2020). However, to maximize what we can learn from the intervention, we asked our research assistants who were conducting regular phone interviews with the entrepreneurs to use a predefined interview protocol (based on 16 items) to assess the level of scientific intensity used by entrepreneurs in making decisions. This protocol led to the determination of a score (on a scale from 0 to 5) that measured the level of scientific intensity of each entrepreneur at each observation point. We use this score to conduct an additional set of analyses using the intervention as an instrument for the level of scientific intensity exhibited by decision makers. This enables us to provide a more precise estimate of the complier average casual effect. In the Online Appendix, in Section 7, we provide more details on how this score was created. The analyses in Tables A21 and A22 of the Online Appendix, show that treated firms demonstrated higher levels of scientific intensity than firms in the control group. Table 2.8 presents the results of a cross-section specification estimated using two-stage least squares. Results on termination, reported in Column (1), show that the increase of one unit in the average scientific intensity increases the probability of terminating of 29.9 percentage points (0.001). In Column (2), we report the results of our analysis on pivot, which show that the increase of one unit in the average scientific intensity increases the probability of pivoting once (versus 0 or more than once) of 25 percentage points (p=0.000). Looking at the effect on performance, results in Column (3) show that an increase of one unit in the average scientific intensity is associated with an increase of EUR 18,703.974 (p=0.056). As a robustness check, we replicate our analyses on termination and pivot using a IV probit specification and using a two-stage least square approach on the longitudinal sample. We report results in the Online Appendix (Tables A23-A29).

Table A8: Instrumenting Scientific Intensity

VARIABLES	(1) Termination 2SLS Cross-Section Full Sample	(2) Pivoting once 2SLS Cross-Section Full Sample	(3) Revenue 2SLS Cross-Section Full Sample
Average Scientific Intensity	0.299***	0.251***	18,703.974*
	(0.001)	(0.000)	(0.056)
Constant	-0.283	-0.391**	-26,351.771
	(0.125)	(0.016)	(0.208)
Observations	754	754	754
R-squared	-0.485	-0.225	0.064
Dummies for mentors and RCT	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT
Time FE	-	-	-

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.4.6 Heterogeneous Treatment Effects

In addition to testing whether our treatment has a main effect on the outcome variables of interest, we are interested in exploring the presence of heterogeneous treatment effects, that is, whether certain groups react differently to the treatment. We are particularly interested in three dimensions that the literature suggests support entrepreneurial decision making: education, work experience (industry and managerial), and age. We measure these constructs at the level of the entrepreneurial team, since our focus on micro-businesses implies that each member of the team is likely to have played a relevant role in decision making. Table 2.9 reports the results of a regression analysis, estimated with OLS, where we regress the key outcome variables in this study (termination, pivot and performance) against Postgraduate (a dummy variable equal to 1 for firms where the average level of highest education degree attained by the member of the team of the entrepreneur was at least the Master degree), the interaction between Intervention and Postgraduate, and the interaction between Intervention and its complement (Non Postgraduate). For entrepreneurs with an average education lower than the postgraduate level, the intervention increases the probability of termination of 9.8 percentage points (p=0.001) and the probability of pivoting once of 10.1 percentage points (p=0.000); it increases revenue of EUR 7,645.550 (p=0.091). For entrepreneurs with a postgraduate degree the intervention does not have a significant impact on the dependent variables.

Table A9: Education OLS Cross-Section

VARIABLES	(1) Termination OLS Cross-Section Full Sample	(2) Pivoting once OLS Cross-Section Full Sample	(3) Revenue OLS Cross-section Full Sample
Postgraduate	0.019	0.040	547.530
	(0.809)	(0.482)	(0.958)
Intervention X Postgraduate	0.141	-0.036	-2,986.407
	(0.197)	(0.679)	(0.796)
Intervention X Non Postgraduate	0.098***	0.101***	7,645.550*
	(0.001)	(0.000)	(0.091)
Constant	0.275***	0.087***	9,906.638***
	(0.000)	(0.000)	(0.007)
Observations	754	754	754
R-squared	0.080	0.085	0.087
Dummies for mentors and RCT	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table 2.10 reports the results of a regression analysis, estimated with OLS, where we regress the key outcome variables in this study against High Industry Experience (a dummy variable equal to 1 for firms where the team of the entrepreneur had on average more than five years of industry experience, corresponding to the 75 percentile of the distribution of industry experience), the interaction between Intervention and High Industry Experience, and with its complement. For entrepreneurs with low industry experience, the intervention increases the probability of termination by 12.3 percentage points (p=0.002) and the probability of pivoting once by 10.8 percentage points (p=0.001). For highly experienced teams the intervention has a non significant effect on the dependent variables.

In Table 2.11 we perform a similar analysis but we look at the level of managerial experience, using four years as the threshold to distinguish between higher and lower experience (corresponding to the 75 percentile of the distribution of managerial experience). Results show that the treatment increases the probability of termination by 12.5 percentage points (p=0.001) for firms with lower experience, while it does not have a statistically significant effect on termination for firms with high managerial experience. Instead, the treatment increases the probability of pivoting once by 6.4 percentage points (p=0.040) for firms with lower experience, while it increases the same probability by 14.3 percentage points (p=0.003) for firms with high managerial experience. The treatment also increases revenue by EUR 8,070.775 (p=0.038) for firms with low managerial experience, but it does not have a statistically significant effect for more experienced firms.

Table A10: Industry Experience OLS Cross-Section

VARIABLES	(1) Termination OLS Cross-Section Full Sample	(2) Pivoting once OLS Cross-Section Full Sample	(3) Revenue OLS Cross-section Full Sample
High Industry Experience	-0.011	0.034	12,559.028*
	(0.835)	(0.512)	(0.066)
Intervention X High Industry Experience	0.053	0.039	21,955.047
	(0.382)	(0.430)	(0.302)
Intervention X Low Industry Experience	0.123***	0.108***	647.409
	(0.002)	(0.001)	(0.906)
Constant	0.299***	0.075**	-204.030
	(0.000)	(0.029)	(0.980)
Observations	754	754	754
R-squared	0.080	0.083	0.103
Dummies for mentors and RCT	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A11: Managerial Experience OLS Cross-Section

VARIABLES	(1) Termination OLS Cross-Section Full Sample	(2) Pivoting once OLS Cross-Section Full Sample	(3) Revenue OLS Cross-section Full Sample
High Managerial Experience	-0.017	-0.031	5,462.768
	(0.704)	(0.516)	(0.357)
Intervention x High Managerial Experience	0.058	0.143***	3,046.304
	(0.368)	(0.003)	(0.620)
Intervention x Low Managerial Experience	0.125***	0.064**	8,070.775**
	(0.001)	(0.040)	(0.038)
Constant	0.294***	0.090**	8,004.405*
	(0.000)	(0.018)	(0.089)
Observations	754	754	754
R-squared	0.081	0.084	0.091
Dummies for mentors and RCT	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Finally, in Table 2.12, we report the results of a regression analysis, estimated with OLS, where we regress the key outcome variables in this study against Mature (a dummy variable equal to 1 for firms where the team of the entrepreneur was on average older than 30 years old), the interaction between intervention and Mature and with complement (Younger). We conduct this analysis using only the data of RCT2, RCT3 and RCT4 because we did not collect the variable age in RCT1. For younger entrepreneurs, the treatment increases the probability of termination by 14.2 percentage points (p=0.000) and of pivoting once by 8.4 percentage points (p=0.015). It increases revenue by 12,579.958 (p=0.094). For more mature entrepreneurs, the intervention increases the probability of pivoting once by 14.6 percentage points (p=0.006), but it does not have a statistically significant effect on termination and revenue.

Table A12: Age OLS Cross-Section

VARIABLES	(1) Termination OLS Cross-Section Full Sample	(2) Pivoting once OLS Cross-Section Full Sample	(3) Revenue OLS Cross-section Full Sample
Mature	0.006	-0.003	10,459.473
	(0.927)	(0.955)	(0.321)
Intervention X Mature	0.025	0.146***	-11,760.632
	(0.726)	(0.006)	(0.110)
Intervention X Younger	0.142***	0.084**	12,579.958*
	(0.000)	(0.015)	(0.094)
Constant	0.283***	0.075*	7,134.708
	(0.000)	(0.052)	(0.119)
Observations	638	638	638
R-squared	0.077	0.075	0.083
Dummies for mentors and RCT	Yes	Yes	Yes
Time FE	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Overall, these results suggest that the impact of a scientific approach on termination, pivoting and revenue is consistently positive and statistically significant for relatively less educated, experienced and mature entrepreneurs. The effect of the intervention for relatively more educated, experienced and mature entrepreneurs is often not statistically significant. We interpret this result to suggest that for these entrepreneurs the intervention has an impact that shows a higher degree of variation. This is an important result that also provides insights for the potential welfare effects of our intervention. We believe that this result does not only speak about the potential of our training, but of the scientific approach as a rigorous tool for decision-making under uncertainty in the contexts that we studied. Given the cost of experience, this suggests that training managers and entrepreneurs to adopt the scientific approach that scientists use in their research, can reach high benefit/cost ratios.

2.5 Conclusion

Entrepreneurs and innovators make choices under conditions of uncertainty. We examine the implications of a scientific approach to decision-making in these cases. Our study underscores the importance of teaching managers and entrepreneurs more than basic business skills (such as accounting or marketing) or soft skills. Our empirical results and model emphasize the importance of teaching them to develop frameworks about what they do and the decisions they have to make, and to test the implications of these frameworks through experiments. In addition to contributing to research on decision-making, we believe that this paper contributes to a larger debate on the design of field experiments. First, it shows that most of our results are supported robustly across contexts, providing confidence in the impact of the intervention. Second, it outlines the limitations of conducting interventions that focus on a small scale and a limited set of contexts. Comparing the results of three of our experiments to Camuffo et al (2020), which focused on the first trial, we observe relevant differences. First, in Camuffo et al. (2020) we do not find a statistically significant effect of the intervention on termination in most of the regressions, while we find a positive and significant effect across all other three RCTs, as also shown by Tables A6-A11 in the online Appendix of this paper. The limited size of the sample in Camuffo et al. (2020) did not produce a sufficient number of terminations to detect this effect. Second, Camuffo et al. (2020) showed a positive effect of the intervention on the number of pivots, which again depends on the fact that in this smaller sample only a few firms pivoted more than once. In the larger sample of this paper we observe more firms that pivot more than once, and discover that the intervention makes pivoting more focused: decision-makers who adopt a scientific approach do not pivot many times. Third, the larger sample size made the results about performance statistically more robust, it enabled us to test the effect of an index of the scientific-intensity of decision-makers instrumented by the intervention, and to study heterogeneity effects. Finally, as shown by the results on the individual RCTs in the Appendix, in all RCTs the estimates of the intervention have the same sign and comparable magnitudes, while of course the statistical significance varies. This suggests that, taken individually, the problem with each trials is statistical significance and precision, not model specification. Our findings about the implications of a scientific approach to managerial decision-making also has practical implications. Research in economics and management has generated many theories and models that prescribe concrete managerial actions. While we teach these theories and models in academic programs, managers and entrepreneurs rarely use them to make decisions, and prefer to rely on their intuitions, experience, gut feelings, or their own logic. This is a serious gap that makes academic research in economics and management less relevant than it could be. While academics may not make their best effort to make their research relevant, the lack of "demand for theory" by decision-makers is also likely to contribute to this gap. Our prediction is that if we nurture a culture of scientific decision-making in firms, the value and the use of theories from academic research in economics and management will also increase. In addition, there have been considerable investments from private and public institutions to create accelerators and programs that teach and support entrepreneurship in contexts ranging from high-tech to basic entrepreneurial activities. A scientific approach, properly adapted to each audience, can improve performance throughout this entire range of entrepreneurial activities, as suggested by the fact that our results are not context- or industry-specific. One of the limitations of this study is that it only covers a limited time period. It would be interesting to explore the treatment effects are in the medium or long term. More in general, we need more research to better understand the implications of a scientific approach to decision-making and how it can, in detail, create opportunities for better innovation and entrepreneurial decisions, and how different types of firms or individuals can take advantage of these opportunities. We hope that future research can shed light on these important micro-foundations of economic performance.

References

Akcigit, Ufuk, and William R. Kerr. 2008. "Growth Through Heterogeneous Innovations." *Journal of Political Economy* 126 (4): 1374-1443.

Allcott, Hunt. "Site Selection Bias in Program Evaluation." 2015. *The Quarterly Journal of Economics* 130 (3): 1117-1165.

Anderson, David R., Dennis J. Sweeney, Thomas A. Williams, Jeffrey D. Camm, and James J. Cochran, 2018. *An Introduction to Management Science: Quantitative Approach.* Cengage learning.

Angrist, Joshua D., and Jörn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con Out of Econometrics." *Journal of Economic Perspectives* 24 (2): 3-30.

Astebro, Thomas. 2003. "The Return to Independent Invention: Evidence of Risk Seeking, Extreme Optimism or Skewness-Loving?" *Economic Journal* 113 (484): 226–39.

Astebro, Thomas, Herz, Holger, Nanda, Ramana, and Weber Robert A. 2014. "Seeking the Roots of Entrepreneurship: Insights from Behavioral Economics." *Journal of Economic Perspectives* 28 (3): 49–70

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries." *Science* 348 (6236).

Bertrand, Marianne, and Antoinette Schoar. 2003. "Managing with Style: The Effect of Managers on Firm Policies." *The Quarterly Journal of Economics* 118 (4): 1169-1208.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *The Quarterly Journal of Economics* 122 (4): 1351-1408.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *The Quarterly Journal of Economics* 128 (1): 1-51.

Bowers, Jake, Nathaniel Higgins, Dean Karlan, Sarah Tulman, and Jonathan Zinman. 2017. "Challenges to Replication and Iteration in Field Experiments: Evidence from Two Direct Mail Shots." *American Economic Review* 107 (5): 462-65.

Camerer, Colin, and Dan Lovallo. 1999. "Overconfidence and excess entry: An experimental approach." *American Economic Review* 89 (1): 306-318.

Campos, Francisco, Michael Frese, Markus Goldstein, Leonardo Iacovone, Hillary C. Johnson, David McKenzie, and Mona Mensmann. 2017. "Teaching Personal Initiative Beats Traditional Training in Boosting Small Business in West Africa." *Science* 357 (6357): 1287-1290.

