Ashoka University Economics
Discussion Paper 64

"Believe... and you're there." On self-confidence and feedback

July 2021

Zeinab Aboutalebi, European University Institute
Ayush Pant, Ashoka University

https://ashoka.edu.in/economics-discussionpapers

Plot No. 2, Rajiv Gandhi Education City,
National Capital Region,
P.O.Rai, Sonepat, Haryana 131029

"Believe ... and you're there." On self-confidence and feedback*

Zeinab Aboutalebi
Department of Economics,
European University Institute †

Ayush Pant
Department of Economics,
Ashoka University ‡

June 30, 2021

Abstract

Employees are often assigned tasks comprising an idea generation phase followed by an idea implementation phase. Furthermore, it is common for supervisors to give feedback to their employees during this process. Giving feedback involves the following tradeoff: while honest feedback encourages employees to discard bad ideas, it can also be demotivating. We derive results on how the agent's self-confidence influences the supervisor's feedback and her performance. First, the supervisor only gives honest feedback to agents who believe in their ability to succeed. Second, receiving honest feedback leads such high self-opinion agents to succeed more frequently. Third, overconfidence is potentially welfare improving. Our results suggest that reducing the gender confidence gap through confidence-building of women can increase welfare and improve performance.

Keywords: Experimentation; Feedback; Dynamic cheap-talk

JEL Classification Numbers: C73, D82, D83, O32, M5

---

*This work was supported by the Economic and Social Research Council [grant number ES/J500203/1]. We would like to thank Robert Akerlof, Motty Perry, Kobi Glazer and Ilan Kremer for their time and insightful comments. We are also grateful to Dan Bernhardt, Costas Cavounidis, Daniel Habermacher, Qingmin Liu, Phil Reny, Federico Trombetta, and seminar participants at Warwick Microeconomic Theory Work-in-Progress seminars, European Winter Meeting of Econometric Society 2017 and Israel Game Theory Conference 2018 for additional comments and clarifications. All errors are our own.

†Email: zeinab.aboutalebi@eui.eu

‡Email: ayush.pant@ashoka.edu.in

"Believe you can and you're halfway there."
— Theodore Roosevelt

1 Introduction

Consider a University professor's dilemma who supervises and provides feedback to his graduate student on the ideas she produces. On the one hand, honest feedback encourages the student to discard bad ideas. On the other hand, such feedback can be demoralizing and discourage both idea generation and effort implementation. We build a model to describe how this tradeoff shapes the supervisor's feedback, the student's trust in the supervisor, her effort, and the overall success in her project.

Similar situations frequently arise in organizations where supervisors hold only soft authority over the employees. While supervisors may not directly order their employees to take specific actions, they may still direct employees' actions due to some informational advantage (Bolton and Dewatripont (2012)). For instance, a partner in a law firm supervises an associate developing a litigation strategy, a project manager in a technology firm supervises an engineer solving a bug in app development, and a senior designer in an architecture firm supervises a junior designer looking for a design solution. In all these examples, an experienced supervisor is better-informed about whether the agent's ideas are more or less likely to succeed.

In Section 2, we present a simple supervisor-agent model with two phases – experimentation and implementation. In the experimentation phase, the agent sequentially generates ideas at a cost, receives feedback from the supervisor on her ideas, and selects an idea to implement. In the implementation phase, the agent decides how much effort to put into completing her chosen idea. The agent's ability is initially unknown, and the agent and supervisor may share different priors. Importantly, we assume the supervisor does not internalize the agent's cost of effort. This misalignment of preferences means that dishonesty is a possibility.

Ability plays a central role in our model. We assume a high-ability agent both generates and implements ideas better than a low-ability agent. As a result, the agent's self-opinion (or self-confidence), belief about her ability affects both the agent's decision regarding how much to experiment and her choice of implementation's effort. Both of these effects, in turn, impact the supervisor's feedback.

After conducting benchmark analysis in Section 3, we present our key results in Section 4. These deal with the differing behavior of the supervisor towards low and high self-opinion agents and the resulting differences in performances. The supervisor never gives a low self-opinion agent honest feedback because doing so is demotivating; it discourages effort in both the experimentation and implementation phases. When negative feedback discourages further experimentation, the supervisor prefers to falsely encourage the agent to induce her to put a higher effort into implementation instead. Therefore, negative feedback is only forthcoming for a high self-opinion agent.

Consequently, receiving supervisor feedback magnifies performance differences between high and low self-opinion agents. Because high self-opinion agents receive honest feedback, they have confidence in their ability and the quality of their ideas, leading to high effort. Receiving more honest feedback with a higher self-opinion allows the agent to experiment more and exert an optimal effort in implementing her chosen idea. In this sense, the supervisor can lead a halfway there self-confident high self-opinion agent all the way to success more often in expectation. Such an opportunity might not be available to a slightly lower self-opinion agent. She has lower confidence in her ideas. Therefore, she might exert too much effort on a bad idea and too little effort on a good idea.

Our paper, therefore, provides new intuition on how a confidence gap between men and women in research-driven industries translates into a performance gap. A vast literature documents such a gender confidence gap. Goodman, Cunningham, and Lachapelle (2002), for instance, emphasize that the main cause of female dropouts from STEM fields is their lack of confidence in their abilities to pursue degrees in these fields.[1] We emphasize that the pathway to underperformance of women with a lower self-confidence is through the lack of willingness of the supervisors to be critical. McCarty (1986) looks at the role of confidence in the reception of feedback in creative tasks and finds that women are less confident than men and that feedback does not reduce this gap.

Such performance differences between less and more confident agents have essential welfare and policy implications. Notably, confidence-building exercises before starting the project that may *incorrectly* raise employees' confidence in their ability can be welfare improving. The discontinuous change in the supervisor's feedback strategy as the agent incorrectly goes from a low self-opinion to a high self-opinion gives rise to this possibility. The cost of overconfidence in their ability is that it leads to too much effort exertion. However, the benefit of overconfidence is that it can lead to honest feedback. This benefit may outweigh the cost.

The examples illustrated above deal with early-career employees who also gain experience through supervision. In fact, one reason why organizations may create such supervisor-agent relationships is to facilitate the agent's learning (Fudenberg and Rayo (2019)). In Section 5, we extend our model to include agent's learning-on-the job with honest supervision. We show here that the magnifying effect of higher self-opinion is still larger. More confident agents may experiment to their maximum limit, while those with lower confidence might not even experiment once.

Related Literature. Our paper connects three strands of literature: experimentation, dynamic feedback, and dynamic communication games. Within experimentation, our work falls under models of motivating experimentation. Broadly, the literature assumes supervisor commitment to ex-ante information policy when analyzing such questions. Ely and Szydlowski (2017), Smolin (2017) and Ali (2017) all require a principal to

[1] Niederle and Vesterlund (2007), Gneezy and Rustichini (2004), Gnther, Arslan Ekinci, Schwierenc, and Strobeld (2010) and Shurchkov (2012) attribute women's underperformance and unwillingness to participate in competitive tasks to confidence gap.

design an optimal information policy to balance the positive effect of good news with the discouraging effect of no or bad news. Unlike these papers, our model produces nontrivial dynamics even with strategic communication. To the best of our knowledge, we are the first to study such settings without commitment.

The literature on dynamic feedback has also primarily been studied with commitment and without experimentation. Orlov (2013) considers a setting in which providing information to the agent might benefit the principal in the short-run but may lead to long-term agency costs. The principal designs an optimal information sharing rule along with a compensation scheme. Boleslavsky and Lewis (2016) also study long-term advisory relationships with commitment in which the agent has new information every period. The principal makes sequential decisions, after which he observes a private signal of the state.[2]

Among literature on dynamic communication games, we are concerned with persuasion in two phases.[3] In this respect, Honryo (2018) and Henry and Ottaviani (2019) are similar to our setting. In these papers, a sender (entrepreneur or researcher) tries to persuade a receiver (venture capitalist or publisher) to take a favorable action by sequentially disclosing some verifiable or costly information. However, we can generate a tradeoff for the sender without assuming verifiability or costly information transmission. In our model, when the supervisor persuades the agent to experiment again, he inadvertently persuades her to exert lower effort in implementation.

Finally, our result on the importance of beliefs in final performance is related to some of the older research starting with Bénabou and Tirole (2002). This vast line of economics research is itself based on Bandura (1977) in the field of psychology. However, such research usually looks at the importance of belief absent any external supervision.

[2] Orlov, Skrzypacz, and Zryumov (2018) is an exception. They look at commitment and no commitment cases in a setting where an agent tries to convince the principal to wait before exercising a real option. Again, however, their model does not have experimentation.

[3] Golosov, Skreta, Tsyvinski, and Wilson (2014) and Renault, Solan, and Vieille (2013) look at situations where the receiver decides after every round of communication. However, neither has the feature of persuasion in two phases.

The presence of a supervisor drives our results on the effect of higher self-opinion and overconfidence.[4]

2 A model of feedback on ideas

An agent (she) works on a project with a supervisor (he). The project involves two distinct stages that occur sequentially. The first stage involves generating or experimenting with ideas, and the second stage requires implementation of the chosen idea. The agent exerts effort toward success in the project in both stages, while the supervisor provides feedback to the agent in the first stage. The supervisor has no commitment power and provides feedback based on what he observes. Success generates positive benefits to both the agent and the supervisor. Only the agent pays the cost of effort.

Stage 1**: Experimentation with ideas.** The process of idea generation involves multiple rounds $r = 1, 2, \ldots$. In each round r, the agent decides whether she wants to draw a new idea. An idea is defined by its quality, q, which could be either good, g, or bad, b, with $0 < b < g \leq 1$. The ability, a, of the agent determines the quality of the idea drawn, which could be either "high," $a = 1$, or "low," $a = 0$. We let $\Pr(q_r = g|a = 1) = \theta \in (0, 1)$ and $\Pr(q_r = g|a = 0) = 0$. Therefore, we assume that only a high ability agent can come up with a good idea in any round.

The ability, unlike the idea, remains persistent throughout the game and is initially unobserved to both players. β_r and B_r respectively denote the agent's and the supervisor's belief about the agent being high-ability at the beginning of round r. Both are non-degenerate and there is common knowledge about the heterogeneous priors.

We assume that the agent possesses a bad outside option idea at the beginning denoted by $q_0 = b$.[5]

[4] Koellinger, Minniti, and Schade (2007) and Hirshleifer, Low, and Teoh (2012) are two papers that empirically show the importance of overconfidence in the context of innovation and creativity.

[5] This assumption is not necessary for the analysis, but helps make relevant comparisons for the agent at any stage. In particular, the first stage decision is the same as all potential future decisions

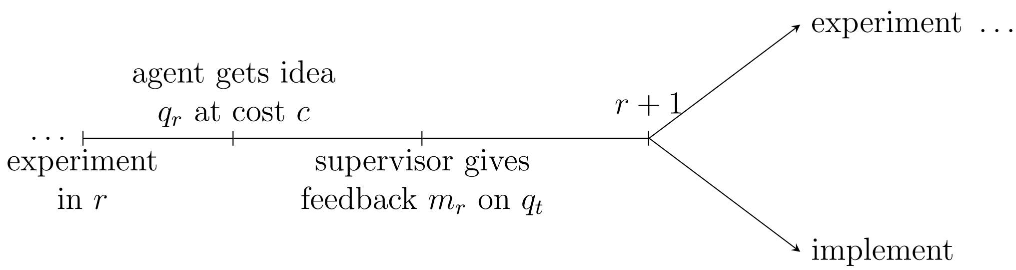


Figure 1: Summary of timing when the agent chooses d_r = experiment

Stage 1 actions and timing. At the beginning of each round of experimentation, the agent decides whether to experiment again or to implement her last idea, i.e., $d_r \in \{\text{experiment}, \text{implement}\}$.

d_r = experiment denotes the agent's decision to stay in Stage 1 and experiment with another idea in round r. Doing so costs the agent c and a new idea is realized.[6] The supervisor bears no cost from experimentation and privately observes the quality q_r of the idea drawn.[7] He then sends a costless message, $m_r \in \{b, g\}$ about what he observed.[8]

Alternately, the agent may choose d_r = implement to move to the second stage in round r. The move is permanent and the agent cannot return to experimenting with ideas again. Figure 1 summarizes timing and actions in Stage 1.

***Stage 2*: Implementation of idea.** Choosing to implement in round $r+1$ means that the agent must implement her last idea, q_r. Thus, we assume that the agent cannot implement an idea abandoned in any of the previous rounds.

Stage 2 actions and timing. When implementing in round $r+1$, the agent exerts

– implement a bad quality idea or experiment again.

[6]Such time and effort costs could arise from seeking inspiration, making online searches, looking up for data, reading material and exploring the literature.

[7]In Section 5 we relax the assumption that quality gets revealed only to the supervisor. We extend the analysis to the case were the agent gradually gains the ability to understand the quality of her own ideas without needing a supervisor. In general, however, it is reasonable to assume that the supervisor is better equipped to understand the quality of ideas that early-career agents may generate. This may be the case due to more experience. Further, we show in Appendix D how the results change when the supervisor also bears cost of providing feedback.

[8]This restriction is without loss of generality as the type space is binary and what matters are the equilibrium mappings from the supervisor type to the message space, i.e., what is the meaning of the messages. Here, messages b and g have their natural meaning and are understood as the idea quality.

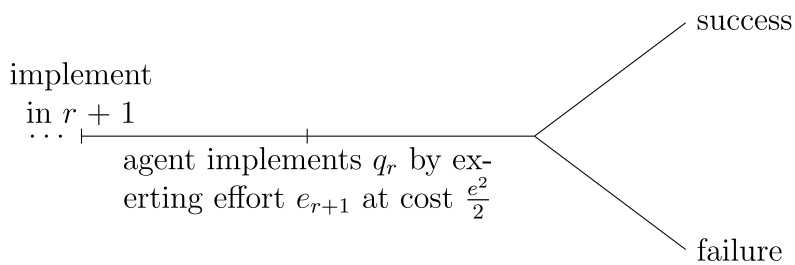


Figure 2: Summary of timing when agent chooses $d_{r+1} =$ implement

effort $e_{r+1} \in \mathbb{R}^+$ at cost $\frac{e_{r+1}^2}{2}$ to complete the project. Completion leads to success or failure of the project such that $\Pr(\text{success}) = aq_r e_{r+1}$. Figure 2 summarizes timing and actions in Stage 2.

