Strategic Learning And Information Transmission*

Alexander Frug†

March 17, 2014

Abstract

It is often the case that an expert needs to explore the state of the world before advising a decision maker. We consider a dynamic cheap talk environment with gradual expert's learning. Both players know what experiments are available to the expert. We show that the gradualness in expert's learning can enhance informativeness in communication even when the learning process is endogenous (part of the expert's strategy), unobservable by the decision maker and non-verifiable. The result suggests that even in the absence of an "objective" reason to hurry with information transmission, putting the expert in some form of "strategic" pressure can be beneficial to both players.

1 Introduction

This paper addresses the problem of information transmission in a cheap talk environment in which the expert's learning is gradual and strategic. At the beginning both the expert and the decision maker are not informed about the state of the world but the expert has the ability to explore it. The exploration of the state may require some time and effort. To emphasize the particular effect of gradualness in the expert's learning we assume that the learning is costless. Also, although gradual, we assume that the learning can be fully accomplished within a short period of time, so we also ignore discounting.

The expert's learning process must follow some general principle, say, certain experiments can be performed over time. Nevertheless, it is often the case that some aspects of the exploration process are left to the expert's discretion; e.g., the selection of a particular collection of experiments to be performed and the ordering of these experiments. The purpose of this analysis is to show that gradualness in expert's learning can be exploited towards a more informative communication even when it is strategic, unobservable and non verifiable.

*I thank Elchanan Ben-Porath, Eddie Dekel and Kfir Eliaz for useful comments and suggestions. I owe a special acknowledgment to my advisor Ady Pauzner for his guidance.

†Tel-Aviv University. Email: frugalex@post.tau.ac.il.

Consider the following simple example: A patient (decision maker) can be healthy (state *0*), or he can have either a mild disorder (state 1) or a severe problem (state 2). Assume that the three states are equally likely. The uninformed patient has to determine the level of his healthcare expenditure, say a, and assume that his first best level of healthcare expenditure at state θ is $a = \theta$. An expert (doctor) can advise to the patient and the talk is cheap. If the patient is healthy, the interests of the players are aligned. Otherwise, the interests are not perfectly aligned and the doctor has a positive bias of 1. In other words, her most desired level of patient's expenditure on healthcare in state θ is $a = \theta + 1$, for $\theta \in \{1, 2\}$. Both players want to minimize the (expected) quadratic loss.

First, note that if the expert is informed, in the most informative equilibrium the patient can learn only whether he is healthy or not. It is impossible to separate $\{1, 2\}$ in equilibrium because in state $\theta = 1$, the expert would prefer to report that "$\theta = 2$".

Now consider the following situation. Assume that the expert is initially uninformed but she has the ability to explore the state of the world as follows. Two experiments are available: experiment(1) and experiment(2). Experiment(i) privately reveals to the expert whether the state is i or not, and each experiment takes one unit of time. Which experiments were performed and whether the expert performed any experiments at all are unobservable and non-verifiable to the decision maker.

In this case, a fully revealing equilibrium (the expert learns and reports the state accurately) exists. At date 1, the expert performs experiment(2). If it reveals that $\theta = 2$, the expert reports this fact *immediately*. Otherwise, she continuous her exploration process by performing experiment(1) at date 2, and reports (truthfully) its outcome.

The expert knows that her only opportunity to induce the highest possible patient's expenditure ($a = 2$), is by reporting that $\theta = 2$ as early as possible. Provided that she can run only one experiment at a time, it is most desirable from her perspective to perform experiment(2) at date 1 and report that $\theta = 2$ only if this is the outcome of the experiment. Otherwise, she has an incentive to conduct experiment(1) at date 2 and report truthfully. This provides the decision maker with some information regarding the expert's information structure. In particular, if a report is submitted at date 1, he learns that the expert could not perform both experiments.

Even if we assume that there is no objective reason to hurry with information transmission, the most informative equilibrium creates a strategic pressure for the expert that turns out to be beneficial to both players.

The literature on costless communication (cheap-talk) between informed experts and uninformed decision makers began with the seminal contributions of Crawford and Sobel (1982) and Green and Stokey (1980)[1]. Many authors have studied a variety of cheap talk environments under different specifications. Sobel (2010) provides a comprehensive literature review on 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 struc-

[1]Earlier literature on costly signaling can be traced back to Spence (1973) and Rothschild and Stiglitz (1976).

ture from the Receiver's perspective for the leading uniform-quadratic case. In a recent contribution by Ivanov (2012) an optimal two stage learning protocol is suggested. The author assumes that the Receiver can perfectly control the dynamic structure of the Sender's information. By coupling distant separable elements[2] at the early stage and by conditioning future experiments on early truthful communication, the Receiver can successfully elicit full information. Although the Sender is initially uninformed, her learning is not strategic but designed by the Receiver in these models.

Austen-Smith (1994) proposed a model in which the Sender chooses whether to (perfectly) learn the state and has the ability to prove information acquisition. Like 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 the expert's choice whether to learn depends only on the realization of the cost of learning which is privately observed by her. The author shows that in this environment the range of biases for which informative communication is possible is extended relative to the Crawford and Sobel (1982) benchmark model.

Several recent contributions also involve expert's strategic learning. In Argenziano et al. (2011) the Sender can affect the quality of her information by selecting how many Bernoulli trials, parametrized with the number θ that corresponds to the state of the world, to perform. The trials are costly for the Sender. The authors compare between "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 over-invest in learning. In Kamenica and Gentzkow (2011) the Sender selects a "signal", i.e., information structure to persuade a decision maker. Like in our model the selection of the information structure is costless and it is left to the Sender's discretion. Unlike in our model, 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. Because the information is costly for the Sender, endogenous information will always be disclosed in equilibrium so disclosure requirements have no effect on the set of the equilibrium outcomes. A related result appears in Di Pei (2013). In that paper the Sender gathers costly information before advising. Like in our model, the learning is strategic and unobservable by the Receiver. Unlike in our model, the information is costly. In this environment, the Sender communicates all her information and all equilibria are less informative than the most informative one in Crawford and Sobel (1982).