Camuffo, Arnaldo, Alessandro Cordova, Alfonso Gambardella, and Chiara Spin. 2020. "A

Scientific Approach to Entrepreneurial Decision-Making: Evidence from a Randomized Control Trial." *Management Science* 66 (2), 564-586.

Csaszar, Felipe A., and Daniel A. Levinthal. 2016. "Mental Representation and the Discovery of New Strategies." *Strategic Management Journal* 37 (10): 2031-2049.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda. 2014. "The Role of Entrepreneurship in US Job Creation and Economic Dynamism." *Journal of Economic Perspectives* 28 (3): 3-24.

Eisenhardt, Kathleen M., and Christopher B. Bingham. 2017. "Superior Strategy in Entrepreneurial Settings: Thinking, Doing, and the Logic of Opportunity." *Strategy Science* 2 (4): 246-257.

Ewens, Michael, Ramana Nanda, and Matthew Rhodes-Kropf. 2018. "Cost of Experimentation and the Evolution of Venture Capital." *Journal of Financial Economics* 128 (3): 422-442.

Fairlie, Robert W., and Javier Miranda. 2017. "Taking the Leap: The Determinants of Entrepreneurs Hiring their First Employee." *Journal of Economics and Management Strategy* 26 (1): 3-34.

Felin, Teppo, and Todd R. Zenger. 2009. "Entrepreneurs as Theorists: On the Origins of Collective Beliefs and Novel Strategies." *Strategic Entrepreneurship Journal* 3 (2): 127-146.

Field, Erica, Seema Jayachandran, and Rohini Pande. 2010. "Do Traditional Institutions Constrain Female Entrepreneurship? A Field Experiment on Business Training in India." *American Economic Review* 100 (2): 125-29.

Galasso, Alberto, and Timothy S. Simcoe. 2011. "CEO Overconfidence and Innovation." *Management Science* 57 (8): 1469-1484.

Gelman, Andrew, Jennifer Hill, and Aki Vehtari. 2020. *Regression and Other Stories.* Cambridge University Press.

Gosnell, Greer K., John A. List, and Robert Metcalfe. 2020. The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains. *Journal of Political Economy* 128(4), 1195-1233.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda. 2013. "Who Creates Jobs? Small versus Large versus Young." *Review of Economics and Statistics* 95 (2): 347-361.

Harvard Business Review Analytic Services. 2016. *The Evolution of Decision Making: How Leading Organizations Are Adopting a Data-Driven Culture.*

Jones, John Bailey, and Sangeeta Pratap. 2020. "An Estimated Structural Model of Entrepreneurial Behavior." *American Economic Review* 110 (9): 2859-98.

Kerr, William R., Ramana Nanda, and Matthew Rhodes-Kropf. 2014. "Entrepreneurship as Experimentation." *Journal of Economic Perspectives* 28 (3): 25-48.

Levitt, Steven D., and John A. List. 2009. "Field Experiments in Economics: The Past, the Present, and the Future." *European Economic Review* 53 (1): 1-18.

List, John, Zacharias Maniadis, and Fabio Tufano. 2017. To Replicate or Not To Replicate? Exploring Reproducibility in Economics through the Lens of a Model and a Pilot Study. No. 00646. *The Field Experiments Website.*

Lovallo, Dan, and Olivier Sibony. 2010. "The Case for Behavioral Strategy," *McKinsey Quarterly*, March.

Malmendier, Ulrike, and Geoffrey Tate. 2008. "Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction." *Journal of Financial Economics* 89 (1): 20-43.

Maniadis, Zacharias, Fabio Tufano, and John A. List. 2015. "How to Make Experimental Economics Research More Reproducible: Lessons from Other Disciplines and a New Proposal." In *Replication in Experimental Economics.* Emerald Group Publishing Limited.

Martin, Roger, and Roger L. Martin. 2009. "The Design of Business: Why Design Thinking Is the Next Competitive Advantage." *Harvard Business Press*, 2009.

McCullough, Bruce D. 2009. "Open Access Economics Journals and the Market for Reproducible Economic Research." *Economic Analysis and Policy* 39 (1): 117-126.

Zimmermann, Christian. 2015. "On the Need for a Replication Journal." FRB St. Louis Working Paper 2015-16.

APPENDIX

2.6 Extension of the model

In this Appendix we extend our model by assuming that after the experiment decision-makers observe outcome $\pi \sim F$, with $\pi \in [0, B]$, where B is a finite upper bound such that $B > \pi_0$, and π_0 is the outside option of the decision-maker at the outset of the process. This also implies that we extend our model by allowing for the possibility that decision-makers can pivot even if, at the end of the experiment that they run in the generic stage s, they observe an outcome higher than the outside option.

Before the experiment in stage $s \geq 1$, decision-makers know that if at the end of the experiment they observe $\pi \leq \pi_{s-1}$, where π_{s-1} is the best option they realized up to this stage, they earn π_{s-1}. Therefore, they run the experiment if $\int_{\pi_{s-1}}^{B} \pi dF(\pi) + \pi_{s-1}F(\pi_{s-1}) - k > \pi_{s-1}$, or after integration by parts

$$B - \int_{\pi_{s-1}}^{B} F(\pi)d\pi - \pi_{s-1} - k > 0 \tag{A1}$$

The decision-makers who satisfy this inequality run the experiment. In order to define the share of these decision-makers we can focus on two sources of randomness across decision-makers: the cost of the experiment k or any parameter of the distribution F. We are agnostic about the exact source of randomness and, for $s \geq 1$, we define $G_{s-1} \equiv G_{s-1}(z_{s-1})$ to be the distribution that represents the share of decision-makers that satisfy this inequality, where z_{s-1} is the left-hand side of (eq: condition$_s -$ 1)$.Otherthingsbeingequal, the distribution function$G$_{s-1}$ decreases with k and increases with parameters that lower F over its entire support. For example, suppose that ξ is a parameter that affects F in the sense of first order stochastic dominance. We can think of $G_{s-1} \equiv \Gamma_{s-1}(z_{s-1} \mid \xi)h(\xi)$, where Γ_{s-1} is the distribution function of z_{s-1} conditional on ξ, and $h(\cdot)$ is the marginal probability of ξ. It is also easy to see that G_{s-1} decreases with π_{s-1}.

Thus, at the beginning of stage 1, G_0 decision-makers satisfy (eq: condition$_s-1).Inparticular,\pi_{s-1} =$ π_0 is the outside option. After the experiment, decision-makers observe π and we set $\Phi_0 \equiv \Phi(\pi_0)$ to be the probability that the experiment yields $\pi \leq \pi_0$. We now assume that, like in the text, decision-makers discover new ideas and they can rule out some ideas in the space A using logic and analogies based on their frameworks. Specifically, we assume that, θ_1 is the probability that they discover a new idea. However, we assume that if they observe $\pi \leq \pi_0$, they rule out $\lambda_{10}A$ ideas, while if they observe $\pi > \pi_0$, they rule out $\lambda_{11}A$ ideas, with $\lambda_{10} < \lambda_{11}$. The rate of discovery of new ideas is the same, θ_1, but if they observe $\pi > \pi_0$, it is natural to assume that they rule out more ideas because any new idea viable for experimentation has to overcome a higher threshold.

Let $q_{10} \equiv \theta_1(1 - \lambda_{10})$ and $q_{11} \equiv \theta_1(1 - \lambda_{11})$. Following the same logic of the model in the text, $1 - G_0$ decision-makers terminate before the experiment in stage 1; after the experiment in stage 1, $G_0 [\Phi_0 G_0 q_{10} + (1 - \Phi_0)G_1 q_{11}]$ decision-makers pivot, $G_0\Phi_0G_0(1 - q_{10})$ terminate, and $(1 - \Phi_0)G_1(1 - q_{11})$ commit to the new idea.

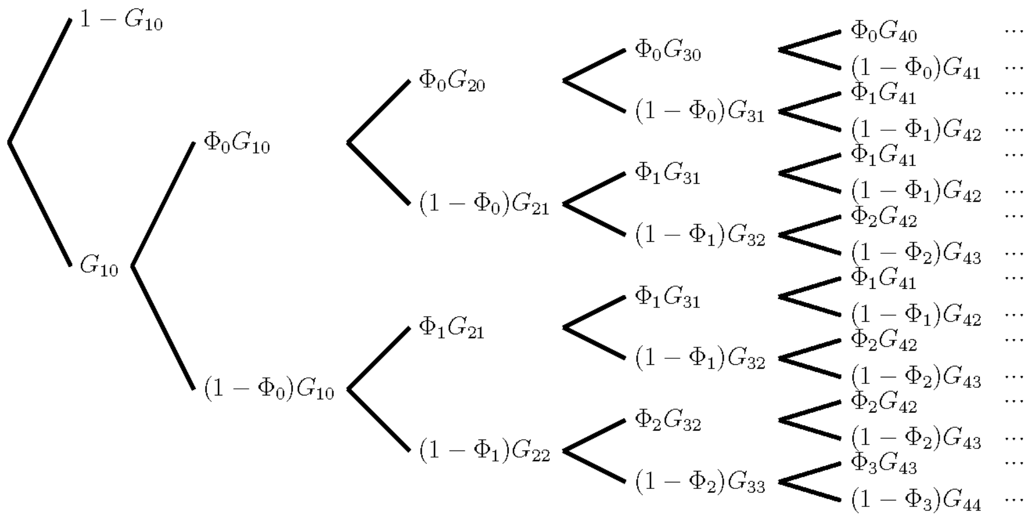
At the end of the experiment in stage 2, we have 4 cases: if in stage 1 decision-makers observed $\pi \leq \pi_0$, at the end of stage 2 they could observe $\pi \leq \pi_0$ or $\pi > \pi_0$; otherwise, they could observe $\pi \leq \pi_1$ or $\pi > \pi_1$, where we now take into account that they have to overcome the higher threshold $\pi_1 > \pi_0$. The probabilities of these 4 cases are, respectively, Φ_0^2, $\Phi_0(1-\Phi_0)$, $(1-\Phi_0)\Phi_1$, $(1-\Phi_0)(1-\Phi_1)$, where $\Phi_1 \equiv \Phi(\pi_1)$. Followng the same logic above and in the text, the probabilities that in these 4 cases they find a new idea they can pivot to are, respectively, $\theta_2(1 - \lambda_{10})(1 - \lambda_{20})$; $\theta_2(1 - \lambda_{10})(1 - \lambda_{21})$; $\theta_2(1 - \lambda_{11})(1 - \lambda_{21})$; $\theta_2(1 - \lambda_{11})(1 - \lambda_{22})$.

More generally, these pattern repeat themselves following the same logic at the end of each stage. In general, the new threshold is the highest observed outcome π in previous experiments. To streamline this representation, consider the end of the generic stage s in which there have been $j \geq s$ updates of π, such that the best available option after the s^{th} experiment is π_j. The probability that decision-makers observe a new idea worth pivoting to in a new experiment in the following stage can be approximated by a geometric

Brownian motion with negative drift $-\gamma_j$. We can then write the expected value of this probability as $q_{sj} = \theta e^{-\gamma_j s}$ such that γ_j increases with j. Intuitively, other things being equal, the higher the number of updates the lower the probability that decision-makers find a new idea. We can then define $G_{sj} \equiv G_s q_{sj}$ as the share of decision-makers who reached stage s, have a history of $j \leq s$ updates, and pivot.

Figure fig: FigureA1 shows how the pivoting decisions unfold. Let $\Phi_j \equiv \Phi(\pi_j)$, $j = 0, 1, 2, ...$ In each branch of the Figure we have the probabilities of pivoting conditional on the history of observations up to that point. In the initial branch, $1 - G_0$ is the probability that decision-makers terminate without running the experiment in stage 1. If they run it, which happens with probability G_0, in the next branch they pivot with probability $\Phi_0 G_{10}$ or $(1-\Phi_0)G_{11}$ depending on whether they observe $\pi_1 \leq \pi_0$ or $\pi_1 > \pi_0$. At the end of stage 2, which is the next branching in Figure fig: FigureA1, the upper branching reports the probability of pivoting if decision-makers observed $\pi_1 \leq \pi_0$ at the end of stage 1 and they observe $\pi_2 \leq \pi_0$ or $\pi_2 > \pi_0$ at the end of stage 2. In the lower branching, we have the equivalent probabilities but we now take into account that the best option they carry on from the previous stage is $\pi_1 > \pi_0$ rather than π_0. The logic of the following steps is the same.

Figure A1: Share of Pivots



Assumptions ass: overconfidence and ass: links extend naturally to this more general case. In particular, scientific decision-makers exhibit a lower G_0 produced by a higher F, and higher θ and γ_j $\forall j \leq s$.

The assumptions relaxed by this extended model do not affect the logic of our model in the text regarding termination. The reason is that termination implies that decision-makers never observe $\pi > \pi_0$. Thus, the probability of termination replicates the one showed in the proof of Proposition 1, and the logic discussed in this proof applies here as well. Thus, in this extended model scientific decision-makers are also more likely to terminate in early stages.

We then focus on the propensity to pivot. The share of decision-makers who pivot at any given stage, conditional upon reaching the stage, is the sum of all the vertical terms in each column corresponding to a given stage in Figure fig: FigureA1. These probabilities change between scientific and non-scientific decision-makers because of G_{sj}. In G_{sj}, G_s only differs between scientific and non-scientific decision-makers because of F. However, differences in F implied by Assumption ass: overconfidence disappear if decision-makers run the experiment of stage 1 because after this experiment all decision-makers satisfy condition (eq: condition$_s - 1).Thisleavesdifferencesin$q$_{sj}$. As shown in the proof of Proposition 2 in the text, these expression eventually decrease for s large enough. Therefore, for s large enough, a sufficient number of vertical terms for a given stage in the columns of Figure fig: FigureA1 will be smaller for scientific decision-makers such that the sum of these vertical terms will also be smaller.