Payoffs. The project yields a payoff of $V \in \{0, 1\}$ with success yielding a payoff of $V = 1$ and zero otherwise. The payoff of the agent is given by $u_A = V - Rc - \frac{e_{R+1}^2}{2}$ where R is the number of rounds the agent experimented. The payoff of the supervisor is given by $u_S = V$. The payoffs highlight the incentive misalignment between the agent and the supervisor. While both players prefer success over failure, the agent alone bears the costs of experimentation and implementation. Once the payoffs are realized, the game ends.

2.1 Discussion of the model

The success probability function. We assume that the success function exhibits complementarities between the quality of the idea, q, effort exerted by the agent, e, and the agent's ability, a. All three need to be simultaneously "large" in order to maximize the probability of success. Such functional form helps generate the encouraging-discouraging tradeoff of feedback in our model.

Observe that a low-ability agent can never succeed and a high-ability agent can turn a bad idea into a success. Thus, a known low-ability agent does not exert effort in either stage. In contrast, a known high-ability exerts maximal effort in both stages. Nevertheless, given that ability is unknown, the agent may want to experiment, learn about her ideas by seeking feedback, and simultaneously learn about her underlying ability. Learning, therefore, produces an encouraging effect of negative feedback – to further learn by experimenting again.[9]

At the same time, the decision d_r will depend on the belief of the agent about her ability due to our success function. A lower belief will usually imply a lower likelihood of subsequent experimentation *and* a lower implementation effort. The two effects together capture the discouraging effect of negative feedback. In this way, our model uniquely ties the first stage of the game, which determines not only the quality of idea but also the belief about ability, with the second stage of the game.

Thus, the supervisor must balance the discouraging effect on the agent from finding out that the current idea is bad with the potentially positive effect of learning.[10]

Self-confidence. We assume that in general the priors of the two players, $\beta_1, B_1 \in (0, 1)$, are not the same. We interpret heterogeneous prior as the difference in confidence on the agent's ability. We will often refer to β_r as agent's *self-confidence* (in her ability). Doing so is in line with Bénabou and Tirole (2003) where the agent has imperfect knowledge of her ability and the belief influences her actions. A principal in their model manages her belief by assigning her to different tasks. However, unlike them, we assume that the principal is also imperfectly informed about the agent's ability, and can manage the agent's belief only using his feedback.

Below we present the main analysis and discuss our results for the case of $g = 1$. The general case is relegated to the Online Appendix.

[9] At the same time, there is an incentive for the agent to implement a bad quality idea by "taking a chance" on it for a high ability agent can succeed with a bad idea.

[10] Assuming substitutability in the success function along any dimension kills part of our tradeoff. For instance, if ability and effort appear additively, then optimal implementation effort does not depend on the belief about ability. The agent could potentially be successful by exerting maximal effort from the get go without "wasting" effort on experimentation. Similarly, if quality and effort were substitutes, then again a bad outcome in experimentation could be made up by exerting a high effort.

3 Benchmark: Two special cases with supervisor commitment

We start our analysis by considering the case of the supervisor's ex-ante commitment to information policies and look at the agent's behavior under two specific policies: no information and full information about the quality of ideas. This preliminary analysis helps us put bounds on the behavior of the agent and supervisor when they interact with each other strategically without supervisor commitment.[11]

3.1 No information policy

If the supervisor provides no information about the ideas, the agent does not learn anything. She neither learns the quality of her idea, q, nor she updates on her ability, β. She may, however, still want to experiment once as a gamble if her self-confidence sufficiently high.[12] The relevant payoff comparison to make the agent experiment is given by

$$-c + \frac{(\beta_1 Q)^2}{2} \geq \frac{(\beta_1 b)^2}{2} \tag{C1}$$

such that the agent chooses an implementation effort of $e = \beta_1 \mathbb{E}(q|a = 1)$, where $\mathbb{E}(.|.)$ is the conditional expectation operator, and $\theta + (1 - \theta)b := Q$. Note that Q is the expected quality of an idea from experimenting when the agent is high ability. See Figure 3.

Lemma 1 *If* $c < \frac{Q^2-b^2}{2}$ *and there is no information about the idea quality* q*, there exists a unique threshold* $N_0 := \left(\frac{2c}{Q^2-b^2}\right)^{\frac{1}{2}}$ *such that*

[11] The purpose of this section is *not* to determine the ex-ante optimal information policy for the supervisor. Doing so is beyond the scope of the paper. Nonetheless, we present some intuitive results on the information design problem in Appendix C.

[12] Observe that the agent does not want to experiment more than once because experimenting is an additional cost without any added benefit.

1. *if the prior belief* $\beta_1 \geq N_0$ *then the agent experiments once and implements her idea by exerting effort* $\beta_1 Q$*; and*

2. *if the prior belief* $\beta_1 < N_0$*, the agent does not experiment and implements her outside option idea* $q_0 = b$ *by exerting effort* $\beta_1 b$*.*

If $c \geq \frac{Q^2 - b^2}{2}$ *and there is no information about the idea quality* q*, the agent does not experiment for any belief. She implements her outside option idea with an effort* $\beta_1 b$*.*

3.2 Full information policy

If the supervisor reveals the idea, then the agent can learn about her ability (besides learning about the quality of her ideas). Using Bayes' rule,

$$\beta_r = \begin{cases} \frac{(1-\theta)\beta_{r-1}}{1-\beta_{r-1}\theta} & \text{if } q_{r-1} = b, \\ 1 & \text{otherwise.} \end{cases}$$

The agent revises her belief downwards each time she learns that she has produced a bad quality idea, but her belief jumps to 1 if she learns that her idea is good. Accordingly, the agent implements her idea with an effort of $\beta_r q$.

Let the value function of the agent at the beginning of round r with belief β_r when her last observed outcome is $q_{r-1} = b$ be $\mathcal{V}^b(\beta_r)$, such that

$$\mathcal{V}^b(\beta_r) = \max\left\{\frac{(\beta_r b)^2}{2}\ ,\ -c + \frac{\beta_r \theta}{2} + (1 - \beta_r \theta)\mathcal{V}^b(\beta_{r+1})\right\}.$$

Note that the agent does not experiment any further after coming up with a good idea. At the same time, she faces the same decision problem as the original one if she comes up with a bad idea, but with a lower belief.

Assuming that the agent wants to start experimenting, we are interested in if and when the agent stops experimenting with repeated bad ideas. Using the one-step look-ahead rule, if the agent finds it optimal to stop experimenting at a belief β, then so does she after another round of experimentation. So,

$$\mathcal{V}^b(\beta) = -c + \frac{\beta\theta}{2} + (1-\beta\theta)\mathcal{V}^b(\beta') < \frac{(\beta b)^2}{2} \text{ where } \beta' = \frac{(1-\theta)\beta}{1-\beta\theta} \text{ and } \mathcal{V}^b(\beta') = \frac{(\beta' b)^2}{2}.$$

Thus, the agent experiments for beliefs that satisfy the condition

$$-c + \frac{\beta\theta}{2} + (1-\beta\theta)\frac{(\beta' b)^2}{2} \geq \frac{(\beta b)^2}{2}. \tag{C2}$$

See Figure 3. Lemma 2 follows from condition (C2) and captures the optimal behavior of the agent under full information policy.

Lemma 2 *If $c < \frac{\theta(1-b^2)}{2}$ and there is full information about the idea quality q, there exists a unique belief threshold F_0 that solves condition (C2) with an equality such that*

1. *if the agent's previous idea was bad, $q_{r-1} = b$, she experiments for any belief $\beta_r \geq F_0$, but implements it with an effort of $\beta_r b$ for any belief $\beta_r < F_0$; and*

2. *if the agent's previous idea was good, $q_{r-1} = g$, the agent implements her idea in round r with a maximal effort of 1.*

If $c \geq \frac{\theta(1-b^2)}{2}$ and there is full information about the idea quality q, the agent does not experiment for any belief and implements her outside option idea with effort $\beta_1 b$.

All proofs that appear in the main text are presented in Appendix A.

3.3 Comparing no information and full information policies

We first compare the belief thresholds induced by the two policies.

Lemma 3 *If $c < \frac{Q^2-b^2}{2}$, then both N_0 and F_0 exist and are unique with $N_0 > F_0$.*

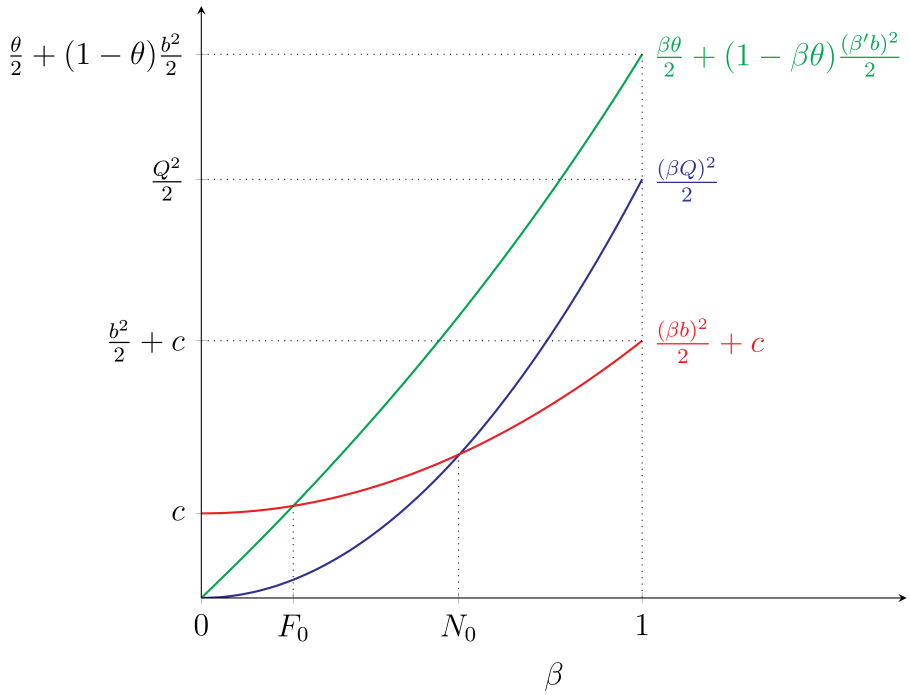


Figure 3: Comparing N_0 and F_0

Figure 3 illustrates why $N_0 > F_0$. It shows that for any belief the value of a final round of experimentation is lower under no information policy owing to no learning. Therefore, as opposed to when she has full information, the agent does not experiment when there is no information for certain beliefs. Accordingly, the cost condition associated with the no information policy, $\frac{Q^2-b^2}{2}$, binds as it is smaller. Unless otherwise stated, we will continue assuming this cost condition for the rest of the analysis.

Next, we consider the welfare implication of the two policies. The welfare is evaluated as the ex-ante expected utility of the players. Observe that the supervisor always prefers full information owing to increased probability of success. Under full information, the agent experiments more often, which increases the overall expected probability of success. Nonetheless, given that the agent also pays for these costs, it is not apparent whether she prefers the full information policy. Our first proposition below highlights that this is indeed the case.

Proposition 1 *For any prior belief $\beta_1 \geq F_0$, the agent (strictly) prefers the full infor-*

mation policy to the no information policy.

The reason is that a higher self-confidence (particularly any belief above F_0) increases the odds of producing a good idea. It makes the agent willing to experiment and pay the cost under full information policy. Such experimentation would not be possible under no information policy despite the agent believing that she can achieve success. Consequently, both the agent and the supervisor are worse-off when the supervisor does not provide information on ideas. We, therefore, seek to determine if and when the supervisor can engage in mutually beneficial truthful communication without commitment.

3.4 An important definition

For any arbitrary belief β_r and $j = 1, 2, \ldots$, let $\phi_j(\beta_r) = \beta_{r+j}$ be the belief after j rounds of application of Bayes' rule on β_r when the ideas were bad in all j rounds. Inverting $\phi_j^{-1}(\beta_{r+j}) = \beta_r$ gives the starting belief β_t that results in β_{t+j} as the terminal belief after j rounds. Define $\phi_j^{-1}(N_0) := N_j$ and $\phi_j^{-1}(F_0) := F_j$ for $j = 1, 2, \ldots$ for the no information and full information belief thresholds, N_0 and F_0 respectively. Therefore, N_j (F_j) is the starting belief which when correctly updated about bad ideas j times leads to the terminal belief N_0 (F_0).

4 Strategic supervisor

4.1 Preliminaries

The game between an agent and a strategic supervisor in the experimentation stage is one of dynamic cheap talk. The supervisor can costlessly provide any feedback independent of the true quality of the idea. Our solution concept is (perfect) Bayesian Equilibrium.

Round r begins for the agent after having observed the last message sent by the supervisor m_{r-1}. Therefore, a realized history for the agent includes the set of all previous messages sent by the supervisor up to and including m_{r-1}, and the sequence of past decisions made. Round r begins for the supervisor after observing the last idea of the agent q_r. Accordingly, a realized history for the supervisor includes, in addition to the history viewed by the agent, the sequence of all the realized ideas from past experimentation attempts.[13]

As mentioned earlier, we are interested in the existence and the properties of the truthful equilibria. Specifically, our aim is to identify the equilibria that induce maximum honesty, i.e., equilibria in which honesty persists for lowest possible beliefs. These are pure strategy equilibria. A pure strategy for the supervisor in round r is a mapping from the realized history to the message space $\{b, g\}$. If the agent expects the supervisor to be honest at belief β_r in round r, the agent's updated posterior in round $r+1$ is as in the full information case; $\beta^b_{r+1} = \frac{(1-\theta)\beta_r}{1-\theta\beta_r}$ if $m_r = b$ and $\beta^g_{r+1} = 1$ otherwise.

