Signaling Games

JOEL SOBEL
Department of Economics, University of California, San Diego, USA

Article Outline

Glossary

Babbling equilibrium An equilibrium in which the sender's strategy is independent of type and the receiver's strategy is independent of signal.

Behavior strategy A strategy for an extensive-form game that specifies the probability of taking each action at each information set.

Behavioral type A player in a game who is constrained to follow a given strategy.

Cheap-talk game A signaling game in which players' preferences do not depend directly on signals.

Condition D1 An equilibrium refinement that requires out-of-equilibrium beliefs to be supported on types that have the most to gain from deviating from a fixed equilibrium.

Divinity An equilibrium refinement that requires out-of-equilibrium beliefs to place relatively more weight on types that gain more from deviating from a fixed equilibrium.

Equilibrium outcome The probability distribution over terminal nodes in a game determined by equilibrium strategy.

Handicap principle The idea that animals communicate fitness through observable characteristics that reduce fitness.

Incomplete information game A game in which players lack information about the strategy sets or payoff functions of their opponents.

Intuitive criterion An equilibrium refinement that requires out-of-equilibrium beliefs to place zero weight on types that can never gain from deviating from a fixed equilibrium outcome.

Nash equilibrium A strategy profile in a game in which each player's strategy is a best response to the equilibrium strategies of the other players.

Neologism-proof equilibrium An equilibrium that admits no self-signaling set.

Pooling equilibrium A signaling-game equilibrium in which each all sender types send the same signal with probability one.

Receiver In a signaling game, the uninformed player.

Self-signaling set A set of types C with the property that precisely types in the set C gain from inducing the best response to C relative to a fixed equilibrium.

Sender In a signaling game, the informed agent.

Separating equilibrium A signaling-game equilibrium in which sender types sent signals from disjoint subsets of the set of available signals.

Signaling game A two-player game of incomplete information in which one player is informed and the other in not. The informed player's strategy is a type-contingent message and the uninformed player's strategy is a message-continent action.

Single-crossing condition A condition that guarantees that indifferent curves from a given family of preferences cross at most one.

Spence-mirrlees condition A differential condition that orders the slopes of level sets of a function.

Standard signaling game A signaling game in which strategy sets and payoff functions satisfy monotonicity properties.

Type In an incomplete information game, a variable that summarizes private information.

Verifiable information game A signaling game with the property that each type has a signal that can only be sent by that type.

Definition of the Subject

Signaling games refer narrowly to a class of two-player games of incomplete information in which one player is informed and the other is not. The informed player's strategy set consists of signals contingent on information and the uninformed player's strategy set consists of actions contingent on signals. More generally, a signaling game includes any strategic setting in which players can use the actions of their opponents to make inferences about hidden information. The earliest work on signaling games

Please note that the pagination is not final; in the print version an entry will in general not start on a new page.

Editor's or typesetter's annotations (will be removed before the final TEX run)

was Spence [73]'s model of educational signaling and Zahari [77]'s model of signaling by animals. During the 1980s researchers developed the formal model and identified conditions that permitted the selection of unique equilibrium outcomes in leading models.

Introduction

The framed degree in your doctor's office, the celebrity endorsement of a popular cosmetic, and the telephone message from an old friend are all signals. The signals are potentially valuable because they allow you to infer useful information. These signals are indirect and require interpretation. They may be subject to manipulation. The doctor's diploma tells you something about the doctor's qualifications, but knowing where and when the doctor studied does not prove that she is a good doctor. The endorsement identifies the product with a particular lifestyle, but what works for the celebrity may not work for you. Besides, the celebrity was probably paid to endorse the product and may not even use it. The phone message may tell you how to get in touch with your friend, but is unlikely to contain all of the information you need to find him – or to evaluate whether you'll meet to discuss old times or to be asked a favor. While the examples share all involve signaling, the nature of the signaling is different. The doctor faces large penalties for misrepresenting her credentials. She is not required to display all of her diplomas, but it is reasonable to assume that degrees are not forged. The celebrity endorsement is costly – certainly to the manufacturer who pays for the celebrity's services and possibly to the celebrity himself, whose reputation may suffer if the product works badly. It is reasonable to assume that it is easier to obtain an endorsement of a good product, but there are also good reasons to be skeptical about the claims. In contrast, although a dishonest or misleading message may lead to a bad outcome, leaving a message is not expensive and the content of the message is not constrained by your friend's information. The theory of signaling games is a useful way to describe the essential features of all three examples.

Opportunities to send and evaluate signals arise in many common natural and economic settings. In the canonical example (due to Spence [73]), a high-ability worker invests in education to distinguish herself from less skilled workers. The potential employer observers educational attainment, but not innate skill, and infers that a better educated worker is more highly skilled and pays a higher wage. To make this story work, there must be a reason that low-ability workers do not get the education expected of a more highly skilled worker and hence obtain a higher wage. This property follows from an assumption that the higher the ability the worker, the easier it is for her to produce a higher signal.

The same argument appears in many applications. For example, a risk-averse driver will purchase a lower cost, partial insurance contract, leaving the riskier driver to pay a higher rate for full insurance (Rothschild and Stiglitz [66] or Wilson [76]). A firm that is able to produce high-quality goods signals this by offering a warranty for the goods sold (Grossman [37]) or advertising extensively. A strong deer grows extra large antlers to show that it can survive with this handicap and to signal its fitness to potential mates (Zahavi [77]).

Game theory provides a formal language to study how one should send and interpret signals in strategic environments. This article reviews the basic theory of signaling and discusses some applications. It does not discuss related models of screening. Kreps and Sobel [44] and Riley [65] review both signaling and screening.

Section "The Model" describes the basic model. Section "Equilibrium" defines equilibrium for the basic model. Section "The Basic Model" limits attention to a special class of signaling game. I give conditions sufficient for the existence of equilibria in which the informed agent's signal fully reveals her private information and argue that one equilibrium of this kind is prominent. The next three sections study different signaling games. Section "Cheap Talk" discusses models of costless communication. Section "Verifiable Information" discusses the implications of the assumptions that some information is verifiable. Section "Communication about Intentions" briefly discusses the possibility of signaling intentions rather than private information. Section "Applications" describes some applications and extensions of the basic model. Section "Future Directions" speculates on directions for future research.

The Model

This section describes the basic signaling game. There are two players, called S (for sender) and R (for receiver). S knows the value of some random variable t whose support is a given set T. t is called the **type** of S. The prior beliefs of R are given by a probability distribution $\pi(\cdot)$ over T; these beliefs are common knowledge. When T is finite, $\pi(t)$ is the prior probability that the sender's type is t. When T is uncountably infinite, $\pi(\cdot)$ is a density function. Player S, learns t, sends to R a signal s, drawn from some set M. Player R receives this signal, and then takes an action a drawn from a set A. (It is possible to allow A to depend on s and S to depend on t.) This ends the game: The payoff to i is given by a function $u^i: T \times M \times A \to \mathbb{R}$.

This canonical game captures the essential features of the classic applications of market signaling. In the labor-market signaling story due to Spence [73] a worker wishes to signal his ability to a potential employer. The worker has information about ability that the employer lacks. Direct communication about ability is not possible, but the worker can acquire education. The employer can observe the worker's level of education and use this to form a judgment about the worker's true level of ability. In this application, S is a worker; R represents a potential employer (or a competitive labor market); t is the student's productivity; s is her level of education; and a is her wage.

Equilibrium

Defining Nash equilibrium for the basic signaling game is completely straightforward when T, S, and A are finite sets. In this case a behavior strategy for S is a function $\mu\colon T \times M \to [0,1]$ such that $\sum_{s \in M} \mu(t,s) = 1$ for all t. $\mu(t,s)$ is the probability that sender-type t sends the signal s. A behavior strategy for R is a function $\alpha\colon M \times A \to [0,1]$ where $\sum_{a \in A} \alpha(s,a) = 1$ for all s. $\alpha(s,a)$ is the probability that R takes action a following the signal s.