In order to establish that they are less likely to pivot after a given stage, it is easy to see from Figure fig: FigureA1 that if in stage s decision-makers reached a node where we observe G_{sj}, the share of these decision-makers who pivot and run the $s+1$ experiment is $\Phi_j G_{s+1j} + (1-\Phi_j) G_{s+1j+1}$. This implies that if in stage s, the share G_{sj} of scientific decision-makers, $\forall j \leq s$, is smaller than that of the non-scientific decison-makers, their probability of pivoting to stage $s+1$ is smaller than that of non-scientific decision-makers because both $G_{s+1j} < G_{sj}$ and $G_{s+1j+1} < G_{sj}$, and both G_{s+1j} and G_{s+1j+1} are smaller for scientific decision-makers than non-scientific decision-makers. As a result, even in this extended model, conditional upon reaching a given stage, scientific decision-makers are eventually less likely to pivot.

The question is whether at the outset of the process scientific decision-makers are more likely to pivot. But this question is not different from what we have discussed in the text, as it involves an assessment of the conditions in the very first two branches of Figure fig: FigureA1 that are the same as the ones in the text. Thus, if scientific decision-makers are more likely to pivot initially, they will eventually be less likely to pivot. If they are not more likely to pivot initially, they will be less likely to pivot from the outset of the process.

2.7 Balance Checks

Table A1: Balance Checks RCT1

Variable Name	Description	Treatment		Control		Difference	
		Mean	SD	Mean	SD	b	p
Currently Employed	Proportion of team members employed at the time of the training	0.68	0.39	0.72	0.42	0.04	(0.567)
Currently Studying	Proportion of team members enrolled in an education program at the time of training	0.19	0.37	0.28	0.42	0.09	(0.250)
Education Level	Highest educational level attained by team members (5=PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise; Team Average)	2.34	0.86	2.12	0.91	-0.22	(0.192)
Experience: Entrepreneurial Founder or Employee	Number of years of experience working with companies other than focal as founder or employee (Team Average)	0.92	2.51	0.32	1.18	-0.59	(0.107)
Experience: Entrepreneurial Mentor	Number of years of experience working with companies other than focal as mentor or consultant (Team Average)	0.02	0.13	0.02	0.13	0.00	(0.981)
Experience: Industry	Number of years of experience in industry (Team Average)	2.55	4.64	2.56	4.78	0.01	(0.991)
Experience: Managerial	Number of years of managerial experience (Team Average)	2.03	3.34	1.22	3.32	-0.81	(0.192)
Idea Stage	Dummy variable assuming value of 1 when the company has one business idea and 0 when the company has started working on the project but has not launched it on the market yet	0.63	0.49	0.65	0.48	0.02	(0.807)
Lombardy	Dummy variable assuming value of 1 when the majority of team members comes from the Italian region of Lombardy and 0 otherwise	0.32	0.47	0.4	0.49	0.08	(0.366)
Sector: Furniture	Dummy variable assuming value of 1 when the company operates in the furniture sector and 0 otherwise	0.25	0.44	0.25	0.43	-0.01	(0.916)
Sector: Internet	Dummy variable assuming value of 1 when the company operates in the internet sector and 0 otherwise	0.44	0.5	0.51	0.5	0.07	(0.467)
Sector: Retail	Dummy variable assuming value of 1 when the company operates in the retail sector and 0 otherwise	0.1	0.3	0.07	0.26	-0.03	(0.548)
Team Size	Number of team members	2.85	1.36	2.72	1.31	-0.13	(0.606)
Observations		59		57		116	

Table A2: Balance Checks RCT2

Variable Name	Description	Treatment Mean	Treatment SD	Control Mean	Control SD	Difference b	Difference p
Age	Age (Team Average)	31.47	8.18	31.41	7.90	-0.06	(0.950)
Analytic Thinking	Agreement on a 1-10 scale with the following statements (Team Average): "Analyzing the situation and looking at the evidence is critical to our company's decision-making", "We carefully assess all the possible alternatives before making a choice for our company", "We prefer to gather all the relevant information before making a decision for our company", "Multiple elements are taken into account when making a decision for our company, pros and cons are carefully evaluated in every situation"	8.38	3.68	8.07	3.28	-0.32	(0.475)
Background: Economics	Team members with an economics background (%)	0.41	0.42	0.31	0.37	-0.10**	(0.046)
Background: Other	Team members with no economics backgrounds (%)	0.22	0.36	0.20	0.33	-0.02	(0.696)
Background: STEM	Team members with a STEM (Science Technology Engineering Mathematics) backgrounds (%)	0.38	0.40	0.49	0.41	0.11**	(0.032)
Certainty	Agreement on a 1-10 scale with the following statements (Team Average): "We are sure about our business model", "We are sure about our strategy"	5.93	1.94	5.61	1.91	-0.32	(0.191)
Consensus	Answer on a 1-10 scale to the following questions (Team Average): "To what extent do you and your team members have consensus on the long term objectives of the firm?", "To what extent do you and your team members have consensus on the short term objectives of the firm?", "To what extent do you and your team members have consensus on the survival strategy of the firm?"	8.85	1.67	8.86	1.66	0.00	(0.990)
Education	Highest educational level attained by team members (5=PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise; Team Average)	1.94	0.74	1.95	0.80	0.00	(0.969)
Experience: Entrepreneurial	Number of years of entrepreneurial experience (Team Average)	1.09	2.19	0.93	1.44	-0.17	(0.480)
Experience: Industry	Number of years of experience in industry (Team Average)	2.84	3.82	2.33	3.62	-0.51	(0.280)
Experience: Managerial	Number of years of managerial experience (Team Average)	2.29	3.69	2.27	4.18	-0.02	(0.971)
Experience: Work	Number of years of work experience (Team Average)	8.73	7.75	9.02	8.85	0.28	(0.788)
Full Time	Percentage of team members working full-time	0.57	0.43	0.62	0.42	0.05	(0.390)
Gender (Female)	Proportion of women in the team	0.27	0.37	0.25	0.36	-0.03	(0.541)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	10.17	9.65	10.96	11.45	0.78	(0.560)
Idea Potential	Independent assessment of the value of the idea	47.22	21.22	47.31	23.25	0.09	(0.975)
Idea Value: Max	Maximum estimated value of the project (0 to 100)	85.08	16.29	85.67	16.16	0.59	(0.773)
Idea Value: Mean	Estimated value of the project (mean, 0 to 100)	65.40	15.53	64.52	16.69	-0.88	(0.668)
Idea Value: Min	Minimum estimated value of the project (0 to 100)	45.71	19.86	43.21	22.93	-2.50	(0.357)
Idea Value: Range	Estimated value of the project (range, 0 to 100)	39.37	18.85	42.46	20.99	3.10	(0.221)
Intuitive Thinking	Agreement on a 1-10 scale with the following statements (Team Average): "We are prone to following our intuitions when making company-related decisions", "We consider feelings and intuitions rather than analysis in our startup decisions", "First impressions are important when making decisions", "It is important to rely on gut feelings and intuition when making decisions"	4.09	1.70	3.83	1.74	-0.25	(0.244)
Lombardy	Dummy variable taking value of 1 when the majority of team members comes from the Italian region of Lombardy, 0 otherwise	0.56	0.47	0.57	0.46	0.01	(0.883)
Months to Revenue	Number of months to revenue	11.52	5.80	11.51	5.85	-0.01	(0.987)
Part Time	Percentage of team members working part-time	0.08	0.18	0.08	0.17	0.00	(0.941)
Probability Termination	Probability of terminating the project	31.64	32.53	32.35	31.60	0.70	(0.863)
Team Size	Number of team members	2.25	1.46	2.28	1.37	0.03	(0.858)
Observations		125		125		250	

Table A3: Balance Checks RCT3

Variable Name	Description	Treatment		Control		Difference	
		Mean	SD	Mean	SD	b	p
Age	Age (Team Average)	30.60	9.29	30.53	7.14	-0.07	(0.963)
Analytic Thinking	Agreement on a 1-5 scale with the following statements (Team Average): "Analyzing the situation and looking at the evidence is critical to our companyÕs decision-making", "We carefully assess all the possible alternatives before making a choice for our company", "We prefer to gather all the relevant information before making a decision for our company" and "Multiple elements are taken into account when making a decision for our company, pros and cons are carefully evaluated in every situation"	4.30	0.63	4.40	0.56	0.11	(0.318)
Background: Economics	Team members with Economics backgrounds (%)	0.18	0.31	0.20	0.36	0.02	(0.701)
Background: Other	Team members with no Economics/STEM backgrounds (%)	0.56	0.43	0.44	0.46	-0.11	(0.152)
Background: STEM	Team members with a STEM (Science Technology Engineering Mathematics) backgrounds (%)	0.26	0.38	0.36	0.45	0.09	(0.223)
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" and "We are sure there is no better business model for our idea"	3.41	0.52	3.32	0.65	-0.09	(0.397)
Currently Studying	Number of team members enrolled in an education program at the time of training	0.26	0.30	0.21	0.30	-0.04	(0.426)
Education	Highest educational level attained by team members (5=PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise; Team Average)	1.85	0.89	2.06	1.09	0.21	(0.240)
Experience: Business Plan	Dummy taking value of 1 if the team had years of experience in business plan design, 0 otherwise	0.26	0.36	0.35	0.43	0.09	(0.228)
Experience: Entrepreneurial	Number of years of entrepreneurial experience (Team Average)	1.65	4.38	1.73	3.37	0.08	(0.908)
Experience: Industry	Number of years of experience in industry (Team Average)	2.77	5.72	3.03	5.04	0.25	(0.792)
Experience: Managerial	Number of years of managerial experience (Team Average)	1.54	2.78	1.76	3.76	0.22	(0.705)
Gender (Female)	Proportion of women in the team	0.31	0.38	0.25	0.36	-0.06	(0.356)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	11.39	10.06	11.76	12.36	0.37	(0.853)
Idea Maturity	Maturity of the idea (in months)	9.32	9.43	11.98	11.63	2.66	(0.158)
Idea Potential	Independent assessment of the value of the idea (two evaluators, average) based on five criteria: innovation, feasibility, sustainability, team competence, market size	49.22	11.99	49.16	12.86	-0.06	(0.978)
Idea Value: Mean	Estimated value of the project (mean)	65.82	18.53	63.30	16.05	-2.52	(0.415)
Intuitive Thinking	Agreement on a 1-5 scale with the following statements (Team Average): "We are prone to following our intuitions when making company-related decisions" and "We consider feelings and intuitions rather than analysis in our startup decisions"	2.74	0.83	2.70	0.99	-0.03	(0.838)
Later Stage	Dummy variable taking value of 1 if the firm is at a more advanced stage than others, 0 otherwise	0.13	0.34	0.11	0.31	-0.03	(0.666)
Locus of Control	Agreement on a 1-7 scale with the following statements (Team Average): "In most jobs you need a lot of luck to excel", "One typically earns what they are worth", "To make money you just need to know the right people", "To get a good position you need luck", "Income is mainly the result of hard work", "There is a direct relationship between a personÕs abilities and the position he/she holds", "Many of the difficulties encountered at work concerns senior colleagues", "Generally, people who work well get rewarded", "Promotions are awarded to people who work well", "To find a good job, having a good network is more important than actual skills", "A well-trained person always finds a satisfycing job" and "To get a really good job you have to have high-level acquaintances"	3.84	0.67	3.79	0.70	-0.05	(0.707)
Months to Revenue	Number of months to revenue	12.69	11.37	14.68	10.58	1.99	(0.310)
Piedmont	Dummy variable taking value of 1 when the majority of team members comes from the Italian region of Piedmont and 0 otherwise	0.55	0.45	0.52	0.48	-0.03	(0.748)
Probability Pivot Idea	Probability of changing the business idea	31.89	22.96	32.53	26.75	0.65	(0.884)
Probability Pivot Other	Probability of changing other components of the business model	52.20	22.97	52.92	26.17	0.73	(0.868)
Probability Pivot Problem	Probability of changing the problem and customer segment	34.57	22.49	34.48	25.20	-0.09	(0.983)
Probability Termination	Probability of terminating the project	13.64	16.53	17.42	21.66	3.78	(0.268)
Risk-averse	Agreement on a 1-7 scale with the following statements (Team Average): "In important matters I never take unnecessary risks, which can be avoided", "In important situations I never deliberately chose to take risks I could have avoided", "I always try to avoid situations that put me at risk of getting into trouble with other people", "I am always very careful and I put safety first" and "I prefer to avoid doing things that expose me to criticism and liability"	4.23	1.03	3.96	1.04	-0.27	(0.151)
Risk-taker	Agreement on a 1-7 scale with the following statements (Team Average): "I can be pretty reckless and take some big risks", "I think I often act boldly and courageously", "I am a brave and daring person and I like to tempt fate in various situations", "There is a direct relationship between a personÕs abilities and the position he/she holds" and "I think I am often less cautious than other people"	4.04	1.13	3.98	0.91	-0.05	(0.766)
Scientific intensity: 1 Theory	Theory development score	2.92	1.32	3.05	1.20	0.13	(0.559)
Scientific intensity: 2 Hypothesis	Hypothesis development score	2.14	1.63	1.98	1.51	-0.16	(0.571)
Scientific intensity: 3 Test	Test score	1.32	1.73	1.29	1.69	-0.03	(0.919)
Scientific intensity: 4 Valuation	Valuation score	0.84	1.49	0.94	1.63	0.09	(0.742)
Self-efficacy	Agreement on a 1-7 scale with the following statements (Team Average): "I think I will always be able to achieve a goal even if I have to perform a difficult task", "Faced with new tasks and challenges, I am always confident that I will be able to complete them", "I am sure I will succeed", "When I have a goal, I almost always get better results than others", "When I take a test or an exam I am sure I can pass it successfully", "I am confident that my results will be recognized and appreciated by others", "I am not worried about difficult situations, because so far I have always managed to get by with my skills", "I never had any problem understanding and facing even the most complicated situations" and "I think I get the crux of the matter first"	5.46	1.07	5.57	0.96	0.11	(0.557)
Self-regulation	Agreement on a 1-7 scale with the following statements (Team Average): "People can count on me to meet the set and planned deadlines", "I can hardly say no", "I change my mind quite often", "Others would describe me as an impulsive person", "I wish I had more self-discipline", "I get carried away by my feelings", "I am not easily discouraged", "Sometimes I canÕt stop but do something, even though I know it is wrong", "I often act without thinking about all the alternatives", "I often do things that seem right in the present, even at the expense of future goals" and "When I pursue a goal I follow the original plan, even when I realize that it is not the best"	4.99	0.82	5.25	0.85	0.25*	(0.090)
Startcup	Dummy variable taking value of 1 if the firm takes part to a local competition, 0 otherwise	0.11	0.32	0.18	0.39	0.07	(0.290)
Team Size	Number of team members	2.51	1.48	2.14	1.36	-0.37	(0.144)
Observations		61		66		127	

