Strategic gradual learning and information transmission

Alexander Frug

September 2016

Strategic Gradual Learning and Information Transmission

Alexander Frug[*]

September 23, 2016

Abstract

Prior to advising a decision maker, the expert needs to gather information about the state of the world. This often takes time and therefore, even if the expert's learning process is unobservable, the timing of the advice is informative in itself. If learning is strategic in that the expert can choose which inspections to perform, the timing of advice may reveal not only the amount but also the type of information available to the expert. This paper studies the expert's covert and strategic process of information acquisition and its effect on the quality of advice. The main result of this paper suggests that, even in the absence of an "objective" reason to expedite information transmission, putting the biased expert under an artificial (or "strategic") pressure, can increase the amount of transmitted information and be beneficial to both players.

Keywords: Gradual Learning, Strategic Pressure, Scheduling of Experiments, Dynamic Information Transmission, Cheap Talk.

JEL codes: D82, D83.

1 Introduction

Strategic interactions with asymmetric information have been a main interest of economics during the past few decades. In many cases, the actual asymmetry in information is not exogenously fixed but changes over time in an endogenous manner. When it takes time to acquire information, even if time is immaterial, the timing of actions becomes informative in itself. Furthermore, if the actual learning process is determined strategically, the timing of actions can reveal not only the amount but also the type of information acquired. This friction of gradualness in learning brings in new strategic considerations and can significantly affect predicted outcomes. However, ongoing strategic information acquisition has not yet received enough attention in the literature.

The present paper addresses the strategic aspect of gradual information acquisition in a sender-receiver environment. I study a dynamic cheap-talk game where, at the start, both players are symmetrically uninformed but the expert has the ability to explore the state. Learning is gradual as it is carried out under some technological limitations; it is impossible to perform many experiments at the same time and only certain experiments can be performed (for example, a "perfect experiment" that instantaneously reveals all relevant information does not exist). Learning is strategic in that the selection of a particular collection of experiments to be performed and the ordering of these experiments are left to the expert's discretion. The actual learning process cannot be observed or verified by the decision maker. To emphasize the particular effect of gradualness in the expert's learning I assume that learning is costless and that the time available for learning places no direct restrictions on the expert's ability to become fully informed.

Consider the following scenario. A patient (decision maker) arrives at a physiotherapy clinic. If his problem falls within the list of conditions treated in the clinic, the problem can either be mild (state 1) or severe (state 2). In addition, it can be the case that the patient's problem is unrelated to the treatments provided in the clinic (state 0). Assume that the three states are equally likely. The patient has to determine the level of his expenditure on physiotherapy a (for example, depending on the number of treatments). His first-best such expenditure at state θ is $a = \theta$. The therapist (expert), who stands to benefit from the treatment, can advise the patient. Assume that, if the patient's problem is unrelated to the treatments offered in the clinic, the therapist prefers not to pursue any active treatment and so the interests of the players are aligned at $\theta = 0$. Otherwise, the therapist is biased towards more expensive treatments and her most-desired level of patient expenditure in state $\theta \in \{1, 2\}$ is $a = \theta + 1$. For the concreteness of example, assume that both players minimize the expected quadratic loss (that is, the squared distance between each player's ideal point and the actual patient expenditure).

The case in which the therapist advises when she is perfectly informed and the patient is uninformed is similar to the model of cheap talk analyzed in Crawford and Sobel (1982). Under these assumptions on the players' information, in the "most informative" equilibrium, the patient only discovers whether his condition can be treated in the clinic ($\theta \in \{1, 2\}$) or not ($\theta = 0$).

To illustrate the key element of the model I now add to the standard cheap-talk environment an explicit description of the process according to which the information asymmetry between players arises.

Assume that both players are initially uninformed but the therapist has the ability to explore the patient's condition as follows. Two examinations are available: ϵ_1 and ϵ_2. Examination ϵ_1 (ϵ_2) privately reveals to the therapist whether $\theta = 1$ ($\theta = 2$) or not. The patient knows what examinations are available but does not observe (or understand) which, if any, are performed. It is also known that each examination takes one unit of time, the examinations are costless from the therapist's perspective, and it is impossible to perform ϵ_1 and ϵ_2 at the same time (e.g., due to technological limitations or regulations). The patient has to make his decision at period $T > 2$ and so the time available is sufficient to fully explore the patient's condition.

In this case, it turns out that the patient can accurately infer his condition. The following behavior of the therapist is consistent with a fully informative equilibrium. First, she performs examination ϵ_2. If the result is positive, she reports it *immediately*. Otherwise, she proceeds to examination ϵ_1, which takes another unit of time. When the result of ϵ_1 is obtained the therapist reports it truthfully.

To see why this is consistent with equilibrium, note that when the therapist chooses her learning strategy she takes into account that, given the patient's best-response strategy, she has only one opportunity to induce the highest possible expenditure of the patient ($a = 2$). This can be done only by reporting that "$\theta = 2$" at the end of period 1 (claiming that "$\theta = 2$" at the end of period 2 would not be reliable). Since only one examination can be performed at a time, the therapist has to choose which to perform first. It is most desirable from her perspective to begin with ϵ_2 and report that "$\theta = 2$" *only* if the result of ϵ_2 is positive. If ϵ_2 is negative, inducing $a = 2$ is not desirable because, at this point, the therapist does not know whether $\theta = 1$ or $\theta = 0$. In the absence of early report, the patient believes that $\theta \in \{0, 1\}$ and essentially no conflict of interests is left so the expert will report truthfully after performing ϵ_1.

The crucial feature in the above equilibrium is the therapist's strategic use of time. Even though the patient can observe neither the examinations nor their outcomes, he knows that at most one examination could have been performed if a report is submitted at the end of period 1. Notice that there is no objective reason to expedite information transmission; however, in the most informative equilibrium the expert is put under *strategic pressure*, in the sense that, certain reports are reliable only if submitted early. This artificial pressure induces a problem of *strategic scheduling* of experiments for the expert. In this case, the most informative equilibrium is fully revealing and, ex ante, this equilibrium is strictly preferred by both players to any other equilibrium.

Strategic gradual learning, even if unobservable, may significantly affect the economic outcomes in many environments where players are asymmetrically informed. Examples may include, in addition to advising, situations of trade where the quality of an asset can only gradually be learned (this environment is studied in Eliaz and Frug (2016)), and recruiting where the parties only gradually learn the available alternatives and qualities of each other. The objective of the present paper is to illustrate this effect in the environment of cheap-talk advising. To present the basic forces at play in the most transparent way I begin with a convenient discrete version of Crawford and Sobel's model and add to the traditional state space a special state where the players' interests are aligned. Furthermore, I assume that choosing the optimal action in this state is particularly important. Even though this is done mostly for technical and

3

expositional purposes, such a state space can actually fit better some real-life situations. One possible interpretation of this special state could be that of a "categorically different" state where it is commonly agreed that no active actions should be taken. For example, if a defendant is innocent, it can be agreed that he should not be punished, while if he is guilty, different parties may consider distinct punishments as more appropriate for each level of guilt; or, if the patient's problem turns out to be unrelated to the doctor's specialty, the doctor may prefer to direct the patient to another doctor, but if the problem falls within her specialty, she might be biased towards more expensive treatments, etc..

The main part of the paper considers the learning technology, i.e., the set of possible experiments, as in the above patient-therapist example. This learning is termed here *Learning by Identification*. The first result (Proposition 1) characterizes the unique inspection order and reporting policy that can be consistent with a fully informative equilibrium (namely, an equilibrium where the expert acquires full information and transmits it truthfully to the decision maker). The expert inspects the states in a direction opposite to her bias and reports immediately (and truthfully) upon locating the true state of the world. Proposition 2 provides a sufficient condition for fully informative equilibria to exist. This condition relates the sender's prior beliefs with her preferences and replaces the standard supermodularity condition that is often assumed in models of asymmetric information where one party is fully informed. The suggested condition is satisfied in the discrete version of the leading uniform-quadratic constant-bias case that has been the focus of many theoretical and applied papers. However, it also applies to many other specifications where, for example, unlike in the uniform-quadratic constant-bias case, the players do not rank identically the set of equilibria ex ante. Intuitively, this condition guarantees that given the decision maker's belief that the unobserved inspection is carried out in the order identified in Proposition 1, the expert does not benefit from *covertly* changing the order of inspections.

After proving those results for a state space that contains the "categorically different" state I extensively discuss the role of this special state, and show that the results are not an artifact of the assumed environment; similar learning and communication dynamics arise in an equilibrium in a state space without such a special state. While it is impossible to induce full separation in this case, in claim 1 I show that, for the discrete version of the uniform-quadratic constant-bias specification of the model, complete separation is possible above some threshold. This threshold is set such that, given the receiver's beliefs that the sender inspects the states from top to bottom, deviating and performing the experiments in a different order is undesirable for the sender. In addition, I illustrate in an example that even if the expert's bias is not of constant sign, strategic pressure can improve communication and be beneficial to both players.

Not surprisingly, the extent to which the gradualness in expert's learning can be exploited to improve the quality of her advice depends on the learning technology. At this point, I am unable to provide full characterization of optimal equilibria for general learning technologies or to identify the set of all learning technologies that improve communication over the fully informed sender benchmark. Instead, in the continuation of the paper I address two additional, arguably natural, learning technologies: *Criteria Learning* and *Cutoff Learning*. The first refers to a situation where a state of the world can be represented by an aggregation of several distinct criteria (random variables) and the expert acquires information by

sequentially inspecting the realized values of those variables separately. I show that strategic pressure can facilitate informative communication under Criteria Learning and that, to fully exploit its gradualness, the expert will often transmit information more than once - even though no actions are taken by the decision maker between the reports.

Under *Cutoff Learning*, for each element of the state space - a "threshold," there is an experiment that reveals whether the true state of the world is below or above that threshold. In Proposition 3, I show that under very general conditions (for a state space with or without the "categorically different" state) such learning technology does not facilitate information transmission at all. In other words, every receiver's information structure that is consistent with an equilibrium under Cutoff Learning is also attainable in an equilibrium of the fully informed expert benchmark.

Several other works have considered certain aspects of sender's strategic learning, e.g., Austen-Smith (1994), Argenziano et al. (2016), and Di Pei (2015). The main difference is that while these models are essentially static, the key feature of the model studied here is the dynamic nature of the expert's learning.