If the supervisor uses the same message independent of the realized history, the supervisor is said to lie or babble. In this case, the agent's posterior belief is the same as her prior belief. We will assume that the agent does not consult the supervisor when she expects the supervisor to lie. This rules out the possibility of the supervisor privately learning and not revealing to the agent the outcome and the arising deviations.

[13]Let $d^r := (d_1, \ldots, d_r)$ and $m^r := (m_1, \ldots, m_r)$ be the sequence of decisions made by the agent and the public messages given by the supervisor until round r. Define the set of histories for the agent and the supervisor at the beginning of round r by H^A_r and H^S_r respectively. The history for the agent at the beginning of round r is

$$h^A_r = (d^{r-1}, m^{r-1}) \in H^A_r \subset (\{\text{experiment}\}^{r-1} \times \{b, g\}^{r-1}).$$

This is also the public history of the play of the game up to round r. In addition to the public history, the supervisor observes $q^r := (q_1, \ldots, q_r)$ and an extra decision of the agent to experiment $d_r =$ experiment. The history for the supervisor at the beginning of round r is

$$h^S_r = (\theta^r, d_r, h^A_r) \in H^S_r \subset (\{b, g\}^r \times \{\text{experiment}\}^r \times \{b, g\}^{r-1}).$$

4.2 Analysis

Note that babbling is always an equilibrium for any belief β. The agent does not learn about the quality of her idea as the supervisor is expected to give uninformative feedback. This is equivalent to the single agent decision-making problem without information and Lemma 1 applies. Expectations of supervisor babbling are self-fulfilling and neither party can profitably deviate.In what follows we determine if truthful equilibria exist (in addition to babbling) for different ranges of beliefs starting with lower ones.

Proposition 2 *For any belief $\beta < F_0$, any communication strategy is an equilibrium and none induces the agent to experiment.*

The region of beliefs $\beta < F_0$ reflects extreme pessimism. Independent of how much information the supervisor provides, the agent neither experiments with ideas nor does she consult the supervisor. She simply implements her bad outside option idea with an effort βb.

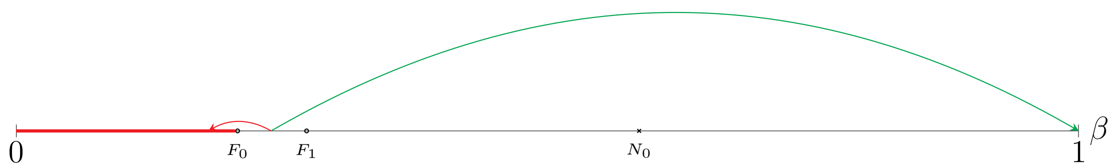
What feedback strategy the supervisor employs for higher beliefs depends on the possibility of the agent experimenting after negative feedback. Our first main result defines the range of beliefs for which babbling is the unique equilibrium owing to the agent abandoning experimenting with ideas. We say that in this region of beliefs the agent holds a *low self-opinion.*

Proposition 3 *If $c < \frac{Q^2-b^2}{2}$, for any belief $F_0 \leq \beta < N_1$, babbling is the unique equilibrium strategy. If $c \geq \frac{Q^2-b^2}{2}$, then babbling is the unique equilibrium strategy for all beliefs.*

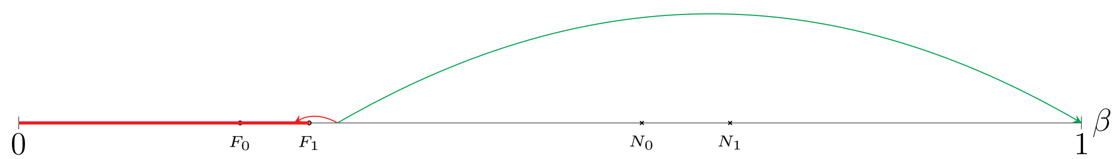
The intuition for this proposition is illustrated in steps using Figure 4.[14] Suppose $c < \frac{Q^2-b^2}{2}$ so that both F_0 and N_0 exist and are below 1. Consider first the range of beliefs

[14]Here we discuss the intuition of why honesty cannot be an equilibrium strategy but the proposition is stronger. The argument will also hold for partially informative strategies. We provide a general proof in Appendix A after describing mixed strategies.

Step 1: Babbling is unique for $F_0 \leq \beta < F_1$



Step 2: Babbling is unique for $F_1 \leq \beta < N_0$



Step 3: Babbling is unique for $N_0 \leq \beta < N_1$

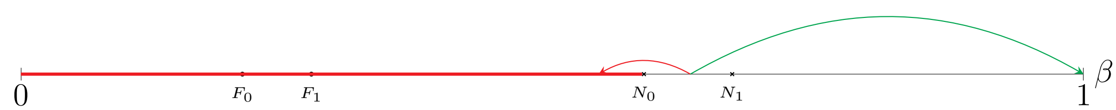


Figure 4: Uniqueness of babbling equilibria for priors $F_0 \leq \beta < N_1$

$F_0 \leq \beta < F_1$ in Step 1 of Figure 4. If the agent expects negative feedback in equilibrium, her posterior belief following such feedback falls below F_0. Here, the agent abandons experimentation and chooses a lower implementation effort of $\beta^b_2 b$ with her updated belief. By deviating and providing positive feedback, however, the supervisor induces a maximal implementation effort of 1. Thus, the supervisor always finds it beneficial to encourage the agent at this stage and truth-telling will not survive. Babbling is unique in this range of priors.

As a result of babbling, the agent best responds by not experimenting because this is identical to a situation with no information and the prior is below N_0 (from Lemma 1). Now, the previous argument applies to the range of beliefs which when updated negatively lead to posteriors below F_1 (see Step 2). In fact, the same logic can now be extended to all the beliefs that lead to posteriors below N_0 after a negative feedback. Such is the case for all prior beliefs beliefs below N_1 (illustrated in Step 3).

Finally, it is straightforward to see that if $c \geq \frac{Q^2 - b^2}{2}$, then an interior N_0 does not exist. As a result, the agent always responds to no information by implementing the outside option. Given the agent's best response, the supervisor always wants to encourage the student instead.

This region of low self-opinion beliefs reflect inefficiencies in the supervisor-agent relationship. From Proposition 1 we know that both the supervisor and the agent prefer full information to no information. Yet, the perceived conflict of interest arising due to a combination of costs of experimentation and low self-opinion is too large to sustain communication.

Moving to higher beliefs, the possibility of informative feedback opens up due to the agent best response to the previous babbling equilibrium when $c < \frac{Q^2-b^2}{2}$. From Lemma 1, we know that the agent experiments once in the region between N_0 and N_1. The previous threat point for the supervisor now potentially disappears. Our next result determines whether this one extra round of experimentation (without the supervisor's assistance) is sufficient for the supervisor to be honest. We will say the agent in this region of beliefs has a *high self-opinion.*

Proposition 4 *Let* $c < \frac{Q^2-b^2}{2}$. *If* $b < \left(\frac{\theta}{1-\theta}\right)^2$ *so that* $b < Q^2$, *a threshold belief level* $\beta^{truth} := \frac{b}{qb+(1-q)Q^2} < 1$ *exists such that*

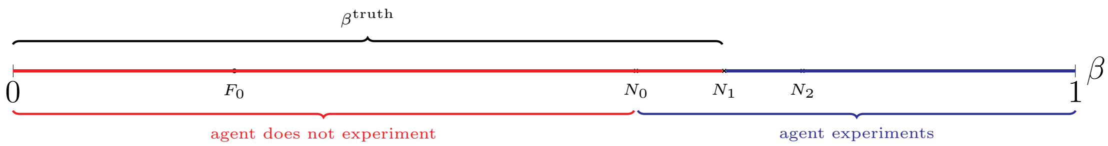
1. *the supervisor provides honest feedback to the agent for all beliefs greater than* N_1 *for* $\frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right) \leq c < \frac{Q^2-b^2}{2}$ *where* $\beta^{truth} < N_1$, *and*

2. *the supervisor provides honest feedback to the agent for all beliefs greater than* β^{truth} *for* $c < \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right) < \frac{Q^2-b^2}{2}$ *where* $\beta^{truth} > N_1$.

If $b \geq \left(\frac{\theta}{1-\theta}\right)^2$ *so that* $b \geq Q^2$, *there are no honest equilibria for any belief* $\beta < 1$ *in which the agent experiments once without feedback in the following round.*

See Figure 5. The first part of Proposition 4 shows that a sufficiently high cost of experimentation generates honest equilibria in the *entire* region of high self-opinion

1. Honest equilibria when $c \geq \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$



2. Honest equilibria when $c < \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$

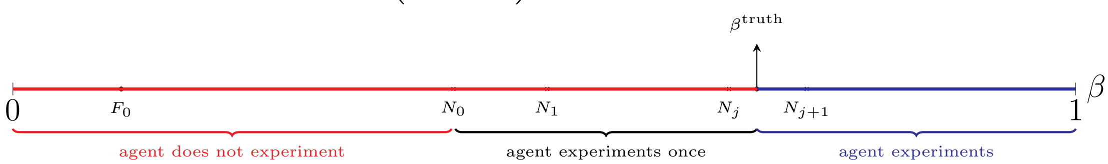


Figure 5: Honest equilibria for different c ranges

beliefs. To see this, let us look at the supervisor's incentives to be honest in the region of beliefs between N_1 and N_2.

A supervisor who has observed a bad idea expects the project to be successful with probability $B_2^b \beta_2^b Q^2$ from being honest when the agent's effort on implementation is $e = \beta_2^b Q$. On the other hand, deviating leads to an expected probability of success of $B_2^b b$ such that the agent exerts an effort of 1 on her bad idea. For such a conjectured strategy to be an equilibrium, we must have that

$$B_2^b \beta_2^b Q^2 \geq B_2^b b \iff \beta_1 \geq \frac{b}{\theta b + (1-\theta)Q^2} := \beta^{\text{truth}} \in (0,1) \text{ if } b < Q^2. \tag{1}$$

The agent's high prior ensures that she exerts a higher effort in implementing her idea of unknown quality, which "buys" supervisor honesty. We call this truth-telling threshold β^{truth}.

Note that the supervisor does not directly care about the agent's cost of experimentation in so far as she attempts to experiment again with an idea. So, β^{truth} does not depend on c. However, the agent experiments without supervision only when her belief is larger than N_0. In addition, we want to identify the condition that guarantees honest equilibria for all beliefs above N_1. The two conditions are simultaneously satisfied if

$\beta^{\text{truth}} \leq N_1$, which can be simplified as $c \geq \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$.

What happens when $c < \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$? In this case, $\beta^{\text{truth}} > N_1$ and can lie between any N_j and N_{j+1}. We can then again construct an honest equilibrium above β^{truth} and a babbling one below. That all of these beliefs are above N_0 ensures that the agent experiments once more when a bad idea is revealed to her and makes such a strategy an equilibrium.

It is important to not confuse our result as stating that the belief region for truthful equilibria expands due to an increase in cost. An increase in c only moves the belief threshold N_1 to match β^{truth}, which is not dependent on the cost. In fact, the belief region for truthful equilibria remains the same as long as $c \leq \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$. An increase in cost beyond this threshold reduces the region of beliefs for truthful equilibria.

It is worth emphasizing that the equilibria we outlined in all the previous propositions are *not* contingent on the supervisor's belief B_1. They are only a function of the agent's self-opinion, β_1. The reason is twofold. The first is the choice of our success function, and the second is that the supervisor does not bear any costs. Together they imply that the supervisor's belief does not matter in his truth-telling incentive constraint. Intuitively, this makes sense because all that the supervisor cares about is maximizing the probability of success. This probability is determined solely by the agent's decisions, which in turn are a function of her self-opinion.[15]

We conclude this section by presenting an important implication of the equilibrium analysis and a key result.

Corollary 1 *The expected probability of success of the agent is larger under a higher self-opinion.*

To see this, note first that the supervisor induces a weakly higher number of rounds

[15]There is no reason in our model for the agent to consider the supervisor's belief. He does not possess any additional prior information that may lead her to believe that the supervisor's belief is more precise. The two players simply agree to disagree on the prior belief.

of experimentation under a higher self-opinion. More rounds of experimentation with more honest feedback increase the chances of securing a good idea.

Second, at the same time, more honest feedback allows the agent to match her implementation effort more closely to the actual idea, thereby maximizing the probability of success. An agent who abandons seeking supervision with lower self-opinion will exert an inefficient implementation effort on her idea. On the other hand, the higher self-opinion agent at the same stage will either exert optimal implementation effort (guaranteeing success) or generate a new idea. Therefore, there is a magnifying effect of a higher belief that results from the combined effect of better experimentation and better implementation.

4.3 Welfare effect of "overconfidence"

We are interested in the welfare analysis arising from a specific policy that organizations may implement, namely to invest in the agent's confidence building. Suppose the agent enters the relationship with the supervisor, or joins the organization, with an *exante prior* belief of β_1. However, before the project starts, the organization can influence the agent's prior belief.[16] We call this new influenced belief the agent's *interim prior* belief, $\tilde{\beta}_1$. Assume that the agent starts working on the project with her interim belief.