Proposition 1 *Behavior strategies (α^*, μ^*) form a Nash Equilibrium if and only if for all $t \in T$*

$$\mu(t,s) > 0 \text{ implies } \sum_{a \in A} U^S(t,s,a)\alpha(s,a) = \max_{s' \in S} \sum_{a \in A} U^S(t,s',a)\alpha(s',a) \quad (1)$$

and, for each $s \in S$ such that $\sum_{t \in T} \mu(t,s)\pi(t) > 0$ and, if $\sum_{t \in T} \mu(t,s)\pi(t) > 0$, then

$$\alpha(s,a) > 0 \text{ implies } \sum_{t \in T} U^R(t,s,a)\beta(t,a) = \max_{a' \in A} \sum_{t \in T} U^R(t,s,a')\beta(t,a')\,, \quad (2)$$

where

$$\beta(t,s) = \frac{\mu(t,s)\pi(t)}{\sum_{t' \in T} \mu(t',s)\pi(t')}\,. \quad (3)$$

Condition (1) states that the S places positive probability only on signals that maximize expected utility. This condition guarantees that S responds optimally to R's strategy. Condition (2) states that R places positive probability only on actions that maximize expected utility, where is taken with respect to the distribution $\beta(\cdot, s)$ following the signal s. Condition (3) states that $\beta(\cdot, s)$ accurately reflects the pattern of play. It requires that R's beliefs be determined using S's strategy and the prior distribution whenever possible. Equilibrium refinements also require that R has beliefs following signals s that satisfy

$$\sum_{t \in T} \mu(t,s)\pi(t) = 0\,, \quad (4)$$

that is are sent with probability zero in equilibrium. Specifically, sequential equilibrium permits $\beta(\cdot, m)$ to be an arbitrary distribution when Eq. (4) holds, but requires that Eq. (2) holds even for these values of s. This restriction rules out equilibria in which certain signals are not sent because the receiver responds to the signal with an action that is dominated.

The ability to signal creates the possibility that R will be able to draw inferences about S's type from the signal. Whether he is able to do so is a property of the equilibrium. It is useful to define two extreme cases.

Definition 1 An equilibrium (α^*, μ^*) is called a separating equilibrium if each type t sends different signals. That is, M can be partitioned into sets M_t such that for each t, $\mu_{s \in M_t}(t,s) = 1$. An equilibrium (α^*, μ^*) is called a pooling equilibrium if there is a single signal s^* that is sent by all types with probability one.

In a separating equilibrium, R can infer S's private information completely. In a pooling equilibrium, R learns nothing from the sender's signal. This definition excludes other possible situations. For example, all sender types can randomize uniformly over a set of two or more signals. In this case, the receiver will be able to draw no inference beyond the prior from a signal received in equilibrium. More interesting is the possibility that the equilibrium will be partially revealing, with some, but not all of the sender types sending common signals.

It is not difficult to construct pooling equilibria for the basic signaling game. Take the labor market model and assume S sends the message s^* with probability one and that the receiver responds to s^* with his best response to the prior distribution and to all other messages with the best response to the belief that t is the least skilled agent. Provided that the least skilled agent prefers to send s^* to sending the cheapest alternative signal, this is a Nash Equilibrium outcome.

The Basic Model

The separating equilibrium is a benchmark outcome for signaling games. When a separating equilibrium exists, then it is possible for the sender to share her information fully with the receiver in spite of having a potential conflict of interest.

Existence of separating equilibria typically requires a systematic relationship between types and signals. An appropriate condition, commonly referred to as the single-crossing condition, plays a prominent role in signaling games and in models of asymmetric information more generally.

In this section I limit attention to a special class of signaling game in which there is a monotonic relationship between types and signals. In these models, separating equilibria typically exist.

I begin by stating the assumption in the environment most commonly seen in applications. Assume that the sets T, S, and A are all real intervals.

Definition 2 $U^S(\cdot)$ satisfies the single-crossing condition if $U^S(t, s, a) \leq U^S(t, s', a')$ for $s' > s$ implies that $U^S(t', s, a) < U^S(t', s', a')$ for all $t' > t$.

In a typical application, $U^S(\cdot)$ is strictly decreasing in its second argument (the signal) and increasing in its third argument (R's response) for all types. Consequently indifference curves are well defined in $M \times A$ for all t. The single-crossing condition states that indifference curves of different sender types cross once. If a lower type is indifferent between type signal-action pairs, then a higher type strictly prefers to send the higher signal. In this way, the single-crossing condition links signals to types in such a way as to guarantee that higher types send weakly higher signals in equilibrium.

Note two generalizations of Definition 2. First, the assumption that the domain of $U^S(\cdot)$ is the product of intervals can be replaced by the assumption that these sets are partially ordered. In this case, weak and strict order replace the weak and strict inequalities comparing types and actions in the statement of the definition. Second, it is sometimes necessary to extend the definition to mixed strategies. In this case, the ordering of A induces a partial ordering of distributions of A through first-order stochastic dominance.

When one thinks of the single-crossing condition geometrically, it is apparent that it implies a ranking of the slopes of the indifference curves of the Sender. Suppose that $U^S(\cdot)$ is smooth, strictly increasing in actions and strictly decreasing in signals so that indifference curves are well defined for each t. Writing the indifference curve as $\{(s, \bar{a}(s; t))\}$, it must be that $U^S(t, s, \bar{a}(s; t)) \equiv 0$, so that the slope of the indifference curve of a type t Sender is

$$\bar{a}_1(s; t) = -\frac{U_2^S(t, s, a)}{U_3^S(t, s, a)}\,, \tag{5}$$

where $\bar{a}_1(s; t)$ is the partial derivative of $\bar{a}(s; t)$ with respect to the first argument, and $U_k^S(\cdot)$ denotes the partial derivative of $U^S(\cdot)$ with respect to its kth argument. Under these conditions, the single-crossing condition is implied by the requirement that the right-hand side of Eq. (5) is decreasing in t. The differentiable version of the single-crossing condition is often referred to as the Spence–Mirrlees condition. Milgrom and Shannon [58] contains general definitions of the single-crossing and Spence–Mirrlees conditions and Edlin and Shannon [26] provide a precise statement of when the conditions are equivalent.

To provide a simple construction of a separating equilibrium, limit attention to a standard signaling game in which the following conditions hold.

1. $T = \{0, \ldots, K\}$ is finite.
2. A and M are real intervals.
3. Utility functions are continuous in action and signal.
4. $U^S(\cdot)$ is strictly increasing in action and strictly decreasing in signal.
5. The single-crossing property holds.
6. The Receiver's best-response function, is uniquely defined, independent of the signal, and strictly increasing in t so that it can be written $BR(t)$.
7. There exists $\bar{s} \in S$ such that $U^S(K, \bar{s}, BR(K)) < U^S(K, s_0^*, BR(0))$.

Conditions 1 and 2 simplify exposition, but otherwise are not necessary. It is important that T, M, and A be partially ordered so that some kind of single-crossing condition applies. Conditions 4–6 impose a monotone structure on the problem so that higher types are more able to send high signals, and that higher types induce higher (and uniformly more attractive) actions. These conditions imply that in equilibrium higher types will necessarily send weakly higher signals. Condition 7 is a boundary condition that makes sending high signals unattractive. It states that the highest type of Sender would prefer to be treated like the lowest type rather than use the signal $\bar{s}$. These properties hold in many standard applications and certainly would be satisfied if $U^R(t, s, a) = -(a - t)^2$.

Separating Equilibrium

To illustrate these ideas, consider a construction of a separating equilibrium.

Proposition 2 *The standard signaling game has a separating equilibrium.*

One can prove the proposition by constructing a possible equilibrium path and confirming that the path can be part of a separating equilibrium.

Step 1. t_0 selects the signal s_0^* that maximizes $U^S(t_0, s, BR(t_0))$.