Notably, communication with a sender who becomes better informed over time has recently been addressed from a different perspective - that of dynamic informational control. In Ivanov (2015, 2016), the receiver has the ability to determine the sender's information structures during the interim stages of the game but cannot observe the information itself. The dynamics where an upwardly biased sender reports on high states before she reports on low states is reminiscent of the one in Ivanov (2015). However, Ivanov's model and the model addressed in the present paper are built on two substantially different sets of assumptions. In the present paper, the process of information acquisition is *covert* and strategic on the sender's side. Therefore the key step is to ensure that, given the receiver's *belief* about sender's learning, the sender does not want to deviate to a different plan of learning - without being detected. In contrast, in Ivanov (2015) the sender is not strategic with respect to information acquisition but rather it is the receiver who specifies and, in particular, observes the sender's information structure at every stage of the game. The identity of the player who specifies the actual learning process is important, especially when sender's learning is unobservable. This not only changes the equilibrium requirements but also affects the results. This point is illustrated in more details in section 3.1. Finally, the apparently similar communication dynamics offered in the two models is not generally optimal in the environment considered in Ivanov (2015). For example, in Frug (2016b) it is shown that if the sender is not too biased, controlling only the first-period sender's information structure and letting the sender become perfectly informed at the end of period two allows full information extraction in the uniform-quadratic constant-bias specification. A more extensive discussion of these and other related papers is deferred to Section 6.

The rest of the paper is organized as follows. Section 2 presents the model where the sender's learning is done by sequential inspections of the different states of the world. In Section 3, I analyze the model, and discuss the interpretation of "strategic pressure." Sections 4 and 5 illustrate how the ideas presented apply to other environments: Section 4 studies an example where the expert is biased towards extreme actions and Section 5 discusses alternative learning technologies. Section 6 discusses related literature and Section 7 concludes.

2 Model

There are two players, sender and receiver, $N = \{S, R\}$.

State Space I begin the analysis with a convenient state space $\Theta = \{\phi, 1, 2, \ldots, n\}$, where ϕ represents a "categorically different" state (intuitively, ϕ corresponds to the state in which the patient's problem is unrelated to the treatments provided by the therapist in the example in Introduction). In Section 3.1, I discuss the role of ϕ and develop a related result in a conventional state space that does not contain ϕ. Let P denote the common prior on Θ.

Learning by Identification For each $k \in \{1, \ldots, n\}$, there is an experiment $\epsilon_k = \mathbf{1}_{\theta=k}$ that reveals whether $\theta = k$ or not. All of the experiments entail no direct costs for any of the players; however, conducting an experiment takes one unit of time and parallel experimentation is not possible. Only the sender has the ability to perform experiments. The experiments she conducts, the order in which she conducts them, and the outcomes cannot be observed or verified. Notice that there does not exist an experiment ϵ_ϕ. In other words, the realization $\theta = \phi$ can be learned by the sender only by means of ruling out all other states.

Information Transmission The sender is allowed to submit only one report, but she can choose when to submit it. Formally, let M be a set of available reports and let $\sigma \in M$ denote the "empty report" corresponding to the sender's choice to "remain silent." Assume that the set M is rich enough to transmit any of the information that was available to the sender.[1] Let $\mathbf{T} = \{1, 2, \ldots, T\}$ be a discrete timeline and assume $T > n$. At each $t \in \mathbf{T}$, the sender may remain silent (i.e., submit an "empty report" $m_t = \sigma$) or submit an "active report" $m_t \in M - \{\sigma\}$. Any terminal history of reports contains at most one "active report." The reports are *cheap talk* in the sense that they have no direct effect on players' utilities.

It is worthwhile to note that the main result of the paper does not rely on the selection of this particular reporting protocol. In other words, every equilibrium under the selected single-report protocol can be recovered in an environment that places no restrictions on the sender's reporting opportunities. The assumed reporting protocol highlights the main aspect of the reporting strategy – the timing of the report's submission.

Receiver's Action At period T, the receiver has to choose an action $a \in \{\phi\} \cup \mathbb{R}_+$. It is convenient to interpret the *action ϕ* as "the choice of not taking an active action." The assumption that $T > n$ implies that, in principle, the time available for learning places no restrictions on the sender's ability to become fully informed.

Utility Functions Player i's utility function is

$$U^i(a, \theta) = \begin{cases} u^i(a, \theta) & a \neq \phi, \, \theta \neq \phi \\ 0 & a = \phi, \, \theta = \phi \\ -\infty & (a = \phi \, and \, \theta \neq \phi) \, or \, (a \neq \phi \, and \, \theta = \phi). \end{cases}$$

---

[1]The richness assumption guarantees that the result is not driven by the limited language. However, since timing of report will play a central role in communication, up until section 5, one available message will suffice for all results.

The assumption that in the latter case the payoff is negative infinity captures in a stylized way the intuition that any mismatch with respect to ϕ is very costly for the players (intuitively, providing irrelevant treatment or not treating an ill patient are both strongly undesirable). Looking forward, this simplifying assumption will prevent the sender from making active recommendations before she learns the true state of the world. This part of the payoff structure is reminiscent of Aghion and Tirole (1997) who assumed that among the set of available projects, at least one is associated with a "significantly negative" payoff which prevents the agent from recommending a specific project when he is uninformed. The simplifying assumption of the infinite cost facilitates the exposition and is clearly stronger than required. It is discussed at the end of Section 3.

The functions $u^i : \mathbb{R}_+ \times \{1, 2, .., n\} \to \mathbb{R}$ for $i \in \{S, R\}$ satisfy the following assumptions. For each $\theta \in \{1, 2, ..., n\}$, the function $u^i(\cdot, \theta) \leq 0$ represents single-peaked preferences; player i's most desired action at state θ, $a^i(\theta) = argmax_a\, u^i(a, \theta)$, is (strictly) increasing in θ; and the sender is upwardly biased in the sense that for all $\theta < n$,

$$u^S(a^R(\theta), \theta) < u^S(a^R(\theta + 1), \theta),$$

that is, if the true state of the world is θ, the sender prefers the action that the receiver would choose under the belief that the state of the world is $\theta + 1$ rather than the action he would choose believing that the state of the world is θ.

Strategies and Equilibrium The applied solution concept is perfect Bayesian equilibrium. The receiver's strategy is an action rule $\alpha(m) \in \Delta A$ that assigns to any terminal history of reports $\{m_t\}_{t \in \mathbf{T}}$ a distribution over actions (recall that the receiver does not observe which experiments are performed). The receiver's pure strategy is denoted by $a(m)$.

The sender's strategy has two components. She chooses both how to acquire information and how to report. A possible representation of a typical sender's (private) history h_t at the end of time t consists of a sequence of experiments that were performed in the first t periods and their outcomes. After every such history, the sender's strategy specifies a (possibly mixed) report choice $m_t(h_t) \in \Delta(M)$ and a (possibly mixed) experiment to be performed next $\epsilon(h_t) \in \Delta(\{\epsilon_{null}, \epsilon_1, ..., \epsilon_n\}).$[2] This approach requires a direct specification of the sender's next move after any possible sequence of experiments. However, in the present environment it is natural to adopt a more succinct and intuitive description of the sender's behavior.

Let $f : \mathbf{T} \to \{null, 1, 2, ..., n\}$ be the sender's learning plan (or, the ordering of experiments) that specifies the unique experiment $\epsilon_{f(t)}$ to be performed at each period t. Such a learning plan induces a sequence of information structures. If the sender follows the learning plan f, her information structure at period t is given by,

$$\mathcal{F}_t(f) = \left\{ \{f(s)\}_{s \leq t}, \Theta - \underset{s \leq t}{\cup} \{f(s)\} \right\}$$

(i.e., a collection of singletons that correspond to elements of the state space that have been inspected and an information set of all uninspected states).

---

[2]The game ends after a non-empty report, thus, adding the sender's reports to the description of her private history is redundant.

The flexibility of the reporting component of the sender's strategy depends on its learning component as the report at each period can depend only on the information collected up to that period. Formally, a (possibly mixed) reporting plan, $m = \{m_t\}_{t \in \mathbf{T}}$, is "$f$-measurable" if m_t is measurable with respect to $\mathcal{F}_t(f)$ for all t. A sender's plan of behavior with a pure learning plan is a pair (f, m^f) where m^f is $f - measurable$.

To obtain a full description of the sender's strategy, it is necessary to specify her future behavior after histories where the learning is inconsistent with f. However, recall that the receiver does not observe which experiments are performed, thus his off-equilibrium beliefs for such events need not be specified. In addition, note that the sender's preferences at each state of the world do not change over time and the continuation of learning is essential *only* if neither of the previously performed experiments revealed the true state of the world. Thus, to specify the sender's behavior in an equilibrium it is sufficient to provide a plan of behavior (possibly mixed), which is consistent with the sender's best response to the receiver's strategy at the beginning of the interaction and not after any possible history of inspection.

3 Analysis

In the main part of the analysis I illustrate how "strategic pressure" can increase the amount of information transmitted by showing that, under some conditions, there exists an equilibrium in which the receiver accurately infers the state of the world. An equilibrium in which the sender acquires full information and transmits it truthfully is termed here a *fully informative equilibrium*. Proposition 1 characterizes the essentially unique sender's behavior in any fully informative equilibrium. Qualitatively, this proposition has two implications. First, the sender will perform her exploration as fast as possible. The fully informative equilibrium has the feature of "strategic pressure," which implies that certain actions can be induced only during the early stages of the game. Therefore, as learning is private and costless, to decide whether to induce a particular action, the sender would prefer to proceed as much as she can with the most effective learning plan, from her perspective. Second, in a fully informative equilibrium, an upwardly biased sender will necessarily explore the states downwards (against the bias direction). Actions towards which she is "most biased" are the first to become no longer inducible. In Sections 4 and 5, it is shown that the insights of strategic pressure and exploration direction also apply to other situations in which the sender's information is acquired through a different learning technology and the bias is not of constant direction.

