Strategic Learning And Information Transmission

Alexander Frug

October 3, 2013

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 conducted over time. Nevertheless, it is often the case that some aspects of the exploration 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.

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,

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

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 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 an 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 and parallel to this work 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 (2013) suggested a model of gradual exogenous learning and communication. In this model, by selecting 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

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

learning even when the expert's learning is strategic (endogenous), unobservable and non verifiable.

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 make two modifications on the standard state space. First, we assume discrete rather than continuous state space. Second, we assume that there exists a particular state ϕ which is, in some sense, "apart" from other states and often can be interpreted as "nothing happened". We discuss later exactly where this helps and how this designated state of the world affects the results.

The addition of the state ϕ to the state space has clear technical implications. Even though we will discuss the role of such a state and advocate that the general tone of the results remains even in its absence, we claim that such a state may be quite appropriate in some environments. A fugitive can hide in each floor $\theta \in \{1, 2, ..., n\}$ of an $n-$floor building, or not being inside the building at all - ϕ. A person who were left alone near n golden coins can either be honest ϕ, or steal any number $\theta \in \{1, 2, .., n\}$ of coins and being accounted, in this case, as a thief.

Imagine a patient approaching a doctor with a headache. It can be quite natural to model this situation as if the patient *might* have various problems ranging from very basic disorders to most severe ones $\theta \in \{1, 2, ..., n\}$. Each θ has its best treatment, and applying another one, which is optimal for some "close" disorder, will do only slightly worse. It is also known that, sometimes, the headache turns out to be nothing at all. Namely, it is possible that non of the disorders $\{1, 2, ..., n\}$ happened and, in this case, applying any medical treatment is worse than just "doing nothing".

We will assume that $\Theta = \{\phi, 1, 2, .., n\}$ throughout the main part of the analysis.

2.2 Model

There are two players $N = \{Sender, Receiver\}$. A state $\theta \in \Theta = \{\phi, 1, 2, .., n\}$ is distributed according to a common prior P. Action $a \in A = \{\phi, 1, 2, .., n\}$ must be selected by the Receiver. Player i's utility function is

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

where $c < \underset{a,\theta,i}{min}\, u^i(a, \theta)$ and for each $\theta \neq \phi$, $u^i(\cdot, \theta) \leq 0$ is a single-peaked function and

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

The assumption that $a = \theta$ is optimal for the Receiver is for convenience, the interpretation of the inequality is that the Sender is positively biased.

Sender's Learning For each $k \in \{1, .., n\}$, there is an experiment $e_k = 1_{\theta=k}$ that reveals whether $\theta = k$ or not. Conducting any experiment e_k takes one unit of time. Only the Sender has access to the set of experiments $\{e_k | k \in \{1, .., n\}\}$ and whether she has conducted any experiment is unobservable and non verifiable. Notice that we assume that there does not exist an experiment e_ϕ. Put differently, ϕ can be learned by the expert 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 $\mathbf{T} = \{1, 2.., T\}$ be a sufficiently long[3] ($T > n$) discrete time-line. For each $t \in \mathbf{T}$, let $M_t(h_t)$ be a set of available reports given the history of previous reports $\{m_\tau\}_{\tau<t}$ such that the empty report, ϕ, is an element of M_t, and whenever $m_s \neq \phi$ for some $s \in \mathbf{T}$, $M_t = \{\phi\}$ for all $t > s$. We assume that whenever $M_t(h_t) \neq \{\phi\}$ this set is rich enough to transmit any information available for the Sender[4].

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 proper sequence $\{m_t\}$, 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 depends on the learning component of the strategy. To make it precise let $f : T \to \{0, 1, .., n\}$ be a given plan of learning, specifying which experiment $e_{f(t)}$ is to be conducted at date $t \in T$, where e_0 denotes "do nothing". Sender's report at date $t \in \mathbf{T}$ must be measurable with respect to the information collected up to this date. 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 an $\mathcal{F}_t(f)$ - measurable function. Since the continuation of learning is relevant only if previous experiments have not yet 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

$$\sigma \in \Delta\{(f, \{m_t(\theta)\}_{t\in T}) | f \in \{0, 1, .., n\}^{\mathbf{T}}, m_t(\theta) \in M_t(h_t) \,\text{is}\, \mathcal{F}_t(f) - \textit{measurable} \,\text{for all}\, t \in \mathbf{T}\}$$

Whenever f is not random, we simply denote $\sigma = (f, \{\tilde{m}_t\})$, where $\{\tilde{m}_t\}$ is $\mathcal{F}_t(f)$-adapted and possibly mixed.