Step 2. Suppose that s_i^* have been specified for $i = 0, \ldots, k-1$ and let $U^*(t_i) = U^S(t_i, s_i^*, BR(t_i))$. Define s_k^* to solve:

$$\max U^S(t_k, s, BR(t_k)) \text{ subject to}$$
$$U^S(t_{k-1}, s, BR(t_k)) \leq U^*(t_{k-1}) .$$

Provided that the optimization problems in Steps 1 and 2 have solutions, the process inductively produces a signaling strategy for the Sender and a response rule for the Receiver defined on $\{s_0^*, \ldots, s_K^*\}$. When $BR(\cdot)$ is strictly increasing, the single-crossing condition implies that the signaling strategy is strictly increasing. To complete the description of strategies, assume that the Receiver takes the action $BR(t_k)$ in response to signals in the interval $[s_k, s_{k+1})$, $BR(t_0)$ for $s < s_0^*$, and $BR(t_K)$ for $s > s_K^*$. By the definition of the best-response function, the receiver is best responding to the sender's strategy. When the boundary condition fails, a fully separating equilibrium need not exist, but when M is compact, one can follow the construction above to obtain an equilibrium in which the lowest types separate and higher types pool at the maximum signal in M (see Cho and Sobel [22] for details).

In the construction, the equilibrium involves inefficient levels of signaling. When $U^S(\cdot)$ is decreasing in the signal, all but the lowest type of sender must make a wasteful expenditure in the signal in order to avoid being treating as having a lower quality. The result that expenditures on signals are greater than the levels optimal in a full-information model continue to hold when $U^S(\cdot)$ is not monotonic in the signal. The sender inevitably does no better in a separating equilibrium than she would do if R had full information about t. Indeed, all but the lowest type will do strictly worse in standard signaling games. On the other hand, the equilibrium constructed above has a constrained efficiency property: Of all separating equilibria, it is Pareto dominant from the standpoint of S. To confirm this claim argue inductively that in any separating equilibrium if t_j sends the signal s_j, then $s_j \geq s_j^*$, with equality only if all types $i < j$ send s_i^* with probability one.

Mailath [50] provides a similar construction when T is a real interval. In this case, the Spence–Mirrlees formulation of the single-crossing condition plays an important role and the equilibrium is a solution to a differential equation.

Multiple Equilibria and Selection

Section "Equilibrium" ended with the construction of a pooling equilibrium. A careful reconsideration of the argument reveals that there typically are many pooling equilibrium outcomes. One can construct a potential pooling outcome by assuming that all sender types send the same signal, that the receiver best responds to this common signal, and responds to all other signals with the least attractive action. Under the standard monotonicity assumptions, this strategy profile will be an equilibrium if the lowest sender type prefers pooling to sending the cheapest available out-of-equilibrium message. Section "Separating Equilibrium" ended with the construction of a separating equilibrium. There are also typically many separating equilibrium outcomes. Assume that types $t = 0, \ldots, r-1$ send signals $s^*(t)$, type r sends $\tilde{s}(k) > s^*(k)$, and subsequent signals $\tilde{s}^*(t)$ for $t > r$ solve:

$$\max U^S(t_k, s, BR(t_k)) \text{ subject to}$$
$$U^S(t_{k-1}, s, BR(t_k)) \leq U(t_{k-1}, \tilde{s}, BR(t_{k-1})) .$$

In both of these cases, the multiplicity is typically profound, with a continuum of distinct equilibrium outcomes (when M is an interval). The multiplicity of equilibria means that, without refinement, equilibrium theory provides few clear predictions beyond the observation that the lowest type of sender receives at least $U^*(t_0)$, the payoff it would receive under complete information and the fact that the equilibrium signaling function is weakly increasing in the sender's type. The first property is a consequence of the monotonicity of S's payoff in a and of R's best response function. The second property is a consequence of the single-crossing condition.

This section describes techniques that refine the set of equilibria. Refinement arguments that guarantee existence and select unique outcomes for standard signaling games rely on the Kohlberg–Mertens [43] notion of strategic stability. The complete theory of strategic stability is only available for finite games. Consequently the literature applies weaker versions of strategic stability that are defined more easily for large games. Banks and Sobel [8], Cho and Kreps [21], and Cho and Sobel [22] introduce these arguments.

Multiple equilibria arise in signaling games because Nash equilibrium does not constrain the Receiver's response to signals sent with zero probability in equilibrium. Specifying that R's response to these unsent signals is unattractive leads to the largest set of equilibrium outcomes. (In standard signaling games, S's preferences over actions does not depend on type, so the least attractive action is well defined.) The equilibrium set shrinks if one restricts the meaning of unsent signals. An effective restriction is condition D1, introduced in Cho and Kreps [21]. This condition is less restrictive than the notion of universal divinity introduced by Banks and Sobel [8], which in finite games is less restrictive than Kohlberg and Mertens's notion of strategic stability.

Given an equilibrium (α^*, μ^*), let $U^*(t)$ be the equilibrium expected payoff of a type t sender and let $D(s, t) = \{a\colon u(t, s, a) \geq U^*(t)\}$ be the set of pure-strategy responses to s that lead to payoffs at least as great as the equilibrium payoff for player t. Given a collection of sets, $X(t)$, $t \in T$, $X(t^*)$ is maximal if it not a proper subset of any $X(t)$.

Definition 3 Behavior strategies (α^*, μ^*) together with beliefs β^* satisfy D1 if for any unsent message s, $\beta(\cdot, s)$ is supported on those t for which $D(s, t)$ is maximal.

In standard signaling games, $D(s, t)$ is an interval: all actions greater than or equal to a particular action will be attractive relative to the equilibrium. Hence these sets are nested. If $D(s, t)$ is not maximal, then there is another type t' that is "more likely to deviate" in the sense that there exists out-of-equilibrium responses that are attractive to t' but not t. Condition D1 requires that the receiver place no weight on type t making a deviation in this case. Notice if $D(s, t)$ is empty for all t, then D1 does not restrict beliefs given s (and any choice of action will support the putative equilibrium). Condition D1 is strong. One can imagine weaker restrictions. The intuitive condition (Cho and Kreps [21]) requires that $\beta(t, s) = 0$ when $D(t, s) = \phi$ and at least one other $D(t', s)$ is non empty. Divinity (Banks and Sobel [8]) requires that if $D(t, s)$ is strictly contained in $D(t', s)$, then $\beta(t', s)/\beta(t, s) \geq \pi(t')/\pi(t)$, so that the relative probability of the types more likely to deviate increases.

Proposition 3 *The standard signaling game has a unique separating equilibrium outcome that satisfies Condition D1.*

In standard signaling games, the only equilibrium outcome that satisfies Condition D1 is the separating outcome described in the previous section. Details of the argument appear in Cho and Sobel. The argument relies on two insights. First, types cannot be pooled in equilibrium because slightly higher signals will be interpreted as coming from the highest type in the pool. Second, in any separating equilibrium in which a sender type fails to solve Step 2, deviation to a slightly lower signal will not lower R's beliefs.

The refinement argument is powerful and the separating outcome selected receives prominent attention in the literature. It is worth pointing out that the outcome has one unreasonable property. The separating outcome described above depends only on the support of types, and not on the details of the distribution. Further, all types but the lowest type must make inefficient (compared to the full-information case) investments in signal in order to distinguish themselves from lower types. The efficient separating equilibrium for a sequence of games in which the probability of the lowest type converges to zero does not converge to the separating equilibrium of the game in which the probability of the lowest type is zero. In the special case of only two types, the (efficient) pooling outcome may be a more plausible outcome when the probability of the lower type shrinks to zero. Grossman and Perry [38] and Mailath, Okuno-Fujiwara, and Postlewaite [51] introduce equilibrium refinements that select the pooling equilibrium in this setting. These concepts share many of the same motivations of the refinements introduced by Banks and Sobel and Cho and Kreps. They are qualitatively different from the intuitive criterion, divinity, and Condition D1, because they are not based on dominance arguments and lack general existence properties.

Cheap Talk

Models in which preferences satisfy the single-crossing property are central in the literature, but the assumption is not appropriate in some interesting settings. This section describes an extreme case in which there is no direct cost of signaling.