Proposition 1 *In any fully informative equilibrium, for all $t < n$, $f(t) = n + 1 - t$ and $m_t^f \neq \sigma$ if and only if $\epsilon_{f(t)} = 1$.*

Proof Fix a fully informative equilibrium. For each $k \in \Theta - \{\phi\}$, denote by $t(k)$ the last period in which the information that $\theta = k$ can be transmitted to the receiver in this equilibrium. A necessary condition for a fully informative equilibrium is that for every $k > 1$, the singleton $\{k-1\}$ is not an element of the sender's information structure at period $t(k)$. Otherwise, upon learning the fact that $\theta = k - 1$, the

sender is better off pretending that $\theta = k$. Thus, it follows that $t(n) < t(n-1) < ... < t(2) < t(1)$, which implies $t(k) \geq n + 1 - k$ for all $k \in \{1, 2, ..., n\}$. Moreover, if $t(k) > n + 1 - k$ for some $k > 1$, from the above strict inequalities, it follows that $t(2) \geq n$. In this case, the sender has a profitable deviation to a strategy by which she performs all the experiments by $t = n$ (and no experiment is postponed relative to the original learning plan), and induces the receiver's action $argmax_a \, u^R(a, \theta = 2)$ at state $\theta = 1$. Thus, in any fully informative equilibrium, $t(k) = n + 1 - k$ for all $k > 1$, which ensures that the learning plan f satisfies $f(t) = n + 1 - t$ for all $t < n$.

Notice that, distinct realizations of ϵ_k (an experiment that is performed at $t(k)$) must induce distinct reports at $t(k)$. Choosing $m_{t(k)} \neq \sigma$ terminates the game. Thus, in a fully informative equilibrium, $\epsilon_k = 0$ must be followed by $m_{t(k)} = \sigma$ for all $k > 1$. $\square$

Next, a sufficient condition for a fully informative equilibrium is provided. This condition relates the sender's preferences to the prior distribution. It replaces the standard supermodularity condition that is often assumed in information economics when considering interactions with informed agents.

The sender's preferences satisfy *p-supermodularity* if, for every $\bar{a} > \underline{a}$ and $\bar{\theta} > \underline{\theta}$, the following inequality holds:

$$p(\bar{\theta})[u^S(\bar{a}, \bar{\theta}) - u^S(\underline{a}, \bar{\theta})] > p(\underline{\theta})[u^S(\bar{a}, \underline{\theta}) - u^S(\underline{a}, \underline{\theta})].$$

To understand the intuition behind this condition it is instructive to consider first the discrete counterpart of the leading uniform-quadratic constant-bias specification of the model: the states $\{1, 2, ..., n\}$ are equally likely and the players' utility functions are $u^R = -(\theta - a)^2$ and $u^S = -(\theta + b - a)^2$, where the constant b measures the degree of interest divergence between the players. The above condition amounts to the standard supermodularity of u^S which is satisfied under this specification. Therefore, a fully revealing equilibrium exists. In this equilibrium, by Proposition 1, the sender inspects the states "from top to bottom" and reports immediately as the true state of the world is identified. In particular, at every stage of the game (as long as the true state has not been found), the sender considers it optimal to spend one unit of time inspecting the highest element of the state space that has not been inspected yet. Now consider a general prior instead of the uniform distribution on $\{1, 2, ..., n\}$. To support a fully revealing equilibrium, the sender must inspect the states from top to bottom (and the receiver must believe the sender follows this order of inspection). However, if, for example, the probability $p(n)$ is very low the sender may find it beneficial to "skip" the experiment ϵ_n and inspect other, more likely, states earlier. Identifying the true state of the world earlier, would allow the sender to induce a higher receiver's action, which is desirable from her perspective. The above condition ensures that the probability of relatively higher states is sufficiently high so that skipping an experiment is non-profitable. What probability is considered to be high enough depends on the relative gains from increasing the receiver's induced action at different states, and thus, the condition involves both the sender's utility function and her beliefs. As time goes by, the sender's beliefs change. However, only the relative likelihood of the uninspected states

is important, hence, it is sufficient to consider the sender's prior beliefs.

Proposition 2 *If the sender's preferences satisfy p-supermodularity, a fully informative equilibrium exists.*

Proof Assume the sender follows the learning plan $f(t) = n + 1 - t$, and submits an active report at period τ, $m_\tau^f \neq \sigma$, if and only if $\epsilon_{f(\tau)} = 1$. In addition, assume that if $m_\tau^f \neq \sigma$, the actual report is chosen randomly (with full support) from $M - \{\sigma\}$. Notice that, if such (f, m^f) is played in an equilibrium, from the receiver's perspective, off-equilibrium events do not exist, and thus, the whole system of the receiver's beliefs is pinned down by the Bayes' rule. Accordingly, the receiver's best response is to select $a^R(n + 1 - \tau)$ if $m_\tau \neq \sigma$ for some $\tau \leq n$, and the action ϕ if $m_t = \sigma$ for all $t \leq n$. It is left to show that (f, m^f) is consistent with the sender's best response to the receiver's strategy.

First, I show that m^f is optimal given f. Since $\phi \in \Theta$, it cannot be optimal for the sender to submit a "false-positive" report, $m_\tau \neq \sigma$, in case $\epsilon_{f(t)} = 0$ for all $t \leq \tau$, because this would induce $a \neq \phi$ while the possibility that $\theta = \phi$ is not yet ruled out. On the other hand, once θ is identified, she cannot benefit from delaying the report because $u^i(\cdot, \theta)$ represents single-peaked preferences, and $a^R(\theta) < a^S(\theta)$ for each θ.

I now show that there does not exist an alternative learning plan g, and a "*g-measurable*" reporting plan m^g such that (g, m^g) constitutes a profitable deviation.

Since the learning is unobservable and unverifiable, and it consists of deterministic and costless experiments, it is without loss of generality to restrict attention to deviations in which the sender's learning plan can be represented as a permutation in[3] S_n. Fix an arbitrary permutation $g \in S_n$ such that $g \neq f$. Let k be the minimal integer with $g(k) > f(k)$. Define a new permutation $g' = (g(k), g(k-1)) \circ g$ that is obtained from g by swapping $g(k)$ and $g(k-1)$, and let m^g and $m^{g'}$ denote some optimal reporting plans given g and g' respectively. I now show that the sender strictly prefers $(g', m^{g'})$ over (g, m^g). This will conclude the proof because, since g is chosen arbitrarily, it follows that *every* $g \neq f$ is suboptimal. Given the receiver's strategy, at least one pure learning plan in S_n must be optimal for the sender. Thus, (f, m^f) is consistent with the sender's best response.

Since $\phi \in \Theta$, it must be the case that m^g and $m^{g'}$ induce identical receiver's actions whenever $\theta \notin \{g(k), g(k-1)\}$. To see this, note that the sender never induces $a \neq \phi$ as long as $\theta = \phi$ is not ruled out, and every $\theta \notin \{g(k), g(k-1)\}$ is learned by the sender at identical times under both learning plans g and g'. Thus, to compare g and g' it suffices to compare the sender's expected utility under g and g' given that $\theta \in \{g(k), g(k-1)\}$.

From the definition of k, it follows that $g(k) > g(k-1)$. Otherwise, $g(k-1) \geq g(k) + 1 > f(k) + 1 = f(k-1)$, a contradiction to the minimality of k. Let $\bar{a}$ and $\underline{a}$ denote the actions that are taken by the receiver after $m_{k-1} \neq \sigma$ and $m_k \neq \sigma$ respectively. Note that that $\bar{a} > \underline{a}$. As $g(k) > f(k)$, expediting $\epsilon_{g(k)}$

---
[3] S_n denotes the set of all possible permutations on the set $\{1, 2, ..., n\}$. Each of the first n periods, $\{1, 2, ..., n\}$ is mapped to a distinct experiment which also corresponds to a number in $\{1, 2, ..., n\}$. For example, $f \in S_n$.

from $t = k$ to $t = k - 1$ generates an (expected) improvement of

$$p(\theta = g(k)) \cdot [u^S(\bar{a}, g(k)) - u^S(\underline{a}, g(k))].$$

On the other hand, a delay of $\epsilon_{g(k-1)}$ from $t = k - 1$ to $t = k$ can have one of two possible consequences. If $u^s(\bar{a}, g(k-1)) < u^s(\underline{a}, g(k-1))$, such a delay does not affect the sender's utility. To see this note that the optimality of m^g implies that, in case $\epsilon_{g(k-1)} = 1$, the sender chooses not to report immediately. This means that if she identifies that $\epsilon_{g(k-1)} = 1$ one period later (as happens under g'), she can induce the same receiver's action under both g and g'.

Alternatively, if $u^s(\bar{a}, g(k-1)) \geq u^s(\underline{a}, g(k-1))$ a delay of $\epsilon_{g(k-1)}$ from $t = k-1$ to $t = k$ amounts to the expected loss of

$$p(\theta = g(k-1)) \cdot [u^S(\underline{a}, g(k-1)) - u^S(\bar{a}, g(k-1))].$$

In this case, if $\epsilon_{g(k-1)} = 1$ (which is revealed at $t = k$ under g'), then $m_k^{g'} \neq \sigma$, because $u^s(\cdot, g(k-1))$ represents single-peaked preferences.

To summarize: the net value of the swap is

$$p(\theta = g(k)) \cdot [u^S(\bar{a}, g(k)) - u^S(\underline{a}, g(k))] + p(\theta = g(k-1)) \cdot max\{0, [u^S(\underline{a}, g(k-1)) - u^S(\bar{a}, g(k-1))]\} \geq$$

$$p(\theta = g(k)) \cdot [u^S(\bar{a}, g(k)) - u^S(\underline{a}, g(k))] + p(\theta = g(k-1)) \cdot [u^S(\underline{a}, g(k-1)) - u^S(\bar{a}, g(k-1))] > 0,$$

where the last inequality follows from the *p-supermodularity* of u^S. $\square$

I interpret this result as follows. Even though there is no "objective" reason to expedite information transmission (e.g., costs or discounting), creating an artificial pressure can actually improve the overall quality of the expert's advice. In particular, for the sender not to have an opportunity to inflate the report in the direction of her bias, it is important that the set of inducible actions will shrink sufficiently fast with time in the direction opposite to the bias. Relatively high actions can be induced by the sender only during the early stages of the game. Of course, for the communication to be efficient, it is also important that the set of inducible actions will shrink at a pace that is in keeping with the expert's physical ability to acquire information. This coordination between the sender's learning abilities and the set of actions that can be induced at different periods arises endogenously in this equilibrium.

It is worthwhile to note that the assumption that $U^i(a, \theta) = -\infty$ if $(a = \phi \, and \, \theta \neq \phi)$ or $(a \neq \phi \, and \, \theta = \phi)$ is clearly stronger than required. Obviously, this can be replaced with a sufficiently large cost which has to be selected properly given the specification of the model. As the example in Introduction suggests, when a particular function u^S is given, typically, a much weaker assumption will suffice to support the fully informative equilibrium. I include this stronger assumption in order to clear the proof of additional elements that depend on the fine specifications of the model. It simplifies the exposition and highlights the general aspect and the main force at work.