Definition 1 *The agent is overconfident about her ability (relative to her exante prior belief) if $\tilde{\beta}_1 > \beta_1$, i.e. her interim belief is larger than the ex-ante belief.*

We want to determine whether the agent would ex-ante prefer to be overconfident in the interim. Said another way, would an agent prefer the organization to build her self-confidence? To answer this question, we evaluate the ex-ante expected utility of the agent under a different interim belief, $W(\tilde{\beta}_1; \beta_1)$. The function $W(.)$ assumes that the

[16] The organization may achieve this through workshops that help build the agent's self-confidence, or by showing her evidence that would lead her to update her ex-ante belief.

agent behaves and makes decisions according to the interim belief but it measures the weights of different outcomes using the ex-ante belief. With some abuse of notation, denote $W(\beta_1; \beta_1)$ by $W(\beta_1)$.

In most contexts, there would be no reason for $W(\tilde{\beta}_1; \beta_1) > W(\beta_1)$. Since this problem is akin to choosing your prior belief, one would not find it beneficial to choose a prior belief different from their actual (or correct) belief.[17] Yet, we show that this is not always the case in our model.

Proposition 5 *Suppose* $\min\left\{\left(\frac{b}{1-b}\right)^2, \frac{\sqrt{b}}{1+\sqrt{b}}\right\} < \theta \leq .5$ *so that* $b < Q^2 < \theta \leq 0.5$*. Then there exists an ex-ante belief threshold* $\bar{\beta}$ *such that for all beliefs* $\beta \geq \bar{\beta}$*, the agent would be strictly better off under a higher interim belief that allows her at least one extra round of experimentation with supervisor feedback. Moreover,* $F_0 < \bar{\beta} < N_0$*.*

The intuition follows from the nature of the equilibrium. There is a discontinuous change in the supervisor's equilibrium strategy at N_1 – he starts being honest. Thus, there exists an agent who at her current ex-ante belief less than N_1 would not receive any feedback but will do so after incorrectly getting her belief bumped up to N_1. Supervisor's honest feedback makes it worthwhile for the agent to be overconfident. A simple corollary follows resulting from the magnifying effect of higher beliefs.

Corollary 2 *An agent with a higher ex-ante belief weakly prefers larger increases in her interim beliefs, i.e. she prefers to be "more" overconfident.*

4.4 An alternate setting

Our model fits more generally in settings where a less informed receiver (agent) seeks feedback from an informed sender (supervisor) after costly effort to decide her future course of action. Consider, for instance, an entrepreneur who works on a project

[17]Consider the problem of $\max_e \beta e - \frac{e^2}{2}$ for belief β. The optimal e is given by $e^* = \beta$, where β is the correct prior belief. Making belief β' a choice variable the maximum would occur at $\beta' = \beta$.

experimenting with ideas, *privately* observing their quality, and implementing one of them. However, he relies on a venture capitalist (VC) to finance such experimentation and implementation of the project. The VC, in turn, takes decisions based on the entrepreneur's recommendations. While the entrepreneur would prefer to continue experimenting until he receives a good idea, the VC would like to cut funding for experimentation when she is sufficiently pessimistic. In such a setting, the entrepreneur is the supervisor, while the VC is the agent.

Our model provides answers to the following questions: When can the entrepreneur credibly communicate her information? How many chances of experimentation does the VC provide the entrepreneur? Notably, our inefficiency result shows that even though the VC would like to continue financing the entrepreneur's experimentation, she calls off the project too early. However, there are benefits from the VC incorrectly believing that the project is good. In this sense, it may be worthwhile for the entrepreneur to invest in building the VC's confidence in her abilities before starting the relationship.

5 Learning-by-doing with more feedback

Our model naturally extends to an environment in which the agent better acquires the ability to understand the quality of her ideas with more feedback. Suppose the agent receives a private signal s_r of the quality of her idea after every experimentation. This signal could either correctly reflect her idea, i.e., $s_r = q_r$, or it may be empty, $s_r = \varnothing$, Let $\pi_r = \Pr(s_r = q_r)$. To capture learning, assume that at the beginning $\pi_1 = \bar{\pi} \in [0, 1)$, and conditional on receiving honest feedback in $r-1$, $\pi_r > \pi_{r-1}$, and otherwise $\pi_r = \pi_{r-1}$. Naturally, there exists a $r = \bar{r}$ where $\pi_{\bar{r}} = 1$ with repeated honest feedback. Until $\bar{r}$, it is always beneficial to consult the supervisor if the agent's own signal is empty and the supervisor provides honest feedback.

As before, start with the agent's decision-making problem for a fixed $\pi \geq 0$ without supervision. The agent may now find herself in one of the two situations following experimentation – either she observes $q_r = b$ or her signal is empty.[18] We ignore the latter decision-making problem as it does not play any role in the strategic supervisor case. Let $\mathcal{V}^b(\beta_r)$ the value function of the agent at the beginning of round r with belief β_r when her last observed outcome is $q_{r-1} = b$.

$$\mathcal{V}^b(\beta_r) = \max\left\{\frac{(\beta_r b)^2}{2}\ ,\ -c + \pi\left(\frac{\beta_r\theta}{2} + (1-\beta_r\theta)\mathcal{V}^b(\beta_{r+1})\right) + (1-\pi)\mathcal{V}^{\varnothing}(\beta_r)\right\}. \quad (2)$$

The optimal decision can again be summarized by a partial information belief threshold P_0^π. The threshold specifies the minimum belief below which the agent does not experiment further with a bad idea for a fixed π. Using the one-step look-ahead rule, it is easy to characterize the condition that will give P_0^π. Accordingly, the agent experiments with bad ideas if

$$-c + \pi\left(\frac{\beta\theta}{2} + (1-\beta\theta)\frac{(\beta' b)^2}{2}\right) + (1-\pi)\frac{(\beta Q)^2}{2} \geq \frac{(\beta b)^2}{2}. \quad \text{(C3)}$$

Lemma 4 *Let $c < \frac{Q^2-b^2}{2}$. There exists a unique belief threshold $F_0 \leq P_0^\pi \leq N_0$ associated with every $0 \leq \pi \leq 1$ which solves the condition (C3) with equality such that*

1. *if the agent's previous idea was bad, $q_{r-1} = b$, she experiments for any belief $\beta_r \geq P_0^\pi$, but implements it with an effort of $\beta_r b$ for any belief $\beta_r < P_0^\pi$, and*
2. *if the agent's previous idea was good, $q_{r-1} = g$, the agent implements her idea in round r with a maximal effort of 1.*

Further, for $\pi > \pi'$, $P_0^\pi < P_0^{\pi'}$, i.e. the threshold P_0^π is decreasing in π.

Once we have P_0^π, define $\phi_j^{-1}(P_0^\pi) := T_j^\pi$ for $j = 1, 2, \ldots$ and for each π as in Section 3.4. Therefore, each P_j^π is the starting belief which when (correctly) updated about

[18] The decision following a $q_r = g$ is to always implement the idea with the maximum effort. The decision to start experimentation with a bad outside option idea is embodied in the first problem.

bad ideas j times leads to the terminal belief P_0^{π}.

We can now present the version of Proposition 3 incorporating the learning-by-doing effect.

Proposition 6 *For any prior belief $F_0 \leq \beta_1 < P_1^{\pi_2}$, babbling is the unique equilibrium strategy.*

While the main intuition of the above proposition is the same as in Proposition 3, there are two points of departure. First, this proposition only deals with the prior belief, not any generic belief. The supervisor's equilibrium behavior differs under the situation when the agent starts with a belief in $[F_0, P_1^{\pi_2})$ from when it ends up in this region through some past experimentation. The reason is that with more experimentation and feedback the agent's belief threshold for further experimentation declines. As a result, supervisor's incentives to provide honest feedback to the agent can also change with more experimentation.

It is for the same reason that, second, the upper bound of the region of prior beliefs for which supervisor babbling is unique is $P_1^{\pi_2}$. The supervisor incorporates the effect of one round of feedback and the increase of π from π_1 to π_2 when providing negative feedback. If the updated belief β_2 falls below $P_0^{\pi_2}$, the agent does not experiment further in the second round. The supervisor in this case has reason to always encourage the agent. Thus, babbling is unique for $\beta_1 < P_1^{\pi_2}$.

Next, we identify a sufficient condition to get supervisor honesty in the region of prior beliefs above $P_1^{\pi_2}$. In the relevant worst case incentive constraint, the supervisor is honest one last time in round one and not thereafter. Specifically, when $P_1^{\pi_2} \leq \beta_1 < P_2^{\pi_2}$, negative feedback pushes the agent to experiment once again in round two after which she stops. Thus, the supervisor is honest if

$$\pi_2(B_2^b\theta + (1 - B_2^b\theta)B_3^b\beta_3^b b^2) + (1 - \pi_2)B_2^b\beta_2^b Q^2 > B_2^b b$$
$$\iff \pi_2(\theta + (1 - \theta)\beta_3^b b^2) + (1 - \pi_2)\beta_2^b Q^2 > b \quad (3)$$

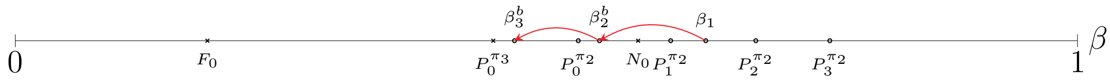


Figure 6: Supervisor truth-telling for $P_1^{\pi_2} \leq \beta_1 < P_2^{\pi_2}$

See Figure 6. The next proposition identifies a sufficient condition that makes condition (3) hold true for any prior belief β_1, and therefore, guarantees honest equilibria.

Proposition 7 *If* $b < \min\left\{\frac{1}{2}, \left(\frac{\theta}{1-\theta}\right)^2\right\}$ *and* $\frac{b-b^2}{2} \leq c < \frac{Q^2-b^2}{2}$*, the supervisor is honest for any prior belief* $\beta_1 \geq P_1^{\pi_2}$*.*

There are two effects of a negative feedback. While the agent is less willing to experiment on account of a lower belief about ability, she is also now *more* willing to experiment owing to learning-by-doing. More honest feedback increases π and reduces the threshold P_0^{π}. The supervisor incorporates this effect of reduced P_0^{π} when providing feedback. Particularly, condition (3) is the worst case scenario precisely because it is assumed $\beta_3 < P_0^{\pi_3}$ so that he does not benefit from providing honest feedback in round two. Consequently, the agent experiments one last time in round two without supervisor feedback. Surely, if condition (3) is satisfied for any belief, so should the weaker condition when $\beta_3 \geq P_0^{\pi_3}$.

The above discussion shows exactly how supervisor behavior at a belief may differ in the situations of the belief being prior or posterior. Supervisor honesty at a belief $F_0 \leq \beta < P_1^{\pi_2}$ is possible when $\beta_1 \geq P_1^{\pi_2}$ but not when $F_0 \leq \beta_1 < P_1^{\pi_2}$. In fact, the agent with a prior belief just above $P_1^{\pi_2}$ may get honest feedback and experiment all the way to F_0, while an agent just below does not. The corollary follows.

Corollary 3 *The expected probability of success of the agent is weakly better under a higher self-opinion in learning-by-doing.*

In fact, there is a potentially larger multiplier effect of a higher self-opinion when the agent can learn on the job through honest feedback. In general, more honest feedback is still more forthcoming for higher self-opinion agents. The agent now has the added benefit of more learning and more experimentation in such a case. An agent with a belief sufficiently higher than $P_1^{\pi_2}$ may receive adequate feedback for $\bar{r}$ rounds to increase her π to 1. On the other hand, an agent just below $P_1^{\pi_2}$ remains there, experimenting only once.

6 Conclusion

In this paper, we showed how an employee responds to criticism influences whether she receives feedback or not. Supervisors may not provide honest feedback to employees who do not believe in their ability. In turn, this hurts their performance and potentially their future careers. Moreover, it also hurts organizations as the supervisors provide inefficiently low levels of honest feedback. Our model highlighted the importance of confidence-building exercises for young creative professionals. Particularly, our model showed that to improve the long-term professional outcomes of women, organizations should invest in their confidence-building.

In the appendices, we first show the proofs of the statements appearing in the main text (Appendix A). Then, we show some omitted proofs. These include comparative statics of F_0 and N_0 (Appendices B.1 and B.2), and the case of generic $g < 1$ (Appendix B.3). Next, we show some basic results on the information design problem in Appendix C. Finally, we discuss in Appendix D the extent to which our results hold in the presence of supervisor costs of providing feedback.

However, our analysis focused only on honest equilibria and we showed limited results in the information design setting. Our work shows the further scope of looking at mixed communication strategies (without commitment) and at information designing (with commitment) in multistage dynamic environments.

References

Ali, O. (2017). Disclosure in multistage projects. *Working paper*.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review*, *84*(2), 191.

Bénabou, R., & Tirole, J. (2002). Self-confidence and personal motivation. *The Quarterly Journal of Economics*, *117*(3), 871–915.

Bénabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The Review of Economic Studies*, *70*(3), 489–520.

Boleslavsky, R., & Lewis, T. R. (2016). Evolving influence: Mitigating extreme conflicts of interest in advisory relationships. *Games and Economic Behavior*, *98*, 110–134.

Bolton, P., & Dewatripont, M. (2012). 9. authority in organizations. In R. Gibbons & J. Roberts (Eds.), *The handbook of organizational economics* (pp. 342–372). Princeton University Press.

Chow, Y., Robbins, H., & Siegmund, D. (1971). Great expectations: The theory of optimal stopping. In (pp. 52–55). Houghton Mifflin.

Ely, J. C., & Szydlowski, M. (2017). Moving the goalposts. *Working paper*.

Ferguson, T. S. (2006). *Optimal stopping and applications.*

Fudenberg, D., & Rayo, L. (2019, November). Training and effort dynamics in apprenticeship. *American Economic Review*, *109*(11), 3780-3812.

Gneezy, U., & Rustichini, A. (2004). Gender and competition at a young age. *American Economic Review*, *94*, 377–581.

Golosov, M., Skreta, V., Tsyvinski, A., & Wilson, A. (2014). Dynamic strategic information transmission. *Journal of Economic Theory*, *151*, 304–341.

Goodman, I., Cunningham, C. M., & Lachapelle, C. (2002). Womens experiences in college engineering. *A Comprehensive Evaluation of Women in Engineering Programs*.