In general, a cheap-talk model is a signaling model in which $u^i(t, s, a)$ is independent of s for all (t, a). Two facts about this model are immediate. First, if equilibrium exists, then there always exists an equilibrium in which no information is communicated. To construct this "babbling" equilibrium, assume that $\beta(t, s)$ is equal to the prior independent of the signal s. R's best response will be to take an action that is optimal conditional only on his prior information. Hence R's action can be taken to be constant. In this case, it is also a best response for S to send a signal that is independent of type, which makes $\beta(t, s)$ the appropriate beliefs. Hence, even if the interests of S and R are identical, so that it there are strong incentives to communicate, there is a possibility of complete communication break down.

Second, it is clear that non-trivial communication requires that different types of S have different preferences over R's actions. If it is the case that whenever some type t prefers action a to action a' then so do all other types, then (ruling out indifference), it must be the case that in equilibrium the receiver takes only one action with positive probability. To see this, note that otherwise one type of sender is not selecting a best response. The second observation shows that cheap talk is not effective in games, like the standard labor-market story, in which the sender's preferences are monotonic in the action of the receiver. With cheap communication, the potential employee in the labor market will always select a signal that leads to the higher possible wage and consequently, in equilibrium, all types of workers will receive the same wage.

A Simple Cheap-Talk Game

There are natural settings in which cheap talk is meaningful in equilibrium. To describe examples, I follow the development of Crawford and Sobel [24] (Green and Stokey [35] independently introduced a similar game in an article circulated in 1981). In this paper, A and T are the unit interval and M can be taken to be the unit interval without loss of generality. The sender's private information or type, t, is drawn from a differentiable probability distribution function, $F(\cdot)$, with density $f(\cdot)$, supported on $[0, 1]$. S and R have twice continuously differentiable von Neumann–Morgenstern utility functions $U^i(a, t)$ that are strictly concave in a and have a strictly positive mixed partial derivative. Let $i = R, S$, $a^i(t)$ denotes the unique solution to $\max_a U^i(a, t)$ and further assume that $a^S(t) > a^R(t)$ for all t. (The assumptions on $U^i(\cdot)$ guarantee that $U^i(\cdot)$ is well defined and strictly increasing.)

In this model, the interests of the sender and receiver are partially aligned because both would like to take a higher action with a higher t. The interests are different because S would always like the action to be a bit higher than R's ideal action. In a typical application, t represents the idea action for R, such as the appropriate expenditure on a public project. Both R and S want actual expenditure to be close to the target value, but S has a bias in favor of additional expenditure.

For $0 \le t' < t'' \le 1$, let $\bar{a}(t', t'')$ be the unique solution to $\max_a \int_{t'}^{t''} U^R(a, t)\,dF(t)$. By convention, $\bar{a}(t, t) = a^R(t)$.

Without loss of generality, limit attention to pure-strategy equilibria. The concavity assumption guarantees that R's best responses will be unique, so R will not randomize in equilibrium. An equilibrium with strategies (μ^*, α^*) **induces** action a if $\{t\colon \alpha^*(\mu^*(t)) = a\}$ has positive prior probability. Crawford and Sobel [24] characterize equilibrium outcomes.

Proposition 4 *There exists a positive integer N^* such that for every integer N with $1 \le N \le N^*$, there exists at least one equilibrium in which the set of induced actions has cardinality N, and moreover, there is no equilibrium which induces more than N^* actions. An equilibrium can be characterized by a partition of the set of types, $t(N) = (t_0(N), \dots, t_N(N))$ with $0 = t_0(N) < t_1(N) < \dots < t_N(N) = 1$, and signals m_i, $i = 1, \dots, N$, such that for all $i = 1, \dots, N-1$* TS2

$$U^S(\bar{a}(t_i, t_{i+1}), t_i)) - U^S(\bar{a}(t_{i-1}, t_i), t_i)) = 0\,, \qquad (6)$$

$$\mu(t) = m_i \text{ for } t \in (t_{i-1}, t_i]\,, \qquad (7)$$

and

$$\alpha(m_i) = \bar{a}(t_{i-1}, t_i)\,. \qquad (8)$$

Furthermore, essentially all equilibrium outcomes can be described in this way.

In an equilibrium, adjacent types pool together and send a common message. Condition (6) states that sender types on the boundary of a partition element are indifferent between pooling with types immediately below or immediately above. Condition (7) states that types in a common element of the partition send the same message. Condition (8) states that R best responds to the information in S's message.

Crawford and Sobel make another monotonicity assumption, which they call condition (M). (M) is satisfied in leading examples and implies that there is a unique equilibrium partition for each $N = 1, \dots, N^*$, the ex-ante equilibrium expected utility for both S and R is increasing in N, and N^* increases if the preferences of S and R become more aligned. These conclusions provide justification for the view that with fixed preferences "more" communication (in the sense of more actions induced) is better for both players and that the closer are the interests of the players the greater the possibilities for communication.

As in the case of models with costly signaling, there are multiple equilibria in the cheap-talk model. The multiplicity is qualitatively different. Costly signaling models have a continuum of Nash Equilibrium outcomes. Cheap-talk models have only finitely many. Refinements that impose restrictions on off-the-equilibrium path signals work well to identify a single outcome in costly signaling models. These refinements have no cutting power in cheap-talk models because any equilibrium distribution on type-action pairs can arise from signaling strategies in which all messages are sent with positive probability. To prove this claim, observe that if message m' is unused in equilibrium, while message m is unused, then one can construct a new equilibrium in which R interprets m' the same way as m and sender types previously sending m randomize equally between m and m'.

TS2 I added a comma. Please check.

Editor's or typesetter's annotations (will be removed before the final TEX run)

In the basic model messages take on meaning only through their use in an equilibrium. Unlike natural language, they have no external meaning. There have been several attempts to formalize the notion that messages have meanings that, if consistent with strategic aspects of the interaction, should be their interpretation inside the game. The first formulation of this idea is due to Farrell [28].

Definition 4 Given an equilibrium (α^*, σ^*) with sender expected payoffs $u^*(\cdot)$, the subset $G \subset T$ is self signaling if $G = \{t\colon U^S(t, BR(G)) > u^*(t)\}$.

That is, G is self signaling if precisely the types in G gain by making a statement that induces the action that is a best response to the information that $t \in G$. (When $BR(t)$ is not single valued it is necessary to refine the definition somewhat and the possibility that $U^S(t, BR(G)) = u^*(t)$ for some t.) See Matthews, Okuno-Fujiwara, and Postlewaite [52].) TS3 Farrell argues that the existence of a self-signaling set would destroy an equilibrium. If a subset G had available a message that meant "my type is in G," then relative to the equilibrium R could infer that if he were to interpret the message literally, then it would be sent only by those types in G (and hence the literal meaning would be accurate). With this motivation, Farrell proposes a refinement.

Definition 5 An equilibrium (α^*, σ^*) is neologism proof if there exist no self-signaling sets relative to the equilibrium.

Rabin [63] argues convincingly that Farrell's definition rules out too many equilibrium outcomes. Indeed, for leading examples of the basic cheap-talk game, there are no neologism-proof equilibria. Specifically, in the Crawford–Sobel model in which S has a bias towards higher actions, there exist self signaling sets of the form $[t, 1]$. On the other hand, Chen, Kartik, and Sobel [20] demonstrate that if one limits attention to equilibria (μ^*, α^*) that the N^*-step equilibrium always satisfies the no incentive to separate (NITS) condition:

$$U^S(\alpha^*(\mu^*(0)), 0) \geq U^S(a^R(0), 0)\,, \tag{9}$$

and that under condition (M) this is the only equilibrium that satisfies Condition (9).

NITS states that the lowest type of Sender prefers her equilibrium payoff to the payoff she would receive if the Receiver knew her type (and responded optimally). [41] introduced and named this condition. The NITS condition can be shown to rule equilibria that admit if self-signaling sets of the form $[0, t]$. Chen [19] and Kartik [41] show that the condition holds in the limits of perturbed versions of the basic cheap-talk game.