11

In Sections 4 and 5, I show that the insights on strategic pressure and exploration manner (in terms of direction) also apply to other situations in which the sender's information is acquired through a different learning technology and the bias is not of constant direction. In the next section, I discuss the role of the special state ϕ and develop a related result in a state space not containing ϕ for the discrete uniform-quadratic constant-bias case.

3.1 The Role of ϕ

When the state space does not include the special state ϕ, fully informative equilibria do not exist and the receiver's most-desired equilibrium depends on the fine details of the game, what severely complicates its derivation in general. On the other hand, it is shown below that the effect of ϕ on the informativeness of communication becomes insignificant if the state space is large relative to the divergence between the players' interests.

Consider again the fully informative equilibrium in Proposition 2. As long as the true state of the world is not identified, the mere possibility of it being ϕ prevents the sender from inducing an action $a \neq \phi$. Assume now that $\Theta = \{1, 2, .., n\}$ and that the sender follows the top-to-bottom learning plan as described in Proposition 1. If the divergence in interests is not too great, during the early periods the sender will not have an incentive to submit a false-positive report. By submitting such a report, she will induce a high action and benefit if the true state of the world is *sufficiently* high (e.g., just below the induced receiver's belief). On the other hand, if the true state of the world turns out to be very low, submitting a false-positive report will induce an action which is too high even from the perspective of an upwardly biased sender. Thus, during the early stages of learning, the sender's behavior as in Proposition 1 seems plausible in this environment as well.

As time goes by, more and more high states are ruled out. Thus, the risk of "far low states" is reduced over time. For example, if $\epsilon_k = 0$ for all $k > 1$, the sender infers that $\theta = 1$ without performing ϵ_1. Rather than waiting another period, she is better off submitting a "false-positive" report that corresponds to "$\theta = 2$." Thus, if $\Theta = \{1, .., n\}$ one can proceed with the logic of the equilibrium as in Proposition 2, but not all the way down the state space. At some point, the players must engage in another reporting phase which will not be fully revealing. The exact form of the optimal equilibrium from the receiver's perspective depends on fine details of the model specification.

I now formalize the above discussion in a discrete version of the canonical uniform-quadratic constant-bias specification of the model.[4]

Let $\Theta = \{1, .., n\}$, $p(k) = \frac{1}{n}$ for each $k \in \Theta$, $u^R(a, \theta) = -(\theta - a)^2$, and $u^S(a, \theta, b) = -(\theta + b - a)^2$, where $b \in \mathbb{N}$ measures the sender's bias. The set of available experiments is $\{\epsilon_k | k \in \Theta\}$ and the sender is

---

[4]The tractability of the uniform quadratic specification of the model and the possibility to derive parametric results (for example as a function of the level of interest divergence) have made this specification of the model the focus of many theoretical as well as applied papers. For example, see, Blume et al. (2007), Goldsman et al. (2009), Ivanov (2010), and Krishna and Morgan (2004).

allowed to submit only one report but she can choose the timing of the report's release.

In Claim 1, I prove that there is a threshold that depends on b (and does not depend on n), such that, in equilibrium, the receiver can discover every state above this threshold. In this sense, the effect of ϕ on the optimal equilibrium becomes negligible as the state space becomes large relative to the sender's bias.

Claim 1 *There exists an equilibrium that attains complete separation of every state $\theta \geq 10b$.*

Proof See Appendix. □

Claim 1 does not characterize the optimal equilibrium but places a lower bound on the amount of information that can be transmitted under this learning technology.[5] The main purpose of Claim 1 is to illustrate how the logic of the equilibrium in Proposition 2 can be applied in a state space without ϕ. Essentially, by giving up communication about states below a certain threshold, an artificial substitute for ϕ is created.

The proof of Claim 1 is more involved than that of Proposition 2 for the following reasons. An important feature in the proof of Proposition 2 is that, due to ϕ, we could limit attention to potential deviations in which a report is submitted *only* if the true state of the world is identified. In the proof of Claim 1, however, deviations to strategies in which both the learning plan is changed and a false-positive report is submitted should also be considered. Moreover, for the state space with ϕ, it was assumed that ϵ_ϕ is not available. In contrast, now every state can be inspected directly.

The main intuition of strategic pressure remains - the sender knows that she has a limited amount of time before choosing whether to induce each action. However, she may consider investing this time in inspecting a number of states below the threshold and upon ruling them out to induce a relatively high action, by submitting a false-positive report. Dealing with these kinds of deviations is central in the proof of Claim 1 (see step 3 in the proof).

In brief, the logic of the key step of the proof can be described as follows. Let $\hat{k}$ denote the threshold above which we attempt to support full revelation and assume the receiver believes that inspections are performed as in Proposition 1. The general structure of the best deviation for the sender is the following: The sender starts with the top-to-bottom inspection but follows it only until she gets to some state l which satisfies $\hat{k} < l < 2\hat{k}$. Then, in the next $l - \hat{k}$ units of time she inspects the lowest states $\{1, ..., l - \hat{k}\}$. If $\epsilon_k = 1$ for some $k \leq l - \hat{k}$, the sender will wait and induce a low action in the future. Otherwise, if $\epsilon_k = 0$ for all $k \leq l - \hat{k}$, she submits a false positive report and induces $a = \hat{k}$ - the highest inducible receiver's action at that period. To make this deviation non-profitable, $\hat{k}$ has to be sufficiently high, so that it will take the sender too long (and hence be too costly) to rule out a sufficient amount of low states. The coefficient selected in the claim is minimal such that full separation above the threshold is consistent with

[5]Sometimes the players can benefit if they coordinate on partially informative communication below the threshold (for example, playing an equilibrium à la Crawford and Sobel (1982) in this region, after all experiments are performed).

equilibrium for all $n, b \in \mathbb{N}$.

High-quality information transmission above some threshold and the early arrival of information on higher states also appear in Ivanov (2015). However, Proposition 2 and Claim 1 illustrate two significant differences in both the strategic considerations and results between the case where information acquisition is strategic and covert on the sender's part, and the case considered in Ivanov (2015) where the receiver specifies and, in particular, observes the evolution of sender's information structures during the game. Firstly, as shown in Proposition 2, the prior distribution of the states plays an important role in incentivizing the sender to inspect the states from top to bottom (and thus supporting the dynamics of the equilibrium), while if it is the receiver who specifies the learning the distribution of the states is immaterial.

Secondly, as shown in Claim 1, the thresholds above which full revelation is possible in equilibrium are very different in the two models. In Ivanov's model, this threshold is determined by the shortest sub-interval on which informative communication is possible in equilibrium. Therefore, having in mind Ivanov's result for the standard continuous state space, it is easy to see that if it is the receiver who controls the sender's information in the present discrete state-space model, full revelation of all states above $4b$ can be supported in equilibrium. However, due to sender's ability to covertly deviate to a different order of inspections (as illustrated in the brief discussion above), this cannot be part of an equilibrium under sender's covert learning. Such considerations are completely absent in a model where the learning is observable or where the receiver controls the sender's information.

Interestingly, the relative quality of information transmission on different regions of the state space is significantly different from the one obtained in the fully informed sender benchmark. In contrast to the result in Claim 1, under the uniform-quadratic constant-bias specification of the model, in all equilibria, higher states are pooled in larger intervals in the resulting receiver's information partition, and so the receiver acquires coarser information on higher states.

4 Bias towards Extreme Actions

One important feature of Crawford and Sobel's model and the model presented above is the unidirectional sender's bias. In some cases, other forms of interest divergence take place. The present section illustrates by way of example how, under a similar learning technology, strategic pressure facilitates information extraction if the sender is biased towards extreme actions.[6]

Let $\Theta = \{-2, -1, 0, 1, 2\}$ and assume that $p(\theta) = \frac{1}{5}$ for each $\theta \in \Theta$. Let $u^R(a, \theta) = -(\theta - a)^2$ and $u^S(a, \theta) = -(\theta + sign(\theta) - a)^2$ be the players' utility functions.[7] These functions capture the idea that, relative to the receiver, the sender is biased towards the extreme positive or negative actions. In particular, within each region (positive/negative), the sender has a constant bias of 1.

---

[6]This type of sender's bias is studied in Gordon (2010).

[7]Recall that $sign(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases}$ for $x \in \mathbb{R}$.

First consider the fully informed sender benchmark. It is easy to see that under the optimal equilibrium the receiver learns whether the state is negative, positive, or 0 (the unique point of agreement).

Now I incorporate gradual learning. Assume that for each $k \in \Theta$, there is an experiment ϵ_k which reveals to the sender whether $\theta = k$. As before, each experiment takes one unit of time, the sender can choose when to submit a report, and at most one report is allowed.

Basic intuition can suggest that each of the regions (above and below zero) can be addressed separately. It turns out that this intuition is wrong and to get full information extraction the inspection plan must alternate between the two regions. Uninspected states in one region serve as a disciplining device for better communication in the other region because they increase the risk in submitting false-positive reports.

Claim 2 *A fully informative equilibrium exists.*

Proof See Appendix. $\square$

The sender's behavior in a fully informative equilibrium is to acquire information according to the following learning plan,

$$\begin{pmatrix} t: & 1 & 2 & 3 & 4 \\ f(t): & -2 & 2 & -1 & 1 \end{pmatrix},$$

and report immediately (and truthfully) upon identifying the state of the world.

Again, compared to the fully informed sender benchmark, the players manage to attain a mutually preferred outcome in which the sender is put under strategic pressure. One feature of the fully informative equilibrium is that the extreme actions ($a = -2$, and $a = 2$) can be induced only during the early stages of the game, therefore ϵ_2 is performed before ϵ_1 and ϵ_{-2} is performed before ϵ_{-1}. This is a direct consequence of the logic of Proposition 1, applied to positive and negative regions separately. The second, and novel, feature of this problem is that it cannot be separated into two sub-problems where the sender explores, say, the positive region first and afterwards turns to the negative region. To achieve complete separation in equilibrium, it is important that states in a given region (negative/positive) not be completely ruled out as long as there is still more than one state to be examined within the opposite region.