A (Perfect Bayesian) equilibrium consists of a Sender's strategy σ^*, a Receiver's action rule $\alpha^*(\{m_t\})$ and a belief $\mu_t(\theta | \{m_s\}_{s\leq t})$ for each $t \in T$, such that μ_t is updated according to the Bayesian rule whenever

[3]This assumes that, in principle, the time available for exploration places no restrictions on the expert's ability to become fully informed.

[4]For example, let $M_t(h_t) = \{\phi, 1, 2, .., 2^n\}$.

possible, $\alpha^*(\{m_t\})$ is optimal given the beliefs, and σ^* is optimal given $\alpha^*(\{m_t\})$. We call an equilibrium e "responsive" 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 responsive equilibrium e such that both equilibria identically map states into actions. In this sense these equilibria are essentially the same and therefore we confine attention to responsive equilibria.

Observation 1 *For each $\alpha(\{m_t\})$, there is a Sender's best response in which $f^{-1}(k) \leq n$ for any $k \in \{1, .., n\}$.*

Since our experiments are deterministic, there is no point in conducting an experiment $e(k)$ more than once, so we can assume without loss that $|f^{-1}(k)| = 1$. Moreover, if $f^{-1}(k) > n$ for some $k \in \{1, .., n\}$, there exists $l < n$ with $f(l) = 0$. Replacing e_0 with e_k at $t = l$ (and avoiding the experiment later) can not harm the Sender as her learning is costless and private information.

Corollary 1 *It is without loss to assume that the Sender's learning is a permutation $f \in S_n$.*

3 Example - The Uniform Quadratic Case

Let the state space and the Receiver's action space be $\Theta = A = \{\phi, 1, 2, .., 9\}$. Assume that $p(\theta) = \frac{1}{10}$ for each $\theta \in \Theta$. Player i's utility function is,

$$U^i(a, \theta) = \begin{cases} -(\theta + 1_{i=S} - a)^2 & a, \theta \neq \phi \\ 0 & a = \theta = \phi \\ -100 & a = \phi \oplus \theta = \phi \end{cases}$$

Proposition 1 *There exist an equilibrium in which the Receiver learns the state of the world.*

Proof Consider the following Sender's strategy: $f(t) = 10 - t$ and $m_t \neq \phi$ iff $\theta = 10 - t$. Notice that this Sender's strategy is of full support (whenever $m_t \neq \phi$ let the Sender report a random message in $M_t - \{\phi\}$). Therefore, the whole system of beliefs μ_t is pinned down by the Bayes rule. Accordingly, the Receiver's best response is to select $a = 10 - t$ if $m_t \neq \phi$ for $t \leq 9$, and $a = \phi$ if $m_t = \phi$ for all $t \leq 9$. It is left to show that $(f, \{m_t\})$ is the Sender's best response to the above specified Receiver's action rule.

Clearly, given f, the reporting component $\{m_t\}$ is optimal. So we have to show that there does not exist other learning plan g, and an $\mathcal{F}_t(g)$-adapted reporting strategy $\{m'_t\}$ such that $(g, \{m'_t\})$ is a profitable deviation.

By corollary 1, let $g \in S_9$. We construct a transition $s(g)$ from f to g, that will consist of a sequence of swaps, such that each swap is a worsening (under the assumption that we adjust the reporting strategy optimally).

Step 1 "Take care of $g(1)$": There exist $k_1 \geq 1$ such that $g(1) = f(k_1)$. Namely, the experiment $e_{g(1)}$ is conducted at date k_1 according to f. We now perform a sequential promotion of $e_{f(k_1)}$ until it occupies the first position in the learning plan. Let,

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

Notice that $f_1 \in S_9$ such that $f_1(1) = g(1)$.

Let us 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 $f_{j-1} \in S_9$ satisfies $f_{j-1}(i) = g(i)$ for each $i < j$, there exist $k_j \geq j$ such that $g(j) = f_{j-1}(k_j)$. Namely, the experiment $e_{g(j)}$ is conducted at date k_j according to f_{j-1}. We now perform a sequential promotion of $e_{f_{j-1}(k_j)}$ until it occupies the jth position in the learning plan. Let,

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

Observe that $f_j \in S_9$ 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 $f_8 = g$ and $s(g) =< s_1, s_2, ..., s_8 >$ is the complete transition sequence from f to g.