All the above models that contain strategic learning are essentially static while the key feature of the model we study is the dynamic nature of the expert's learning. When the expert reports to the decision maker she also provides a signal about her information quality. Frug (2014) considers a model of gradual exogenous learning and communication. In this model, by designing an appropriate reporting protocol, the players manage to exploit the gradualness of the Sender's information arrival towards a more informative communication. The current study suggests that players can benefit from the gradualness of the expert's learning even when the expert's learning is strategic (endogenous), unobservable and non verifiable.

[2]Similar idea also appears in Golosov et al. (2013).

2 Environment

2.1 State Space

The platform of this study is the Crawford and Sobel (1982) model of cheap talk. To present the forces behind our main result in the most transparent way we begin with a particular, most convenient, state space. It differs from the state space in Crawford and Sobel (1982) in two manners. First, we assume discrete rather than continuous state space. Second, we assume that there exists a particular state ϕ which is, in some sense, disjointed from other states and often can be interpreted as "nothing happened".

The inclusion of ϕ has clear technical implications, it simplifies the exposition and highlights the main intuition behind the result. This was our primary motivation. Later, we discuss exactly where and how this affects the result and develop a related result in a more standard state space that does not contain ϕ.

2.2 Model

There are two players - Sender and Receiver, $N = \{S, R\}$. A state $\theta \in \Theta = \{\phi, 1, 2, .., n\}$ is distributed according to a common prior P. Action $a \in \{\phi\} \cup \mathbb{R}$ must be selected by the Receiver at time $T > n$. Player i's utility function is

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

where $u^i(\cdot, \theta) \leq 0$ is a single-peaked function, $argmax_a\, u^i(a, \theta)$ is (strictly) increasing in θ, and the Sender is positively biased for all θ:

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

Assume that the Sender's preferences satisfy *p-supermodularity* by which we mean the following condition: $p(\bar{\theta})[u^S(\bar{a}, \bar{\theta}) - u^S(\underline{a}, \bar{\theta})] > p(\underline{\theta})[u^S(\bar{a}, \underline{\theta}) - u^S(\underline{a}, \underline{\theta})]$ whenever $\bar{a} > \underline{a}$ and $\bar{\theta} > \underline{\theta}$. This property replaces the standard supermodularity condition that is often assumed in information economics when considering interactions with informed agent.

Next we address the manner in which the expert (Sender) explores the state of the world and the reporting protocol.

Sender's Learning For each $k \in \{1, .., n\}$, there is an experiment $\epsilon_k = \mathbf{1}_{\theta=k}$ that reveals whether $\theta = k$ or not. Conducting any experiment ϵ_k takes one unit of time. Only the Sender has access to the set of experiments $\{\epsilon_k | k \in \{1, .., n\}\}$ and whether she had conducted any experiment is unobservable and non verifiable. Notice that we assume 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 a complete elimination of other states.

Sender's Reporting We assume that the Sender is allowed to submit only one report but she can choose the report timing.

Formally, let M be a set of available messages[3] and $\mathbf{T} = \{1, 2.., T\}$ be a discrete time-line[4]. For each $t \in \mathbf{T}$, the Sender may remain silent (i.e. submit an 'empty report' $m_t = \phi$) or submit a report $m_t \in M$. Any terminal history of reports contains at most one non-empty report.

2.3 Strategies and Equilibrium

The Receiver chooses an action rule $\alpha(\{m_t\}) \in \Delta A$ that attaches to any terminal history of reports a distribution over actions. If $|supp(\alpha(\{m_t\}))| = 1$ for every terminal history of reports, that is, the Receiver's strategy is never mixed, we denote the corresponding pure strategy by $a(\{m_t\})$.

The Sender's strategy is twofold. She chooses both how to explore θ and how to report. Clearly, the flexibility of the reporting component of the strategy, $m = \{m_t\}_{t \in \mathbf{T}}$, depends on its learning component f. To make it precise let $f : \mathbf{T} \to \{0, 1, .., n\}$ be a given plan of learning, specifying which experiment $\epsilon_{f(t)}$ to be conducted at date $t \in \mathbf{T}$, where ϵ_0 denotes "do nothing". The reporting component of the Sender's strategy, $m = \{m_t\}_{t \in \mathbf{T}}$, must be "$f - measurable$". Formally, for each $t \in \mathbf{T}$ a sigma algebra $\mathcal{F}_t(f) = \sigma\{\{f(s)\} | s \leq t\}$ is formulated and a (pure) report $m_t(\theta)$ must be $\mathcal{F}_t(f)$ - measurable (can depend only on the information collected up to this date). Since the continuation of learning is relevant only if previous experiments have not revealed θ (and a non empty report has not been submitted), it is without loss to assume that the whole plan of learning f is selected at the beginning of the interaction and is not periodically revised.

To summarize, a Sender's strategy is $\sigma \in \Delta\{(f, m) | f \in \{0, 1, .., n\}^{\mathbf{T}}, m = \{m_t(\theta)\} \, \text{is} \, \mathcal{F}_t(f) - adapted\}$. Whenever f is not random, we simply denote $s = (f, m^f)$, where m^f is $f - measurable$ and possibly mixed.

A (Perfect Bayesian) equilibrium consists of a Sender's strategy s^*, a Receiver's action rule $\alpha^*(\{m_t\})$ and a belief $\mu_t(\theta | \{m_s\}_{s \leq t})$ for each $t \in \mathbf{T}$, such that μ_t is updated according to the Bayesian rule whenever possible, $\alpha^*(\{m_t\})$ is optimal given the beliefs, and s^* is optimal given $\alpha^*(\{m_t\})$. We call an equilibrium e "*non-redundant*" if the Receiver attaches a distinct action to any element in his information partition induced by e. It is easy to see that for any equilibrium e', there exists a non-redundant equilibrium e such that both equilibria identically map states into actions. In this sense these equilibria are essentially equivalent and therefore we confine attention to non-redundant equilibria.

Observation 1 *Since the learning is unobservable and non verifiable, and it consists of deterministic and costless experiments, we can assume without loss of generality that the Sender's (pure) learning plan is a permutation $f \in S_n$.*

[3]Assume that this set is rich enough to transmit any information available for the Sender.

[4]Recall that we assume that $T > n$ to capture the idea that, in principle, the time available for the exploration places no restrictions on the expert's ability to become fully informed.