5   Alternative Learning Technologies

To demonstrate how the gradualness of the sender's learning can enhance informativeness in communication I assumed a particular learning technology - *Learning by identification*. Clearly, the result of having a fully revealing equilibrium in Proposition 2 is an extreme case, which relies on this particular learning. In this section I address other learning technologies and elaborate on the intuition of strategic pressure.

I begin with introducing a slightly more general model which accounts for a more general sender's experimentation. This model will serve as a unified framework for what has been presented earlier and all of the following results.

An *experiment* $\mathbf{Q} = \{Q(\theta) : \theta \in \Theta\}$ is a partition of Θ with the standard interpretation that upon performing $\mathbf{Q}$, the sender learns that $\theta \in Q(\theta)$. A *learning technology* is characterized by the set of available experiments. This set, denoted by $EX = \{\mathbf{Q}_i | i \in I\}$, where I is a (finite) set of indices, is commonly known. Each experiment takes one unit of time and it is impossible to conduct more than one experiment at a time. The experiments performed and their outcomes are unobservable and nonverifiable.

At the end of each period, the sender submits a cheap-talk report regarding her progress in exploration. (In previous sections, it was convenient and without loss of generality to restrict the sender's reporting opportunities. As illustrated below in section 5.1, with general learning technologies, allowing for only one report is with loss of generality.) An action must be selected by the receiver at time $T > |\Theta|$.

Clearly, the completely uninformative equilibrium always exists. Also, if the set EX is *rich* in the sense that upon performing all of the experiments in EX the sender becomes completely informed, it is immediate that any equilibrium of the fully informed sender benchmark can be reproduced in the present environment. The assumption that $T > |\Theta|$ implies that, in principle, the time available for exploration places no restrictions on the expert's ability to explore the state. I now turn to analyze two alternative learning technologies that may be natural under some circumstances, and then conclude this section with several general observations.

5.1 Criteria Learning

In many cases, a sufficient representation of a state of the world is a value that is determined by several different criteria of varying importance. Exploring these criteria separately one after another is a natural way of learning in such cases. Regardless of the order in which these criteria are explored, the information can only be acquired gradually because it is impossible to inspect the different attributes at the same time.

Consider the following example in which the state θ is determined by the realizations of two different random variables ζ and η, and, for convenience, the total value $\theta(\zeta, \eta)$ is uniformly distributed. Specifically, let $\zeta, \eta \sim U\{0, .., k\}$ and let[8] $\theta = (k+1)\zeta + \eta$. Intuitively, this implies that ζ is the "more important" criterion. Assume again $u^R(a, \theta) = -(\theta - a)^2$, and $u^S(a, \theta, b) = -(\theta + b - a)^2$, where $b \in \mathbb{N}$ is the sender's bias.

Claim 3 *Let $b \geq 1$. Any equilibrium under the fully informed sender benchmark splits Θ into at most $k + 1$ "intervals" of different lengths.*

Proof See Appendix. □

I now introduce strategic gradual learning. Assume that two experiments are available,

$$EX = \{check_\zeta, check_\eta\},$$

---

[8]To verify that $\theta \sim U\{0, k^2 + 2k\}$, write every $x \in \{0, .., k^2 + 2k\}$ in terms of its unique representation $x = (k+1)x_1 + x_2$, where $x_i \in \{0, .., k\}$. Then, $Pr(\theta \leq x) = Pr(\zeta < x_1) + Pr(\zeta = x_1) \cdot Pr(\eta \leq x_2) = \frac{x_1}{k+1} + \frac{1}{k+1} \cdot \frac{x_2+1}{k+1} = \frac{1}{(k+1)^2} \cdot (x+1)$.

such that each experiment reveals (to the sender) the realized value of the corresponding random variable.

If a decision maker has to explore the state himself and he has only one period to acquire information, it is natural that he would focus on the most important criterion. On the other hand, if a decision maker has sufficient time to check all the attributes before taking an action, the ordering of the experiments would not matter. In the present model, the decision maker relies on a biased expert's advice. In this case, even in the absence of direct time constraints, desired equilibria again involve indirect, strategic pressure that forces the expert to submit an early report. Before submitting this report, the expert also chooses to focus on the most important attribute, but for completely different strategic considerations.

Claim 4   *Let $b \leq \left\lceil \frac{k}{2} \right\rceil$. There exists an equilibrium that is strictly preferred to any equilibrium under the fully informed sender benchmark.*[9]

Proof   See Appendix. $\square$

The equilibrium that outperforms the fully informed sender benchmark involves the sender focusing her learning on the most important criterion, ζ, and reporting its realized value truthfully before she could become completely informed. The main part of the proof is to show that the sender does not want to secretly deviate and perform $check_\eta$ instead of $check_\zeta$.

In this case, even one round of communication does strictly better than any equilibrium in the fully informed sender benchmark. However, it worth mentioning that, contrary to previous sections, the players can benefit from allowing for more than one round of communication. Specifically, if $4b \leq k + 1$, after revealing the value of ζ at $t = 1$ the players can have another report regarding the residual state space in which they play an informative equilibrium à la Crawford and Sobel (1982). For example, if $k = 3$ and $b = 1$, the resulting receiver's information partition under the optimal equilibrium is[10]

$$\{\underbrace{\{0\}, \{1,2,3\}}_{\zeta=0}, \underbrace{\{4\}, \{5,6,7\}}_{\zeta=1}, \underbrace{\{8\}, \{9,10,11\}}_{\zeta=2}, \underbrace{\{12\}, \{13,14,15\}}_{\zeta=3}\}.$$

Intuitively, in such an equilibrium the expert first provides "general" information and later, as she becomes more informed, she adds finer details to her previous, general report. No actions are taken by the receiver between the reports. However, having the first report submitted at an early stage of the expert's learning, even though additional report is expected before an action is taken, is essential for improving the overall quality of the expert's advice.

---

[9]For $r \in \mathbb{R}$, $\lceil r \rceil$ is the smallest integer such that $\lceil r \rceil \geq r$.

[10]For comparison, the receiver's information partition in the optimal equilibrium under the fully informed sender benchmark is $\{\{0\}, \{1,2,3\}, (4,5,6,7,8), \{9,10,11,12,13,14,15\}\}$. For details see Frug (2016a).

17

5.2 Cutoff Learning

The aim of the present section is to illustrate the richness of the problem under consideration by presenting a natural gradual learning technology that cannot be exploited to improve the informativeness of communication. It is shown below that if the sender can, at each period, select an element $x \in \Theta$ and privately learn whether $\theta \geq x$, gradual learning does not increase the amount of information that can be transmitted in equilibrium, relative to the fully informed sender benchmark. This result is given under fairly general assumptions.

Let Θ be a finite, linearly ordered, discrete state space. Assume that for each θ, $u^R(\cdot, \theta)$ is concave and $u_1^R(\cdot, \theta) = 0$ for some a. This ensures that, given any information, there is a unique optimal action from the receiver's perspective. Also assume that $u^i(a, \theta)$ is supermodular.

A "cutoff k" experiment, $c_k = \mathbf{1}_{\theta \geq k}$, reveals to the sender whether $\theta \geq k$. The set of available experiments in the cutoff learning regime is given by the collection $EX = \{c_k : k \in \Theta\}$. In each period the sender privately chooses a cutoff k_t, which may depend on previous experiments and their results.[11] For a state space like that presented in Section 2, let $\phi < \theta$ for every $\theta \neq \phi$. Adding ϕ to the conventional state space and allowing the receiver to avoid active actions as in Section 2, does not affect the result given in the next proposition.

Proposition 3 *Every equilibrium under the cutoff learning regime is equivalent (in terms of the transmitted information) to some equilibrium under the fully informed sender benchmark.*

Proof Consider an equilibrium under cutoff learning. Every sequence of sender's reports generates a receiver's belief $\mu \in \Delta(\Theta)$, which induces a unique receiver's action $a(\mu)$. More generally, the sender's strategy induces a receiver's belief structure $\tilde{\mu} \in \Delta(\Delta(\Theta))$, that is, a distribution of beliefs over states. It is without loss of generality to assume that, in equilibrium, every belief in $supp\,\tilde{\mu}$ induces a distinct receiver's action.[12] Let $a_1 < a_2 < ... < a_k$ denote the receiver's actions that are induced with positive probability and let μ_i be the receiver's belief that supports a_i.

I now show that at each state θ, the sender induces a belief $\mu \in supp\,\tilde{\mu}$ with positive probability only if, from her perspective, given the state θ, the action $a(\mu)$ is at least as good as $a(\mu')$ for all $\mu' \in supp\,\tilde{\mu}$. Assume by contradiction there exist beliefs $\mu, \mu' \in supp\,\tilde{\mu}$, and a state $\theta \in \Theta$ such that, $u^s(a(\mu), \theta) < u^s(a(\mu'), \theta)$ and $a(\mu)$ is induced by the sender with positive probability at state θ.

First consider the case $a(\mu) < a(\mu')$. Let i be the maximal index such that there exists $\theta' \in supp\,\mu_i$ with $u^s(a_i, \theta') < u^s(a_j, \theta')$ for some $j > i$. From the supermodularity of the sender's utility function it follows that $u^s(a_i, max\,supp\,\mu_i) < u^s(a_{i+1}, max\,supp\,\mu_i)$. In particular, it follows that $max\,supp\,\mu_i < max\,\Theta$. For $\theta \in \Theta$, let $\theta + 1$ ($\theta - 1$) denote the lowest (highest) element in Θ, which is greater (lower) than θ. Since

---

[11] Note that given any history of experiments, the experiment $c_{min\,\Theta}$ is completely uninformative and so it can be interpreted as "doing nothing," that is, the sender is not forced to acquire information.

[12] For any equilibrium in which two different beliefs induce the same action there exists an equilibrium that spares the receiver all the redundant information.

$max\,supp\,\mu_i \in supp\,\mu_i$ and $max\,supp\,\mu_i + 1 \notin supp\,\mu_i$, the experiment $c_{max\,supp\,\mu_i+1}$ is part of the sender's inspection strategy and it is performed before the receiver's beliefs μ_i,μ_{i+1} are induced. Replacing the experiment $c_{max\,supp\,\mu_i+1}$ with the experiment $c_{max\,supp\,\mu_i}$, and maintaining the original reporting strategy as a function of experiments outcomes, constitutes a profitable deviation as, at state $\theta = max\,supp\,\mu_i$, the sender would induce the action a_{i+1} (which is preferred to a_i), and at every other state of the world, the original and the modified strategies coincide.