Inequality (9) holds automatically in any perfect bayesian equilibrium of the standard signaling model. This follows because when R's actions are monotonic in type and S's preferences are monotonic in action, the worst outcome for S is to be viewed as the lowest type. This observation would not be true in Nash Equilibrium, where it is possible for R to respond to an out-of-equilibirum message with an action $a < BR(0)$.

Variations on Cheap Talk

In standard signaling models, there is typically an equilibrium that is fully revealing. This is not the case in the basic cheap-talk model. This leads to the question of whether it is possible to obtain more revelation in different environments.

One possibility is to consider the possibility of signaling over many dimensions. Chakraborty and Harbaugh [18] consider a model in with T and A are multidimensional. A special case of their model is one in which the components of T are independent draws from the same distribution and A involves taking a real-valued action for each component of T. If preferences are additively separable across types and actions, Charkraborty and Harbaugh provide conditions under which categorical information transmission, in which the S transmits the order of the components of T, is credible in equilibrium even when it would not be possible to transmit information across if the dimensions were treated in isolation. It may be credible for S to say "$t_1 > t_2$," even if she could not credibly provide information about the absolute value of either component of t.

Effective communication requires that different types of preferences have different preferences over outcomes. In standard signaling models, the heterogeneity arises because different sender types have different costs of sending messages. In cheap-talk models, the heterogeneity arises with one-dimensional actions if different sender types have different ideal actions. With multi-dimensional actions, heterogeneity could come simply from different sender types having different preferences over the relative importance of the different issues. Another simple variation is to assume the existence of more than one sender. In the two-sender game, nature picks t as before, both Senders learn t and simultaneously send a message to the receiver, who makes a decision based on the two messages. The second sender has preferences that depend on type and the receiver's action, but not directly on the message sent. In this environment, assume that $M = T$, so that the set of available messages (this is essentially without loss of generality). One can look for equilibria in which the senders report honestly. Denote by $a^*(t, t')R$'s response to the pair of messages (t, t'). If an equilibrium in which both senders report honestly exists, then R's response to identical messages, $a^*(t, t) = a^R(t)$, and it must be the case that there exists a specification of $a(t, t')$ for $t \neq t'$ such that for all $i = 1$ and 2 and $t \neq t'$,

$$U^{S_i}(t, a^*(t, t)) \geq U^{S_i}(t, a^*(t, t')) \,. \qquad (10)$$

It is possible to satisfy Condition (10) if the biases of the senders are small relative to the set of possible best responses. Krishna and Morgan [46] studies a one-dimensional model of information transmission with two informed players. Ambrus and Takahashi [1] and Battiglini [9] provide conditions under which full revelation is possible when there are two informed players and possibly multiple dimensions of information.

In many circumstances, enriching the communication structure either by allowing more rounds of communication (Aumann and Hart [2] and Forges [29]), mediation (Ben-Porath [10]), or exogenous uncertainty (Blume and Board [16] or Kawamura [42]) enlarges the set of equilibrium outcomes.

Verifiable Information

Until now, the focus has been on situations in which the set of signals available does not depend on the true state. There are situations in which this assumption is not appropriate. There may be laws that ban false advertisement. The sender may be able to document details about the value of t. Models of this kind were first studied by Grossman [37] and Milgrom [57]. For example, if t is the sender's skill at playing the piano, then if there is a piano available t could demonstrate that she has skill at least as great as t (by performing at her true ability), but she may not be able to prove that her skill is no more than t (the receiver may think that she deliberately played the piano badly).

To model these possibilities, suppose that the set of possible messages is the set of all subsets of T. In this case, messages have "literal" meanings: When the sender uses the message $s = C \in T$, this can be interpreted as a statement of the form: "my type is in C." If senders cannot lie, then $M(t)$ must be the set of subsets of T that contain t. If type t is verifiable, then $\{t\} \in M(t')$ if and only if $t' = t$. If there are no additional costs of sending signals, this model can be viewed as a variation of cheap talk models in which the message space depends on t. In general, one can treat verifiable information models as a special case of the general signaling game in which the cost of sending certain signals is so large that these signals can be ruled out. Lying is impossible if $M(t) = \{C \subset 2^T : t \in C\}$. In this setting, it is appropriate to require equilibria to be consistent with the signaling structure.

Definition 6 The equilibrium (σ^*, α^*) is rationalizable if

$$\alpha(C, a) > 0 \text{ implies } \sum_{t \in T} U^R(t, s, a)\beta(t, a) = \max_{a' \in A} \sum_{t \in T} U^R(t, s, a')\beta(t, a') \,, \qquad (11)$$

where $\beta(t, a) = 0$ if $t \notin C$.

Compared to (2), (11) requires that beliefs place positive probability only on types capable of sending the message "my type is an element of C."

Proposition 5 *Suppose that A and T are linearly ordered, that the Receiver's best response function is increasing in type, and that all Sender types prefer higher actions. If lying is not possible, then any rationalizable equilibrium* (σ^*, α^*).

Grossman [37] and Milgrom [57] present versions of this proposition. Seidman and Winter [69] generalize the result.

Provided that the Receiver responds to the signal $\{t\}$ with $BR(t)$, each type can guarantee a payoff of $BR(t)$. On the other hand, if any type receives a payoff greater than $BR(t)$, then some higher type must be doing worse. Another way to make the same point is to notice that the highest type $\{\bar{t}\}$ has a weakly dominant strategy to reveal her type by announcing $\{\bar{t}\}$. Once this type is revealed, the next highest type will want to reveal herself and so on. Hence verifiable information will be revealed voluntarily in an environment where cheap talk leads to no revealing and costly signaling will be compatible with full revelation, but at the cost of dissipative signaling.

The full-revelation result depends on the assumption that the sender and receiver share a linear ranking over the quality of information. Giovannini and Seidmann [31] discuss more general settings in which the ability to provide verifiable information need not lead to full revelation.

Communication About Intentions

In a simple signaling game, signals potentially provide information about private information. Another possibilities is to add a round of pre-play communication to a given game. Even if the game has complete information, there is the possibility that communication would serve to select equilibria or permit correlation that would otherwise be infeasible. Farrell and Rabin [64]'s review article discusses this literature in more detail.

Aumann [3] argues that one cannot rely on pre-play communication to select a Pareto-efficient equilibrium. He considers a simple two-player game with Pareto-ranked equilibria and argues that no "cheap" pre-play signal would be credible.

Ben-Porath and Dekel [11] show that adding a stage of "money burning" (a signal that reduces all future payoffs by the same amount) when combined with an equilibrium refinement can select equilibria in a complete information game. Although no money is burned in the selected equilibrium outcome, the potential to send costly signals creates dominance relationships that lead to a selection.

Vida [75] synthesizes a literature that compares the set of equilibrium outcomes available when communication possibilities are added to a game to the theoretically larger set available if there is a reliable mediator available to collect information and recommend actions to the players.

Applications

Economic Applications

There is an enormous literature that uses signaling models in applications. Riley's [65] survey contains extended discussion of some of the most important applications. What follows is a brief discussion of some central ideas.

In a simple signaling game, one informed agent sends a single signal to one uninformed decision maker. This setting is reach enough to illustrate many important aspects of signaling, but it plainly limited. Interesting new issues arise if there are many informed agents, if there are many decision makers, and if the interaction is repeated. Several of the models below add some or all of these novel features to the basic model.

Advertising Advertisements are signals. Models similar to the standard model can explain situations in which higher levels of advertisement can lead consumers to believe the quality of the good is higher. In a separating equilibrium, advertising expenditures fully reveal quality. As in all costly signaling models, it is not important that there be a direct relationship between quality and signal, it is only necessary that firms with higher quality have lower marginal costs of advertising. Hence simply "burning money" or sending a signal that lowers utility by an amount independent of quality and response can be informative. The consumer may obtain full information in equilibrium, but someone must pay the cost of advertising. There are other situations where it is natural for the signal to be linked to the quality of the item. Models of verifiable information are appropriate in this case. When the assumptions of Proposition 5 hold, one would expect consumers to obtain all relevant information through disclosures without wasteful expenditures on signaling. Finally, cheap talk plays a role in some markets. One would expect costless communication to be informative in environments where heterogeneous consumers would like to identify the best product. Cheap talk can create more efficient matching of product to consumer. Here communication is free although will in leading models separating equilibria do not exist.