Table A4: Balance Checks RCT4

Variable Name	Description	Treatment		Control		Difference	
		Mean	SD	Mean	SD	b	p
Age	Age (Team Average)	35.77	8.56	36.37	9.20	0.60	(0.590)
Background: Economics	Team members with Economics backgrounds (%)	0.15	0.29	0.15	0.29	0.00	(0.940)
Background: Other	Team members with no economics backgrounds (%)	0.08	0.11	0.09	0.16	0.01	(0.410)
Background: STEM	Team members with a STEM (Science Technology Engineering Mathematics) backgrounds (%)	0.30	0.39	0.36	0.43	0.06	(0.260)
Business Age	Age of the business (years)	2.48	3.22	3.28	5.17	0.80	(0.140)
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	0.70	3.34	0.76	-0.07	(0.440)
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	-0.10	(0.340)
Experience: Enterpreneurial	Number of years of entrepreneurial experience (Team Average)	3.85	3.49	4.64	5.95	0.79	(0.200)
Experience: Industry	Number of years of experience in industry (Team Average)	6.75	6.47	7.70	7.56	0.95	(0.280)
Experience: Managerial	Number of years of managerial experience (Team Average)	5.96	5.29	6.22	6.16	0.26	(0.730)
Experience: Work	Number of years of work experience (Team Average)	13.02	7.98	13.53	8.59	0.51	(0.620)
Gender (Female)	Proportion of women in the team	0.42	0.42	0.50	0.44	0.08	(0.150)
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	(0.540)
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	(0.140)
Hours: Total Weekly	Weekly hours dedicated to the company (Team Average)	31.55	18.57	29.61	17.18	-1.94	(0.390)
Idea Value: Mean	Estimated value of the project (mean, 0 to 100)	66.73	17.05	66.62	20.22	-0.11	(0.960)
Idea Value: Range	Estimated value of the project (range, 0 to 100)	39.26	22.03	38.00	21.94	-1.26	(0.650)
Probability Expansion	Probability of expanding the business outside of the current industry or market	68.25	27.40	66.59	28.12	-1.67	(0.630)
Probability Pivot Idea	Probability of making a radical change to the business	45.85	28.18	42.12	26.99	-3.72	(0.280)
Probability Pivot Problem	Probability of changing the problem and customer segment	38.18	26.16	40.55	26.26	2.38	(0.470)
Scientific Intensity	Scientific intensity	2.61	1.18	2.41	1.25	-0.20	(0.180)
Team Size	Number of team members	2.14	1.95	2.31	2.14	0.18	(0.490)
Turnover Annual	Annual turnover (2018) £	50616.11	145448.79	71977.35	195899.81	21361.24	(0.320)
Turnover Monthly	Monthly turnover (January 2019) £	5113.83	17734.76	6099.50	24490.47	985.67	(0.710)
Observations		133		128		261	

Table A5: Balance Checks Full Sample

Variable Name	Description	Treatment		Control		Difference	
		Mean	SD	Mean	SD	b	p
Business Age	Age of the business (years)	0.87	2.24	1.12	3.39	0.24	(0.244)
Education	Highest educational level attained by team members (5=PhD, 4=MBA, 3=MSc, 2=BA, 1=high school, 0=otherwise; Team Average)	2.24	0.88	2.21	0.91	-0.04	(0.585)
Experience: Managerial	Number of years of managerial experience (Team Average)	3.36	4.52	3.32	5.17	-0.04	(0.915)
Experience: Entrepreneurial	Number of years of entrepreneurial experience (Team Average)	2.12	3.43	2.21	4.21	0.09	(0.749)
Experience: Industry	Number of years of experience in industry (Team Average)	4.11	5.60	4.30	6.11	0.19	(0.663)
Team Size	Number of team members	2.34	1.65	2.33	1.67	-0.01	(0.924)
Turnover: Annual	Annual turnover EUR	20266.69	102520.24	28300.88	137456.09	8034.19	(0.364)
Observations		378		376		754	

2.8 Termination

Table A6: Termination OLS Cross-Section

VARIABLES	(1) Termination OLS Cross-Section Full Sample	(2) Termination OLS Cross-Section RCT1	(3) Termination OLS Cross-Section RCT2	(4) Termination OLS Cross-Section RCT3	(5) Termination OLS Cross-Section RCT4
Intervention	0.104***	0.035	0.096**	0.194**	0.097**
	(0.001)	(0.647)	(0.044)	(0.037)	(0.035)
Constant	0.283***	0.316	0.364***	0.761**	0.287***
	(0.000)	(0.219)	(0.001)	(0.011)	(0.002)
Observations	754	116	250	127	261
R-squared	0.078	0.183	0.034	0.158	0.026
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A7: Termination Probit Cross-Section

VARIABLES	(1) Termination Probit Cross-Section Full Sample	(2) Termination Probit Cross-Section RCT1	(3) Termination Probit Cross-Section RCT2	(4) Termination Probit Cross-Section RCT3	(5) Termination Probit Cross-Section RCT4
Intervention	0.299***	0.105	0.249**	0.613***	0.295**
	(0.000)	(0.635)	(0.023)	(0.008)	(0.015)
Constant	-5.038***	-5.154***	-0.279	1.342**	-0.581***
	(0.000)	(0.000)	(0.114)	(0.050)	(0.002)
Observations	754	111	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A8: Termination OLS Panel

VARIABLES	(1) Termination OLS Panel Full Sample	(2) Termination OLS Panel RCT1	(3) Termination OLS Panel RCT2	(4) Termination OLS Panel RCT3	(5) Termination OLS Panel RCT4
Intervention	0.024***	0.003	0.049***	0.015**	0.016**
	(0.000)	(0.633)	(0.005)	(0.029)	(0.020)
Constant	0.027	0.050***	0.034	0.059***	0.011
	(0.211)	(0.000)	(0.288)	(0.002)	(0.396)
Observations	8,508	1,606	3,178	1,955	1,769
Number of id	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A9: Termination Probit Panel

VARIABLES	(1) Termination Probit Panel Full Sample	(2) Termination Probit Panel RCT1	(3) Termination Probit Panel RCT2	(4) Termination Probit Panel RCT3	(5) Termination Probit Panel RCT4
Intervention	0.164***	0.084	0.118*	0.311**	0.203**
	(0.000)	(0.458)	(0.059)	(0.024)	(0.014)
Constant	-5.639***	-6.891***	-1.665***	-5.171***	-2.163***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	8,508	1,606	3,178	1,955	1,769
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A10: Hazard of Termination

VARIABLES	(1) Hazard of termination Survival Full Sample	(2) Hazard of termination Survival RCT1	(3) Hazard of termination Survival RCT2	(4) Hazard of termination Survival RCT3	(5) Hazard of termination Survival RCT4
1.intervention	0.375***	0.101	0.334**	0.512	0.416**
	(0.000)	(0.664)	(0.012)	(0.158)	(0.014)
Observations	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A11: Week of Termination

VARIABLES	(1) Week of termination OLS Full Sample	(2) Week of termination OLS RCT1	(3) Week of termination OLS RCT2	(4) Week of termination OLS RCT3	(5) Week of termination OLS RCT4
Intervention	-2.322**	-1.114	-4.137*	-1.407	-1.606*
	(0.012)	(0.551)	(0.062)	(0.452)	(0.090)
Constant	32.446***	42.557***	51.376***	31.697***	32.103***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	754	116	250	127	261
R-squared	0.242	0.176	0.042	0.101	0.041
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.9 Pivot

Table A12: Number of Pivots

VARIABLES	(1) # Pivots OLS Cross-Section Full Sample	(2) # Pivots OLS Cross-Section RCT1	(3) # Pivots OLS Cross-Section RCT2	(4) # Pivots OLS Cross-Section RCT3	(5) # Pivots OLS Cross-Section RCT4
Intervention	-0.032	0.261**	0.012	-0.370	-0.038
	(0.654)	(0.021)	(0.841)	(0.311)	(0.588)
Constant	0.432***	0.536	1.238***	1.107	0.435***
	(0.000)	(0.217)	(0.000)	(0.448)	(0.000)
Observations	754	116	250	127	261
R-squared	0.120	0.105	0.070	0.068	0.019
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A13: Pivot OLS

VARIABLES	(1) Pivoting once OLS Cross-Section Full Sample	(2) Pivoting once OLS Cross-Section RCT1	(3) Pivoting once OLS Cross-Section RCT2	(4) Pivoting once OLS Cross-Section RCT3	(5) Pivoting once OLS Cross-Section RCT4
Intervention	0.087***	0.027	0.123***	0.080	0.083*
	(0.001)	(0.726)	(0.001)	(0.288)	(0.079)
Constant	0.083***	-0.013	0.336***	0.807**	0.085**
	(0.007)	(0.736)	(0.004)	(0.023)	(0.030)
Observations	754	116	250	127	261
R-squared	0.082	0.147	0.064	0.059	0.040
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A14: Pivot Multinomial Probit Full Sample

	(1)	(2)	(3)
	Pivoting once	Pivoting twice	Pivoting more than twice
	Multinomial Probit	Multinomial Probit	Multinomial Probit
	Cross-Section	Cross-Section	Cross-Section
VARIABLES	Full Sample	Full Sample	Full Sample
Intervention	0.370***	0.148	-0.287
	(0.010)	(0.397)	(0.117)
Constant	-1.374***	-2.104***	-2.438***
	(0.000)	(0.000)	(0.000)
Observations	754	754	754
Dummies for mentors and RCT	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

For RCT 1, a Multinomial Probit model does not converge due to the fact that only a few firms have pivoted more than once.

Table A15: Pivot Multinomial Probit RCT2

	(1)	(2)	(3)
	Pivoting once	Pivoting twice	Pivoting more than twice
	Multinomial Probit	Multinomial Probit	Multinomial Probit
	Cross-Section	Cross-Section	Cross-Section
VARIABLES	RCT2	RCT2	RCT2
Intervention	0.465**	0.107	-0.229
	(0.014)	(0.744)	(0.156)
Constant	-0.371	-1.192*	-0.936**
	(0.447)	(0.052)	(0.016)
Observations	250	250	250
Dummies for mentors and RCT	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A16: Pivot Multinomial Probit RCT3

	(1)	(2)	(3)
	Pivoting once	Pivoting twice	Pivoting more than twice
	Multinomial Probit	Multinomial Probit	Multinomial Probit
	Cross-Section	Cross-Section	Cross-Section
VARIABLES	RCT3	RCT3	RCT3
Intervention	0.117	0.027	-0.892
	(0.795)	(0.941)	(0.190)
Constant	1.062	-1.680	-1.757
	(0.485)	(0.254)	(0.405)
Observations	127	127	127
Dummies for mentors and RCT	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A17: Pivot Multinomial Probit RCT4

	(1)	(2)	(3)
	Pivoting once	Pivoting twice	Pivoting more than twice
	Multinomial Probit	Multinomial Probit	Multinomial Probit
	Cross-Section	Cross-Section	Cross-Section
VARIABLES	RCT4	RCT4	RCT4
Intervention	0.422*	-0.070	-0.314
	(0.100)	(0.819)	(0.269)
Constant	-1.259***	-1.415***	-1.916***
	(0.000)	(0.000)	(0.000)
Observations	261	261	261
Dummies for mentors and RCT	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor RCT	Intervention Mentor RCT

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.10 Performance

Table A18: Performance OLS Cross-Section

VARIABLES	(1) Revenue OLS Cross-section Full Sample	(2) Revenue OLS Cross-section RCT1	(3) Revenue OLS Cross-section RCT2	(4) Revenue OLS Cross-section RCT3	(5) Revenue OLS Cross-section RCT4
Intervention	6,504.108**	10,799.493	1,514.605	263.431	12,227.935
	(0.046)	(0.125)	(0.136)	(0.269)	(0.164)
Constant	9,039.968***	-4,899.747	-445.998	-594.302	6,297.301
	(0.006)	(0.403)	(0.859)	(0.484)	(0.344)
Observations	754	116	250	127	261
R-squared	0.086	0.220	0.023	0.052	0.036
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A19: Performance OLS Panel

VARIABLES	(1) Revenue (Flow) OLS Panel Full Sample	(2) Revenue (Flow) OLS Panel RCT1	(3) Revenue (Flow) OLS Panel RCT2	(4) Revenue (Flow) OLS Panel RCT3	(5) Revenue (Flow) OLS Panel RCT4
Intervention	677.342*	674.968*	84.145	71.107**	1,528.492
	(0.075)	(0.086)	(0.106)	(0.017)	(0.133)
Constant	820.151*	-601.024	-134.552	-163.533	3,025.476
	(0.086)	(0.196)	(0.341)	(0.278)	(0.124)
Observations	10,730	1,856	4,500	2,286	2,088
Number of id	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.11 Scientific Intensity

In this section, we provide additional details about the coding scheme used to assess the extent to which entrepreneurs use a scientific approach to decision-making. A team of research assistants conducted regular phone calls with all entrepreneurs taking part to our programs. Calls followed a detailed protocol with a script including a number of open-ended questions which are used to measure scientific decision-making. In using open-ended questions, we follow an approach similar to the one employed by Bloom and Van Reenen (2007, 2010) for their World Management Survey. Like Bloom and Van Reenen, we asked open-ended questions until an accurate assessment of the decision-making practices could be made by the research assistants. This allowed them to gather detailed information about decision-making approaches rather than ask directly about respondents' perception and aspirations. In addition, respondents were not aware that their responses were being scored against a predefined coding scheme, which helped ensure the collection of unbiased information. Decision-making practices were scored from 0 (lowest score) to 5 (highest score), across four key areas: 1) Theory, 2) Hypotheses, 3) Tests, 4) Evaluation.

When entrepreneurs formulate a theory, they elaborate a set of core ideas (and the key relationship between them) that explains why their business proposition should be viable. A theory generates a firm-specific point of view (Felin and Zenger, 2017) that flashes out what key assumptions decision-makers hold. These assumptions are then articulated as hypotheses or predictions that flow logically from the theory (Popper, 1972). Hypotheses provide the basis for a data gathering process that provides evidence in support or against of such hypotheses. Data can be gathered through tests of various nature, including qualitative (interviews, observations, etc.) and quantitative (data collection through surveys, A/B testing, etc.) data gathering techniques. Following the data collection process, entrepreneurs carefully analyze the results of their tests and re-evaluate their theory in light of these results.