The case $a(\mu) > a(\mu')$ is similar. Let a_i be the lowest action such that $u^s(a_i, \theta = min\,supp\,\mu_i) < u^s(a_{i-1}, \theta = min\,supp\,\mu_i)$. The experiment $c_{min\,supp\,\mu_i}$ is performed before the receiver's beliefs μ_i, μ_{i-1} are induced. By replacing the experiment $c_{min\,supp\,\mu_i}$ with an experiment $c_{min\,supp\,\mu_i+1}$ and maintaining the original reporting strategy as a function of experiments outcomes, the sender improves upon the original strategy as, at state $\theta = min\,supp\,\mu_i$ she induces a_{i-1} (an improvement by assumption), and at every other state, the original and the modified strategies coincide.

To complete the proof, observe that what has been shown is that among all the elements of $supp\,\tilde{\mu}$, the fully informed sender is willing to induce the same receiver's beliefs as she induces in the equilibrium with cutoff learning. Therefore, the belief structure $\tilde{\mu}$ can be supported by an equilibrium in the fully informed sender benchmark. □

I conclude this section with two general observations on rich collections of experiments. It is trivial that adding an experiment to a rich collection of experiments can make the players worse off (to see this, simply add to any collection of experiments the "perfect experiment" that is fully revealing). On the other hand, Observation 1 shows that adding experiments to rich collections can also make the players *better off*. Observation 2 shows that dispensing with experiments that are unused in the optimal equilibrium can also make the players *better off*. In the following observations, I apply Proposition 3 and use a result on optimal equilibria in the discrete uniform quadratic specification of Crawford and Sobel's model as given in Frug (2016a).

Observation 1   *Adding experiments to a rich collection of experiments can make the players better off.*

Example   Consider the discrete uniform-quadratic constant-bias case with $\Theta = \{1, .., 10\}$ and $b = 1$. By Proposition 3, the optimal equilibrium under the *cutoff learning* regime induces a receiver's information partition $\{\{1\}, \{2, .., 4\}, \{5, .., 10\}\}$, which is identical to the one obtained under the optimal equilibrium in the fully informed sender benchmark. Let $\mathbf{Q} = \{\{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8\}, \{9, 10\}\}$. If $\mathbf{Q}$ is added to the set of available experiments the following sender's behavior is consistent with an equilibrium. The sender performs $\mathbf{Q}$ at $t = 1$ and *immediately* reports truthfully the outcome of this experiment, all subsequent reports are completely uninformative. Provided that the sender's information structure is given by $\mathbf{Q}$, it is immediate that truth-telling is incentive compatible and a direct calculation shows that the sender's expected payoff in this equilibrium is $-\frac{5}{4}$. To see that the sender does not have an incentive to deviate to a

different experiment, note that *every* cutoff experiment partitions Θ in two. This implies that performing such an experiment at $t = 1$ will induce at most two receiver's actions. Thus, if an experiment other than $\mathbf{Q}$ is performed at $t = 1$, say c_j, any sender's report at the end of period 1 will induce at least four elements of the state space such that the distance between the sender's most-desired action and the action she induces at this state is at least 2. Therefore, $E[u^S(\cdot, \theta)|c_j] < -\frac{1}{10} \cdot 4 \cdot 2^2 < -\frac{5}{4}$.

Observation 2 *Dispensing with experiments that are unused in the optimal equilibrium can make the players better off.*

Example Consider the discrete uniform-quadratic constant-bias case with $\Theta = \{1, .., 5\}$ and $b = 1$. By Proposition 3, the optimal equilibrium under the *cutoff learning* regime induces a receiver's information partition $\{\{1\}, \{2, .., 5\}\}$, which is identical to the one obtained under the optimal equilibrium in the fully informed sender benchmark. Moreover, this can be attained by a report at $t = 1$ after the sender performs c_2. The set of experiments $EX' = \{c_2, c_5\}$ is obtained by dispensing with some unused experiments. However, with EX', the following sender's behavior is consistent with an equilibrium. Perform c_5 at $t = 1$ and report truthfully. If $\theta < 5$, perform c_2 at $t = 2$ and report truthfully. It is straightforward to verify that the sender does not have a profitable deviation. The partition obtained in this equilibrium refines the receiver's information structure under the optimal equilibrium with $EX = \{c_k : k \in \Theta\}$. Thus, both players are now better off.

6 Related Literature

The literature on costless communication (cheap talk) between informed experts and uninformed decision makers began with the contributions of Crawford and Sobel (1982) and Green and Stokey (1980).[13] Since then, many authors have studied a variety of cheap-talk environments under different specifications. Sobel (2010) provides a comprehensive literature review of the communication literature.

Several previous works consider environments where the sender is imperfectly informed. Fisher and Stocken (2001) showed that, in equilibrium, the accuracy of the receiver's information is not monotonic in the quality of the sender's information. This was extended by Ivanov (2010) who characterized the optimal static information structure from the receiver's perspective for the leading uniform-quadratic constant-bias case.

Dynamic communication where the sender's information changes over time is also addressed in Ivanov (2015, 2016). In these models, it is assumed that the receiver cannot observe information directly but he can perfectly control the sender's information structure and change it during the game. In Ivanov (2016) an optimal two-stage protocol is suggested. By coupling distant separable elements at the early stage and by conditioning the informativeness of the second signal on the truthfulness of the first report, the receiver

---

[13]Earlier literature on costly signaling can be traced back to Spence (1973).

can successfully elicit full information.[14]

In Ivanov (2015), the dynamic information control is restricted so that the receiver cannot alter future sender's information structures as a function of her past reports. The dynamic information control in this model works as follows. The receiver gradually refines the sender's information such that, in each period, the sender learns whether the state is above a given threshold that decreases over time. The information control is designed such that it is incentive compatible for the sender to truthfully report his observed information. Thus, in effect, the sender reports on high states before she reports on lower states, and as the number of periods increases, the receiver can elicit an arbitrarily accurate information about states that are positioned above a given threshold (determined by the level of interest divergence between the players). In both of these papers by Ivanov, the sender is not strategic with respect to information acquisition as it is determined by the receiver who, in particular, observes the sender's information structures. On the other hand, the main focus of the present paper is on the sender's unobservable strategic choice of learning. While the communication dynamics in Ivanov (2015) and the dynamics in Proposition 1 and Claim 1 in the present paper are related, the identity of the party specifying the learning process and the observability of information acquisition make the sender's strategic reasoning significantly different between the models and change the results. Intuitively, the sender's strategic considerations in Ivanov's model are parallel to those in the reporting component of the sender's strategy in the present model, if she could not affect the process of information acquisition. In addition, the dynamics in Ivanov (2015) is not generally optimal. For a different construction that allows for full information extraction in two periods for the uniform-quadratic constant-bias case where the sender is not too biased see Frug (2016b).

Several works have also considered certain aspects of the sender's strategic learning. Austen-Smith (1994) proposed a model in which the sender chooses whether to learn (perfectly) the state and has the ability to prove information acquisition. The choice whether to become informed depends on the realization of the cost of learning that is privately observed. The author shows that, relative to Crawford and Sobel (1982), more information can be transmitted in equilibrium. As in our model, the choice whether to become informed is left to the sender's discretion. Unlike in our model, the learning is costly, verifiable, and non-gradual. In Argenziano et al. (2016) the sender can affect the quality of her information by selecting how many Bernoulli trials to perform. The trials are costly for the sender. The authors compare "covert" and "overt" selection of the number of trials to be performed and show that under the overt regime it is possible to force the sender to overinvest in learning. In Kamenica and Gentzkow (2011) the sender selects a "signal," i.e., an information structure used to persuade a decision maker. While the selection of the information structure is costless and left to the sender's discretion, the receiver observes both the information structure and its realization. In Gentzkow and Kamenica (2012) the sender also publicly selects the information structure but now it is no longer costless. A more informative information structure is associated with a higher cost. Since the information is costly for the sender, endogenous information will always be disclosed in equilibrium and so disclosure requirements have no effect on the set of equilibrium outcomes. A related result appears in Di Pei (2015). In that paper the sender gathers

---

[14]A similar idea of separable pair of states appears in Golosov et al. (2013).

costly information before advising the receiver. The sender communicates all her information and all equilibria are less informative than the most informative one in Crawford and Sobel (1982).

All of the above models that contain strategic learning are essentially static. However, the key feature of the model studied in this paper is the dynamic nature of the sender's learning. The timing of the sender's reports provides a signal to the receiver not only about the quality but also about the type of information available to the sender. Frug (2014) considers a model of gradual exogenous learning and communication. That paper focuses on the design of efficient reporting protocols.

7 Conclusion

This work studied the effect of a gradual and strategic process of information acquisition in a sender-receiver environment. The main ingredients of the considered environment were (1) gradual strategic learning and (2) the time available for learning places no direct restrictions on the overall quality of the expert's information. If the expert is expected to submit a report at the end of her learning (as she becomes fully informed), the problem of experiment scheduling is completely vacuous. On the other hand, the artificial pressure, presented in most of the results in the paper, raised an important issue of *strategic scheduling* of experiments for the expert. It is not only how much information, but also, which type of information is acquired at each period. The only motivation for the scheduling problem in this environment is strategic as, by assumption, the state can be learned at no cost before an action has to be taken. It has been shown that strategic pressure can increase the amount of information transmitted and make both parties better off. This may suggest a broader question of strategic scheduling both in the environment of observable experiments (unobservable outcomes) and unobservable experimentation.

While the potential gains from gradualness in information acquisition depend on the assumed learning technology, this paper has shown that the sender-receiver interaction changes dramatically, both in terms of strategic considerations and the outcomes, if the fixed and exogenous information asymmetry is replaced with endogenous gradual learning. In many strategic interactions with asymmetric information, the information asymmetry is not fixed. Players take measures to improve their information and they can partially control the process according to which they acquire information. As the gradualness in learning can have a significant potential effect on economic outcomes, it may be valuable to consider gradual and strategic attainment of private information in other contexts and in a broader perspective.

8 Appendix

Proof of Claim 1 Consider the sender's plan of behavior $s^* = (f, m^f)$ where $f(t) = n + 1 - t$ for all $t \in \{1, .., n\}$, and $m_\tau^f \neq \phi$ if and only if $\epsilon_{f(\tau)} = 1$ and $\tau \in \{1, .., n + 1 - 10b\}$. The receiver's best response is then to select $a^*(\{m_t\}) = n + 1 - \tau$ if $m_\tau \neq \sigma$ for some $\tau \in \{1, .., n + 1 - 10b\}$ and $a^*(\{m_t\}) = E[\theta | \theta < 10b] = 5b$ if $m_t = \sigma$ for all $t \in \{1, .., n + 1 - 10b\}$. I now prove, in three steps, that $s^* = (f, m^f)$ is consistent with a sender's best response strategy to the receiver's strategy.