Limit Pricing Signaling models offer one explanation for the phenomenon of limit pricing. An incumbent firm have private information about its cost. Potential entrants use the pricing behavior of the firm to draw inferences about the incumbent's cost, which determines profitability of entry. Milgrom and Roberts [55] construct an equilibrium in which the existence of incomplete information distorts prices: Relative to the full information model, the incumbent charges lower prices in order to signal that the market is relatively unprofitable. This behavior has the flavor of classical models of limit pricing, with one important qualification. In a separating equilibrium the entrant can infer the true cost of the incumbent and therefore the low prices charged by the incumbent firm fails to change the entry decision.

Bargaining Several authors have proposed bargaining models with incomplete information to study the existence and duration of strikes (Fudenberg and Tirole [30], Sobel and Takahashi [72]). If a firm with private information about its profitability makes a take-it-or-leave it offer to a union, then the strategic interaction is a simple signaling model in which the magnitude of the offer may serve as a signal of the firm's profitability. Firms with low profits are better able to make low wage offers to the union because the threat of a strike is less costly to a firm with low profits than one with high profits. Consequently settlement offers may reveal information. Natural extensions of this model permit counter offers. The variation of the model in which the uninformed agent makes offers and the uninformed agent accepts and rejects is formally almost identical to the canonical model of price discrimination by a durable-goods monopolist (Ausubel and Deneckere [4] and Gul, Sonnenschein, and Wilson [39]).

Finance Simple signaling arguments provide potential explanations for firms' choices of financial structure. Classic arguments due to Modigliani and Miller [59] and imply that firms' profitability should not depend on their choice of capital structure. Hence this theory cannot organize empirical regularities about firm's capital structure. The Modigliani–Miller theorem assumes that the firm's managers, shareholders, and potential shareholders all have access to the same information. An enormous literature assumes instead that the firm's managers have superior information and use corporate structure to signal profitability.

Leland and Pyle [48] assume that insiders are risk averse they would prefer to diversify their personal holdings rather than maintain large investments in their firm. The value of diversification is greater the lower the quality of the firm. Hence when insiders have superior information than investors, there will be an incentive for the insiders of highly profitable firms to hold maintain inefficiently large investments in their firm in order to signal profitability to investors.

Dividends are taxed twice under the United States tax code, which raises the question of why firms would issue dividends when capital gains are taxed at a lower rate. A potential explanation for this behavior comes from a model in which investors have imperfect information about the future profitability of the firm and profitable firms are more able than less profitable firms to distribute profits in the form of dividends (see Bhattachrya [14]).

Reputation Dynamic models of incomplete information create the opportunity for the receiver to draw inferences about the sender's private information while engaging in an extended interaction. Kreps and Wilson [45] and Milgrom and Roberts [56] provided the original treatments of reputation formation in games of incomplete information. Motivated by the limit pricing, their models examined the interaction of a single long-lived incumbent facing a sequence of potential entrants. The entrants lack information about the willingness of the incumbent to tolerate entry. Pricing decisions of the incumbent provide information to the entrants about the profitability of the market.

In these models, signals have implications for both current and future utility. The current cost is determined by the effect the signal has on current payoffs. In Kreps–Wilson and Milgrom–Roberts, this cost is the decrease in current profits associated with charging a low price. In other models (for example Morris [60] or Sobel [71]) the actual signal is costless, but it has immediate payoff implications because of the response it induces. Signals also have implications for future utility because inferences about the sender's private information will influence the behavior of the opponents in future periods. Adding concern for reputation to a signaling game will influence behavior, but whether it leads to more or less informative signaling depends on the application.

Signaling in Biology

Signaling is important in biology. In independent and almost contemporaneous work, Zahavi [77] proposed a signaling model that shared the essential features of Spence [73]'s model of labor-market signaling. Zahavi observed that there are many examples in nature of animals apparently excessive physical displays. It takes energy to produce colorful plumage, large antlers, or loud cries. Having a large tail may actually make it harder for peacocks to flea predators. If a baby bird makes a loud sound to get his mother's attention, he may attract a dangerous predator. Zahavi argued that costly signals could play a role in sexual selection. In Zahavi's basic model, the sender is a male and the receiver is a female of the same species. Females who are able to mate with healthier males are more likely to have stronger children, but often the quality of a potential mate cannot be observed directly. Zahavi argued that if healthier males could produce visible displays more cheaply than less healthy males, then females would be induced to use the signals when deciding upon a mate. Displays may impose costs that "handicap" a signaler, but displays would persist when additional reproductive success compensates for their costs. Zahavi identifies a single-crossing condition as a necessary condition for the existence of costly signals.

The development of signaling in biology parallels that in economics, but there are important differences. Biology replaces the assumption of utility maximization and equilibrium with fitness maximization and evolutionary stability. That is, their models do not assume that animals consciously select their signal to maximize a payoff. Instead, the biological models assume that the process of natural selection will lead to strategy profiles in which mutant behavior has lower reproductive fitness than equilibrium behavior. This notion leads to static and dynamic solution concepts similar to Nash Equilibrium and its refinements. Fitness in biological models depends on contributions from both parents. Consequently, a full treatment of signaling must take into account population genetics. Grafen [34] discusses these issues and Grafen [33] and Siller [70] provide further theoretical development of the handicap theory. Finally, one must be careful in interpreting heterogeneous quality in biological models. Presumably natural selection will act to eliminate the least fit individuals. Natural selection should operate to eliminate the least fit genes in a population. To the extent that this arises, there is pressure for quality variation within a population to decrease over time. The existence of unobserved quality variations needed for signaling may be the result of relatively small variations about a population norm.

While most of the literature on signaling in biology focuses on the use of costly signals, there are also situations in which cheap talk is effective. A leading example is the "Sir Philip Sidney Game," originally developed by John Maynard Smith [54] to illustrate the value of costly communication between a mother and child. The child has private information about its level of hunger and the mother must decide to feed the child or keep the food for itself. Since the players are related, survival of one positively influences the fitness of the other. This creates a common interest needed for cheap-talk communication. There are two ways to model communication in this environment. The first is to assume that signaling is costly, with hungrier babies better able to communicate their hunger. This could be because the sound of a hungry baby is hard for sated babies to imitate or it could be that crying for food increases the risk of predation and that this risk is relatively more dangerous to well fed chicks than to starving ones (because the starving chicks have nothing to lose). This game has multiple equilibria in which signals fully reveal the state of the baby over a range of values (see Maynard Smith [54] and Lachmann and Bergstrom [47]). These papers look a model in which both mother and child have private information. Alternatively, Bergstrom and Lachmann [13] study a cheap-talk version of the game. Here there may be an equilibrium outcome in which the baby bird credibly signals whether or not he is hungry. Those who signal hunger get fed. The others do not. Well fed baby birds may wish to signal that they are not hungry in order to permit the mother to keep food for herself. Such an equilibrium exists if the fraction of genes that mother and child share is large and the baby is already well fed.

Political Science

Signaling games have played an important role in formal models of political science. Banks [7] reviews models of agenda control, political rhetoric, voting, and electoral competition. Several important models in this area are formally interesting because they violate the standard assumptions frequently satisfied in economic models. I describe two such models in this subsection.

Banks [6] studies a model of agenda setting in which the informed sender proposes a policy to a receiver (decision-maker), who can either accept or reject the proposal. If the proposal is accepted, it becomes the outcome. If not, then the outcome is a fall-back policy. The fall-back policy is known only to the sender. In this environment, the sender's strategy may convey information to the decision maker. Signaling is costly, but, because the receiver's set of actions in binary, fully revealing equilibria need not exist. Refinements limit the set of predictions in this model to a class of outcomes in which only one proposal is accepted in equilibrium (and that this proposal is accepted with probability one), but there are typically a continuum of possible equilibrium outcomes.