In line with key literature in this area, we consider each component of the scientific decision-making approach as a multifaceted construct. For instance, the articulation of a theory rests on a wide variety of aspects, such as its clarity, level of detail, the extent to which it is based on evidence and the extent to which it considers alternative explanations. 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. In addition, each of these sub-elements can greatly vary in quality across entrepreneurs. One entrepreneur might have an extremely clear theory related to how his/her firm generates value for customers, while another might have a very murky explanation for his/her value creation process. All research assistants received extensive training prior to performing calls. The multiple training and practice sessions organized by the research team clarified how to score each sub-component. These sessions also provided clear examples with related scores to create an objective standard research assistants could refer to when coding. We provide an overview of the sub-components of the scientific approach and their related scores in Table A20 below.

Table A20: Scientific Intensity Components

Component	Sub-component	Definition	Score
Theory	Clarity of theory	The extent to which the theory is understandable	1 (not clear at all) to 5 (extremely clear)
Theory	Articulation of theory	The extent to which the theory is detailed	1 (not detailed at all) to 5 (extremely detailed)
Theory	Consideration of alternatives	The extent to which the theory includes alternative possible options	1 (no consideration of alternatives at all) to 5 (careful consideration of many alternatives)
Theory	Theory based on evidence	The extent to which the theory is based on objective evidence	1 (not based on objective evidence at all) 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	1 (not explicit at all) to 5 (extremely explicit)
Hypotheses	Coherence of hypotheses	The extent to which hypotheses are coherent with the theory	1 (not coherent at all) 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	1 (not detailed at all) 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	1 (not falsifiable at all) to 5 (extremely falsifiable)
Tests	Coherence of tests	The extent to which the test is coherent with the hypotheses	1 (not coherent at all) 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	1 (not valid at all) 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	1 (not representative at all) 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	1 (not rigorous at all) to 5 (extremely rigorous)
Evaluation	Data-based assessment	The extent to which the evaluation is based on data	1(not based on data at all) 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	1 (not coherent at all) to 5 (extremely coherent)
Evaluation	Systematic evaluation	The extent to which the evaluation is based on systematically collected and analysed data	1 (not systematic at all) 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	1 (not explanatory at all) to 5 (extremely explanatory)

Table A21: Scientific Intensity OLS Cross-Section

VARIABLES	(1) Scientific intensity OLS Cross-section Full Sample	(2) Scientific intensity OLS Cross-section RCT1	(3) Scientific intensity OLS Cross-section RCT2	(4) Scientific intensity OLS Cross-section RCT3	(5) Scientific intensity OLS Cross-section RCT4
Intervention	0.331***	0.581***	0.206*	0.341	0.321**
	(0.000)	(0.002)	(0.060)	(0.164)	(0.015)
Constant	2.081***	1.155***	2.488***	1.298**	2.086***
	(0.000)	(0.006)	(0.000)	(0.018)	(0.000)
Observations	754	116	250	127	261
R-squared	0.120	0.178	0.092	0.057	0.028
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A22: Scientific Intensity OLS Panel

VARIABLES	(1) Scientific intensity OLS Panel Full Sample	(2) Scientific intensity OLS Panel RCT1	(3) Scientific intensity OLS Panel RCT2	(4) Scientific intensity OLS Panel RCT3	(5) Scientific intensity OLS Panel RCT4
Intervention	0.355***	0.437***	0.286*	0.339*	0.386***
	(0.000)	(0.000)	(0.055)	(0.072)	(0.000)
Constant	1.271***	0.875***	2.079***	1.098***	2.403***
	(0.000)	(0.001)	(0.000)	(0.004)	(0.000)
Observations	10,730	1,856	4,500	2,286	2,088
Number of id	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.12 Instrumenting Scientific Intensity

Table A23: Termination 2SLS Cross-Section

VARIABLES	(1) Termination 2SLS Cross-Section Full Sample	(2) Termination 2SLS Cross-Section RCT1	(3) Termination 2SLS Cross-Section RCT2	(4) Termination 2SLS Cross-Section RCT3	(5) Termination 2SLS Cross-Section RCT4
Average Scientific Intensity	0.299***	0.082	0.344*	0.597	0.253***
	(0.001)	(0.615)	(0.060)	(0.177)	(0.003)
Constant	-0.283	0.345	-0.213	-0.191	-0.188
	(0.125)	(0.289)	(0.559)	(0.853)	(0.290)
Observations	754	116	250	127	261
R-squared	-0.485	0.101	-0.832	-1.737	-0.309
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A24: Termination IV Probit

VARIABLES	(1) Termination IV Probit Cross-section Full Sample	(2) Termination IV Probit Cross-section RCT1	(3) Termination IV Probit Cross-section RCT2	(4) Termination IV Probit Cross-section RCT3	(5) Termination IV Probit Cross-section RCT4
Average Scientific Intensity	0.603***	0.257	0.576***	0.833***	0.611***
	(0.000)	(0.547)	(0.000)	(0.000)	(0.000)
Constant	-5.456	-5.737***	-1.657***	-0.998	-1.896***
	(0.000)	(0.000)	(0.000)	(0.275)	(0.000)
Observations	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A25: Termination 2SLS Panel

VARIABLES	(1) Termination 2SLS Panel Full Sample	(2) Termination 2SLS Panel RCT1	(3) Termination 2SLS Panel RCT2	(4) Termination 2SLS Panel RCT3	(5) Termination 2SLS Panel RCT4
Scientific Intensity	0.055***	0.006	0.206**	0.038	0.047***
	(0.000)	(0.635)	(0.024)	(0.183)	(0.004)
Constant	-0.040*	-0.003	-0.308	-0.011	-0.076**
	(0.079)	(0.862)	(0.142)	(0.873)	(0.030)
Observations	8,508	1,606	3,178	1,955	1,769
Number of id	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A26: Pivot 2SLS

	(1)	(2)	(3)	(4)	(5)
	Pivoting once	Pivoting once	Pivoting once	Pivoting once	Pivoting once
	2SLS	2SLS	2SLS	2SLS	2SLS
VARIABLES	Full Sample	RCT1	RCT2	RCT3	RCT4
Average Scientific Intensity	0.251***	0.063	0.441***	0.248	0.219*
	(0.000)	(0.693)	(0.007)	(0.134)	(0.057)
Constant	-0.391**	-0.119	-0.477	0.334	-0.325
	(0.016)	(0.693)	(0.136)	(0.553)	(0.180)
Observations	754	116	250	127	261
R-squared	-0.225	0.130	-0.846	-0.191	-0.306
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A27: Pivot IV Probit

	Pivoting once	Pivoting once	Pivoting once	Pivoting once	Pivoting once
	IV Probit	IV Probit	IV Probit	IV Probit	IV Probit
VARIABLES	Full Sample	RCT1	RCT2	RCT3	RCT4
Average Scientific Intensity	0.684***	0.283	0.816***	0.605**	0.660***
	(0.000)	(0.658)	(0.000)	(0.021)	(0.000)
Constant	-5.687***	-6.558***	-2.302***	-0.290	-2.236***
	(0.000)	(0.000)	(0.000)	(0.816)	(0.000)
Observations	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A28: Performance 2SLS Cross-Section

	(1)	(2)	(3)	(4)	(5)
	Revenue	Revenue	Revenue	Revenue	Revenue
	Cross-section	Cross-section	Cross-section	Cross-section	Cross-section
VARIABLES	Full Sample	RCT1	RCT2	RCT3	RCT4
Average Scientific Intensity	18,703.974*	25,182.411*	5,437.553	812.963*	32,031.798
	(0.056)	(0.087)	(0.155)	(0.056)	(0.172)
Constant	-26,351.771	-47,474.210*	-10,420.817	-1,979.193*	-53,791.488
	(0.208)	(0.090)	(0.141)	(0.061)	(0.273)
Observations	754	116	250	127	261
R-squared	0.064	0.106	-0.134	0.043	0.009
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	-	-	-	-	-
Clustered Errors	Intervention Mentor RCT	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

Table A29: Performance 2SLS Panel

	(1)	(2)	(3)	(4)	(5)
	Revenue (Flow)	Revenue (Flow)	Revenue (Flow)	Revenue (Flow)	Revenue (Flow)
	Panel	Panel	Panel	Panel	Panel
VARIABLES	Full Sample	RCT1	RCT2	RCT3	RCT4
Scientific Intensity	1,903.092*	1,542.864*	304.581	209.543	3,884.029
	(0.092)	(0.094)	(0.170)	(0.190)	(0.180)
Constant	-1,598.680	-2,367.739	-469.140	-446.874	-7,014.216
	(0.380)	(0.104)	(0.225)	(0.266)	(0.296)
Observations	10,730	1,856	4,500	2,286	2,088
Number of id	754	116	250	127	261
Dummies for mentors and RCT	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Clustered Errors	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor	Intervention Mentor

Robust pval in parentheses, *** p<0.01, ** p<0.05, * p<0.1. All specifications control for the variables that were unbalanced between the treatment and control group despite randomization.

2.13 References

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *The Quarterly Journal of Economics* 122 (4):

1351-1408.

Bloom, Nicholas, and John Van Reenen. 2010. "New approaches to surveying organizations." *American Economic Review*, 100(2), pp.105-09.

Felin, Teppo, and Todd R. Zenger. 2017 "The theory-based view: Economic actors as theorists." *Strategy Science* 2.4: 258-271.

Popper, Karl. 1972. "Objective Knowledge, an Evolutionary Approach" *Clarendon Press.*

Chapter 3

A scientific approach to decision-making: evidence from a lab experiment with a simulation game

This paper tries to answer the question: Can a simulation game replicate real results about the adoption of a more scientific approach to entrepreneurial decision-making? In recent years, simulation games in the economic-managerial field have seen a progressive affirmation and perfectly integrated into the educational process traditional (Chiang et al., 2011). This work therefore intends to prove that a simulator is a valid tool to measure the ability of player-entrepreneurs to make decisions and replicate how they would act in a real context. I compare real a simulated results by running a lab experiment using a simulation game ("Start-up legend") I developed. Findings are in line with previous research in real context about the adoption of a scientific approach to entrepreneurial decision-making (Camuffo, et Al., 2020, 2021).

3.1 Introduction

This paper presents a simulation game ("Start-up legend") about the launch of a start-up that allows players to adopt (or not) a scientific approach to entrepreneurial decision-making (Camuffo et al 2020, 2021). The motivation behind this paper lies in the surge of field experiments in the managerial and entrepreneurial research and the need of replication in science and in the managerial field (Goldfarb and King, 2016, Astebro and Hoos, 2020). Running large field experiments in the entrepreneurial and managerial field requires significant efforts, resources and time and this might significantly impact the possibility to replicate previous results to corroborate theories and expand rapidly the knowledge. In this light, simulation games can help to scale the size of results and replicate findings and, once validated, help to find other interesting results. This is what is currently already done in several fields: we can easily think to car pilots simulating a gran-prix or airplane pilots simulating flights. What is a simulation game in our context? Consistently with the definitions of Siemer Angelides (1995) and Tennyson Jorczak (2008), computer-based simulation games are characterized by being constituted by a simulation of the decision-making process in an artificial environment, to study the consequences of one's decisions, and ultimately learn from them (Sitzmann, 2011). Business simulation games, focused on the management of economic processes, are effective methods for replicating the business challenges that managers might face before confronting the real world (Jerman et al, 2010). Players can choose the actions to be taken and can gain experience regarding the consequences of those actions. Furthermore, as they simulate the real-world system, From players' perspective a clear advantage of these simulation games lies in the possibilities that they can test themselves in a simulated context before facing similar situations in real life. In this light, I developed a simulation game replicating the launch of start-up where players are asked to set-up a business idea to provide a service to solve the sustainability of urban mobility in a Milan-like city. Players could, during each gaming session, adopt a scientific approach to decision-making: they could ask for more information build a compelling theory, formulate hypotheses, test their assumptions through interviews and surveys and decide whether to pivot, i.e. change partially the idea, launch the idea or terminate it. This is in line with previous research about the scientific approach (Camuffo et Al. 2020, 2021). The simulation game is based on the business case of a real-life startup "MiMoto" which operates an electric scooter sharing service in three main Italian cities. This choice is since MiMoto founders participated in the Camuffo et al. (2020) field experiment and tuned out to be an excellent case to build the game on because, during the field experiment, founders received a comprehensive training on the scientific approach and launched the idea with a good success on the market. This allowed me to benchmark the simulation game with a best practice in terms of adoption of the scientific approach. Indeed, the game has been structured with a tree-path logic. There is one best scenario that player could develop, that is the exact replication of what MiMoto founders launched on the market, regardless they used a scientific approach. All the other scenarios are less performing declinations of this best scenario. Moreover, players could also opt to solve the problem of urban sustainable mobility by launching a mini e-car sharing service or a e-bike sharing service. Within each solution (e-scooter, mini e-car, e-bike) the logic is the same: there is one best scenario. In total, there are more than two millions scenarios that players could develop. E-scooter (MiMoto-like) dominates the other two solutions, while mini e-car is the least performing. To validate the simulation game, I ran a pilot field experiment with 125 master business students, where 75 students attended a deep academic course (48 hours) about the scientific approach, the remaining 50 not. What I find is that students treated with a scientific approach show higher level of adoption of the scientific approach in the game and treated students terminate more their ideas and show more conservative beliefs about the success of their start-up. These results are in line with previous research about the scientific approach (Camuffo et Al. 2020, 2021, Messinese 2021) from field experiments with real start-up. This is an interesting result that encourages to keep exploring this innovative way to test theories even in the strategic and entrepreneurial field. The paper proceeds as follows. In the next section I introduce simulation games and describe the "Start-up legend". Then I present results from a pilot field experiment.

Then I compare them with previous finding in a real context and conclude.