Note 1: Every swap (l, h) in $s(g)$, with $h > l$, involves a postponement of the experiment e_h by one unit of time and a promotion of e_l by one unit of time.

Note 2: Even before performing the swap (l, h), from the Sender's perspective, e_h was scheduled "too late". Only "relatively low" actions are inducible when h is inspected.

Due to the existence of ϕ, the Sender prefers not to induce an action $a \neq \phi$ as long as $\theta = \phi$ 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 e_h from t to $t + 1$ and bringing e_l forward from $t + 1$ to t.

By notes 1 and 2, the net value from postponing e_h is

$$\frac{1}{10}[u^s(9 - t, h) - u^s(10 - t, h)]$$

as the Sender will induce the highest possible action immediately upon she finds out $e_h = 1$ (e_h is conducted too late both before and after the swap).

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 $u^s(10-t,l) < u^s(9-t,l)$. The net value of bringing e_l forward from t to $t+1$ is therefore,

$$\frac{1}{10} \cdot max\{0, u^s(10-t,l) - u^s(9-t,l)\}$$

Clearly, the swap (l,h) is a worsening if $u^s(10-t,l) < u^s(9-t,l)$, as the Sender postponed e_h and gained nothing. So consider the case where $u^s(10-t,l) \geq u^s(9-t,l)$. The total value of the swap in this case is

$$\frac{1}{10} \cdot [u^s(9-t,h) - u^s(10-t,h) + u^s(10-t,l) - u^s(9-t,l)] < 0$$

because of the supermodularity of Sender's preferences. QED

4 General Result

The uniform quadratic case illustrated above is central in the literature of cheap talk because of its tractability and the ability to derive closed form results. In this section we generalize proposition 1 to a more general case. The proposition in this section is aimed to emphasize the sufficient condition on the interdependence between the prior distribution over the states and the expert's preferences, to allow for separating equilibrium.

Definition *Local-weighted-supermodularity* is satisfied if for any pair $\underline{\theta} \neq \phi, \overline{\theta} \neq \phi$ with $\underline{\theta} < \overline{\theta}$ and any $a > 1$,

$$p(\overline{\theta})[u^S(a,\overline{\theta}) - u^S(a-1,\overline{\theta})] \geq p(\underline{\theta})[u^S(a,\underline{\theta}) - u^S(a-1,\underline{\theta})]$$

Proposition 2 If local-weighted-supermodularity is satisfied then there exist an equilibrium in which the Receiver learns the state of the world.

Proof *see Appendix.*

The local-weighted-supermodularity property replaces the standard supermodularity that is often assumed in an environment where the Sender is informed. Intuitively, it aims to assume that "higher" states are sufficiently likely, or, that there do not exist experiments that it is worthwhile to give up spending time on them as they might reveal the state with only a negligible probability.

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, it is important that the set of inducible actions will shrink with time in the direction opposite to the bias. It is important that the pace of this shrinkage will be sufficiently fast for the expert not to have any opportunity to inflate the report in the direction of her bias, provided that it respects her physical ability to acquire information.

5 The Role of ϕ

In this section we discuss the role of the distinguished state ϕ in this environment. We assert that, 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 our main example. The possibility that $\theta = \phi$ (with certain assumptions) prevents the expert from inducing an "active" action ($a \neq \phi$) as long as she hadn't learned θ. Assume now that $\Theta = \{1, 2, .., 9\}$. The following observation rises when examining a profile of strategies that is similar to the above equilibrium: provided that the expert follows the same learning plan as in proposition 1, during the early periods, as in the result above, she does not have an incentive to report "false positive" (induce an action if 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 $e_k = 0$ for all $k > 1$, the expert can be better off by inducing $a = 2$ rather than waiting another period to find out (what she already knows!) that $e_1 = 1$ and to induce $a = 1$.

This suggests that if $\Theta = \{1, .., n\}$ one can proceed with the logic of the equilibrium as in proposition 2, but not all the way down. Instead, at some point, the players must engage into another reporting phase, which, in general, will be only partially revealing. While this principle is pretty general, the exact tailoring of the second phase of communication can have "ad-hoc" features (that may depend on the exact formulation of the game) and, in general, proving that a given equilibrium is the most informative one can be much more involved. To our view, this modification of the state space cleans the ad-hoc aspects of most informative equilibria and facilitates the exposition of the general part of the message.