3 Main Result

Proposition 1 *There exists a separating equilibrium.*

Proof Consider the following Sender's strategy $s^* = (f, m^f)$: $f(t) = n+1-t$ and $m^f_\tau \neq \phi$ iff $\epsilon_{f(\tau)} = 1$. Notice that this Sender's strategy is of full support (whenever $m_t \neq \phi$ let the Sender report a random message in M). Therefore, the whole system of beliefs μ_t is pinned down by the Bayes rule. Accordingly, the Receiver's best response is to select $a(\{m_t\}_{t\in\mathbf{T}}) = argmax_a\, u^R(a, \theta = n+1-\tau)$ if $m_\tau \neq \phi$ for some $\tau \leq n$, and $a(\{m_t\}_{t\in\mathbf{T}}) = \phi$ if $m_t = \phi$ for all $t \leq n$. It is left to show that (f, m^f) is the Sender's best response to the above specified Receiver's action rule.

Since $u^S(a,\theta)$ is single-peaked for each θ and $argmax_a\, u^R(a,\theta) < argmax_a\, u^S(a,\theta)$, the reporting component m^f is optimal given f. So we have to show that there does not exist an alternative learning plan g, and a $g-measurable$ reporting strategy m^g such that (g, m^g) constitutes a profitable deviation given the Receiver's strategy.

Let $g \in S_n$. We construct a transition sequence $s(g)$ from f to g, that consists of swaps, such that each swap is a worsening for the Sender. (under the assumption that we adjust the reporting strategy optimally).

Step 1 *Take care of* $g(1)$: There exists $k_1 \geq 1$ such that $g(1) = f(k_1)$. Namely, the experiment $\epsilon_{g(1)}$ is scheduled for date k_1 according to f. We now consider f and sequentially advance $\epsilon_{f(k_1)}$ until it occupies the first position in the resulting learning plan. Let,

$$f_1 = (f(1), f(k_1)) \circ (f(2), f(k_1)) \circ ... \circ (f(k_1-1), f(k_1)) \circ f$$

Note that $f_1 \in S_n$ such that $f_1(1) = g(1)$. Define the f_1-prefix, s_1, of the transition sequence by

$$s_1 =< (f(k_1-1), f(k_1))\, , \,(f(k_1-2), f(k_1)), ..., (f(2), f(k_1))\, , \,(f(1), f(k_1)) >$$

Step j *Take care of* $g(j)$: Provided that $s_1, .., s_{j-1}$ are defined and $f_{j-1} \in S_n$ satisfies $f_{j-1}(i) = g(i)$ for each $i < j$, there exists $k_j \geq j$ such that $g(j) = f_{j-1}(k_j)$. Again, we sequentially advance $\epsilon_{f_{j-1}(k_j)}$ until it occupies the jth position in the resulting learning plan. Let,

$$f_j = (f_{j-1}(1), f_{j-1}(k_j)) \circ (f_{j-1}(2), f_{j-1}(k_j)) \circ ... \circ (f_{j-1}(k_j-1), f_{j-1}(k_1)) \circ f_{j-1}$$

$f_j \in S_n$ such that $f_j(i) = g(i)$ for all $i \leq j$. Define the f_j-continuation, s_j, of the transition sequence by

$$s_j =< (f_{j-1}(k_j-1), f_{j-1}(k_1))\, , \,(f_{j-1}(k_j-2), f_{j-1}(k_1))\, , .., \,(f_{j-1}(1), f_{j-1}(k_j)) >$$

Notice that $f_{n-1} = g$ and $s(g) =< s_1, s_2, ..., s_{n-1} >$ is the whole transition sequence from f to g.

Remark 1: Every swap (l, h) in $s(g)$, with $h > l$, involves a delay of the experiment ϵ_h by one unit of time while ϵ_l is expediated by one unit of time.

Remark 2: Even before we performed the swap (l, h), ϵ_h was scheduled "too late" from the Sender's perspective (only actions that are "too low" for the Sender are inducible when h is inspected).

Due to ϕ, the Sender prefers not to induce an action $a \neq \phi$ as long as the possibility $\theta = \phi$ is not ruled out. Therefore the swap (l, h) affects the reporting strategy only if $\theta \in \{l, h\}$. Thus, to evaluate the swap (l, h) it is sufficient to calculate the total value from postponing ϵ_h from t to $t+1$ and expediating ϵ_l from $t+1$ to t.

Let $\underline{a} = argmax_a \, u^R(a, \theta = n - t)$ and $\bar{a} = argmax_a \, u^R(a, \theta = n - t + 1)$. Recall that $\underline{a} < \bar{a}$. By remarks 1 and 2 the Sender will induce the highest possible action immediately upon she finds out $\epsilon_h = 1$. The net value from postponing ϵ_h from t to $t+1$ is

$$p(\theta = h)[u^s(\underline{a}, h) - u^s(\bar{a}, h)]$$

Given a learning plan f' that is consistent with applying a prefix of $s(g)$ on f, finding out that $\theta = l$ one period earlier will not affect Sender's reporting strategy in the case that $u^s(\bar{a}, l) < u^s(\underline{a}, l)$. The net value of expediating ϵ_l from $t+1$ to t is therefore,

$$p(\theta = l)max\{0, u^s(\bar{a}, l) - u^s(\underline{a}, l)\}$$

The total value of the swap is therefore

$$p(\theta = h)[u^s(\underline{a}, h) - u^s(\bar{a}, h)] + p(\theta = l)max\{0, u^s(\bar{a}, l) - u^s(\underline{a}, l)\} \leq$$

$$\leq p(\theta = h)[u^s(\underline{a}, h) - u^s(\bar{a}, h)] + p(\theta = l)[u^s(\bar{a}, l) - u^s(\underline{a}, l)] < 0$$

by p-supermodularity of u^S. QED

We interpret this result as follows. Even though there is no "objective" reason to hurry with information transmission, sometimes it is possible to put the expert in an effective form of "strategic" pressure. In particular, for the expert 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, provided that it respects the expert's physical ability to acquire information.

Also, it is worthwhile to remark that the assumption that $U^i(a, \theta) = -\infty$ if $a = \phi \oplus \theta = \phi$ is clearly too strong and, as the simple example in Introduction suggests, a much weaker assumption would suffice to support this equilibrium. However, this assumption simplifies the exposition and helps in the comparison with the benchmark in the next section.