3.2 Simulation game

3.2.1 Introducing business simulation games

The literature on Simulation games is characterized by a plurality of definitions and by a poor consensus on their characteristics (Garris, Ahlers, Driskell, 2002; Hays, 2005). Several definitions on simulation games agree that they are interactive, governed by rules, goal-oriented, stimulate the competitiveness and imagination of players (Driskell Dwyer, 1984; Gredler, 1996; Tobias Fletcher, 2007; Vogel et al., 2006). In line with Tennyson Jorczak (2008) there are no longer clear boundaries between entertainment, a fundamental feature of games, and simulations, oriented to the replication of a real context or situation. It is therefore essential to clarify what is meant by simulation games. Consistent with the literaturef (Siemer Angelides, 1995; Tennyson Jorczak, 2008), computer-based simulator games are characterized by being constituted by a simulation of the decision-making process in an artificial environment, to study the consequences of one's decisions, and ultimately learn from them (Sitzmann, 2011). Business simulation games, focused on the management of economic processes, are effective methods for replicating the business challenges that players might face before confronting the real world (Jerman et al, 2010). Indeed, participants can choose the actions to be taken and can gain experience regarding the consequences of those actions. In recent years simulation games have seen a progressive affirmation and interest regarding their application in a variety of sectors, particularly in education and training. There are different types of simulation games, for our objectives I focused on the so-called tailor-made games. They are games designed to be applied to a specific problem and are designed for a specific use and purpose. The way in which the problem or issue that make up the course of the game is dealt with is referred to a specific situation (Peters Van de Westelaken, 2011). The process of designing and applying simulation games to complex problems is shown in Figure A1.

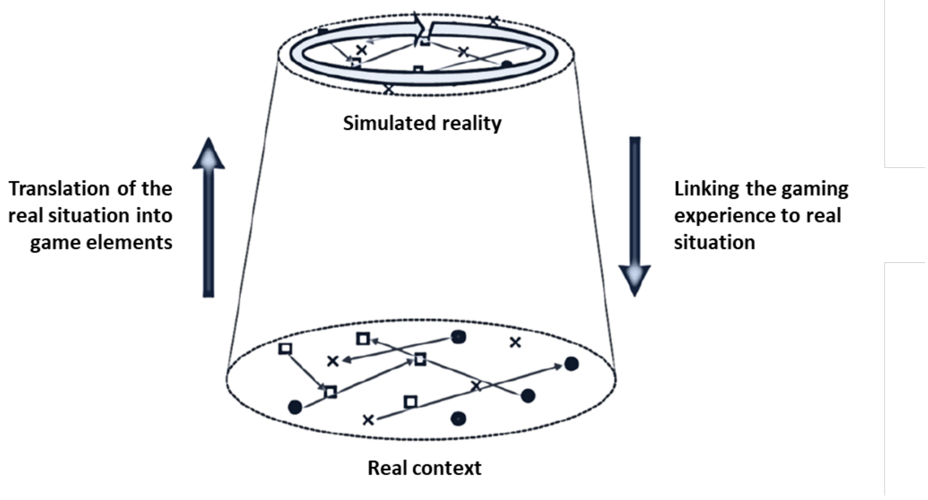


Figure A1: The process of designing and applying simulation games to real problems.

The starting point is a complex problem related to a specific situation or real truth. This reality (also called the reference system) is characterized by numerous aspects and elements of a different nature, and by multiple relationships between them. The goal in designing a simulation game is to reproduce this complex reality in a simpler model. In building this simpler model, three principles play an important role (Peters Van de Westelaken, 2014): 1) Reduction: not all elements, distinct in the real-life situation can be represented in the model, only the most important elements are included. 2) Abstraction: the elements included in the new model are not necessarily represented in as much detail as in the real-life situation, in other words: the system is abstract 3) Symbolization: the elements of the real-life situation are represented in the new model in a new, symbolic aspect. In the process of translating from reality to a reduced model, four phases can be distinguished (Peters Van de Westelaken, 2014). The design specifications: the aim is to clarify the purpose of the simulation game, what the final product should look like and under what conditions it will be used. System analysis: identification of the relevant elements of the reference system and their relationships. The design of the game: realization of the translation of reality into the game, looking for a metaphor and a suitable game format. The construction of the game: effective construction of the game, transformation of ideas into tangible products. This step also includes the testing phase and consequent product adaptation.

3.2.2 Key elements of a simulation game

Simulation games share common elements (Peters Van de Westelaken, 2014). Here I briefly discuss the most important ones.

Scenario. The scenario contains an overview of the situation and the relevant elements to understand the context. In other words, the scenario is the description of the reality of the game that the participants will have to read before starting to play.

Events. Events are important tools for setting the dynamics of the simulation game. Furthermore, they can be used to focus participants' attention on specific elements or to reject unwanted developments. When designing the game, it is determined exactly which event will be introduced at which time. The time and content of the event are exactly scheduled; however, the event comes up unexpectedly for attendees.

Roles. In a simulation game, characteristics and properties are assigned to a role, and these determine the actions of the one called to represent it. Roles typically differ in terms of goals, responsibilities, authority, resources, interests, etc. The clear definition of roles in the game is the most important in relation to the goal for which the simulation game is developed. Players are assigned to a specific role and behave accordingly during the simulation game: all actions and decisions are performed and evaluated from the point of view of this role.

Game cycles and phases. A simulation game consists of a series of steps, which are performed sequentially. In the configuration of a simulation game, two types of cycles can be distinguished: macro and micro. The macro-cycle is about setting up the whole simulation game in such a way that the objectives for which the game was developed can be achieved. The micro-cycle concerns the sequence of activities and actions within the game phases.

Decisions. During the game phases the participants have to make several decisions. These decisions have an influence on the course of the game. It is essential to have an overview of all the decisions that can be made and their consequences in order to have an overview of all the cause-effect relationships to be included in the game. To do this, it is possible to create a mapping during the design phase, in which the game phases are compared with the decisions that can be made.

Data. To be able to carry out what is required correctly, participants must have at their disposal the information necessary to make decisions or estimate the consequences of decisions. The type of data provided can range from raw data, for which participants need to find a good interpretation, to ready-to-use information, in which an interpretation is already provided. The data can be presented already in the scenario or introduced in the following game phases.

Indicators. The performance of players should be assessed on the basis of defined criteria. The indicators can be various; there may be quantitative indicators but also qualitative indicators.

Accounting system. It constitutes the set of rules by which it is possible to calculate a score for each of the indicators. During the game the participants make decisions. A score is calculated through the accounting system which indicates the level of performance of the participants' actions according to the behaviours sought.

3.2.3 Start-up Legend

In this section I present the simulation game used the produce the findings of this paper. I developed the simulation game "Start-up Legend" at Bocconi University (which holds all the right on the game) that financed and supported the whole project through the BUILT (Bocconi University Innovations in Learning and Technology) department and with the support of professional game developers. This effort comes has been justified by a previous basic pilot version game I develop in HTML (Figure A2) and tested with

Bocconi students.

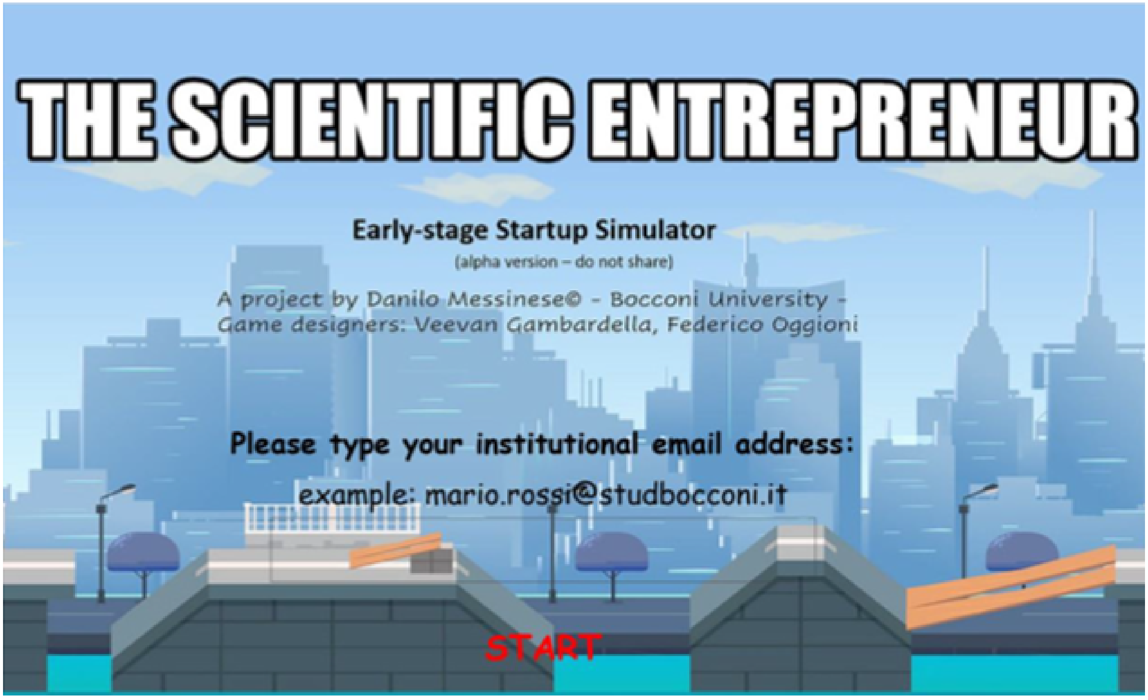


Figure A2: Pilot version of the simulation game.

"Start-up Legend" reproduces the early-stage phase of a start-up in the sustainable urban mobility sector (Figure A3). Its macro-cycle consists in 1) fill a simplified version of a Business Model Canvas by interaction with two characters that act as co-founders of the player (Figure A4). It is important to remark that these two co-founders do not incentivise any decision-making process; 2) decide whether to formulate hypothesis or not (Figure A5) 3) Conduct an interview or a survey to test the idea or hypotheses, if selected (Figure A6) 4) Analyse test results (Figure A7) 5) make a decision: explore more and eventually pivot by changing some components of the Business Model Canvas, launch the start-up or terminate it (Figure A8).

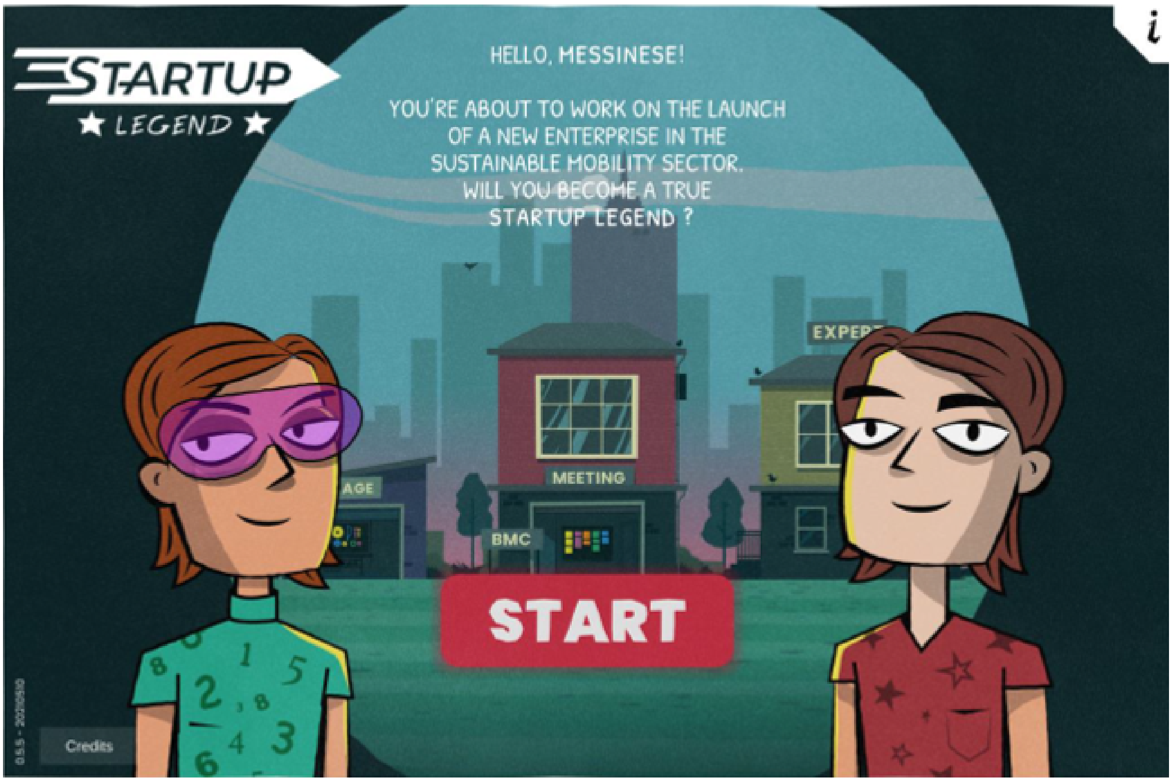


Figure A3: The process of designing and applying simulation games to real problems.