Matthews [53] develops a cheap-talk model of veto threats. There are two players, a Chooser (C), who plays the role of receiver, and a Proposer (P), who plays the role of sender. The players have preferences that are represented by single-peaked utility functions which depend on the real-valued outcome of the game and an ideal point. P's ideal point is common knowledge. C's ideal point is her private information, drawn from a prior distribution that has a smooth positive density on a compact interval, $[\underline{t}, \bar{t}]$. The game form is simple: C learns her type, then sends a cheap-talk signal to P, who responds with a proposal. C then either accepts or rejects the proposal. Accepted proposals become the outcome of the game. If C rejects the proposal, then the outcome is the status quo point.

As usual in cheap-talk games, this game has a babbling outcome in which C's message contains no information and P makes a single, take-it-or-leave-it offer that is accepted with probability strictly between 0 and 1. Matthews shows there may be equilibria in which two outcomes are induced with positive probability (size-two equilibria), but size $n > 2$ (perfect Bayesian) equilibria never exist. In a size-two equilibrium, P offers his ideal outcome to those types of C whose message indicates that their ideal point is low; this offer is always accepted in equilibrium. If C indicated that his ideal point is high, P makes a compromise offer that is sometimes accepted and sometimes rejected.

Future Directions

The most exciting developments in signaling games in the future are likely to come from interaction between economics and other disciplines.

Over the last ten years the influence of behavioral economists have led the profession to rethink many of its fundamental models. An explosion of experimental studies have already influenced the interpretation of signaling models and have led to a re-examination of basic assumptions. There is evidence that economic actors lack the strategic sophistication assumed in equilibrium models. Further, economic agents may be motivated by more than their material well being. Existing experimental evidence provides broad support for many of the qualitative predictions of the theory (Banks, Camerer, and Porter [5] and Brandts and Holt [17]), but also suggests ways in which the theory may be inadequate.

The driving assumption of signaling models is that when informational asymmetries exist, senders will attempt to lie for strategic advantage and that sophisticated receivers will discount statements. These assumptions may be reconsidered in light of experimental evidence that some agents will behave honestly in spite of strategic incentives to lie. For example, Gneezy [32] and Hurkens and Kartik [40] present experimental evidence that some agents are reluctant to lie even when there is a financial gain from doing so. There is evidence from other disciplines that some agents are unwilling or unable to manipulate information for strategic advantage and that people may be well equipped to detect these manipulations in ways that are not captured in standard models (see, for example, Ekman [27] or Trivers [74]). Experimental evidence and, possibly, results from neuroscience may demonstrate that the standard assumption that some agents cannot manipulate information for their strategic advantage (or that other agents have ability to see through deception) will inform the development of novel models of communication in that include behavioral types. Several papers study the implications of including behavioral types into the standard paradigm. The reputation models of Kreps and Wilson [45] and Milgrom and Roberts [55] are two early examples. Recent papers on communication by Chen [19], Crawford [23], Kartik [41], and Olszewski [62] are more recent examples. New developments in behavioral economics will inform future theoretical studies.

There is substantial interest in signaling in philosophy. Indeed, the philosopher David Lewis [49] (first published in 1969) introduced signaling games prior to the contributions of Spence and Zahavi. Recently linguists have been paying more attention to game-theoretic ideas. Benz, Jäger and Van Rooij [12] collects recent work that attempts to formalize ideas from linguistic philosophy due to Grice [36]. While there have been a small number of contributions by economists in this area (Rubinstein [67] and Sally [68] are examples), there is likely to be more active interaction in the future.

Finally, future work may connect strategic aspects of communication to the actual structure of language. Blume [15], Cucker, Smale, Zhou [25], and Nowak and Krakauer [61] present dramatically different models on how structured communication may result from learning processes. Synthesizing these approaches may lead to fundamental insights on how the ability to send and receive signals develops.

Acknowledgments

I thank the Guggenheim Foundation, NSF, and the Secretaría de Estado de Universidades e Investigación del Ministerio de Educación y Ciencia (Spain) for financial support. I am grateful to the Departament d'Economia i d'Història Econòmica and Institut d'Anàlisi Econòmica of the Universitat Autònoma de Barcelona for hospitality and administrative support.

Bibliography

Primary Literature

1. Ambrus A, Takahashi S (2006) Multi-sender cheap talk with restricted state space. Technical report. Harvard University
2. Aumann R, Hart S (2003) Long cheap talk. Econometrica 71(6):1619–1660
3. Aumann RJ (1990) Nash equilibria and not self-enforcing. In: Gabszewicz JJ, Richard J-F, Wolsey LA (eds) Economic Decision Making: Games, Econometrics and Optimisation. Elsevier, CE4, pp 201–206
4. Ausubel LM, Deneckere RJ (1989) Reputation in bargaining and durable goods monopoly. Econometrica 57(3):511–531
5. Banks J, Camerer C, Porter D (1994) An experimental analysis of nash refinements in signaling games. Games Econ Behav 6(1):1–31
6. Banks JS (1990) Monopoly agenda control and asymmetric information. Q J Econ 105(2):445–464
7. Banks JS (1991) Signaling Games in Political Science. Routledge, CE4
8. Banks JS, Sobel J (1987) Equilibrium selection in signaling games. Econometrica 55(3):647–661
9. Battaglini M (2002) Multiple referrals and multidimensional cheap talk. Econometrica 70(4):1379–1401
10. Ben-Porath E (2003) Cheap talk in games with incomplete information. J Econ Theory 108:45–71
11. Ben-Porath E, Dekel E (1992) Signaling future actions and the potential for sacrifice. J Econ Theory 57:36–51
12. Benz A, Jäger G, Van Rooij R (eds) (2005) Game Theory and Pragmatics. Palgrave MacMillan, CE4
13. Bergstrom CT, Lachmann M (1998) Signalling among relatives. III. talk is cheap. Proc Natl Acad Sci USA 95:5100–5015
14. Bhattacharya S (1979) Imperfect information, dividend policy, and the bird in the hand fallacy. Bell J Econ 10(1):259–270
15. Blume A (2000) Coordination and learning with a partial language. J Econ Theory 95:1–36
16. Blume A, Board O (2006) Noisy talk. Technical report. University of Pittsburgh
17. Brandts J, Holt CA (1992) An experimental test of equilibrium dominance in signaling games. Am Econ Rev 82(5):1350–1365
18. Chakraborty A, Harbaugh R (2007) Comparative cheap talk. J Econ Theory CE5
19. Chen Y (2005) Perturbed communication games with honest senders and naive receivers. Technical report. Arizona State University, CE4

CE4 Please provide the publisher location.
CE5 Please provide the volume and page-range.