4 The Informed Sender Benchmark

At this point, it is natural to compare the result in proposition 1 with the most desirable information partition can be attained in equilibrium a la Crawford and Sobel (1982) - CS equilibrium. To do so, we will focus on the discrete uniform quadratic case. Assume that the utility functions are given by the negative quadratic loss $u^R(a, \theta) = -(\theta - a)^2$ and $u^S(a, \theta, b) = -(\theta + b - a)^2$ where $b \in \mathbb{N}$ is the Sender's constant bias.

In the appendix, we characterize the most desirable CS equilibrium (for both the state space with ϕ and without ϕ). In brief, one can easily derive the most desirable equilibrium information partition as follows. Define $x_1 = 1$ and $x_{j+1} = x_j + 4b - 2$ for $j \in \mathbb{N}$. Next, find the index i^* that

$$\sum_{i=1}^{i^*} x_i \leq n < \sum_{i=1}^{i^*+1} x_i$$

and define

$$y = n - \sum_{i=1}^{i^*} x_i$$

Let $\{y_i\}_{i=1}^{i^*}$ be the maximal (in lexicographic sense) weakly increasing sequences of length i^* such that $y_i \in \{0\} \cup \mathbb{N}$ and $\sum_{i=1}^{i^*} y_i = y$. The most desirable CS equilibrium in this environment is given by separating the state ϕ (if exists) and partitioning the states $\{1, 2, ..., n\}$ into monotone subsets of magnitude $x_i + y_i$.

For example, let $\Theta = \{\phi, 1, 2, .., 10\}$ and $b = 1$. Then, $x_1 = 1$, $x_2 = 3$, $x_3 = 5$, $x_4 = 7$. Accordingly $i^* = 3$ and $y = 1$. The maximal weakly increasing sequence $\{y_i\}_{i=1}^{3}$ is then given by $\{0, 0, 1\}$. Recall that under the assumptions in previous section a fully revealing equilibrium exists, while the structure of the most desirable CS equilibrium is

ϕ	1	2,3,4	5,6,7,8,9,10

A standard feature of this equilibrium is that the higher "intervals" contain more elements. Namely, holding b fixed, as n increases the Receiver obtains less precise information regarding high states of the world. This is, of course, not the case in proposition 1 and as we show below, this distinction remains true also in the standard discrete state space, namely $\{1, 2, ..., n\}$.

5 The Role of ϕ and The Standard State Space

In this section we discuss the role of the special state ϕ and develop a related result in a more standard state space. The absence of ϕ might severely complicate the derivation of most informative equilibria. On the other hand, its effect on the equilibrium is significant only if the bias is large relative to the range of the possible states. When the bias is relatively small, the main force of "strategic pressure" is dominant but one can not push it too far since a critical feature of the equilibrium is that the "residual" space shrinks with time, inflating the relative significance of the bias.

Consider again the discrete uniform quadratic case with ϕ and assume for simplicity that $b = 1$. As long as the Sender hadn't learned θ, the possibility that $\theta = \phi$ prevents her from inducing an "active" action $(a \neq \phi)$. Assume now that $\Theta = \{1, 2, .., n\}$ and assume that the Sender follows the same learning plan as in proposition 1. During the early periods she does not have an incentive to submit a "false positive" report (induce an action even though the experiments did not reveal θ). By inducing a high action, she benefits in the case that θ is just below the induced action but she loses if θ turns out to be very low. Thus, it seems that, during the early stages, the Sender's behavior as in proposition 1 is plausible in this environment as well. But, as we proceed, the risk of "far low states" reduces with time. For example if $\epsilon_k = 0$ for all $k > 1$, the expert can be better off by inducing $a = 2$ (submitting a false positive report that corresponds to $\epsilon_2 = 1$) rather than waiting another period to find out (what she already knows) that $\theta = 1$ and to induce $a = 1$.

This suggests that if $\Theta = \{1, .., n\}$ one can proceed with the logic of the equilibrium as in proposition 1, but not all the way down. Instead, at some point, the players must engage into another reporting phase, which will not be fully revealing. In the next proposition, we prove that for the discrete uniform quadratic case, there exist 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. Essentially, we create an artificial substitute for ϕ to prevent the Sender from submitting false positive reports and put her under similar strategic pressure by shrinking the set of inducible actions over time.

Let $\Theta = \{1, .., n\}$ and assume that the states are equally likely, $p(k) = \frac{1}{n}$ for each $k \in \Theta$. Let $u^R(a, \theta) = -(\theta - a)^2$ and $u^S(a, \theta, b) = -(\theta + b - a)^2$ where $b \in \mathbb{N}$ denotes the Sender's bias.

Sender's Learning For each $k \in \{1, .., n\}$, there is an experiment $\epsilon_k = \mathbf{1}_{\theta=k}$ that reveals whether $\theta = k$ or not. Conducting any experiment ϵ_k takes one unit of time. Only the Sender has access to the set of experiments $\{\epsilon_k | k \in \{1, .., n\}\}$ and whether she had conducted any experiment is unobservable and non verifiable. Similarly to observation 1, we can assume without loss of generality that a learning plan is a permutation of $\{1, .., n\}$.

Sender's Reporting As before, we assume that the Sender is allowed to submit only one report but she can choose the report timing.

Proposition 2 *There exists an equilibrium that attains complete separation of every state $\theta \geq 10b$.*

Proof Consider the following profile of strategies. Sender's strategy $s^* = (f, m^f)$: $f(t) = n+1-t$ for all $t \in \{1, .., n\}$, $m^f_\tau \neq \phi$ iff $\epsilon_{f(\tau)} = 1$ and $\tau \in \{1, .., n+1-10b\}$. Receiver's strategy: $a^*(\{m_t\}) = n+1-\tau$ if $m_\tau \neq \phi$ for some $\tau \in \{1, .., n+1-10b\}$ and $a(\{m_t\}) = 5b$ if $m_t = \phi$ for all $t \in \{1, .., n+1-10b\}$. We will prove that this strategy profile constitutes an equilibrium.