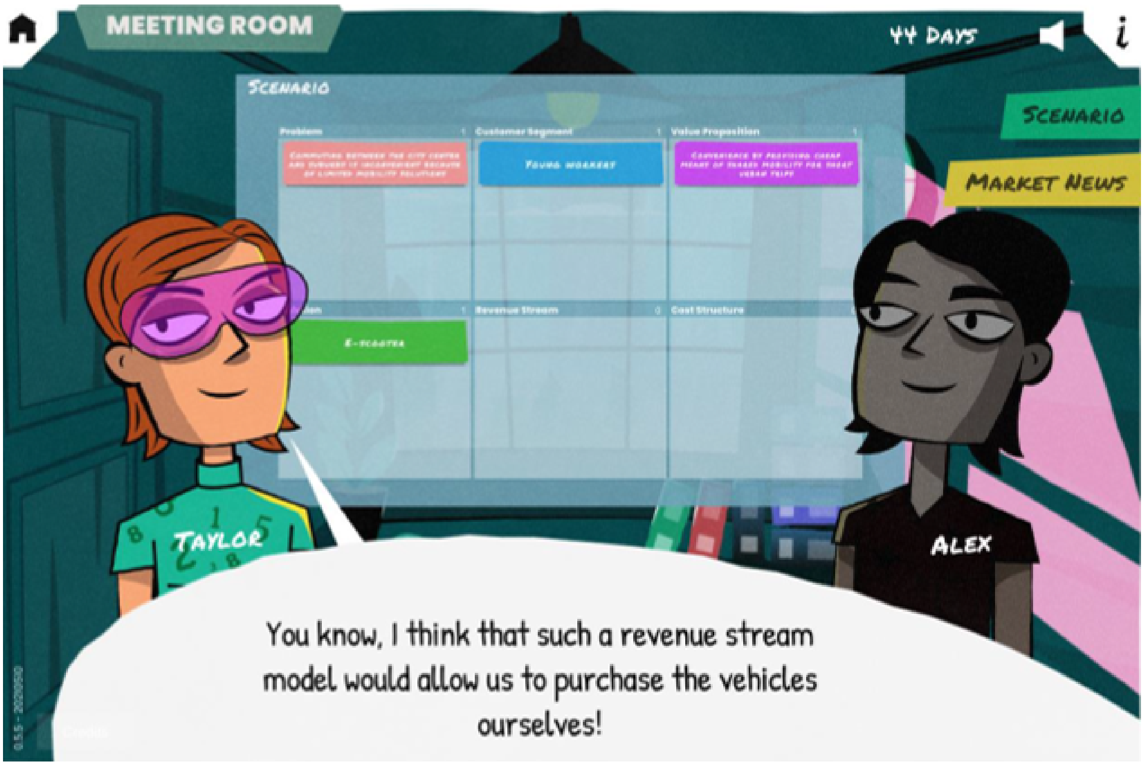


Figure A4: Welcome screen of the new version

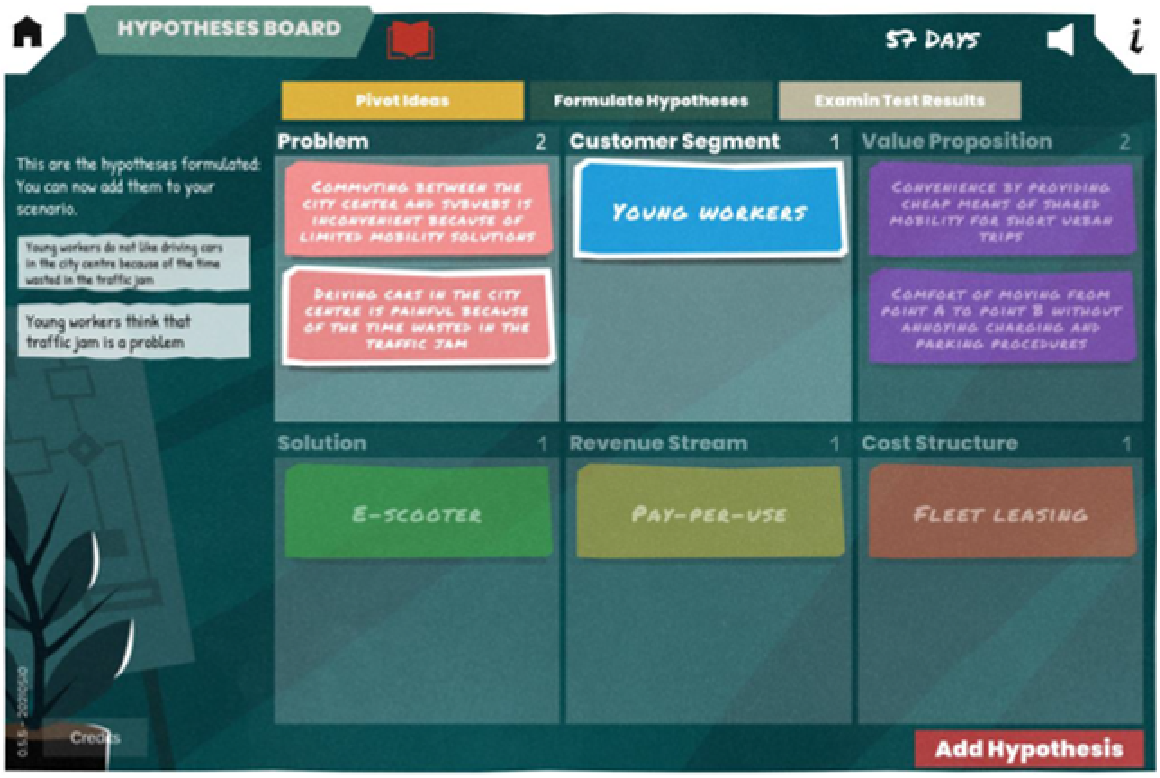


Figure A5: Formulating hypotheses

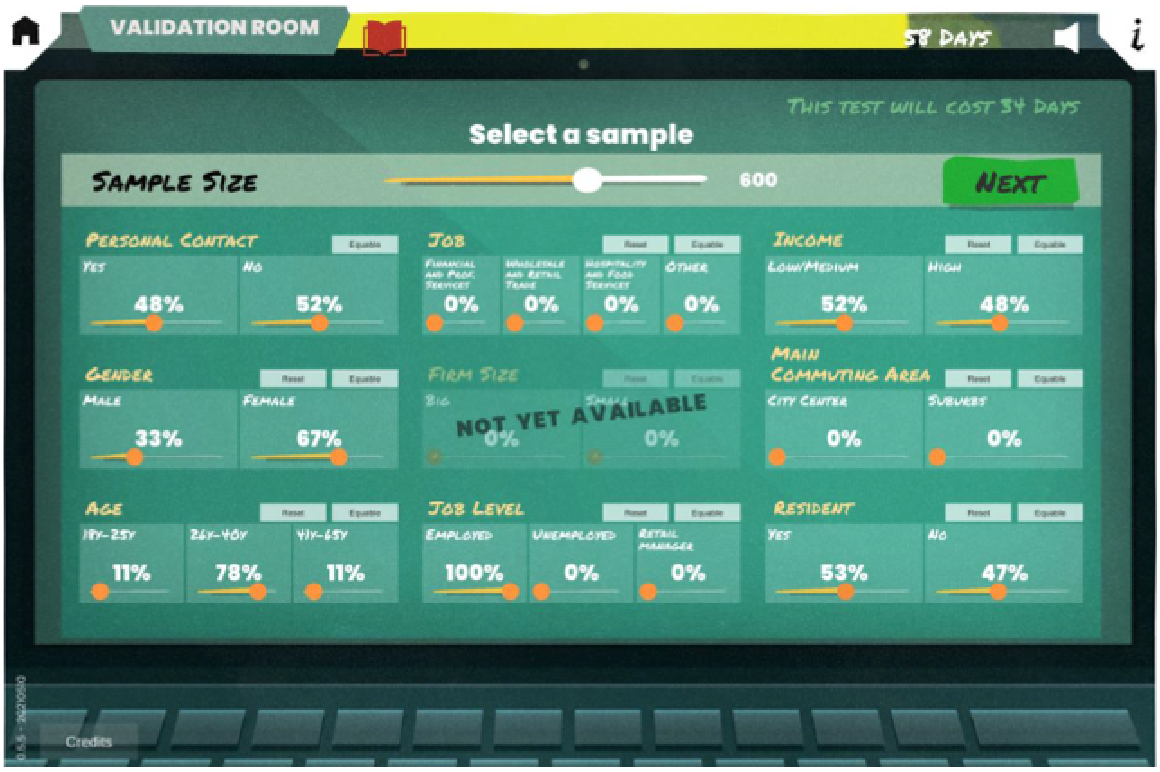


Figure A6: Selecting a sample for the test

Figure A7: Results of tests

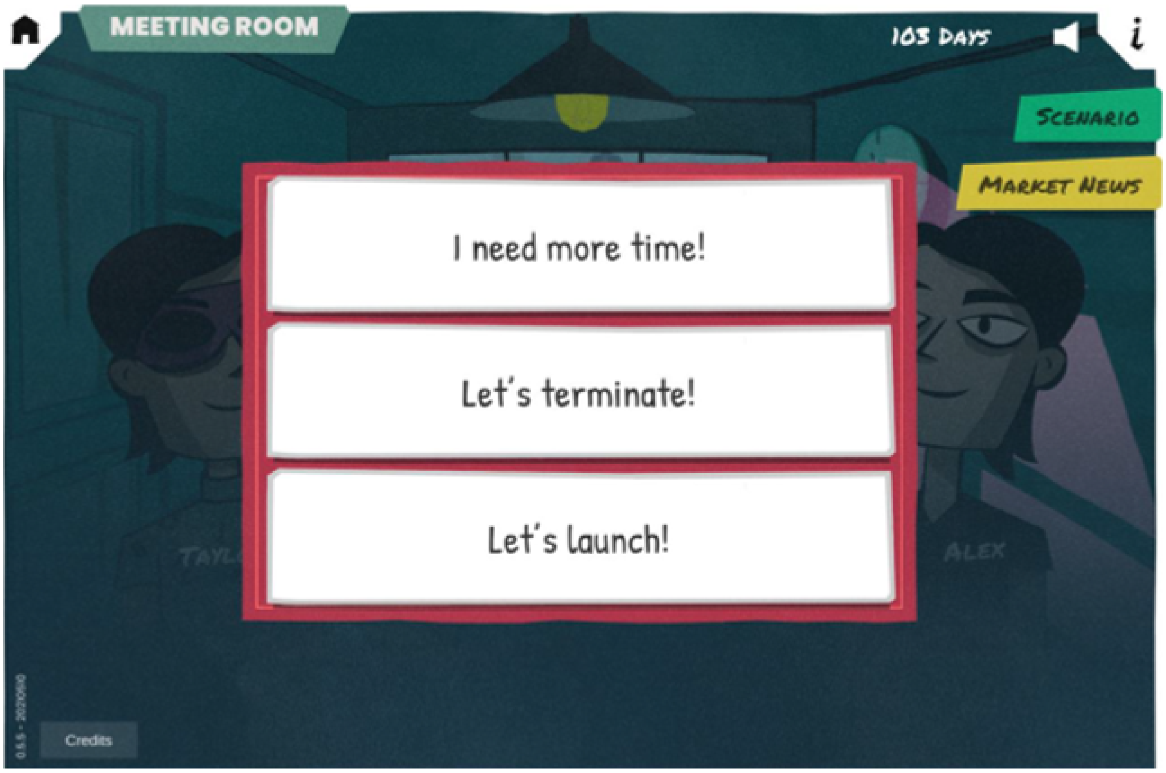


Figure A8: The decision screen

The simulation game "Start-up Legend" has been structured with the aim to simulate the launch of a start-up and, at the same time, to measure the impact of the adoption of a more scientific approach to decision's outcomes. To achieve this goal, a tree-path logic has been used to develop this game. Each of the six components of the Business Model Canvas could be filled with at maximum 3 items (minimum 1), but the "Solution" component that allowed to pick only one possible solution between E-scooter, E-bike and Mini E-car. All the possible combinations have been associated to a score ranging from 0 to 100, where 100 is the exact Business Model Canvas that MiMoto founders provided after have been treated with scientific training in the RCT studied by Camuffo, Cordova, Gambardella and Spina (2020) and that succeeded on the market. Figure A9 shows the values associated to scenarios, where. I also introduced the possibility to impose two random shocks: 1) E-bikes and E-scooter changed values symmetrically and 2) all the three solutions perform with low or even negative values. This is useful to avoid biases from players. Overall, there are more than 2 million of possible combinations.

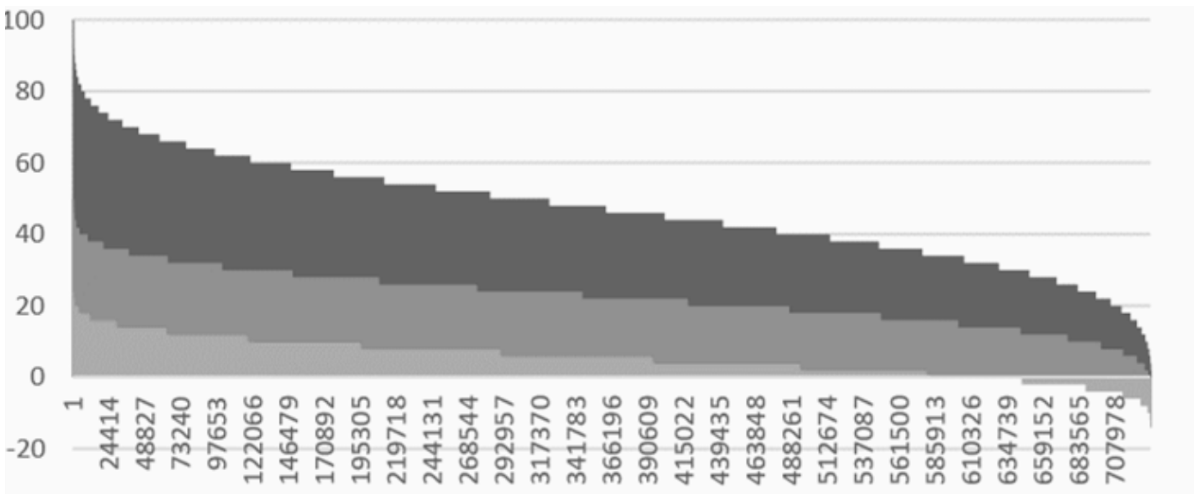


Figure A9: The dark grey region represents all the possible combinations of BMC about E-scooter, the middle region all the possible combinations of BMC for E-bikes and the lighter grey region represents the combinations for mini-E-cars

Players receive also news from the market to build initial evidence about what is the most performing solution when filling the Business Model Canvas and can eventually pivot to other solutions if they realize that it is the case. They have the possibility to test their Business Model Canvas by formulating (or not) hypotheses. To test their business idea they can either interview or submit a survey of a sample of potential customers. To do so, they can choose what questions to ask and to whom, using open-like and close-ended questions. Answers have been built to provide signals in line with the value of the test combination of the Business Model Canvas. Selecting a wrong sample, i.e. not in line with the correct customer segment, would provide noisy signals. After generating test results, players can pivot or conduct other tests or decide whether it is time to terminate or launch the idea. At the beginning and at the end of the gaming sessions, players are asked what is, on a scale from 0 to 100, their assessment of the value of the idea. In other words, they provide a prior and posterior belief about the success of the business idea. The game also measures the extent to which players adopt a scientific approach during the gaming session. This is done by asking questions measuring the logic underlying the choice of a certain combination of the Business Model Canvas (Theory), measuring the usage of hypotheses, both the number and the quality, (Hypotheses) and the quality of the sample used to test, both in terms of size and target (Test).