For the particular example discussed above, it turns out that in the most informative equilibrium the Receiver learns the state accurately whenever $\theta \notin \{2, 3\}$. In proposition 3 we prove it more generally for $\Theta = \{1, .., n\}$ to illustrate that the effect of ϕ turns negligible as the state space becomes large relative to the bias.

Standard State Space

Consider the standard (discrete) state space $\Theta = \{1, 2, ..., n\}$.Assume that each state is equally likely and,

$$u^i(a, \theta) = -(\theta + 1_{i=S} - a)^2$$

where the indicator $1_{i=S}$ captures the Sender's positive bias of one. Assume that the Sender's learning technology and the reporting protocol are similar to those in previous section.

Lemma 1 *Full information revelation is inconsistent with equilibrium.*

Proof We say that an action a is inducible at date τ if, given a history $m_s = \phi$ for all $s < \tau$, there exists a "suffix" sequence $\{m_s\}_{s\geq\tau}$ such that $\alpha^*(\{m_s\}_{s=1}^T)(a) = 1$. First notice that in any fully revealing equilibrium it is necessary that an action $a > 1$ be inducible only strictly before the Sender learns whether $\theta = a - 1$. Otherwise, upon learning that $\theta = a - 1$, the Sender can pretend that $\theta = a$ and improve over her candidate equilibrium payoff.

In any fully revealing equilibrium, by the time of the last experiment the state is perfectly learned by the Sender. Thus at most one experiment e_k is planned not to be performed. In this case whether $\theta = k$ is learned by the expert by ruling out other possibilities. Notice that $a = k$ is inducible when the expert is perfectly informed about θ because the equilibrium is fully revealing. If $k > 1$, the Sender's type $\theta = k-1$ can pretend $\theta = k$, a contradiction. Otherwise, if $k = 1$ (or if all experiments $\{e_k | k \in \{1, .., n\}\}$ are conducted), denoting by $t(e_k)$ the date at which k is conducted, we must have $t(e_n) < t(e_{n-1}) < .. < t(e_2)$. By $t(e_2)$, a date when $a = 2$ is inducible, whether $\theta = 1$ is known to the Sender and with this information she can be better off by pretending $\theta = 2$. QED

Proposition 3 *In the most informative equilibrium the Receiver learns all $\theta \notin \{2, 3\}$.*

Proof Following Lemma 1, it is sufficient to prove that the partition $\{\{2, 3\}, \{k\} | k \in \Theta - \{2, 3\}\}$ is consistent with equilibrium. Consider the following profile of strategies.

Sender: $\sigma = (f, \{m_t\})$ such that $f(t) = \begin{cases} n+1-t & t \leq n-3 \\ 1 & t = n-2 \end{cases}$ and $m_t \neq \phi$ iff $e_{f(t)} = 1$.

Receiver: $a(\{m_t\}) = \begin{cases} n+1-t & m_t \neq \phi \wedge t \leq n-3 \\ 1 & m_{n-2} \neq \phi \\ 3 & otherwise \end{cases}$

Note that Receiver's beliefs are completely pinned down by the Bayes rule, and it is easy to see that $a(\{m_t\})$ is a best response. We will show now that σ is a best response to $a(\{m_t\})$.

Denote by $u^S[\theta|\sigma]$ the Sender's utility in state θ given that she plays according to σ (and the Receiver plays $a(\{m_t\})$). Assume by way of contradiction that there exists an improvement over σ and denote it by σ'. In this case, there exist $k \in \Theta$ such that $u^S[k|\sigma'] > u^S[k|\sigma]$. Define

$$k_0 := min\{k | u^S[k|\sigma'] > u^S[k|\sigma]\}$$

The following observations follow from the definition of k_0:

I. $k_0 \geq 3$; as σ reaches the highest possible payoff at $\theta \in \{1, 2\}$ according to $a(\{m_t\})$.

II. $E[u^S[k|\sigma']|k \geq k_0] > u^S[k|\sigma]|k \geq k_0]$; since σ' is assumed to be preferred over σ.

III. The Sender induces the action $a = k_0 + 1$ at state k_0.

We say that $a = k_0+1$ is *directly* induced at the state k_0 if it is induced after conducting the experiment e_{k_0}. The action $a = k_0 + 1$ is induced *indirectly* if e_{k_0} has not been conducted before.