First, it is clear that the Receiver's strategy is a best response to the Sender's strategy. Also, it is easy to see that given the Receiver's strategy and the learning plan f, the reporting component of the Sender's strategy m^f is optimal: The Sender can not benefit from postponing a report once θ is revealed because this will induce a lower Receiver's action which is worse for the Sender. Also, the Sender can not benefit from a "false positive" report $m_t \neq \phi$ in the case that $\epsilon_{f(\tau)} = 0$ for all $\tau \leq t$, for some t. To see this let $\bar{a} \geq 10b$ be the action that is induced by a non-empty report at date t, we prove below that

$$E[u^S(\bar{a}, \theta)] = -\frac{1}{n}\sum_{\theta=1}^{\bar{a}-1}(\theta + b - \bar{a})^2 < -\frac{1}{n}[\sum_{\theta=1}^{10b-1}(\theta + b - 5b)^2 + (\bar{a} - 10b)b^2]$$

where the LHS is the expected utility from a "false positive" report that induces the action $\bar{a}$ and the RHS is the Sender's expected utility from continuing to follow the strategy s^*. The inequality holds because

$$\sum_{\theta=1}^{\bar{a}-1}(\theta + b - \bar{a})^2 > \sum_{\theta=1}^{\bar{a}-10b}(b)^2 + \sum_{\theta=\bar{a}-10b+1}^{\bar{a}-1}(\theta + b - \bar{a})^2 = \sum_{\theta=1}^{10b-1}(\theta + b - 10b)^2 + (\bar{a} - 10b)b^2$$

and it remains to note that $\sum_{\theta=1}^{10b-1}(\theta + 9b)^2 > \sum_{\theta=1}^{10b-1}(\theta + 4b)^2$ which follows from the observation that $4b$ is "closer" than $9b$ to the average of the summation range and the fact that we use a quadratic loss function.

The intuition is simple. Although the Sender is positively biased, inducing an action $\bar{a} > 10b$ for all $\theta < \bar{a}$ is not profitable because this action is "too high". This will indeed be profitable for the Sender if it turns out that θ is sufficiently close to $\bar{a}$ but she will bear a significant loss relative to s^* for low realizations of θ.

It is left to show that there does not exist an alternative learning plan $g : \{1, ..n\} \to \Theta$ and a $g - measurable$ reporting policy m^g such that $s = (g, m^g)$ is a profitable deviation for the Sender. We show this in two steps but we first provide some definitions and prove the following Lemma.

For any Sender's strategy $s = (g, m^g)$ and state $\theta \in \Theta$, we define $a_s(\theta)$ to be the Receiver's action induced by the Sender's strategy s at state θ. Also, for each θ let $d_s(\theta) = |argmax_a u^S(a, \theta) - a_s(\theta)|$, namely, the distance between the s-induced action and the Sender's first best in the state θ.

Lemma 1 Let $a \geq 10b$ and let $s = (g, m^g)$ be a Sender's strategy that does not induce the action a in the case that $\epsilon_{g(t)} = 0$ for all $t \leq n + 1 - a$. Then for any $\theta \geq a$, $\frac{1}{n+1-\theta}\sum_{\theta'=\theta}^{n} d_s(\theta') \geq b$.

Proof Assume $\underline{\theta} \geq a$ satisfies $d_s(\underline{\theta}) < b$. Then $\underline{\theta} < n$ and $\epsilon_{\underline{\theta}}$ is expediated relative to f (it is performed before $t = n+1-\underline{\theta}$). Expediating $\epsilon_{\underline{\theta}}$ by one unit of time necessarily means that there exist $\theta > \underline{\theta}$ such that ϵ_θ is postponed by (at least) one unit of time relative to f. According to the Receiver's strategy and the Sender's preferences, $d_s(\theta)$ is increased by (at least) 1 relative to $d_{s^*}(\theta)$ where $s^* = (f, m^f)$. Note that for any $\theta \geq 10b$, $d_{s^*}(\theta) = b$ and this proves the Lemma.

Step 1: First we show that there does not exists a profitable deviation to a strategy $s = (g, m^g)$ that induces an action $a \geq 10b$ only if $\epsilon_{g(t)} = 1$ for some $t \leq n + 1 - a$.

Let $s = (g, m^g)$ be such a strategy. By Lemma 1, $\frac{1}{n+1-10b} \sum_{\theta=10b}^{n} d_s(\theta) \geq b$. Therefore, since $d_{s^*}(\theta) = b$ for any $\theta \geq 10b$ and we assume the uniform quadratic case, the Sender's expected utility on states $\theta \geq 10b$ under the strategy s is bounded by her expected utility on these states under the strategy s^*.

Consider now the states $\theta < 10b$. If $d_s(\theta) \geq d_{s^*}(\theta)$ for all $\theta < 10b$, s^* is at least as good as s and step 1 follows. Assume that there exists $\underline{\theta} < 10b$ with $d_s(\theta) < d_{s^*}(\theta)$. In this case, $a_s(\underline{\theta}) \geq 10b$ which means that $\epsilon_{\underline{\theta}}$ is performed no later than $t = n+1-10b$. As a result, there must exists a state $\bar{\theta} \geq 10b$ such that $\epsilon_{\bar{\theta}}$ is not performed by the Sender at $t \leq n + 1 - 10b$. Thus, $a_s(\bar{\theta}) = 5b$.

Assume first that $a_s(\underline{\theta}) \leq \bar{\theta}$. In this case, at state $\bar{\theta}$, the Sender's utility is decreased by at least $(6b)^2 - b^2$ (the smallest decrease in utility is obtained at $\bar{\theta} = 10b$). The maximal gain at state $\underline{\theta}$ (obtained if $\underline{\theta} = 10b - 1$) is $(6b - 1)^2 - (b - 1)^2$. For any $b \in \mathbb{N}$, $(6b)^2 - b^2 > (6b - 1)^2 - (b - 1)^2$ so the Sender is not better off. If we dismiss the assumption that $a_s(\underline{\theta}) \leq \bar{\theta}$, the maximal gain at $\underline{\theta}$ can be increased up to $(6b-1)^2$ which is obtained if the Sender's type $\underline{\theta}$ induces her first best action ($a_s(\underline{\theta}) = 11b-1$). However, since $a_s(\underline{\theta}) > \bar{\theta}$, additional experiments must be postponed. Moreover, a decrease in $d_s(\underline{\theta})$ by one unit (below $b - 1$) leads to an increase by one unit above b at some other state, which is not profitable as the states are equally likely and the loss function is quadratic. This proves step 1.