3.3 Theoretical background

Creating a new venture typically takes place under conditions of high uncertainty (Kirzner, 1973; Gans et al., 2019). According to existing research, entrepreneurs tend to use one of these two approaches. On the one hand, entrepreneurs can adopt a trial-and-errors approach (Nicholls-Nixon et al. 2000, Dencker et al. 2009) so that they test sequentially until they reach a satisfactory solution. On the other hand, entrepreneurs can adopt a more structured approach to decision-making. This implies a clear course of actions (Delmar and Shane, 2003; Blank 2006; Ries 2011). The scientific approach (Camuffo et al. 2019) is part of this second category. Entrepreneurs using a scientific approach to entrepreneurial decision-making apply a set of steps – similar to those applied by scientists – to develop their business idea. When using this approach, entrepreneurs start with the definition of a mental representation or a "theory" (Csaszar and Ostler 2019; Felin and Zenger 2009) that frames the business problem that entrepreneurs wish to solve and logically links the components of the business model. They then explicitly formulate falsifiable hypotheses to validate or confute the theory. To falsify their hypotheses, entrepreneurs design and execute well-tailored experiments and tests. Experiments and tests should be conducted by designing them coherently with the theory, by targeting the correct sample and evaluate the results. Previous research shows that entrepreneurs adopting a scientific approach make more precise and unbiased decisions that translate in higher probability of terminating their idea. In other words, they realize if they were overestimating the value of their business idea and terminate it to avoid false positive. (Camuffo et Al. 2020, 2021). These previous studies produce their empirical findings by running large field experiments with early-stage entrepreneurs. They randomize entrepreneurs between a treatment and a control group and train start-ups in the treatment group on how to adopt a more scientific approach to decision-making. As mentioned in previous sections, running RCTs require a huge effort a takes a lot of time to produce results. This in contrast with the goal to replicate research findings to corroborate them. In this light, I ran a pilot lab experiment aimed to validate the simulation game, by comparing simulated results with real results from previous RCTs in real context.

3.3.1 Empirical context and Results

The experiment was conducted in Milan in 2021, engaging 125 Master Students with major in Data Science and Business Analytics (DBSA) and Economics of Innovation and Technologic Management (EMIT) enrolled in their second year. 75 students have participated a course held by Alfonso Gambardella about the scientific approach to decision-making. The remaining 50 students did not attend any course about scientific decision-making. Table 3.1 show descriptive statistics about the main variables collected.

Table 3.1

Variable	Definition	Obs	mean	sd	min	max
Treatment	Dummy equal to 1 if players were treated with a scientific training, 0 otherwise	125	0.6	0.49	0	1
Termination	Dummy equal to 1 if players terminated the idea, 0 otherwise	125	0.2	0.4	0	1
Days	Day at which players launched or terminated the idea	125	179.34	82.94	61	481
Idea value	Idea value (range from 0 to 100)	125	23.06	26.62	-5.63	76.32
Scientific intensity	Variable ranging from 0 to 1, measuring the extent to which players use the scientific approach	125	0.47	0.27	0.12	1
Prior	Player's evaluation of the value of the idea at the end of the brainstorming (range 0 to 100)	125	57.87	16.18	0	100
Posterior	Player's evaluation of the value of the idea at the end of the gaming session (range 0 to 100)	125	61.27	24.76	5	100
Condition	Dummy equal to 1 if E-scooter or E-bike most performing solution, 0 if no solution was performing	125	0.74	0.44	0	1
EMIT	Dummy equal to 1 if student enrolled in the EMIT master program, 0 otherwise	125	0.38	0.49	0	1

I focus on the most robust result of previous research in the field to validate the simulation game: the termination rate of entrepreneurs treated with a scientific approach. From past studies (Camuffo et Al. 2021), we know that entrepreneurs adopting a more scientific approach to decision-making are more likely to terminate their idea than other. This is mainly due to an improvement in gathering and interpreting signals and test results.

Table 3.2 show results from a linear probability model and a probit model. Column (1) shows that students treated with a scientific training are more likely to terminate by 14.8 percentage points (p=0.056). The marginal effect in Column (2) show that the scientific treatment increases the probability to terminate by 11.8 percentage points (p=0.099). I control for their Prior assessment of the value of the idea and for the condition of the gaming session.

Table 3.2

	(1)	(2)
VARIABLES	OLS Cross-Section	Probit Cross-Section
Treatment	0.145*	0.470
	(0.056)	(0.107)
Prior	-0.007***	-0.024***
	(0.001)	(0.005)
Constant	0.555***	0.357
	(0.000)	(0.514)
Observations	125	125
R-squared	0.098	
Dummy for Condition	Yes	Yes

pval in parentheses*** p<0.01, ** p<0.05, * p<0.1

My analysis also includes an intent-to treat effect, in line with previous research (Camuffo et Al. 2021). This reflects the actual impact of the course about the scientific approach on the decision outcomes. Table 3.3 show how students that attended the course about the scientific approach turn out to adopt more the scientific approach in the simulation game. This reflects the fact that the simulation game, even though it simplifies the real context, allows to apply what has been learn during a real training. Table 3.4 presents the results of a linear probability model and a probit model, using two-stage least squares. The regression in column (1) reports a non-significant, but positive impact of the adoption of the scientific approach to the probability of terminating the idea. While, the instrumented scientific intensity in probit model in column (2) increases significantly the probability of termination.

Table 3.3

VARIABLES	(1) OLS Cross-Section
Treatment	0.098**
	(0.048)
Constant	0.406***
	(0.000)
Observations	125
R-squared	0.031

pval in parentheses*** p<0.01, ** p<0.05, * p<0.1

Table 3.4

VARIABLES	(1) 2SLS OLS Cross-Section	(2) 2SLS Probit Cross-Section
Scientific Intensity	3.540	3.757***
	(0.443)	(0.000)
Prior	-0.013	-0.014
	(0.187)	(0.178)
Treatment		
Constant	-0.230	-0.719
	(0.835)	(0.131)
Observations	125	125
R-squared	-4.877	
Dummy for Condition	Yes	Yes

pval in parentheses *** p<0.01, ** p<0.05, * p<0.1

The last step of this analysis concerns the change of beliefs of players. Table 3.5 show the impact of the scientific treatment on the posterior assessment of the value of the idea, before making a final decision. In line with previous findings analysing real contexts, the scientific treatment makes players more conservative by 22.58 points on a 0-100 scale (p=0.000), after controlling for their prior belief.

Table 3.5

	(1)
VARIABLES	OLS Cross-Section
Treatment	-22.580***
	(0.000)
Prior	0.079
	(0.533)
Constant	69.453***
	(0.000)
Observations	125
R-squared	0.193
Dummy for Condition	Yes

pval in parentheses*** p<0.01, ** p<0.05, * p<0.1

Overall, these results show that, when looking at termination rate, the simulation game seems to replicate the results found in previous RCTs with real start-ups. This is a first promising result that could encourage further research in this direction. Simulation games could be a powerful tool to scale research findings in real settings.

3.4 Conclusion

This paper presents an innovative way to test theories about the adoption of decision-making approaches in entrepreneurship. I used a simulation game "Start-up Legend" 75 that I developed with the financial, technological, and intellectual support of Bocconi (which holds all the rights on the game) to replicate past research findings on the scientific approach. I ran a lab experiment with 125 master students, where 75 of them attended a deep course on scientific decision-making. I asked them to play the game where they could simulate the launch of a start-up, from the ideation phase. They could brainstorm with virtual co-founders and test their idea. They were exposed to the possibility to use some "scientific" tools, such as formulate hypotheses, conduct precise test, set thresholds before making a decision. The output of the gaming session was twofold: they could either terminate or launch the idea. Findings are in line with previous research in real contexts (Camuffo et Al. 2020, 2021). This corroborates the intuition that real managerial practices and decisions can be replicated and simulated in a virtual reality. Of course, my findings have several limitations. The experiment I ran entails students, while previous studies engage real entrepreneurs. The allocation of students between the two groups might be impacted by other behavioural and demographic characteristics for which I do not control. The game might induce players treated with a scientific approach to be more likely to adopt it, because some tool, such as hypotheses, can be visualized and they do not require an additional mental effort. I think that these limitations can be a starting point to further improve the game and run more precise experiments with real entrepreneurs. Using simulation games to run lab experiments to replicate results from field experiments could turn out to be a turning point for research in the entrepreneurial and strategic field.

References

Astebro, T. (2003). The return to independent invention: Evidence of unrealistic optimism, risk seeking or skewness loving? The Economic Journal, 113(484), 226-239.

Baker, T., Miner, A., Eesley, D. (2003). Improvising firms: Bricolage, account giving and improvisa-tional competencies in the founding process. Research Policy, 32(2), 255-276.

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

Bowers J., Higgins N., Karlan D., Tulman S., and Zinman J. (2017). "Challenges to Replication and Iteration in Field Experiments: Evidence from Two Direct Mail Shots." American Economic Review 107 (5): 462-65.

Camuffo, A., Cordova, A., Gambardella, A., Spina, C. (2019). Scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. Management Science, 66(2).

Camuffo, A., Gambardella, A., Spina, C. (2020). Small changes with impact: Experimental evidence of a scientific approach to the decision-making of entrepreneurial firms. Working Paper.

Camuffo, A., Gambardella, A., Messinese, D., Novelli, E., Paolucci, E., Spina, C. (2021). A scientific approach to innovation management. evidence from four field experiments. NBER

Cassar, G. (2014). Industry and startup experience on entrepreneur forecast performance in new firms. Journal of Business Venturing, 29(1), 137-151.

Chesbrough, H. (2010). Business model innovation: opportunities and barriers. Long Range Planning, 43(2-3), 354-363.

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

Davenport, T. (2009). How to design smart business experiments. Strategic Direction, 25(8).

Dencker, J. C., Gruber, M., Shah, S. K. (2009). Pre-entry knowledge, learning, and the survival of new firms. Organization Science, 20(3), 516-537.

Driskell, J., Dwyer, D. (1984). Microcomputer videogame based training. Educational Technology(24), 11-17.

Eisenmann, T., Ries, E., Dillard, S. (2012). Hypothesis-driven entrepreneurship: the lean startup. Harvard Business School Case(9-812-095).

Farkas, G. (2007). Social software in libraries: building collaboration, communication, and community online. Information Today, Inc.

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

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

Felin, T., Gambardella, A., Stern, S., Zenger, T. (2019). Lean startup and the business model: Experimentation revisited. Long Range Planning.

Fried, J., Hansson, D., Linderman, M. (2009). Getting Real: the smarter, faster, easier way to build a successful web application. 37signals.

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

Garris, R., Ahlers, R., Driskell, J. (2002). Games, motivation, and learning: A research and practice model. Simulation Gaming(33), 441-467.

Gredler, M. (1996). Educational games and simulations: A technology in search of a research paradigm. In D. Jonassen, Handbook of research for educational communications and technology. New York: MacMillan.

Hays, R. (2005). The effectiveness of instructional games: A literature review and discussion. Orlando, FL: Naval Air Warfare Center, Training Systems Division.

Jerman, B., Ribeiro, C., Fernandes, J., Pereira, J., Arh, T. (2010). Analysing the required properties of business simulation games to be used in e-learning and education. Intelligent Informaton Management, 4(6), 348-356.

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

Kirzner, Israel M. (1973). Competition and Entrepreneurship. Chicago: University of

Chicago Press.

Koning, R., Hasan, S., Chatterji, A. (2019). Experimentation and startup performance: Evidence from A/B testing (No. w26278). National Bureau of Economic Research.

List J., Maniadis Z., and Tufano F.. 2017. To Replicate or Not To Replicate? Exploring Reproducibility in Economics through the Lens of a Model and a Pilot Study. No. 00646. The Field Experiments Website

Maniadis, Zacharias, Fabio Tufano, and John A. List. 2015. "How to Make Experimental Economics Research More Reproducible: Lessons from Other Disciplines and a New Proposal." In Replication in Experimental Economics. Emerald Group Publishing Limited.

McKinsey Company. (2020). McKinsey Problem Solving Game. https://www.mckinsey.com/careers/m digital-assessment

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

Nicholls-Nixon, C. L., Cooper, A. C., Woo, C. Y. (2000). Strategic experimentation: Understanding change and performance in new ventures. Journal of Business Venturing, 15(5), 493-521.

O'Neil, H., Wainess, R., Baker, E. (2005). Classification of learning outcomes: Evidence from the computer games literature. The Curriculum Journal(16), 455-474.

Osterwalder, A., Pigneur, Y. (2010). Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers. John Wiley Sons.

Osterwalder, A., Pigneur, Y., Tucci, C. (2005). Clarifying business models: origins, present, and future of the concept. Commun. Assoc. Inf. Syst.(16), 15-37.

Parker, S. C. (2006). Learning about the unknown: How fast do entrepreneurs adjust their beliefs? Journal of Business Venturing, 21(1), 1-26.

Peters, V., Van de Westelaken, M. (2011). Simulation games and a management approach: Thinking in systems. Nimega: Samenspraak Advies.

Peters, V., Van de Westelaken, M. (2014). Simulation games: A concise introduction to the design process. Nimega: Samenspraak Advies.

Ries, E. (2011). The Lean Startup: how today's entrepreneurs use continuous innovation to create radically successful businesses. Crown Books.

Shepherd, D. A., Haynie, J. M., McMullen, J. S. (2012). Confirmatory search as a useful heuristic? Testing the veracity of entrepreneurial conjectures. Journal of Business Ventures, 27(6), 637-651.

Siemer, J., Angelides, M. (1995). Evaluating intelligent tutoring with gaming simulations. Proceedings of the 1995 Winter Simulation Conference, (p. 1376- 1382).

Silva, J. (2004). Venture capitalists' decision-making in small equity markets: A case study using participant observation. Venture Capital, 6(2-3), 125-145.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. Personnel Psychology(64), 489-528.

Tobias, S., Fletcher, J. (2007). What research has to say about designing computer games for learning. Educational Technology(47), 20-29.

Vogel, J., Vogel, D., Cannon-Bowers, J., Bowers, C., Muse, K., Wright, M. (2006). Computer gaming and interactive simulations for learning: A meta-analysis. Journal of Educational Computing Research(34), 229-243.

Womack, J., Jones, D., Roos, D. (1990). Machine that changed the world. Simon ans Schuster.

Zenger, T. (2016). Beyond competitive advantage: How to solve the puzzle of sustaining growth while creating value. Harvard Business Review Press.