Gnther, C., Arslan Ekinci, N., Schwierenc, C., & Strobeld, M. (2010). Women cant jump? an experiment on competitive attitudes and stereotype threat. *Journal of Economic Behavior and Organization*, *75*(3), 3959–401.

Henry, E., & Ottaviani, M. (2019, March). Research and the approval process: The organization of persuasion. *American Economic Review*, *109*(3), 911-55.

Hirshleifer, D., Low, A., & Teoh, S. H. (2012). Are overconfident ceos better innovators? *The Journal of Finance*, *67*(4), 1457–1498.

Honryo, T. (2018). Dynamic persuasion. *Journal of Economic Theory*, *178*, 36 - 58.

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *The American Economic Review*, *101*(6), 2590–2615.

Koellinger, P., Minniti, M., & Schade, C. (2007). i think i can, i think i can: Overconfidence and entrepreneurial behavior. *Journal of economic psychology*, *28*(4), 502–527.

McCarty, P. A. (1986). Effects of feedback on the self-confidence of men and women. *The Academy of Management Journal*, *29*(4), 840–847.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *Quarterly Journal of Economics*, *122*, 1067–10101.

Orlov, D. (2013). Optimal design of internal disclosure. *Working paper*.

Orlov, D., Skrzypacz, A., & Zryumov, P. (2018). Persuading the principal to wait. *Working paper*.

Renault, J., Solan, E., & Vieille, N. (2013). Dynamic sender–receiver games. *Journal of Economic Theory*, *148*(2), 502–534.

Shurchkov, O. (2012). Under pressure: gender differences in output quality and quantity under competition and time constraints. *Journal of the European Economic Association*, *10*(5), 1189–1213.

Smolin, A. (2017). Dynamic evaluation design. *Working Paper*.

Appendices

A Proofs from main text

We present general proofs in mixed strategies wherever we can. The first section provides some new mathematical notation for this purpose.

Mathematical notation for mixed strategies

We focus attention on limited recall of past ideas so that when the agent experiments one more round, she does not recall the previous ideas she has worked on. In other words, the last idea get destroyed as the agent experiments another round. As a result, the supervisor does not need to make back dated messages about all the previous ideas. A strategy for the agent ρ_r in round r is a mapping from the last observed message to a possible mixed decision to continue experimenting with ideas or implementing the last one. Let $\rho_r^{m_{r-1}} = \Pr(d_r = \text{implement} \mid m_{r-1})$ be the probability that the agent decides to implement the project following the last message.

Similarly, a strategy for the supervisor, σ_r in round r, is a mapping from the last idea to a possible mixed message about its quality. Let $\sigma_r^{q_r} = \Pr(m_r = q_r \mid q_r)$ be the probability of the supervisor honestly revealing the quality of the observed idea. Depending on the expected strategy of the supervisor, the agent conditions her action only on the last message received.

Let the sequence $\hat{\boldsymbol{\sigma}} = \{(\hat{\sigma}_r^g, \hat{\sigma}_r^b)\}_{r=1}^R$ denote the conjectured strategy of the supervisor, and let $\hat{\boldsymbol{\rho}} = \{(\hat{\rho}_r^g, \hat{\rho}_r^b\}_{r=1}^R$ denote the conjectured strategy of the agent. Given the conjectured strategy of the supervisor, the agent updates beliefs about the two unknowns – her ability and the quality of the last idea produced. The belief about being high ability is β_r. Let the belief about her idea matching the supervisor message be denoted by λ_r.

Observe that the public history, the one observed by the agent, h_r^A at the beginning of round r can be summarized by the current public beliefs β_r and λ_r. The private history of the supervisor h_r^S at the beginning of round r can be summarized by the current private belief B_t and q_r.

We can now informally describe the notion of equilibrium. We say that a pair of sequences of conjectured strategies $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$ constitute an equilibrium if (1) they are both the best responses to each other given the beliefs β_r, λ_r and B_r for each t, and (2) the beliefs β_r, λ_r and B_r are consistent with what the players are conjectured to do, i.e. $\boldsymbol{\sigma}$ and $\boldsymbol{\rho}$. Strategies expressed in the text without a hat constitute an equilibrium.

When both the messages are expected in equilibrium, either the two messages induce different beliefs, one higher and the other lower than β_r, or β_r remains the same with both the messages. We will call the former *informative* strategy and the latter *babbling* strategy.

The supervisor is expected to babble in equilibrium in round $r-1$ if $\hat{\sigma}^g_{r-1} = 1-\hat{\sigma}^b_{r-1}$. Whenever the supervisor babbles, it might be useful to think of babbling in mixed strategies rather than in pure strategies. However, there are also babbling equilibria in pure strategies. Say the agent conjectures that the supervisor only says $m_r = g$ on path. While there is no update of beliefs on path, the message $m_r = b$ is off path and we would need to specify beliefs in the information set following this message. Such an equilibrium is supported by any belief $\beta^{\text{offpath}} \in [0, \beta_{r-1})$.

When the supervisor is expected to be informative, we will assume without loss of generality that messages have their natural meaning – the posterior after $m_{r-1} = g$ increases, and it fall after $m_{r-1} = b$. So, we assume that $\hat{\sigma}^g_{r-1} > 1 - \hat{\sigma}^b_{r-1}$ for informativeness.

We restrict attention here to informative strategies in which $\sigma^g = 1$, i.e., the supervisor is always honest about a good idea. The supervisor cannot credibly commit to lying when $q_r = g$. When the supervisor sees $q_r = g$, discouraging the agent is not optimal. If discouragement leads to another round of experimentation, then the supervisor faces the risk of abandoning the current good idea and never getting a new one. Alternately, if discouragement leads to implementation then the agent exerts a lower effort. Going forward, we assume $\sigma_r^g = 1$, and replace σ_r^b with σ_r. Then the posterior beliefs about ability is

$$\beta_r^b = \frac{(1-\theta)\beta_{r-1}}{1-\theta\beta_{r-1}} \text{ and } \beta_r^g = \frac{(1-\hat{\sigma}_{r-1}(1-\theta))\beta_{r-1}}{1-\hat{\sigma}_{r-1}(1-\theta\beta_{r-1})} \tag{4}$$

where $\beta_r^{m_{r-1}} = \Pr(a=1|m_{r-1})$, and

$$\lambda_r^b = 1 \text{ and } \lambda_r^g = \frac{\beta_{r-1}\theta}{1-\hat{\sigma}_{r-1}(1-\theta\beta_{r-1})}. \tag{5}$$

where $\lambda_r^{m_{r-1}} = \Pr(q_{r-1} = m_{r-1}|m_{r-1})$. Thus, the value of a negative message under any informative strategy is the same as in a truth-telling strategy. When an agent receives $m_r = b$ then she can be sure that $q_r = b$. However, a positive message $m_r = g$ cannot be trusted.

Proof of Lemma 2

Proof.

Part 1: Existence and uniqueness of F_0 using one-step look-ahead rule

For a given set of parameters, there is no straightforward closed form solution to the equation in condition (C2) that we get from the one-step look-ahead rule. We therefore need to establish the existence of belief threshold(s). First, it can be verified that both the LHS and RHS of condition (C2) are monotonically increasing and convex in β.

Second, we show that if $2c < \theta(1-b^2)$ then the threshold belief F_0 exists, and is unique. Consider the range of beliefs $0 \le \beta \le 1$. Since $c > 0$ and LHS at $\beta = 0$ is zero, RHS cuts the LHS from above at least once. It can further be verified that RHS at $\beta = 1$ is lower than LHS at $\beta = 1$. Since both LHS and RHS are monotonically increasing, they must intersect at exactly one point. Call that belief F_0. Thus, F_0 exists and is unique. It can similarly be shown that if a unique threshold belief F_0 exists, then $2c < \theta(1-b^2)$.

Finally, we need to show that the agent does not experiment when $2c \geq \theta(1-b^2)$. This is so because then the RHS is always above the LHS, so that even experimentation once is not beneficial. Under $2c \geq \theta(1-b^2)$, $\text{LHS}|_{\beta=1} \leq \text{RHS}|_{\beta=1}$. Given that both LHS and RHS of condition (C2) are increasing convex functions, a concern is that there might be two points of intersection. However, it is easy to verify that the slope of the RHS is lower than the slope of the LHS at both $\beta = 0$ and $\beta = 1$.

Part 2: Optimality of one-step look-ahead decision rule

If we show that our optimal stopping problem is monotone, Corollary of Theorem 3.3 (Chow, Robbins, and Siegmund (1971, p. 54)) readily establishes the optimality of the one-step look-ahead rule.[19] The problem is monotone if whenever the one-step look-ahead rule calls the agent to implement in round R, then so does it for all future rounds no matter what ideas are generated. Let

$$-c + \frac{\beta_R \theta}{2} + (1-\beta_R\theta)\frac{(\beta_{R+1}b)^2}{2} < \frac{(\beta_R b)^2}{2} \text{ where } \beta_{R+1} = \frac{(1-\theta)\beta_R}{1-\beta_R\theta}. \tag{6}$$

Given our discussion in Part 1, $\beta_R < F_0$. We want to show that equation (6) is also true for a $\beta < \beta_R < F_0$. This is also immediate from the discussion and Figure 3. In addition, note that the benefit is bounded above by $\frac{1}{2}$ and the cost of experimentation is fixed at c. Thus, it is optimal to implement the project with $q = b$ for $\beta < F_0$ and continue experimenting otherwise. Finally, we have already shown the proof of the choice of e^F in the main text. ■

[19] See also Ferguson (2006) for a description of the sufficient conditions when we have a maximization problem.

Proof of Lemma 3

Proof. Fix the parameters such that $2c < Q^2 - b^2$. Since, $\theta(1-b^2) > Q^2 - b^2$, both N_0 and F_0 exist and are unique. To compare N_0 and F_0, we only need to compare the LHS of conditions (C1) and (C2), which are both increasing and convex in β. Since LHS of both conditions as $\beta \to 0$ also tends to zero, and the LHS of condition (C1) is smaller than LHS of condition (C2) at $\beta = 1$, we can conclude that $F_0 < N_0$. ∎

Proof of Proposition 1

Proof.

Let $W^F(\beta_1)$ denote the ex-ante expected utility of the agent under the full information policy when her prior is β. Suppose $F_j \leq \beta_1 < F_{j+1}$, then

$$W^F(\beta_1) = \beta_1 \frac{\theta}{2} \sum_{r=0}^{j} (1-\theta)^r - \beta_1 c \sum_{r=0}^{j} (1-\theta)^r + \beta_1 (1-\theta)^j \left[b e^F - \frac{(e^F)^2}{2} \right] + \\ - (1-\beta) \left[(j+1)c + \frac{(e^F)^2}{2} \right] \quad (7)$$

where $e^F = \beta_{1+j+1} b$ with $\beta_{1+j+1} = \frac{(1-\theta)^{j+1}\beta_1}{1-\beta_1+(1-\theta)^{j+1}\beta_1}$ from Bayes' rule. By making these substitutions, we can simplify (7) as

$$W^F(\beta_1) = \beta_1 \frac{\theta}{2} \sum_{r=0}^{j} (1-\theta)^r - c\beta_1(1-\theta) \sum_{r=0}^{j-1} (1-\theta)^r + -c(1-\beta_1)j - c + \\ + \beta(1-\theta)^{j+1} \beta_{1+j+1} \frac{b^2}{2} \quad (8)$$

We want to show that $W^F(\beta_1) > \frac{(\beta_1 b)^2}{2} > \frac{(\beta_1 Q)^2}{2} - c$ for $F_0 \leq \beta < N_0$ and $W^F(\beta_1) > \frac{(\beta_1 Q)^2}{2} - c \geq \frac{(\beta_1 b)^2}{2}$ for $\beta_1 \geq N_0$. But note that from Lemma 3,

$$\beta_1 \frac{\theta}{2} + (1-\beta_1\theta)\frac{(\beta_2 b)^2}{2} - c > \max\left\{ \frac{(\beta_1 b)^2}{2}, \frac{(\beta_1 Q)^2}{2} - c \right\} \quad (9)$$

for $\beta_1 \geq F_0$. Therefore, for any β it will be sufficient to show that $W^F(\beta_1) \geq \beta_1 \frac{\theta}{2} + (1 - \beta_1\theta)\frac{(\beta_2 b)^2}{2} - c$. Proof by induction. The above is trivially true for $j = 0$. For $j = 1$, we can show that

$$\begin{aligned} &\beta_1 \frac{\theta}{2}(1 + 1 - \theta) + \beta_1(1-\theta)^2\beta_3\frac{b^2}{2} - (1-\beta_1\theta)c - c > \beta_1\frac{\theta}{2} + (1-\beta_1\theta)\frac{(\beta_2 b)^2}{2} - c \\ \iff &\beta_2\frac{\theta}{2} + (1-\beta_2\theta)\frac{(\beta_3 b)^2}{2} > \frac{(\beta_2 b)^2}{2} + c, \end{aligned}$$

which is true from condition (C2) since $F_0 \leq \beta_2 < F_1$. Suppose the statement is true for some arbitrary $j = k$. We need to show that the statement is also true for $j = k+1$. We can show that

$$\begin{aligned} &\beta\frac{\theta}{2}\sum_{r=0}^{k+1}(1-\theta)^r - c\beta(1-\theta)\sum_{r=0}^{k}(1-\theta)^r - c(1-\beta)(k+1) - c + \beta(1-\theta)^{k+2}\beta_{1+k+2}\frac{b^2}{2} > \\ &\quad \beta\frac{\theta}{2}\sum_{r=0}^{k}(1-\theta)^r - c\beta(1-\theta)\sum_{r=0}^{k-1}(1-\theta)^r - c(1-\beta)k - c + \beta(1-\theta)^{k+1}\beta_{1+k+1}\frac{b^2}{2}, \end{aligned}$$

which simplifies to

$$\beta_{1+k+1}\frac{\theta}{2} + (1-\beta_{1+k+1}\theta)\frac{(\beta_{1+k+2}b)^2}{2} > \frac{(\beta_{1+k+1}b)^2}{2} + c, \tag{10}$$

which is true since $F_0 \leq \beta_{1+k+1} < F_1$ and equation (10) holds from condition (C2). It is now easy to see that $W^F(\beta_1)$ is increasing in β_1 with a lower bound of $\beta_1\frac{\theta}{2} + (1 - \beta_1\theta)\frac{(\beta_2 b)^2}{2} - c$ for $\beta_1 = F_0$. ■

Proof of Proposition 3

Proof. We prove this statement in steps by considering different regions of starting prior β_1. There exists a $j \in \{0, 1, 2, \ldots\}$ where belief F_j is such that $F_j < N_0 \leq F_{j+1}$. The value that j takes depends on the parameters.