Step 2: Now we show that a deviation to a strategy in which an action $a \geq 10b$ is induced after $\epsilon_{g(t)} = 0$ for all $t \leq n + 1 - a$, is not profitable.

Let $s = (g, m^g)$ be a Sender's strategy that induces an action $a = 10b$ if the learning plan g does not reveal θ until date $t = n + 1 - 10b$. In this case there are at least $10b - 1$ states, $\theta_1 < \theta_2 < ... < \theta_{10b-1}$, in which the Sender induces the action $a = 10b$. First, assume that $\theta_1 \geq 3b$. The vector of smallest possible distances $(d_s(\theta_j))_{j=1}^{10b-1}$ is given by

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

On the other hand, the vector of distances $(d_{s^*}(\theta_j))_{j=1}^{10b-1}$ is bounded by

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

We assume the uniform quadratic case. Therefore, to compare between s and s^* on $\{\theta_1, ..., \theta_{10b-1}\}$ we can sum up the square distances and compare the costs (negative utility) at these states. The following inequality holds[5] for any $b \in \mathbb{N}$,

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

therefore, s^* is better than s on the states $\{\theta_1, ..., \theta_{10b-1}\}$.

Similarly to the considerations at the beginning of step 1, by Lemma 1 the Sender's expected utility on states $\theta \geq \theta_{10b-1}$ under the strategy s is bounded by her expected utility on these states under the strategy s^*.

It is also impossible that s does better than s^* at states $\theta < 5b$ because s^* induces the best possible action in the support of $a^*(\cdot)$ for these states. This shows that s^* is preferred over s by the Sender. To complete the proof two comments are needed:

1. By replacing one of the states in $\{\theta_1, ..., \theta_{10b-1}\}$ with a state $\theta < 3b$ the LHS of the strict inequality in $(\star)$ will increase by, at least, $(\theta + b - 10b)^2 - (5b-1)^2$ and the RHS will increase by, at most, $(\theta + b - 5b)^2 - (6b-1)^2$. For any $b \in \mathbb{N}$ and $\theta < 3b$, we have $(\theta + b - 10b)^2 - (5b-1)^2 > (\theta + b - 5b)^2 - (6b-1)^2$. Therefore, the Sender is better off under s^* rather than under s even if we dismiss the assumption that $\theta_1 \geq 3b$.

2. If we replace the action $a = 10b$, with another action $\hat{a} > 10b$, we necessarily add more states which induce $\hat{a}$ (The Sender had less time to explore the state). This will add distances $d_s \geq 5b$ to the LHS of the strict inequality in $(\star)$. On the other hand, we will only supplement more distances $d_{s^*} = b$ to the RHS of the strict inequality in $(\star)$. Clearly, the inequality remains true. QED

Proposition 2 does not present the most informative equilibrium. Its main purpose is to illustrate how the logic of the equilibrium in proposition 1 can be applied for a more standard state space. Importantly, the threshold in proposition 2 is independent of n. Moreover, for the construction of equilibrium as in proposition 2, this threshold is pretty tight: selecting $c \leq 8$ will not suffice for any $b \in \mathbb{N}$ and for sufficiently small values of b, the profile of strategies that corresponds to $c = 9$ also does not constitute an equilibrium. However, for sufficiently large values of b, it is possible to separate a few additional states below the suggested threshold. Also, choosing a more complicated equilibrium below the threshold (for example, playing a CS equilibrium in this region) can also enhance informativeness.

6 Concluding Remarks

In this short paper, we demonstrated how the players can benefit from the fact that the expert's learning is gradual. The main intuition is that "strategic pressure" can be a valuable tool to elicit more information in an environment where the expert's learning is gradual, strategic and unobservable. In particular, it is important that the set of inducible actions will shrink sufficiently fast over time, respecting the expert's

[5] By using the formula for the sum of squares of the first K naturals: $\sum_{k=1}^{K} k^2 = \frac{K^3}{3} + \frac{K^2}{2} + \frac{K}{6}$

physical ability to acquire information.

6.1 Cut-off Learning

In order to demonstrate how Sender's gradual learning can lead to a more informative communication we assumed a particular learning technology. Clearly, the result that we obtained a fully revealing equilibrium is an extreme case. Of course, not every gradual learning can be effectively exploited. In what follows we briefly present an alternative learning technology that might seem natural but, contrary to the above, it admits the opposite extreme result. We show below that if the expert can, at each period, select an element $x \in \Theta$ and privately learn whether $\theta \geq x$, the gradualness in expert's learning can not be exploited to increase the equilibrium informativeness relative to the static communication with a perfectly informed Sender, i.e., CS equilibrium.

Cut-Off Learning A "cut-off k" experiment, c_k reveals to the Sender whether $\theta \geq k$. Formally, $c_k(\theta) = 1_{\theta \geq k}$. The set of available experiments is given by the collection $\{c_k : k \in \Theta\}$. Each experiment takes one unit of time and it is impossible to conduct more than one experiment at a time. At each date the Sender privately chooses a cut-off k_t, which may depend on the results of previous experiments. Thus, a Sender's strategic learning is a plan $(k_t(h_t))_{t \in T}$ where $h_t = (c_{k_s})_{s<t}$.

Proposition 3 *Under the cut-off learning, every Receiver's equilibrium partition is consistent with some CS equilibrium.*

Proof Let e be a equilibrium. First, observe that the Sender's information structure at each date $t \in \mathbf{T}$ has the following property: if $k_1, k_2 \in J_t$, where $k_1, k_2 \in \Theta$ and $J_t \subset \Theta$ is an element of the Sender's information partition at date t, then every $k_1 \leq k \leq k_2$ satisfies $k \in J_t$. Otherwise there exist $i < j$ such that $c_i = 0$ and $c_j = 1$, a contradiction. As the Sender's preferences satisfy super modularity, e is monotonic. Also, note that under e, the Receiver's information structure is a coarsening of the Sender's information, at each date.