In the case of *direct* inducement, e_{k_0} is brought forward at σ', relative to σ, by (at least) one period. Therefore there exist $l_0 > k_0$ such that e_{l_0} is postponed at σ', relative to σ, by (at least) one period. By the *supermodularity of* u^S, the net value from exploring e_{k_0} earlier and postponing e_{l_0} by one period is negative. Therefore, there must be another $k_1 > k_0$ such that $u^S[k_1|\sigma'] > u^S[k_1|\sigma]$. At state k_1 the Sender must induce $a = k_1 + 1$ *directly* so, again, there must exist $l_1 > k_1$ such that e_{l_1} is postponed by (at least) one period. Notice that l_1 need not be different from l_0 (if $l_1 = l_0$, it is necessary to postpone e_{l_0} again because of the promotion of e_{k_1}). Again, by the supermodularity of u^S, the improvement at $\theta = k_1$ and the worsening at $\theta = l_1$ sum up to another negative change in Sender's utility. The series $\{k_j\}$ is finite (strictly increasing sequence of natural numbers bounded by n) so II is impossible if the the action in in III is *induced direct*ly.

Assume now that the Sender induces the action $a = k_0 + 1$ in III *indirectly*, i.e., at date $n - k_0$ the Sender submits a report as if $e_{n-k_0} = 1$ (possibly a "false positive"). Observe that in this particular utility function, the cost associated with any postponement of e_l for $l > k_0$ is greater than the benefit from bringing *two* experiments e_k forward, relative to f[5]. Therefore, any change in σ before $t = n - k_0$ must contradict II (note that actions greater than k_0+1 may be induced only directly). Thus, σ' and σ coincide for $t = 1, .., n - k_0 - 1$. Note that in this case, we have $E[u^S[k|\sigma']|k > k_0] = u^S[k|\sigma]|k > k_0]$ and a "false positive" $m_{n-k_0} \neq \phi$ is consistent with II.

Finally, from the measurability of the reporting strategy, the action $a = k_0 + 1$ is induced by all types $\theta \in \{1, .., k_0\}$. Notice that $u^S[\theta|\sigma'] > u^S[\theta|\sigma]$ only for $\theta = k_0$ and that a strict worsening is borne by every $\theta \in \{1, ..k_0 - 2\}$, which is non empty according to I. It is easy to see that

$$1 = u^S[k_0|\sigma'] - u^S[k_0|\sigma] < 3 \leq u^S[1|\sigma] - u^S[1|\sigma'] \leq \sum_{\theta=1}^{k_0-1} (u^S[\theta|\sigma] - u^S[\theta|\sigma'])$$

and therefore

$$E[u^S[k|\sigma'] = E[u^S[k|\sigma']|k \leq k_0]+E[u^S[k|\sigma']|k > k_0] < E[u^S[k|\sigma]|k \leq k_0]+E[u^S[k|\sigma]|k > k_0] = E[u^S[k|\sigma]]$$

This contradicts the assumption on σ'. QED

[5]By bringing forward two experiments the Sender saves with probability $\frac{2}{n}$ a cost of 1, while postponing an experiment by one period leads to an increase in loss from 1 to 4, with probability $\frac{1}{n}$.

6 Cut-Off Learning

A particular learning technology that consists of a collection of experiments, each of which is capable to indicate a specific realization of the state of the world, suggests that gradualness in learning can be exploited towards a more informative communication in this environment. This learning technology is an extreme case where complete learning and full revelation are consistent with equilibrium. In what follows we present an alternative learning technology that might seem realistic but, contrary to the above, it admits the opposite extreme result. We show below that if the expert can, at each date, privately learn whether $\theta \geq x$, for any $x \in \Theta$, the gradualness in experts learning can not be exploited to increase the equilibrium informativeness relative to the static communication with a perfectly informed Sender. To present it in the most striking and succinct way, we prove it for the original C-S environment [6].

Recall the standard C-S setting. The state of the world θ is 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 $u_{12}^i(\cdot) > 0$.

Cut-Off Learning Let $T \in \mathbb{N}$ and let $\mathbf{T} = \{1, 2, .., T\}$ be the time-line. A "cut-off p" experiment, $c(p)$ reveals to the Sender whether $\theta \geq p$. Formally, $c(p) = 1_{\theta \geq p}$. The set of available reports 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. At each date the Sender chooses a cut-off, which may depend on the results of previous experiments[7]. Thus, a Sender's strategic learning is a plan $(p_t(h_t))_{t \in T}$ where $h_t = (c(p_s))_{s<t}$.