Step 1: Babbling is the unique equilibrium for $F_0 \leq \beta_1 < F_1$

Consider any informative strategy $\hat{\sigma}_1 \in (0, 1]$ including the truth-telling strategy. In any such strategy, the agent best responds to $m_1 = b$ by experimenting once in round one and then implementing with effort $e = \beta_2^b b$ (since $\beta_2^b < F_0$, $\lambda_2^b = 1$, Proposition 2). A message $m_1 = g$ instead leads to a higher belief $\beta_2^g \in (\beta_1, 1]$, which can either push the agent to implement her idea with a higher effort or to experiment again.

If the agent best responds to $m_1 = g$ by implementing her idea, the supervisor is better off deviating since $B_2^b b \beta_2^g(\lambda_2^g + (1 - \lambda_2^g)b) > B_2^b \beta_2^b b^2$. Alternately, the agent may best respond by experimenting again. Consider the worst case of only experimenting once and then implementing. The relevant payoff comparison is between payoff from honesty, $B_2^b \beta_2^b b^2$, and the payoff from lying, $B_2^b \theta + (1 - B_2^b \theta) B_3^b \beta_3^b(\beta_2^g(\beta_1))b^2$ where $\beta(\beta')$ is the updated belief starting from β'. It can now be verified that the latter is larger than the former because $\beta_3^b \geq \beta_2^b$ as β_3^b is negatively updated from $\beta_2^g > \beta_2^b$. Further, it follows that for any more experimentation, deviating must be strictly better for the supervisor. Thus, only the babbling strategy remains which is always an equilibrium. The agent's equilibrium strategy is to implement her outside information idea, i.e. $d_1 =$ implement with $e = \beta_1 b$ since $\beta_1 < N_0$ (see Lemma 1).

Step 2: Proving babbling is a unique equilibrium for $F_1 \leq \beta_1 < N_0$

If $j = 0$, then we are already done. If $j = 1$ then it is enough to show that babbling is the unique equilibrium in the range $F_1 \leq \beta_1 < N_0$. Here, if the posterior $\beta_2 < F_1$ then the supervisor babbles (from Step 1 above) and that the agent best responds by implementing with effort $e = \beta_2^b b$. As before now, the supervisor is better off deviating to induce the agent to either implement with a higher effort or experiment again. Thus, babbling is the unique equilibrium strategy of the supervisor.

If $j \in \{2, 3, \dots\}$, then it needs to be shown that babbling is a unique equilibrium strategy in the ranges $F_1 \leq \beta_1 < F_2$, $\dots$, $F_{j-1} \leq \beta_1 < F_j$ and $F_j \leq \beta_1 < N_0$. Doing so is immediate using the above logic sequentially starting from $F_1 \leq \beta_1 < F_2$. Therefore, babbling is the unique equilibrium strategy of the supervisor and the agent does not experiment, i.e. $d_1 =$ implement and $e = \beta_1 b$.

Step 3: Proving babbling is a unique equilibrium for $N_0 \leq \beta_1 < N_1$

For $j \geq 1$, the reasoning is exactly as in Step 2 for $N_0 \leq \beta_1 < N_1$. For $j = 0$, we have already shown that babbling is a unique equilibrium strategy for $F_0 \leq \beta_1 < N_0 < F_1$ or $F_0 < N_0 \leq \beta_1 < F_1$. Note that since $F_0 < N_0$, it must be the case that $F_1 < N_1 < F_2$. Using the same argument as above, it can be shown that babbling is also unique for $F_1 \leq \beta_1 < N_1$. ■

Proof of Proposition 4

Proof. We prove the proposition in steps.

Step 1: $\sigma_1 = 1$ is an equilibrium for $N_1 \leq \beta_1 < N_2$ if $c \geq \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$ where $b < \left(\frac{\theta}{1-\theta}\right)^2$ so that $b < Q^2$.

Consider the conjectured strategy $\hat{\sigma}_1 = 1$ for $N_1 \leq \beta_1 < N_2$. When the supervisor observes $q_1 = b$, his expected probability of success by sending message $m_1 = b$ is $B_2^b \beta_2^b Q^2$. On the other hand by sending a message $m_1 = g$ when the gent expects supervisor to be honest leads her to exert $e = 1$ in implementation ($\hat{\lambda}_2^g = 1$ and $\hat{\beta}_2^g = 1$). The expected probability of success is then $B_2^b b$. Truth-telling is an equilibrium if

$$B_2^b \beta_2^b Q^2 \geq B_2^b b \iff \beta_1 \geq \frac{b}{\theta b + (1-\theta)Q^2} := \beta^{\text{truth}} \in (0,1) \text{ for } b < Q^2. \tag{11}$$

$\hat{\sigma}_1 = 1$ is an equilibrium if for all $N_1 \leq \beta_1 < N_2$, it is also the case that $\beta_1 \geq \beta^{\text{truth}}$, or equivalently $\beta^{\text{truth}} \leq N_1$. Using β^{truth} from (11), $N_0 = \left(\frac{2c}{Q^2-b^2}\right)^{\frac{1}{2}}$ (from Lemma 1), and $N_1 = \frac{N_0}{1-q(1-N_0)}$ we get $c \geq \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$. Thus, $\sigma_1 = 1$ is an equilibrium under the conditions outlined.

Step 2: If $\sigma_1 = 1$ is an equilibrium for $N_1 \leq \beta_1 < N_2$, then it must be an equilibrium for all $N_2 \leq \beta_1 < 1$.

Consider the region of priors $N_2 \leq \beta_1 < N_3$. We check whether $\hat{\sigma}_1 = 1$ is an equilibrium. The relevant incentive constraint for the supervisor to reveal the truth at this stage is

$$B_2^b b \leq B_2^b \theta + (1 - B_2^b \theta) B_3^b \beta_3^b Q^2 \iff b \leq \theta + (1-\theta)\beta_3 Q^2. \tag{12}$$

Observe from Step 1 that $b < N_0 Q^2$, where N_0 is the lowest value that β_3^b can take in equation (11). Naturally, N_0 is also the lowest value that β_3^b can take in condition (12). Since $b < \beta_3^b Q^2 < 1$, condition (12) is always satisfied. It can similarly be checked that for any higher β_1 the truth-telling incentive constraint is satisfied.

Step 3: The proof of what happens when $c < \frac{1}{2}\left(\frac{b^2}{Q^2} - \frac{b^4}{Q^4}\right)$ follows from the discussion in the text and Step 2 above. Finally, if $b \geq (\frac{\theta}{1-\theta})^2$ so that $b > Q^2$, the supervisor's truth-telling incentive constraint cannot be satisfied for any interior belief β_1. Thus, there is no truthful equilibrium in which the agent experiments once without feedback following a bad message. Note that this does not preclude truthful equilibria followed by supervisor and agent mixing. ■

Proof of Proposition 5

Proof. To prove the statement, first we consider different ranges of ex-ante beliefs and determine the sufficient condition that makes exactly one additional round of experimentation with feedback welfare-improving.

Case 1: $F_0 \leq \beta_1 < N_0$.

First, note that an increase to $F_0 \leq \tilde{\beta}_1 < N_1$ cannot be welfare-improving. Here, $W(\beta_1) = \frac{(\beta_1 b)^2}{2}$. When $F_0 \leq \tilde{\beta}_1 < N_0$, then $W(\tilde{\beta}_1; \beta_1) = \beta_1 b^2 \tilde{\beta}_1 - \frac{(\tilde{\beta}_1 b)^2}{2}$, which is maximized at $\tilde{\beta}_1 = \beta_1$. When $N_0 \leq \tilde{\beta}_1 < N_1$, then $W(\tilde{\beta}_1; \beta_1) = \beta_1 Q^2 \tilde{\beta}_1 - \frac{(\tilde{\beta}_1 Q)^2}{2} - c$, which is maximized at $\tilde{\beta}_1 = \beta_1$ giving a maximized value of $\frac{(\beta_1 Q)^2}{2} - c$. But from $F_0 \leq \beta_1 < N_0$, it must be $\frac{(\beta_1 b)^2}{2} > \frac{(\beta_1 Q)^2}{2} - c$.

Second, consider an increase to $N_1 \leq \tilde{\beta}_1 < N_2$. In this case,

$$W(\tilde{\beta}_1; \beta_1) = \frac{\beta_1 \theta}{2} + \beta_1(1-\theta)Q^2\tilde{\beta}_2 - (1-\beta_1\theta)\left(c + \frac{(\tilde{\beta}_2 Q)^2}{2}\right) - c, \tag{13}$$

where $\tilde{\beta}_2$ follows from Bayesian updating from $\tilde{\beta}_1$. Note that $\frac{\partial W(\tilde{\beta}_1;\beta_1)}{\partial \tilde{\beta}_1} < 0$. Thus, we limit ourselves to the most beneficial (or minimum) increase of $\tilde{\beta}_1 = N_1$ so that $\tilde{\beta}_2 = N_0$. In addition, $W(\tilde{\beta}_1; \beta_1)$ linearly increases in β_1. Thus, we wish to identify the condition that provides us the following inequality at $\beta_1 = N_0$,

$$W(\tilde{\beta}_1 = N_1; \beta_1 = N_0) > W(\beta_1 = N_0) = \frac{(N_0 b)^2}{2}. \tag{14}$$

Substituting from equation (13) and using $c = \frac{(N_0 Q)^2 - (N_0 b)^2}{2}$ from condition (C1) and simplifying, we get

$$\frac{N_0\theta}{2} - \frac{(N_0 Q)^2}{2}(1 + 2\theta(1 - N_0)) > -(1 - N_0\theta)\frac{(N_0 b)^2}{2}. \tag{15}$$

The RHS of equation (15) is strictly less than 0. Thus, it will be sufficient to show that LHS > 0, which can be simplified to

$$\theta > Q^2(1 + 2\theta(1 - N_0))N_0. \tag{16}$$

Assuming $\theta > Q^2$, the condition highlighted in equation (16) is true if $(1 + 2\theta(1 - N_0))N_0 \leq 1$. After substituting for N_0, we can simplify further as

$$c \leq \frac{Q^2 - b^2}{2}\frac{1}{4\theta^2}. \tag{17}$$

Since we are already in the space of parameters in which $c \leq \frac{Q^2 - b^2}{2}$, condition (17) is satisfied whenever $\theta \leq 0.5$. Condition (17) and equivalently condition (14) hold for $\theta \leq 0.5$. Thus, there exists an ex-ante belief threshold $F_0 < \bar{\beta} < N_0$ such that the agent ex-ante prefers to hold an interim belief of N_1 for all beliefs $\beta \geq \bar{\beta}$.

Case 2: $N_0 \leq \beta_1 < N_1$.

Again, an increase to $N_0 \leq \tilde{\beta}_1 < N_1$ cannot be welfare-improving. In case of an increase in belief to $N_1 \leq \tilde{\beta}_1 < N_2$,

$$W(\tilde{\beta}_1;\beta_1) = \frac{\beta_1\theta}{2} + \beta_1(1-\theta)Q^2\tilde{\beta}_2 - (1-\beta_1\theta)\left(c + \frac{(\tilde{\beta}_2 Q)^2}{2}\right) - c. \tag{18}$$

As before, $\frac{\partial W(\tilde{\beta}_1;\beta_1)}{\partial \tilde{\beta}_1} < 0$, so that we look for the minimum increase in belief to $\tilde{\beta}_1 = N_1$. We wish to compare $W(\tilde{\beta}_1;\beta_1)$ from equation (18) with $W(\beta_1) = \frac{(\beta_1 Q)^2}{2} - c$. Note first that $\frac{\beta_1\theta}{2} > \frac{(\beta_1 Q)^2}{2}$. It is now easy to check that $\beta_1(1-\theta)Q^2\tilde{\beta}_2 - (1-\beta_1\theta)\left(c + \frac{(\tilde{\beta}_2 Q)^2}{2}\right) > 0$ after replacing $\tilde{\beta}_2 = N_0$ and $c = \frac{(N_0 Q)^2 - (N_0 b)^2}{2}$. Thus, $W(\tilde{\beta}_1;\beta_1) > W(\beta_1)$. Further, as $W(\beta_1) > \beta_1 Q^2\tilde{\beta}_2 - \frac{(\tilde{\beta}_2 Q)^2}{2} - c$, we conclude that

$$\frac{\beta_1\theta}{2} - (1-\beta_1\theta)c - \beta_1\theta\left(Q^2\tilde{\beta}_2 - \frac{(\tilde{\beta}_2 Q)^2}{2}\right) > 0. \tag{19}$$

Case 3: $N_j \leq \beta_1 < N_{j+1}$ for $j \geq 1$.