Assume that the Receiver's partition in e, is not consistent with CS equilibrium. Then, there are two adjacent elements I_1, I_2 of the Receiver's information partition under e such that either $u^S(a(I_1), min\, I_2) > u^S(a(I_2), min\, I_2)$ or $u^S(a(I_2), max\, I_1) > u^S(a(I_1), max\, I_1)$ where $a(I)$ is the action selected by the Receiver once he learns that $\theta \in I$. As the Receiver's information is a coarsening of Sender's information, both actions $a(I_1)$ and $a(I_2)$ are yet inducible at date in which $c_{min\, I_2}$ is performed. The inequalities above imply that the Sender can benefit from replacing $c_{min\, I_2}$ with either $c_{min\, I_2 - 1}$ or $c_{min\, I_2 + 1}$, contradicting the assumption that e is an equilibrium. QED

Comment: The proposition is valid both for the case of $\Theta = \{1, 2, ...n\}$ and for the case $\Theta = \{\phi, 1, 2, ...n\}$, under the natural interpretation that $\phi < k$ for each $k \neq \phi$. Also, note that, given any history, $c_{min\,\Theta}$, whether it be c_1 or c_ϕ, is uninformative to the expert so it can be interpreted as "doing nothing". In this sense, our definition of a learning plan is without loss of generality.

Since this learning technology may seem plausible for the traditional continuous state space, in the appendix, we provide a supplementary proof for this case as well.

Proposition 3' *Consider the standard (continuous) Crawford and Sobel (1982) environment and assume that the expert can explore the state via "cut-off learning". Every Receiver's equilibrium partition in this environment is consistent with CS equilibrium.*

Proof see Appendix.

6.2 Experiments Scheduling

The artificial pressure that is created by the equilibrium in propositions 1 and 2 drew an important problem of experiments scheduling for the expert. This problem is completely vacuous in the case that the expert is expected to submit a report after she had a sufficient amount of time to become fully informed. We showed that, sometimes, it is possible to perform the available experiments in a particular order that allows for more information to be transmitted. This may suggest a broader question of experiments scheduling. One may ask, given a general collection of available experiments (say an experiment is a subset of Θ), how should the (biased) expert schedule these experiments in order to be more reliable. This question can be addressed both in the environment of observable experiments (unobservable outcomes) and unobservable experimentation, as considered in this paper.

7 Appendix

A Discrete Uniform Quadratic Case

Let $\Theta = \{1, 2, .., n\}$ .Assume that the utility functions are given by the negative quadratic loss $u^R(a, \theta) = -(\theta - a)^2$ and $u^S(a, \theta, b) = -(\theta + b - a)^2$, where $b \in \mathbb{N}$ is the Sender's constant bias.

Lemma 2 The (ex-ante) most desirable CS equilibrium in this environment is given by partitioning the states $\{1, 2, ..., n\}$ into monotone subsets $J_1, J_2, .., J_{i^*}$ such that $min\, J_{i+1} = max\, J_i + 1$ for $i = 1, ..i^* - 1$ and $|J_i| = x_i + y_i$, where i^*, $\{x_i\}$ and $\{y_i\}$ are defined as follows:

1. $x_1 = 1$ and $x_{j+1} = x_j + 4b - 2$,
2. i^* satisfies $\sum_{i=1}^{i^*} x_i \leq n < \sum_{i=1}^{i^*+1} x_i$,

3. $\{y_i\}_{i=1}^{i^*}$ is the lexicographically maximal weakly increasing sequence such that $y_i \in \{0\} \cup \mathbb{N}$ and $\sum_{i=1}^{i^*} y_i = n - \sum_{i=1}^{i^*} x_i$

Proof First, note that because of the quadratic loss functions, given any information set the Receiver selects a pure action and any (non-redundant) equilibrium in this environment is monotonic. Let e be an equilibrium and let Θ_e denote the Receiver's information partition, corresponding to e. For $I \in \Theta_e$, let $a(I)$ denote the action chosen by the Receiver once he reveals that $\theta \in I$. Notice that $a(I) = argmax_a\, E_\theta[u^R(a,\theta)|\theta \in I] = \frac{min\, I + max\, I}{2}$.

Pick two two adjacent elements $I, J \in \Theta_e$, that is, $min\, J - 1 = max\, I$. Then, from the considerations of type $max\, I$, in equilibrium, we must have $u^S(a(I), max\, I) \geq u^S(a(J), max\, I)$. This is equivalent to $max\, I + b - a(I) \leq a(J) - max\, I - b$. Plugging $a(I)$, $a(J)$ into this inequality and rearranging terms we obtain the following necessary condition for equilibrium

$$|J| \geq |I| + 4b - 2 \tag{1}$$

Similarly, from the considerations of $min\, J$ we obtain the condition

$$|J| \leq |I| + 4b + 2 \tag{2}$$

As usual, the supermodularity ensures that the incentive compatibility constraints of internal types follow from satisfying incentive compatibility constraints for the edge types. Accordingly, We say that a sequence of "lengths" $l_1, ..., l_k$ is compatible with equilibrium if for each $1 < j \leq k$,

$$l_{j-1} + 4b - 2 \underset{(1')}{\leq} l_j \underset{(2')}{\leq} l_{j-1} + 4b + 2$$

and $\sum_{i=1}^{k} l_i = n$.

Define $x_1 = 1$ and $x_{j+1} = x_j + 4b - 2$ for $j \in \mathbb{N}$. If there exists i^* such that $\sum_{i=1}^{i^*} x_i = n$, it is immediate that the partition $\{J_1, .., J_{i^*}\}$ of Θ, where $J_i = \{1 + \sum_{k=1}^{i-1} x_k, .., \sum_{k=1}^{i} x_k\}$, constitutes the (ex-ante) most preferred Receiver's information partition that is consistent with equilibrium. In other words, x_i for $i = 1, .., i^*$ is the optimal equilibrium compatible sequence. It follows from the fact that this partition consists of smallest possible elements J_i (containing the smallest possible number of adjacent states) that satisfy condition (1).

Assume that there does not exist $j \in \mathbb{N}$ such that $\sum_{i=1}^{j} x_i = n$ and let i^* be the integer with $\sum_{i=1}^{i^*} x_i < n < \sum_{i=1}^{i^*+1} x_i$. Clearly, from (1) it is impossible to have an equilibrium that will partition Θ into more than i^* elements.