Editor's or typesetter's annotations (will be removed before the final TEX run)

20. Chen Y, Kartik N, Sobel J (2007) On the robustness of informative cheap talk. Technical report, UCSD
21. Cho I-K, Kreps DM (1987) Signaling games and stable equilibria. Q J Econ 102(2):179–221
22. Cho I-K, Sobel J (1990) Strategic stability and uniqueness in signaling games. J Econ Theory 50(2):381–413
23. Crawford VP (2003) Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions. Am Econ Rev 93(1):133–149
24. Crawford VP, Sobel J (1982) Strategic information transmission. Econometrica 50(6):1431–1451
25. Cucker F, Smale S, Zhou D-X (2004) Modeling language evolution. Found Comput Math 4(3):315–343
26. Edlin AS, Shannon C (1998) Strict single-crossing and the strict spence-mirrlees condition: A comment on monotone comparative statics. Econometrica 66(6):1417–1425
27. Ekman P (2001) Telling Lies. W.W. Norton, New York
28. Farrell J (1993) Meaning and credibility in cheap-talk games. Games Econ Behav 5(4):514–531
29. Forges F (1990) Equilibria with communication in a job market example. Q J Econ 105(2):375–398
30. Fudenberg D, Tirole J (1983) Sequential bargaining with incomplete information. Rev Econ Stud 50(2):221–247
31. Giovannoni F, Seidmann DJ (CE6) Secrecy, two-sided bias and the value of evidence. Games Econ Behav. forthcoming
32. Gneezy U (2005) Deception: The role of consequences. Am Econ Rev 95(1):384–394
33. Grafen A (1990) Biological signals as handicaps. J Theor Biol 144:517–546
34. Grafen A (1990) Sexual selection unhandicapped by the fisher process. J Theor Biol 144:473–516
35. Green JR, Stokey NL (CE6) A two-person game of information transmission. J Econ Theory. forthcoming
36. Grice HP (1991) Studies in the Way of Words. Harvard University Press, CE4
37. Grossman S (1981) The role of warranties and private disclosure about product quality. J Law Econ 24:461–483
38. Grossman S, Perry M (1987) Perfect sequential equilibria. J Econ Theory 39:97–119
39. Gul F, Sonnenchein H, Wilson R (1986) Foundations of dynamic monopoly and the coase conjecture. J Econ Theory 39:155–190
40. Hurkens S, Kartik N (2006) (When) Would I lie to you? Comment on Deception: The role of consequences. Technical report, UCSD
41. Kartik N (2005) Information transmission with almost-cheap talk. Technical report, UCSD
42. Kawamura K (2006) Anonymity, garbling, and overconfidence in communication. Technical report. Nuffield College, Oxford
43. Kohlberg E, Mertens J-F (1986) On the strategic stability of equilibria. Econometrica 54(5):1003–1037
44. Kreps DM, Sobel J (1994) Signalling. In: Aumann RJ, Hart S (eds) Handbook of game theory: with economics applications. Handbooks in Economics, no 11, vol 2. Elsevier, CE4, chap 25, pp 849–868
45. Kreps DM, Wilson R (1982) Reputation and imperfect information. J Econ Theory 27:253–277
46. Krishna V, Morgan J (2001) A model of expertise. Q J Econ 116(2):747–775
47. Lachmann M, Bergstrom CT (1998) Signalling among relatives. II. beyond the tower of babel. Theor Popul Biol 54:146–160
48. Leland HE, Pyle DH (1977) Informational asymmetries, financial structure, and financial intermediation informational asymmetries, financial structure, and financial intermediation. J Finance 32(2):371–387
49. Lewis D (2002) Convention: A Philosophical Study. Blackwell, CE4
50. Mailath GJ (1987) Incentive compatibility in signaling games with a continuum of types. Econometrica 55(6):1349–1365
51. Mailath GJ, Okuno-Fujiwara M, Postlewaite A (1993) On belief based refinements in signaling games. J Econ Theory 60(2):241–276
52. Mathews SA, Okuno-Fujiwara M, Postlewaite A (1991) Refining cheap-talk equilibria. J Econ Theory 55(2):247–273
53. Matthews SA (1989) Veto threats: Rhetoric in a bargaining game. Q J Econ 104(2):347–369
54. Maynard Smith J (1991) Honest signalling: the Philip Sidney game. Anim Behav 42:1034–1035
55. Milgrom P, Roberts J (1982) Limit pricing and entry under incomplete information: An equilibrium analysis. Econometrica 50(2):443–459
56. Milgrom P, Roberts J (1982) Predation, reputation, and entry deterence. J Econ Theory 27:280–312
57. Milgrom PR (1981) Good news and bad news: Representation theorems. Bell J Econ 21:380–391
58. Milgrom PR, Shannon C (1994) Monotone comparative statics. Econometrica 62(1):157–180
59. Modigliani F, Miller MH (1958) The cost of capital, corporation finance and the theory of investment. Am Econ Rev 48(3):261–297
60. Morris S (2001) Political correctness. J Political Econ 109:231–265
61. Nowak MA, Krakauer DC (1999) The evolution of language. Proc Natl Acad Sci 96(14):8028–8033
62. Olszewski W (2004) Informal communication. J Econ Theory 117:180–200
63. Rabin M (1990) Communication between rational agents. J Econ Theory 51:144–170
64. Rabin M, Farrell J (1996) Cheap talk. J Econ Perspectives 10(3):103–118
65. Riley JG (2001) Silver signals: Twenty-five years of screening and signaling. J Econ Lit 39(2):432–478
66. Rothschild M, Stiglitz J (1976) Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. Q J Econ 90(4):629–649
67. Rubinstein A (2000) Economics and language. Cambridge University Press, CE4
68. Sally D (2005) Can I say, bobobo and mean, there's no such thing as cheap talk? J Econ Behav Organ 57(3):245–266
69. Seidmann DJ, Winter E (1997) Strategic information transmission with verifiable messages. Econometrica 65(1):163–169
70. Siller S (1998) A note on errors in grafen's strategic handicap models. J Theor Biol 195:413–417
71. Sobel J (1985) A theory of credibility. Rev Econ Stud 52(4):557–573
72. Sobel J, Takahashi I (1983) A multistage model of bargaining. Rev Econ Stud 50(3):411–426
73. Spence AM (1974) Market Signaling. Harvard University Press, Cambridge
74. Trivers RL (1971) The evolution of reciprocal altruism. Q Rev Biol 46:35–58

CE6 Please update if possible.

75. Vida P (2006) Long Pre-Play Communication in Games. PhD thesis, Autonomous University of Barcelona
76. Wilson C (1977) A model of insurance markets with incomplete information. J Econ Theory 16:167–207
77. Zahavi A (1975) Mate selection- a selection for a handicap. J Theor Biol 53:205–214

Books and Reviews

Admati A, Perry M (1987) Strategic delay in bargaining. Rev Econ Stud 54:345–364
Austen-Smith D (1990) Information transmission in debate. Am J Political Sci 34(1):124–152
Battigalli P, Siniscalchi M (2002) Strong belief and forward induction reasoning. J Econ Theory 106(2):356–391
Bernheim BD (1994) A theory of conformity. J Political Econ 102(5):841–877
Blume A, Kim Y-G, Sobel J (1993) Evolutionary stability in games of communication. Games Econ Behav 5:547–575
Fudenberg D, Tirole J (1991) Game Theory. MIT Press, Cambridge
Gibbons R (1992) Game Theory for Applied Economists. Princeton University Press, CE4
Kartik N, Ottaviani M, Squintani F (2007) Credulity, lies, and costly talk. J Econ Theory 134(1):93–116
Krishna V, Morgan J (2001) A model of expertise. Q J Econ 116(2): 747–775
Lo P-Y (2006) Common knowledge of language and iterative admissibility in a sender-receiver game. Technical report. Brown University
Manelli AM (1996) Cheap talk and sequential equilibria in signaling games. Econometrica 64(4):917–942
Milgrom P, Roberts J (1986) Price and advertising signals of product quality. J Political Econ 94(4):796–821
Noldeke G, Samuelson L (1997) A dynamic model of equilibrium selection in signaling markets. J Econ Theory 73(1):118–156
Noldeke G, Van Damme E (1990) Signalling in a dynamic labour market. Rev Econ Stud 57(1):1–23
Ottaviani M, Sørensen PN (2006) Professional advice. J Econ Theory 126(1):120–142
Rabin M (1994) A model of pre-game communication. J Econ Theory 63(2):370–391
Ramey G (1996) D1 signaling equilibria with multiple signals and a continuum of types. J Econ Theory 69(2):508–531
Rasmusen EB (2006) Games and Information: An Introduction to Game Theory, 4th edn. Blackwell, CE4
Riley JG (1979) Informational equilibrium. Econometrica 47:331–359
Sobel J, Stole L, Zapater I (1990) Fixed-equilibrium rationalizability in signaling games. J Econ Theory 52(2):304–331
Spence AM (1973) Job market signaling. Q J Econ 90:225–243
Swinkels JM (1999) Education signallling with preemptive offers. Rev Econ Stud 66(4):949–970