Step 1: m^f is optimal given the learning plan f.

The sender cannot benefit from postponing a report once θ is revealed because this will induce a lower receiver's action, which is worse for the sender. The main part of this step is to show that the sender cannot benefit from a "false-positive" report that induces an action $a \geq 10b$, provided that none of the previous experiments identified the true state of the world. It is shown below that the sender's expected utility from a "false-positive" report at time $n + 1 - a$ (LHS) is lower than her expected utility from continuing to follow (f, m^f) (RHS):

$$-\frac{1}{a-1}\sum_{\theta=1}^{a-1}(\theta+b-a)^2 < -\frac{1}{a-1}\left(\sum_{\theta=1}^{a-10b}(b)^2 + \sum_{\theta=a-10b+1}^{a-1}(\theta+b-a)^2\right) =$$

$$= -\frac{1}{a-1}\left(\sum_{\theta=1}^{10b-1}(\theta+b-10b)^2 + (a-10b)b^2\right) < -\frac{1}{a-1}\left(\sum_{\theta=1}^{10b-1}(\theta+b-5b)^2 + (a-10b)b^2\right).$$

It is left to show that there does not exist an alternative learning plan $g : \{1, ..n\} \rightarrow \Theta$ and a g-*measurable* reporting policy m^g such that (g, m^g) constitutes a profitable deviation for the sender. Denote by $a(\theta|s)$ the action induced at state θ if the sender plays according to a plan s and let

$$K = \{(g, m^g) : if\, m_t^g \neq \sigma\, and\, t \leq n + 1 - 10b\, then\, \epsilon_{g(t')} = 1\, for\, some\, t' \leq t\}$$

be the set of all plans of behavior in which an active report is submitted only after the true state of the world is identified.

Step 2: The plan $s^* = (f, m^f)$ is optimal in K.

Let $s = (g, m^g) \in K$ such that $g \neq f$ and let k be the minimal integer with $g(k) \neq f(k)$. Then, $g(k) < f(k)$. Let $g' = (g(k), f(k)) \circ g$ be the learning plan in which the experiments $\epsilon_{f(k)}$ and $\epsilon_{g(k)}$ are swapped. If $a(g(k)|s) \leq a(f(k)|s)$, g' is clearly preferred to g, as the sender can induce a higher action at state $f(k)$ and keep the state-to-action mapping unchanged in other states. If $a(g(k)|s) > a(f(k)|s)$, g' allows the sender to induce the *higher* action $a(g(k)|s)$ in the *higher* state $f(k)$ and the *lower* action $a(f(k)|s)$ in the *lower* state $g(k)$ without changing the state-to-action mapping at other states. Supermodularity of the sender's preferences ensures that g' is preferred. At least one pure learning plan must be consistent with an optimal plan in K, thus (f, m^f) is optimal in K.

Step 3. The plan $s^* = (f, m^f)$ is optimal among the plans outside K.

Let $t(a)$ be the period in which, if an active report is submitted, the receiver's action a is induced. The plans of behavior outside K have the property that there is $a \geq 10b$, such that, after performing $t(a)$ experiments whose outcome was negative, action a is induced.

For each $\bar{t} \in \{1, ..., t(10b)\}$, let $N_{\bar{t}} = \{(g, m^g) : if\, \epsilon_{g(t)} = 0\, for\, all\, t < \bar{t},\, then\, m_{\bar{t}} \neq \sigma\}$ be the set of plans according to which, after $\bar{t}$ experiments with negative outcomes, the sender submits an active report at the end of period $\bar{t}$ (and thus, induces an action $a = n + 1 - \bar{t}$). Let $N = \underset{\bar{t} \leq t(10b)}{\cup} N_{\bar{t}}$. In addition, given a plan $s = (g, m^g) \in N_{\bar{t}}$ denote by ζ and η the maximal and the minimal elements of Θ that is not inspected

before or at time $\bar{t}$, respectively.

Step 3.1: Under an optimal plan in $N_{\bar{t}}$, if $\theta \in \{\eta, ..., \zeta\}$, then θ is not inspected before or at time $\bar{t}$.

Let $s = (g, m^g) \in N_{\bar{t}}$ and denote by a the action induced in all states that where uninspected during the first $\bar{t}$ periods. Assume that there exists $\eta < \theta < \zeta$ such that θ is inspected in the first $\bar{t}$ periods under g. If $a(\theta|s) > a$ $(a(\theta|s) < a)$, by supermodularity, the sender is better off performing ϵ_ζ (ϵ_η) instead of ϵ_θ and inducing the action $a(\theta|s)$ in state ζ (η) - and the action a at state θ. If $a(\theta|s) = a$, the sender gains nothing by performing ϵ_θ but can benefit if a different experiment is performed instead. If, at η, the sender prefers one of the inducible actions that are below a, she can be better off if instead of ϵ_θ she performs ϵ_η. Otherwise, it must be the case that $\eta > 1$ and therefore $u^S(a, \zeta) < u^S(a+1, \zeta)$ since $\zeta - \eta \geq a$. Moreover, either $g^{-1}(\eta - 1) < \bar{t}$ or $g^{-1}(\theta) < \bar{t}$. Therefore, the sender can be better off by performing ϵ_ζ instead of $\epsilon_{\eta-1}$ or ϵ_θ.

Step 3.2: Let $s = (g, m^g)$ be an optimal plan in $N_{\bar{t}}$; then, $g^{-1}(\theta_0) > g^{-1}(\theta_1) > g^{-1}(\theta_2)$ for $\theta_0 < \eta$ and $\zeta < \theta_1 < \theta_2$. In words, every inspection of an element below the interval that remains uninspected before the false-positive report is performed after all of the states above ζ are inspected, and moreover, all states above ζ are inspected "from top to bottom." This step follows as a direct consequence of the supermodularity of the sender's preferences.

Step 3.3: Let $s = (g, m^g)$ be an optimal plan in $N_{t(\hat{a})}$; then $\eta \leq \hat{a} - 3b$. Assume by contradiction that $g(t) = \hat{a} - 3b$ for some $t \leq t(\hat{a})$. From 3.1 and the fact that $\hat{a} - 1$ states are not inspected before $t(\hat{a})$, it follows that $\zeta > \hat{a} + 6b$. Hence, the distance between the induced action and the sender's optimal action at state ζ is $d_s(\zeta) \geq 7b$. If the sender performs ϵ_ζ before inspecting states below η, she will be able to induce the action ζ at state ζ. Thus, the loss $(d_s(\zeta))^2$ at state ζ can be replaced with a loss of b^2. On the other hand, giving up $\epsilon_{\hat{a}-3b}$ increases the loss in this state by at most $(2b)^2$.

Step 3.4: The optimal sender's strategy in N is an element of $N_{t(10b)}$. Let $\bar{t} < t(10b)$. Consider an optimal strategy in $N_{\bar{t}}$, and let $a > 10b$ be such that $t(a) = \bar{t}$. I now show that there is a better strategy (from the sender's perspective) in $N_{\bar{t}+1}$. From 3.2, first the sender explores all the states above ζ (from top to bottom) and then turns to the exploration of states below η. Action a is induced at each of the states $\eta, \eta+1, .., \zeta-1, \zeta$. Consider $s' = (g', m^{g'}) \in N_{\bar{t}+1}$, where s and s' differ only in that before exploring states below η, the sender performs ϵ_ζ, $m_{t(\zeta)} \neq \sigma$ if and only if $\epsilon_\zeta = 1$, and the action $a - 1$ is induced at period $\bar{t} + 1$. To see that the sender prefers s' to s, it is sufficient to note that the sender's payoff at ζ under s' is higher than her payoff at η under s. It is immediate that under s and s' the sender obtains identical payoffs at all states above ζ and below η. Also, it is easy to see (illustrated in the figure below) that the sender's payoff at $\theta \in \{\eta+1, .., \zeta\}$ under s is equal to her payoff at $\theta - 1$ under s'.

To conclude, note that the sender's payoff at ζ, under s', is $-b^2$. From 3.3, $\eta \leq a - 3b$, which means that the sender's payoff at η, under s, is at most $-(2b)^2$.

I now conclude the proof by directly showing that the sender is better off under s^* than under the optimal plan of behavior in $N_{t(10b)}$, denoted by $s = (g, m^g)$. Under s, by definition, the action $a = 10b$ is induced at least in $10b - 1$ states, denote these states by $\theta_1 < \theta_2 < ... < \theta_{10b-1}$. First, assume that $\theta_1 > 3b$. The vector of smallest possible distances between the induced and the most desired actions from

24
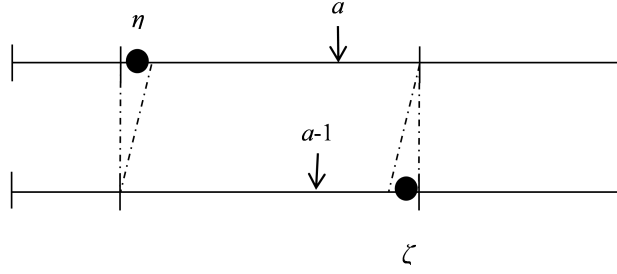
the sender's perspective, $(d_s(\theta_j))_{j=1}^{10b-1}$, is given by

$$(5b-1, 5b-2, ..., 1, 0, 1, ..., 5b-2, 5b-1).$$

On the other hand, by $\theta_1 > 3b$, the vector of the largest such distances under s^* is given by

$$(6b-1, \ldots, b+1, \underbrace{b, \ldots, b}_{5b}).$$

Notice that for any $\theta \geq \zeta$ and $\theta \leq \eta$ we have $a(\theta|s) = a(\theta|s^*)$. As the states are uniformly distributed, s and s^* can be compared by summing up the "losses" (negative utility) given by the squared distances $d_s^2(\theta)$ and $d_{s^*}^2(\theta)$ on $\{\theta_1, ..., \theta_{10b-1}\}$. By using the formula for the sum of squares of the first K naturals: $\sum_{k=1}^{K} k^2 = \frac{K^3}{3} + \frac{K^2}{2} + \frac{K}{6}$, the following inequality holds for any $b \in \mathbb{N}$:

$$\sum_{j=1}^{10b-1} d_s^2(\theta_j) \geq 2\sum_{k=1}^{5b-1} k^2 > (5b)b^2 + \sum_{k=1+b}^{6b-1} k^2 \geq \sum_{j=1}^{10b-1} d_{s^*}^2(\theta_j);$$

therefore, s^* is better than s. By replacing one of the states in $\{\theta_1, ..., \theta_{10b-1}\}$ with a state $\theta \leq 3b$, the LHS of the strict inequality increases by, at least, $(\theta + b - 10b)^2 - (5b-1)^2$ and the RHS increases by, at most, $(\theta + b - 5b)^2 - (6b-1)^2$. For any $b \in \mathbb{N}$ and $\theta \leq 3b$, the increase in LHS is greater than in RHS. Thus, the sender is better off under s^* rather than under s even if we dispense with the assumption that $\theta_1 \geq 3b$. Therefore, s^* is better than s, and the proposition follows. $\square$