Note that $\{x_i\}_{i=1}^{i^*}$ is not an equilibrium compatible sequence because $\sum x_i < n$. But this is the sequence that consists of smallest elements that satisfies satisfies (1'). Define $y = n - \sum_{i=1}^{i^*} x_i$. We look for an optimal "allocation" of the y excess elements $(y_1, y_2, ..., y_{i^*})$ such that $l_i = x_i + y_i$ is an equilibrium compatible sequence.

From (1'), $\{y_i\}_{i=1}^{i^*}$ must be weakly increasing. Because if $y_{j+1} < y_j$ for some $j < i^*$, we get $x_{j+1}+y_{j+1} < x_j + y_j + 4b - 2$ (as by definition $x_{j+1} = x_j + 4b - 2$), a contradiction. On the other hand, due to the quadratic loss function, it is evident that the marginal loss from increasing the length of a large interval is greater than the marginal loss from increasing the length of a small interval. Therefore, if we must add one element to some interval, it has to be added to the smallest possible interval subject to the constraints above. Therefore, $\{y_i\}_{i=1}^{i^*}$ is the lexicographically maximal weakly increasing sequence such that $y_i \in \{0\} \cup \mathbb{N}$ and $\sum_{i=1}^{i^*} y_i = y$. QED

It is trivial that if $\Theta = \{\phi, 1, 2, ..., n\}$ in the (ex-ante) most desirable equilibrium ϕ must be separated from other states and the structure of the partition of $\Theta - \{\phi\}$ remains the same.

B Cut-off Learning in Continuous State Space

Crawford and Sobel (1982) standard (continuous) environment: Let the state of the world $\theta \in [0, 1]$ be randomly drawn from a differentiable probability distribution function $F(\theta)$ with positive and continuous density $f(\theta)$. The utilities of both players are given by $u^R(a, \theta)$ and $u^S(a, \theta)$ and for $i = R, S$, $u_1^i(a, \theta) = 0$ for some $a \in \mathbb{R}$, $u_{11}^i(\cdot) < 0$ and $u_{12}^i(\cdot) > 0$. Similarly to the above, a Sender's strategic learning is a plan $(p_t(h_t))_{t \in T}$ where $h_t = (c_{p_s})_{s<t}$, $c_p(\theta) = 1_{\theta \geq p}$ and the set of available experiments is given by the collection $\{c_p : p \in [0, 1]\}$. Each experiment takes one unit of time and it is impossible to conduct more than one experiment at a time.

Proposition 3' *Consider the standard (continuous) Crawford and Sobel (1982) environment and assume that the expert can explore the state via "cut-off learning". Every Receiver's equilibrium partition in this environment is consistent with CS equilibrium.*

Proof of Proposition 3' Let e be an equilibrium. First, observe that the Sender's information structure at each date $t \in \mathbf{T}$ consists of only convex elements. Otherwise there exist $0 < p < q \leq 1$ such that $c_p = 0$ and $c_q = 1$, a contradiction. As the Sender's preferences satisfy supermodularity and we restrict attention to non-redundant equilibria, the equilibrium is monotonic. The equilibrium Receiver's information partition at each date is, by definition, a coarsening of the Sender's information structure. Thus, it is given by a finite collection of successive and non-degenerate intervals with endpoints $0 = q_0 < q_1 < q_2 < ... < q_m = 1$. Denote by $a(q_j, q_{j+1})$ the action chosen by the Receiver if he learns that θ belongs to the interval with endpoints q_j, q_{j+1}.

Let $t_e(q_j) \in \mathbf{T}$ denote the point in time when the Sender conducts c_{q_j} according to the equilibrium e. Because the Receiver's information is always (weakly) coarser than the Sender's information, in equilibrium, the action $a(q_j, q_{j+1})$ must be induced (weakly) after the Sender performs the experiments c_{q_j} and $c_{q_{j+1}}$. Thus, for each $j \in \{1, .., m-1\}$, both $a(q_{j-1}, q_j)$ and $a(q_j, q_{j+1})$ are inducible at $t_e(q_j)$.

Given q_{j-1} and q_{j+1}, the cut-off q_j must solve,

$$max_q\,[\int_{q_{j-1}}^{q} u^S(a(q_{j-1},q_j),\theta)dF(\theta) + \int_{q}^{q_{j+1}} u^S(a(q_j,q_{j+1}),\theta)dF(\theta)]$$

Thus, a necessary condition for an equilibrium is that for each $j \in \{1,..,m-1\}$,

$$u^S(a(q_{j-1},q_j),q_j) = u^S(a(q_j,q_{j+1}),q_j)$$

But this condition is satisfied by all CS equilibria[6]. QED

References

1. Argenziano, R., Squintani, F., Severinov, S., 2011. Strategic Information Acquisition, mimeo.

2. Austen-Smith, D. 1994. Strategic transmission of costly information. Econometrica, 62, 955-963.

3. Crawford, V., Sobel, J., 1982. Strategic information transmission. Econometrica 50, 579-594.

4. Fisher, P., Stocken, P., 2001. Imperfect information and credible communication. Journal of Accounting Research, 39, 119-134.

5. Frug, A., 2014. Cheap Talk with Gradual Learning, mimeo.

6. Gentzkow, M., Kamenica, A., 2012. Disclosure of Endogenous Information, mimeo.

7. Golosov, M., Skreta, V., Tsyvinsky, A., Wilson, A., 2013. Dynamic Strategic Information Transmission, mimeo.

8. Green, J., Stockey, N., 1980. A two-person Game of information transmission. Harvard university, HIER discussion paper 751.

9. Ivanov, M., 2010. Informational Control and Organizational Design. Journal of Economic Theory, 145, 721-751.

10. Ivanov, M., 2012. Dynamic Learning and Strategic communication, mimeo.

11. Kamenica, E., Gentzkow, M., 2011. Bayesian Persuasion. American Economic Review, 101, 2590-2615.

12. Pei, D., 2013. Communication with Endogenous Information Acquisition, mimeo.

[6]See theorem 1 in Crawford and Sobel (1982)

13. Rothschild, M., & Stiglitz, J., 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. The quarterly journal of economics, 629-649.

14. Spence, M., 1973. Job Market Signaling. Quarterly Journal of Economics 87, 355-374.

15. Sobel, J., 2010. Giving and Receiving Advice. Econometric Society 10th World Congress.