Proposition 4 *There does not exist an equilibrium that is strictly more informative than the most informative C-S equilibrium.*

Proof 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 $0 < p_1 < p_2 \leq 1$ such that $c(p_1) = 0$ and $c(p_2) = 1$, a contradiction. As the Sender's preferences satisfy supermodularity and we restrict attention to responsive 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}.

[6] The result remains the same in the discrete state space that we used above. Incentive constraints of the "edge types" will be required instead of the succinct indifference condition that will follow.

[7] Note that given any history, $c(0)$ is not informative to the expert so conducting this experiment captures the expert's choice not to explore the state at this date.

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, the action $a(q_j, q_{j+1})$ can not be induced before the Sender performed 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,

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

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 C-S equilibria[8]. QED

Comment: By providing the Sender with a sufficient amount of time one can easily support any finite C-S equilibrium partition in the environment of strategic cut-off learning.

7 Appendix

Proof of Proposition 2 Consider the following Sender's strategy: $f(t) = n + 1 - t$ and $m_t \neq \phi$ iff $\theta = n+1-t$. Notice that this Sender's strategy is of full support (whenever $m_t \neq \phi$ let the Sender report a random message in $M_t - \{\phi\}$). Therefore, the whole system of beliefs μ_t is pinned down by the Bayes rule. Accordingly, the Receiver's best response is to select $a = n+1-t$ if $m_t \neq \phi$ for $t \leq n$, and $a = \phi$ if $m_t = \emptyset$ for all $t \leq n$. It is left to show that $(f, \{m_t\})$ 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, $\{m_t\}$ is optimal given f. So we have to show that there is no other learning g, and an $\mathcal{F}_t(g)$-adapted reporting strategy $\{m_t'\}$ such that $(g, \{m_t'\})$ is a profitable deviation.

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

Step 1 "Take care of $g(1)$": There exist $k_1 \geq 1$ such that $g(1) = f(k_1)$. Namely, the experiment $e_{g(1)}$ is conducted at date k_1 according to f. We now perform a sequential promotion of $e_{f(k_1)}$ until it occupies the first position in the learning plan. Let,

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

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

Notice that $f_1 \in S_n$ such that $f_1(1) = g(1)$.

Let us 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 exist $k_j \geq j$ such that $g(j) = f_{j-1}(k_j)$. Namely, the experiment $e_{g(j)}$ is conducted at date k_j according to f_{j-1}. We now perform a sequential promotion of $e_{f_{j-1}(k_j)}$ until it occupies the jth position in the learning plan. Let,

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

Observe that $f_j \in S_n$ 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.

Note 1: Every swap (l, h) in $s(g)$, with $h > l$, involves a postponement of the experiment e_h by one unit of time and a promotion of e_l by one unit of time.

Note 2: Even before performing the swap (l, h), from the Sender's perspective, e_h was scheduled "too late". Only "relatively low" actions are inducible when h is inspected.

Due to the existence of ϕ, the Sender prefers not to induce an action $a \neq \phi$ as long as $\theta = \phi$ 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 e_h from t to $t + 1$ and bringing e_l forward from $t + 1$ to t.

By notes 1 and 2 and the Sender's preferences being single-peaked, the net value from postponing e_h is

$$p(\theta = h)[u^s(n - t, h) - u^s(n - t + 1, h)]$$

as the Sender will induce the highest possible action immediately upon she finds out $e_h = 1$ (e_h is conducted too late both before and after the swap).

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 $u^s(n-t+1, l) < u^s(n-t, l)$. The net value of bringing e_l forward from t to $t + 1$ is therefore,

$$p(\theta = l)max\{0, u^s(n - t + 1, l) - u^s(n - t, l)\}$$

The total value of the swap is therefore

$$p(\theta = h)[u^s(n-t,h) - u^s(n-t+1,h)] + p(\theta = l)max\{0, u^s(n-t+1,l) - u^s(n-t,l)\} \leq$$

$$\leq p(\theta = h)[u^s(n-t,h) - u^s(n-t+1,h)] + p(\theta = l)[u^s(n-t+1,l) - u^s(n-t,l)] \leq 0$$

by the "local-weighted supermodularity". 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., 2013. 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.

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.