Now,

$$\begin{aligned} W(\beta_1) = \frac{\beta_1\theta}{2}\sum_{r=0}^{j-1}(1-\theta)^r - \beta_1 c\sum_{r=0}^{j}(1-\theta)^r + \beta_1(1-\theta)^j\left(Q^2\beta_j - \frac{(\beta_j Q)^2}{2}\right) \\ - (1-\beta_1)\left((j+1)c + \frac{(\beta_j Q)^2}{2}\right), \end{aligned} \tag{20}$$

and if the interim belief is increased to $N_{j+1} \leq \tilde{\beta}_1 < N_{j+2}$ such that $\beta_j = \tilde{\beta}_{j+1}$, then

$$\begin{aligned} W(\tilde{\beta}_1;\beta_1) = \frac{\beta_1\theta}{2}\sum_{r=0}^{j}(1-\theta)^r - \beta_1 c\sum_{r=0}^{j+1}(1-\theta)^r + \beta_1(1-\theta)^{j+1}\left(Q^2\beta_j - \frac{(\beta_j Q)^2}{2}\right) \\ - (1-\beta_1)\left((j+2)c + \frac{(\beta_j Q)^2}{2}\right). \end{aligned} \tag{21}$$

Using (20) and (21), $W(\tilde{\beta}_2; \beta_1) > W(\beta_1)$ condition simplifies to

$$\frac{\beta_j \theta}{2} - (1 - \beta_j \theta)c - \beta_j \theta \left(Q^2 \beta_j - \frac{(\beta_j Q)^2}{2} \right) > 0. \tag{22}$$

We know the above condition (22) is true because $N_0 \leq \beta_j < N_1$ and that condition (19) holds.

To complete the proof, note that $W(\tilde{\beta}_1; \beta_1)$ is increasing in β_1. Thus, an agent with a higher ex-ante belief gains more from an increase in the interim belief that allows her more rounds of experimentation with feedback. Thus, a higher ex-ante belief makes it more beneficial to have a higher interim belief as well. ■

Proof of Lemma 4

Proof. The proof of the first part is the same as that of Lemma 2. To show, that P_0^{π} is decreasing in π note that the RHS of condition (C3) is a convex combination of the RHS of conditions (C1) and (C2). π is the weight on the RHS of condition (C2). Naturally, as π increases, the belief threshold P_0^{π} moves closer to F_0. Figure 7 shows the decreasing relationship arising out of the intersection between the orange plane (RHS) and the blue plane (LHS) of condition (C3). ■

Proof of Proposition 7

Proof. As shown in the text, the condition required for supervisor honesty for beliefs $P_1^{\pi_2} \leq \beta_1 < P_2^{\pi_2}$ is given by equation (3). We are looking for a sufficient condition that ensures the RHS > LHS in equation (3) for any belief.

Note that since $P_0^{\pi_2} \leq \beta_2 < P_1^{\pi_2}$, from the agent's optimal best response in condition

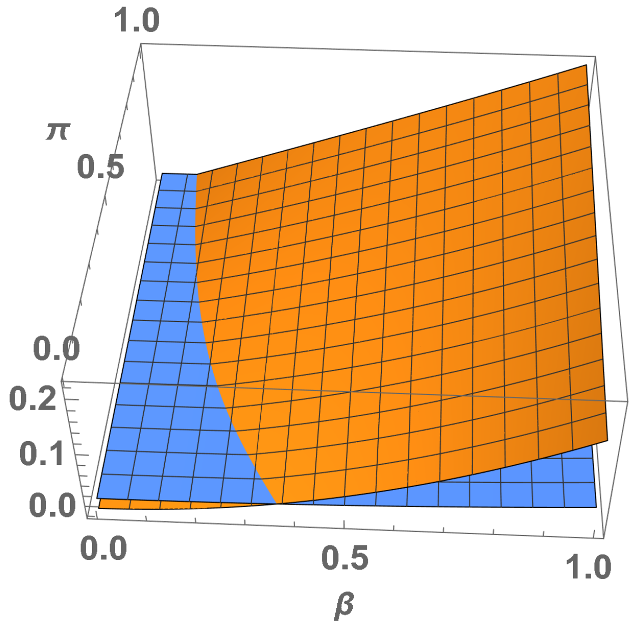


Figure 7: LHS (blue) and RHS (orange) of condition (C3) for $\theta = 0.47$, $b = 0.2$ and $c = 0.02$

(C3) it must be the case that

$$\pi_2(\beta_2\theta + (1-\beta_2\theta)(\beta_3 b)^2) + (1-\pi_2)(\beta_2 Q)^2 > (\beta_2 b)^2 + 2c$$
$$\iff \pi_2(\theta + (1-\theta)\beta_3 b^2) + (1-\pi_2)\beta_2 Q^2 > \beta_2 b^2 + \frac{2c}{\beta_2}. \tag{23}$$

Now, observe that the LHS of equation (23) is the same as that of equation (3). Thus, a sufficient condition to get supervisor honesty is $f(\beta_1) := (\beta_2 b)^2 - \beta_2 b + 2c > 0$. It is easy to verify that $f(\beta_1)$ has a minimum at $\underline{\beta}_1 = \frac{1}{\theta+(1-\theta)2b}$ where the function takes a value of $2c - 0.25$. Naturally, if $c \geq 0.125$, then we are done. However, when $c < 0.125$, we need to determine a new sufficient condition that makes $f(\beta_1) > 0$ for all beliefs. To identify this condition, note that $\underline{\beta}_1 = \frac{1}{\theta+(1-\theta)2b} \geq 1 \iff b \leq \frac{1}{2}$. Also, $f(\beta_1) = 0 \implies \beta_1^* = \frac{1\pm\sqrt{1-8c}}{\theta(1\pm\sqrt{1-8c})+(1-\theta)2b}$. To get our required sufficient condition, it must be that the smaller of the β_1^* solutions is greater than or equal to 1 when $b \leq \frac{1}{2}$. Thus, $\frac{1-\sqrt{1-8c}}{\theta(1-\sqrt{1-8c})+(1-\theta)2b} \geq 1 \iff c \geq \frac{b-b^2}{2}$.

To verify that the set of parameters $b \leq \frac{1}{2}$ and $c \geq \frac{b-b^2}{2}$ is nonempty, we require $b < Q^2 \iff b < \frac{\theta^2}{(1-\theta)^2}$. Thus, we require that $b < \min\left\{\frac{1}{2}, \frac{\theta^2}{(1-\theta)^2}\right\}$. ■

B Proofs omitted from the main text

B.1 Comparative statics of full information belief threshold

Lemma 5 *F_0 is decreasing in c, increasing in b, and decreasing in θ.*

Proof. It is immediate to see that F_0 is decreasing in c as it reduces the value of experimenting again, LHS in condition (C2), for every belief level β.

Second, consider the effect of an exogenous increase in b. We have that $\frac{\partial \text{LHS}}{\partial b} = (1-\beta\theta)(\beta')^2 b$, and $\frac{\partial \text{RHS}}{\partial b} = \beta^2 b$. Since $\beta > \beta'$ and $1 > \beta\theta$, $\frac{\partial \text{LHS}}{\partial b} < \frac{\partial \text{RHS}}{\partial b}$. Thus, the value from implementing increases by more than the value from experimenting, which leads to a higher F_0.

Finally, consider an exogenous increase in θ. The RHS remains unchanged with an increase in θ. For the LHS,

$$\frac{\partial \text{LHS}}{\partial q} = \frac{\beta}{2} - b^2\beta\beta'\left(1 - \frac{\beta'}{2}\right).$$

This is positive if $\frac{1}{2} > b^2\beta'\left(1-\frac{\beta'}{2}\right)$, which is true since $\frac{\partial b^2\beta'\left(1-\frac{\beta'}{2}\right)}{\partial \beta'} = b^2(1-\beta') > 0$ and at the limits the inequality holds. As $\beta' \to 0$, we have that $b^2\beta'\left(1-\frac{\beta'}{2}\right) \to 0$ and as $\beta' \to 1$, $b^2\beta'\left(1-\frac{\beta'}{2}\right) \to \frac{b^2}{2}$. ■

An increase in b makes implementing a bad idea more attractive, and therefore, leads to a higher F_0. The agent finishes the project at a higher belief where she exerts a higher effort in implementing a bad idea. An increase in θ lowers F_0. This is so because conditional on being of high-ability, a higher θ increases the probability of a good idea. Therefore, when ability is unknown it makes experimentation more attractive and pushes the agent to experiment for longer.

B.2 Comparative statics of no information belief threshold

A decrease in the probability of coming up with a good idea, θ, or an increase in the cost of experimentation, c, increases N_0. Finally, an increase in b can have a non-monotonic effect on N_0 depending on the initial value. For $b < \frac{1-\theta}{2-\theta}$, an increase in b decreases N_0. For $b > \frac{1-\theta}{2-\theta}$, an increase in b increases N_0.

The intuition for non-monotonicity in b is as follows. b measures the success rate (for any given effort level) from a bad idea when the agent is of high-ability. When the agent does not observe the quality of her idea from experimentation, then she experiments only once as a gamble. When b increases from a sufficiently low b to begin with, it makes this gamble more attractive – the agent reasons that even if the gamble fails, i.e. $q = b$ is the outcome of the gamble, she is more likely to succeed because of a higher b. On the other hand, when b increases further from an already high level, then the gamble becomes less attractive. This is so because the agent already has an outside option bad idea available which then becomes relatively more attractive to finish.

B.3 The case of $0 < b < g < 1$

We now generalise the likelihood of success of the project with the realisation of a good idea. Suppose now that $g < 1$ and the probability of success with a good idea is *age*. We show below that the equilibrium analysis is identical under this generalization.

No information policy. We can rewrite condition (C1) for the agent experimenting once as

$$-c + \frac{(\beta_1 \tilde{Q})^2}{2} \geq \frac{\beta_1 b^2}{2} \tag{C1}$$

where $\tilde{Q} := \theta g + (1-\theta)b$ is the expected quality of the idea produced. The condition in Lemma 1 to get agent experimentation accordingly changes to

$$c < \frac{\tilde{Q}^2 - b^2}{2}. \tag{24}$$

Therefore, the new no information belief threshold changes to $N_0 = \left(\frac{2c}{\tilde{Q}^2 - b^2}\right)^{1/2}$, and Lemma 1 goes through with a parametric change.

Full information policy. Let the value function of the agent at the beginning of round r with belief β_r when her last observed outcome is $q_{r-1} = b$ be $\mathcal{V}^b(\beta_r)$, such that

$$\mathcal{V}^b(\beta_r) = \max\left\{\frac{(\beta_r b)^2}{2}\ ,\ -c + \frac{\beta_r\theta g^2}{2} + (1-\beta_r\theta)\mathcal{V}^b(\beta_{r+1})\right\}.$$

Note that even with $g < 1$ the agent does not experiment any further after coming up with a good idea, as this $q_r = g$ is the best she can achieve. At the same time, she faces the same decision problem as the original one if she comes up with a bad idea, but with a lower belief.

Once again using the one-step look-ahead rule, if the agent finds it optimal to stop experimenting at a belief β, then so does she after another round of experimentation. So, the following condition is necessary for the agent to experiment:

$$-c + \frac{\beta\theta g^2}{2} + (1-\beta\theta)\frac{(\beta' b)^2}{2} \geq \frac{(\beta b)^2}{2}. \tag{C2}$$

It is now straightforward to see that Lemma 2 holds for $c < \frac{\theta(g^2-b^2)}{2}$ holds. The altered proof is available on request.

The comparison between N_0 and F_0 also holds according to Lemma 3. Both belief thresholds exist for $c < \frac{\tilde{Q}^2-b^2}{2}$ because $\frac{\tilde{Q}^2-b^2}{2} < \frac{\theta(g^2-b^2)}{2}$.

Equilibrium analysis. First, we check that Propositions 2 and 3 hold as before. Note that the current change does not induce truthful equilbria for beliefs agent would stop following a bad message. This is because the probability of success with a good idea is now lower as $g^2 < 1$. Therefore, if stopping was optimal for the agent with $g = 1$, then it should be optimal with $g < 1$ too.

It remains to check if truth-telling is still not an equilibrium strategy for the supervisor. In so far as the agent implements her idea after negative feedback, truth-telling will still not be an equilibrium strategy for the supervisor. The reason is that the supervisor's payoff from truth-telling would be $B^b_2\beta^b_2b^2$, while deviation gives him B^b_2bg. Deviation is optimal because $B^b_1\beta_1b^2 < B^b_1bg$ since $b < g$. So for all beliefs below N_1 babbling would still be the unique equilibrium.

Second, let us look at the beliefs above N_1 and identify the condition for truthful equilibria à la Proposition 4. For all beliefs above N_1 truthfully revealing the quality of a bad idea would lead to at least one more round of experimentation without feedback before stopping. Truth-telling would be an equilibrium strategy for the supervisor if and only if revealing a bad idea would yield a higher payoff for the supervisor. This would be the case if:

$$B^b_2\beta^b_2(g\theta + (1-\theta)b)^2 > B^b_2gb \tag{25}$$

Therefore the new truth telling belief threshold will be

$$\beta_1 \geq \frac{gb}{\theta gb + (1-\theta)\tilde{Q}^2} = \beta^{truth} \in (0,1) \text{ if } gb < \tilde{Q}^2 \tag{26}$$

Therefore a truth telling equilibrium exists if $gb < \tilde{Q}^2$ and Proposition 4 holds.

C Committed supervisor

C.1 A note on the enforcement of commitment

Here we present the case of the supervisor committing to an information policy before the agent starts experimenting with ideas. Before we do so, we should understand how such a commitment may be enforced. An information disclosure policy is a sequence of (potentially mixed) promises about revealing the observed ideas that the agent will produce. There are at least two ways in which one may imagine such promises being enforced.

The first is to interpret the disclosure policy sequence of public tests. In this interpretation, the supervisor may not observe the idea but he designs tests that will reveal to the agent (and to the supervisor) the true quality. Thus, commitment to information policy is akin to commitment to test designs. This interpretation is in the spirit of Kamenica and Gentzkow (2011) and Smolin (2017). Another way in which such a commitment may be enforced is through "presentation" of ideas to multiple supervisors. Many co-supervisors rather than one main supervisor may work to discipline each other. This requires that if the optimal disclosure policy involves mixing by the supervisors then they all should agree on such a mixing and then enforce it (say by punishing deviations with full disclosure). Alternately, one supervisor's recommendation may be cross-examined by another supervisor who has also observed the agent's idea.