Proof of Claim 2 Let $\begin{pmatrix} t: & 1 & 2 & 3 & 4 \\ f(t): & -2 & 2 & -1 & 1 \end{pmatrix}$, and $m_t^f \neq \sigma$ if and only if $\epsilon_{f(t)} = 1$. Moreover, if the sender submits a non-report, assume that it is chosen randomly, and with full support, from $M - \{\sigma\}$. Thus, the receiver's beliefs are derived from the Bayes' rule and the receiver's best reply is

$$a^*(m) = \begin{cases} f(t) & m_t \neq \sigma, \, t \leq 4 \\ 0 & m_t = \sigma, \, \forall t \leq 4. \end{cases}$$

I now show that s^* is the sender's best response to $a^*(m)$. Denote by $d_s(\theta)$ the distance between the sender's most-desirable action and the action induced by the strategy s at state θ. The sender's expected utility if she plays according to s is simply $E[u^S|s] = -\frac{1}{5}\sum_{\theta \in \Theta}(d_s(\theta))^2$. Under s^*, whenever $\theta \neq 0$ the sender experiences a loss of 1, and thus $E[u^S|s^*] = -\frac{4}{5}$.

Clearly, m^f is optimal given f. Note that if $s' = (g, m^g)$ is a profitable deviation, it must be the case that the sender is better off at $\theta = 1$ or at $\theta = -1$.

First, assume that the sender is better off at $\theta = -1$. In this case, under s', the sender induces the action $a = -2$ at the state $\theta = -1$. If $g(1) = -1$, and $m_1 \neq \sigma$ if and only if $\epsilon_{-1} = 1$, the action $a = -2$ is not induced at the state $\theta = -2$. In this case $d_{s'}(-2) \geq 2$, and $E[u^S|s'] \leq -\frac{1}{5}((d_{s'}(-2))^2 + (d_{s'}(2))^2) \leq -1 < E[u^S|s^*]$. Otherwise, it must be the case that, $\epsilon_{g(1)} = 0 \Rightarrow m_1 \neq \sigma$. In this case, there are at least four states in which the action $a = -2$ is induced. Thus, $d_{s'}(\eta) > 2$ for some $\eta > 0$ and $E[u^S|s'] \leq -\frac{1}{5}(d_{s'}(\eta))^2 < -1 < E[u^S|s^*]$.

Now, assume that the sender is better off at $\theta = 1$. In this case, under s', the sender induces the action $a = 2$ at the state $\theta = 1$. If $1 \in \{g(1), g(2)\}$, and $m_2 \neq \sigma$ if and only if $\epsilon_1 = 1$, then either $2 \notin \{g(1), g(2)\}$ or $-2 \notin \{g(1), g(2)\}$. Consequently, $d_{s'}(\eta) \geq 2$ for some $\eta \in \{-2, 2\}$, and so $E[u^S|s'] \leq -\frac{1}{5}((d_{s'}(-2))^2 + (d_{s'}(2))^2) \leq -1 < E[u^S|s^*]$. Otherwise, it must be the case that, if $\epsilon_{g(1)} = 0$ and $\epsilon_{g(2)} = 0$, then $m_2 \neq \sigma$. In this case, the action $a = 2$ is induced in at least three different states, and so there is $\eta \leq 0$ with $d_{s'}(\eta) \geq 2$, and therefore $E[u^S|s'] \leq -\frac{1}{5}((d_{s'}(\eta))^2 + (d_{s'}(2))^2) \leq -1 < E[u^S|s^*]$. $\square$

Proof of Claim 3 Observe that for each $I \in \mathbb{N}$, we have $\sum_{i=1}^{I}(1 + 2(i-1)) = I^2$. According to Procedure 1 in Frug (2016a), if $b = 1$, in the optimal equilibrium the length of the i-th interval is exactly $1 + 2(i-1)$. The claim follows for $b = 1$ as $|\Theta| = (k+1)^2$. For $b > 1$, every equilibrium splits Θ into a smaller number of intervals. $\square$

Proof of Claim 4 I construct an equilibrium. Consider the following strategy profile: the sender performs $check_\zeta$ (explores the most important criterion) and submits a "separating" report at $t = 1$, denoted by $m^*(\zeta)$. Regardless of the report at $t = 1$, no valuable information is transmitted at $t = 2$ (the sender "babbles"). The receiver plays his best response (denoted by $a^*(m)$). The choice of $b \leq \lceil \frac{k}{2} \rceil$ ensures that, provided that the sender's information structure corresponds to knowing only the value of ζ, she is willing to reveal the truth. To show that this is indeed an equilibrium I now show that the sender cannot benefit from performing $check_\eta$ instead of $check_\zeta$. Let $a^S(\eta)$ denote the most-desirable action, from the sender's perspective, given a realization of η (while ζ is unknown). Notice that

$$a^S(\eta) = argmax_a E[u^S(a, \theta)|\theta \in \{(k+1)\zeta + \eta\}] = \frac{k(k+1)}{2} + \eta + b.$$

Substitute and rearrange terms to get

$$E[u^S(a^*(m(\eta)),\theta)] \leq -\frac{1}{(k+1)^2}\sum_{\eta=0}^{k}\sum_{\zeta=0}^{k}(\zeta\cdot(k+1)+\eta+b-a^S(\eta))^2 = -(k+1)\sum_{\zeta=0}^{k}(\zeta-\frac{k}{2})^2 =$$

$$= -\frac{k(k+2)(k+1)^2}{12} \underset{(*)}{<} -(\frac{(k+1)^2}{12}+1) < -\frac{k^2+2k+12}{12} = E[u^S(a^*(m^*(\zeta)),\theta)].$$

The RHS is the sender's expected utility if she follows the equilibrium strategy. The LHS is her utility if she deviates and privately explores η instead of ζ. The inequality denoted by $(*)$ holds for every integer $k \geq 2$. It is straightforward to verify that for $k = 1$, the suggested profile of strategies also constitutes an equilibrium. In the latter case, however, it is a weak equilibrium. To complete the proof, note that the suggested equilibrium splits Θ into $k + 1$ intervals of the same length. Thus, the variance of this partition is necessarily lower than the variance of the unbalanced partition that is obtained under the optimal equilibrium in the fully informed sender benchmark (see claim 3). $\square$

References

1. Aghion, Philippe, and Jean Tirole. "Formal and real authority in organizations." Journal of political economy (1997): 1-29.

2. Argenziano, Rossella, Francesco Squintani, and Sergei Severinov. "Strategic Information Acquisition and Transmission." American Economic Journal: Microeconomics 8.3 (2016): 119–155.

3. Austen-Smith, David. "Strategic Transmission of Costly Information." Econometrica 62, no. 4 (1994): 955-63.

4. Blume, Andreas, Oliver J. Board, and Kohei Kawamura. "Noisy talk." Theoretical Economics 2 (2007): 395-440.

5. Crawford, Vincent P., and Joel Sobel. "Strategie Information Transmission." Econometrica 50, no. 6 (1982): 1431-1451.

6. Di Pei, Harry. "Communication with endogenous information acquisition." Journal of Economic Theory 160 (2015): 132-149.

7. Eliaz, Kfir, and Alexander Frug. When to Learn What in Bilateral Trade. mimeo, (2016).

8. Fischer, Paul E., and Phillip C. Stocken. "Imperfect information and credible communication." Journal of Accounting Research 39.1 (2001): 119-134.

9. Frug, Alexander. Gradual Learning and Efficient Reporting. mimeo, 2014.

10. Frug, Alexander. "A note on optimal cheap talk equilibria in a discrete state space." Games and Economic Behavior 99 (2016): 180–185.

11. Frug, Alexander. "Dynamic cheap talk with static informational control." Economics Letters 143 (2016): 118-120.

12. Gentzkow, Matthew, and Emir Kamenica. Disclosure of endogenous information. mimeo University of Chicago, 2012.

13. Golosov, Mikhail, Vasiliki Skreta, Aleh Tsyvinski, and Andrea Wilson. "Dynamic strategic information transmission." Journal of Economic Theory 151 (2014): 304-341.

14. Goltsman, Maria, Johannes Hörner, Gregory Pavlov, and Francesco Squintani. "Mediation, arbitration and negotiation." Journal of Economic Theory 144, no. 4 (2009): 1397-1420.

15. Gordon, Sidartha (2010). "On Infinite Cheap Talk Equilibria," Technical report, Sciences Po.

16. Green, Jerry R., and Nancy L. Stokey. A two-person game of information transmission. Harvard Institute of Economic Research, 1980.

17. Ivanov, Maxim. "Informational control and organizational design." Journal of Economic Theory 145.2 (2010): 721-751.

18. Ivanov, Maxim. "Dynamic information revelation in cheap talk." The BE Journal of Theoretical Economics (Contributions) 15 (2015): 251–275.

19. Ivanov, Maxim. "Dynamic learning and strategic communication." International Journal of Game Theory 45.3 (2016): 627–653.

20. Kamenica, Emir, and Matthew Gentzkow. "Bayesian Persuasion." The American Economic Review (2011): 2590-2615.

21. Krishna, Vijay, and John Morgan. "The art of conversation: eliciting information from experts through multi-stage communication." Journal of Economic theory 117.2 (2004): 147-179.

22. Spence, Michael. "Job market signaling." The quarterly journal of Economics (1973): 355-374.

23. Sobel, Joel. "Giving and receiving advice." Econometric Society 10th World Congress. 2010.