Nonetheless, these interpretations are not obvious and might not be realistic in many settings. For instance, an apprentice working on a project might only be assigned one expert due to cost concerns. It is also not obvious how a supervisor might commit to a test design that reveals his private information to the agent. Because of this limitation, we present the commitment case as an extension of the model in Section 4. We show here how the supervisor can achieve better outcomes (relative to the equilibrium outcome) both for himself and the agent by committing to an information policy. We maintain the common prior assumption throughout this section.

C.2 Immediate honesty

Consider first the policy in which the supervisor is committed to revealing the true idea quality after each round of experimentation. We call this a policy of *immediate honesty*. It is the same as the full information policy of Section 3.2. As illustrated in Lemma 2 such a policy induces the agent to experiment with continued bad ideas all the way down to the belief F_0. As we have already shown in Proposition 1, this policy is better than the no information policy. It follows from Proposition 1 that this policy is better than the equilibrium policy both for the agent and the supervisor.

Lemma 6 *The immediately honest policy is Pareto superior to the equilibrium policy.*

Thus, both the supervisor and the agent stand to gain if the supervisor commits to honesty. However, as we show below, the supervisor can do better than immediate honesty.

C.3 Delayed honesty

The supervisor's preferred policy will be one that makes the agent experiment every time she has bad idea but implement immediately if she gets a good idea. Thus, while on the one hand he wants to be honest with the agent, he also wants the agent to experiment as often as possible. An intuitive candidate optimal policy is the one that delays information revelation. Such a policy of *delayed honesty* tries to fulfill the two objectives simultaneously. Below we show the limits of the delayed honesty policy and how it is better than immediate honesty.

A delayed honesty policy is a combination of a disclosure time and what to recommend at that disclosure time. A disclosure timing rule is a mapping from the current belief β_r to a choice of round $R+r$ at which the supervisor requires the agent to present her ideas. Equivalently, R is the number of rounds the agent is required to experiment before approaching the supervisor. He then makes a comment about each of the R ideas according to a recommendation policy which is a mapping of $\{b, g\}^R$ onto itself.

A recommendation policy is honest if the supervisor honestly reveals the quality of all the ideas that the agent has produced. We restrict attention to honest recommendation policies for the time being and analyze the optimal disclosure time R^*. Starting with prior β_1, the agent and the supervisor update their belief about the ability sequentially according to Bayes' rule at the disclosure time R. Thus, if the supervisor reveals that any of the ideas are good they both update their belief to 1 and otherwise revise their belief downwards by R times

$$\beta^b_{R+1} = \frac{(1-\theta)^R \beta_1}{1 - \theta\beta_1 \sum_{r=0}^{R-1} (1-\theta)^r}.$$

Fix a prior $\beta_1 \geq F_0$ and consider a disclosure policy that requires the agent to experiment at least R times to receive feedback from the supervisor. We are interested in finding out the *maximum* number of rounds of delay. We are interested in the disclosure policy that induces the agent to quit experimentation and implement any one her ideas after discovering all bad ideas, i.e., $\beta^b_{R+1} < F_0$.[20] We say that such a policy is *implementable* if the agent prefers to experiment R times and receiving feedback to not experimenting and implementing her outside option idea. This yields the following implementability constraint (IC)

$$\frac{1}{2}\beta_1[1 - (1-\theta)^R(1 - (\beta^b_{R+1} b)^2)] \geq \frac{(\beta_1 b)^2}{2} + Rc. \tag{IC}$$

Observe that since the agent is expected to carry out multiple rounds of experimentation without knowing their outcome, she evaluates the possibility of attaining a good

[20] If there is any implementable delayed policy that leads to a posterior above F_0, then the same can be achieved by an immediately honest policy by inducing the same number of rounds of experimentation. We will refer to delayed honesty policy as the one which leads to posteriors below F_0 so that more number of rounds are induced than in immediately honest policy.

idea relative to β_1. Conditional on being high-ability, with probability $(1-\theta)^R$ she expects to attain only bad ideas to implement, and with the remaining probability she expects at least one good idea. Therefore, with probability $\beta_1(1-(1-\theta)^R)$ she receives 1/2 and with probability $\beta_1(1-\theta)^R$ she will revise her belief down to β^b_{R+1} after the supervisor honestly reveals all her R ideas are bad. At this point, she will implement any one of her bad ideas to obtain an expected benefit of $\frac{(\beta^b_{R+1}b)^2}{2}$. Finally, there is no benefit of experimentation if the agent is low-ability. The agent must also pay the cost of experimentation for R rounds. All this is captured in the LHS of (IC) condition as the expected payoff of experimentation.

If the agent instead opts for implementing her bad outside option idea, she expects to receive a payoff of $\frac{(\beta_1 b)^2}{2}$. The expected payoff from implementing is shown in the RHS of the (IC) condition. The (IC) condition thus puts a limit on the maximum number of rounds the agent is willing to experiment when she is at a belief β_1 and the supervisor is committed to revealing all the information after those rounds.

We next analyse the supervisor's incentives under such a policy. The supervisor's ex-ante expected payoff from a R-implementable policy is

$$\beta_1[1-(1-\theta)^R(1-(\beta^b_{R+1}b)^2)]. \tag{27}$$

Does the supervisor benefit from a higher or a lower R? A higher R gives more chances at experimentation to produce a good idea, but it also depresses the effort of the agent in case of such event. The following lemma identifies a sufficient condition under which the first order effect of increased chances at experimentation dominates the second order effect of reduced effort.

Lemma 7 *If $\theta > b$, the supervisor's exante expected payoffs are increasing in the number of rounds the agent experiments R.*

Proof. From (27), the supervisor prefers $R+1$- over R-implementable policy if

$$\begin{aligned} \beta_1[1-(1-\theta)^R(1-(\beta^b_{R+1}b)^2)] &> \beta_1[1-(1-\theta)^{R+1}(1-(\beta^b_{R+2}b)^2)] \\ \iff \theta+(1-\theta)(\beta^b_{R+2}b)^2 &> (\beta^b_{R+1}b)^2 \end{aligned} \tag{28}$$

If $\theta > b$, then it can readily be verified that equation (28) is satisfied. ■

The condition $\theta > b$ guarantees that taking a chance on experimentation leads to a greater probability of success than implementing a bad idea. The supervisor's maximization problem, therefore, reduces to getting the agent to experiment as many rounds as possible. This is solely determined by the (IC) condition. It is immediate that the expected benefit of experimentation to the agent under such a policy, although increasing in β_1, is bounded above by 1/2. Consequently, for a higher β_1 the agent should want to experiment more number of rounds but up to a limit. This limit is a result of the bounded benefits and the increasing cost of experimentation. Our objective is to determine the maximum (β_1, R) combination that is implementable with such a policy.

For this purpose, fix R. A minimum prior belief should make the (IC) bind. The reason is that both LHS and RHS are increasing in β_1 and intersect exactly once. Define this minimum prior belief by $\bar{\beta}^R$. So for any belief $\beta_1 \geq \bar{\beta}^R$ the agent finds it optimal to at least experiment R times. Observe that $\bar{\beta}^R$ must be increasing in R since the agent must have a higher belief to induce her to experiment more often by paying a higher cost. Let $\bar{\beta}^{\bar{R}}$ be the maximum of this increasing sequence so that $\bar{R}$ gives the maximum number of rounds that are implementable and $\bar{\beta}^{\bar{R}}$ is the minimum prior that can induce those many rounds. Proposition 8 follows from the above discussion.

Proposition 8 *The maximum number of rounds R^* the supervisor can delay honestly revealing the outcomes and induce experimentation at prior β_1 is given by $\bar{\beta}^{R^*} \leq \beta_1 < \bar{\beta}^{R^*+1}$ if $\beta_1 \leq \bar{\beta}^{\bar{R}}$, and is equal to $\bar{R}$ if $\beta_1 > \bar{\beta}^{\bar{R}}$.*

We end this section with the following observation.

Observation 1 *The supervisor weakly prefers a policy of delayed honesty to immediate honesty when delayed honesty is implementable, i.e. when* $\beta_1 \leq \bar{\beta}^{\bar{R}}$.

Ali (2017) derives the same result when determining the optimal dynamic disclosure policy in a slightly different environment. In his setting, the agent needs two consecutive successes in order to be successful in the project. The experiments yield success with a positive probability only if the project is of a good type. Ali shows that the more informed party always has an incentive to delay information revelation while the less informed party would prefer early revelation. While we do not solve for the optimal policy here, we showed here delaying may be preferred to immediately revealing the outcome by the supervisor.

D Time-constrained supervisor

In some situations, it is possible that a benevolent supervisor partially internalizes the costs borne by the agent. The expert's (i.e. the supervisor's) prior experience from when he as an apprentice (agent) may be one of the reasons. Another way to think about it is that the supervisor faces some time costs of going through the agent's idea and providing her feedback, and overseeing her implementation of the chosen idea. We adopt this latter interpretation and assume that the supervisor has to only pay the cost of providing feedback when he is honest. However, he must always pay the cost of overseeing the implementation. We further assume common priors in this section.

For the two players $i \in \{A, S\}$, the agent (A) and the supervisor (S), let the cost of experimentation be c_i and the cost of implementation be $\frac{\phi_i e^2}{2}$. The difference between these costs for the two players captures any preference conflict between them. In so far as $c_S < c_A$ and $\phi_S < \phi_A = 1$ (assume), the preference conflict persists. For a given $(c_S, \phi_S) > 0$, there will be a "full information" threshold for the supervisor as well. Call this threshold F_0^S. This reflects the preferences of the supervisor and determines what are the maximum number of rounds the supervisor desires the agent to experiment (or the belief threshold equivalently) with full information.

Consider first the case of $\phi_S = \phi_A = 1$ so that the supervisor fully internalizes the cost of implementation. In the limiting case of $c_S = 0$ studied in the main text, there was no belief threshold for the supervisor. The supervisor wanted the agent to continue experimenting with full information until she succeeded. However, when $c_S < c_A$ and the supervisor is also required to contribute, we have $F_0^S < F_0^A$. Alternately, $F_0^S \geq F_0^A$ if $c_S \geq c_A$.

Proposition 9 *Let $\phi_S = \phi_A = 1$. The supervisor offers honest feedback for beliefs $\beta_1 \geq F_0^A$ if $c_S < c_A$, and for belief $\beta_1 \geq F_0^S$ if $c_S \geq c_A$.*

Proof. Suppose $c_S < c_A$ and consider the belief range $F_0^A \leq \beta_1 < F_1^A$. If the agent expects the supervisor to be honest, this is the last range of beliefs where the agent experiments with a bad idea. The supervisor is honest if $\frac{(\beta_2^b b)^2}{2} > \beta_2^b b - \frac{1}{2}$. This is true because $\frac{(\beta_2^b b)^2}{2} = \max_e \beta_2^b be$ at $e = \beta_2^b b$. Thus, the supervisor is honest.

Suppose $c_S \geq c_A$. Consider the range of beliefs $F_0^S \leq \beta_1 < F_1^S \leq N_1^A$. The supervisor is honest here because $\frac{(\beta_2^b b)^2}{2} > \beta_2^b b - \frac{1}{2}$. Note that it is in the interest of the supervisor to babble in round two because it is costless for him to do so, and the agent best responds by not experimenting further since $\beta_2^b < N_0^A$. Next, consider the range of beliefs $N_1^A \leq F_0^S \leq \beta_1 < F_1^S$. Again, the supervisor is honest because $\frac{(\beta_2^b Q)^2}{2} > \frac{(\beta_2^b b)^2}{2} > \beta_2^b b - \frac{1}{2}$. It is in the interest of the supervisor to not bear the cost of giving honest feedback in round two as $\beta_2^b < F_0^S$. At the same time, it is in the interest of the agent to experiment once without supervisor feedback as $\beta_2^b \geq N_0^A$. It is now straightforward to see that the supervisor is also honest for the range of beliefs $F_0^S \leq \beta_1 \leq N_1^A < F_1^S$. ■

When the supervisor is time-constrained, he cares both about success, and about costly supervision from the agent experimenting in pursuit of success. In turn, this eliminates the fear of discouragement. Notably, now it is more costly for the supervisor to keep offering feedback beyond a certain belief over letting the agent implement a bad idea. We can then get honest equilibria for some additional ranges of beliefs. Thus, a more time-constrained supervisor can potentially offer more honest feedback. In particular, this happens when $F_0^S < N_1^A$ if $c_S \geq c_A$.

We can now look at the case of $\phi_S = 0 < \phi_A = 1$. If the supervisor does not internalize the cost of implementation, all our results from the main text go through. This is the case even if the supervisor partially internalizes the costs of experimentation.

Proposition 10 *If $\phi_S = 0$, then the equilibrium strategies are given by Propositions 2, 3 and 4.*

To understand the intuition, let $c_S = c_A$ and consider whether honesty is an equilibrium strategy for $F_0^A \leq \beta < F_1^A$. At this belief, if the supervisor is expected to be honest, then following a negative message the agent abandons experimentation and exerts a low implementation effort on the idea. If instead, she receives a positive message, she exerts 1 on her idea. Deviating and sending positive feedback does strictly better for the supervisor. This breaks down the honest equilibrium. The issue arises here because the supervisor wants the agent to exert the maximal effort independent of the potential of the idea produced. The supervisor fears discouragement leading to lower effort in implementation which precludes honesty.[21]

[21] It is possible to derive a belief threshold above which the supervisor is expected to be honest in equilibrium for a generic ϕ_S and given c_S and c_A. However, it does not add value to the analysis here. The basic forces remain as described above.