Kellogg School of Management at Northwestern University
Nancy L. Schwartz Memorial Lecture, 19 April 2017

Modeling Strategic Communication: From Rendezvous and Reassurance to Trickery and Puffery

Vincent P. Crawford

University of Oxford, All Souls College, and University of California, San Diego

I owe thanks to many friends, students, and co-authors for valuable conversations and advice on this topic. My research received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 339179. The contents reflect only my views and not the views of the ERC or the European Commission, and the European Union is not liable for any use that may be made of the information contained therein. The University of Oxford, All Souls College, and the University of California, San Diego also provided funding.

Revised 25 April 2017

Introduction

Communication plays an essential role in most human interactions.

Yet economic and game-theoretic models of communication seem to fall short of what's needed for a full understanding of its effects.

In this lecture I draw lessons from existing work and make suggestions for future work, motivated by two main desiderata:

- Models of communication should reflect that it is an all-purpose tool, which people use as much to persuade or deceive as to coordinate

  By contrast, early analyses like Schelling 1960 and Lewis 1969 focused on coordination with common interests; and even recent analyses do not always reflect a sufficiently wide range of strategic purposes

- Models of communication should not reflexively assume Nash equilibrium in applications where people cannot plausibly have learned to play one and equilibrium's strategic thinking justification is strained

  Instead, behavioral assumptions should be evidence-based, replacing equilibrium with empirically better-supported alternatives as needed

This lecture compares the predictions of Nash equilibrium-based models and a class of nonequilibrium behavioral models based on "level-k" thinking, with evidence about how people use communication.

The comparison spans games where communication serves a variety of purposes, from Rendezvous and Reassurance to Trickery and Puffery.

Level-k models come closer to the evidence, but still stop short of what's needed to understand how communication matters in relationships.

The lecture ends by discussing kind of models are needed to do better:

- Models that do full justice to *strategic uncertainty*, the unpredictability of how people respond to strategic settings; and

- Models in which there are substantive differences between models without and with communication, whether abstract or natural-language

Such models should help us to understand better how people bring about and maintain cooperation in relationships.

Terminology:

- In a *game*, two or more *players* choose decisions or *strategies*, which jointly determine their welfares or *payoffs*. Any uncertainty is handled by assuming that players' welfares are represented by their (mathematically) *expected payoffs*

- For simplicity I will assume that players know a game's structure as *common knowledge*, except if there is private information, its *distributions* are what is assumed to be common knowledge

  Thus players face mainly uncertainty about other players' responses to the game: *strategic uncertainty*

Rendezvous and Reassurance, without communication

The most familiar Rendezvous game is Battle of the Sexes.

Its players choose simultaneously where to try to meet, with a strong preference to meet *somewhere* but different preferences about *where*.

	Fights	Ballet
Fights	2, 1	0, 0
Ballet	0, 0	1, 2

Battle of Sexes

The main issue that players face is finding a way to break the symmetry of their roles, as is required for efficient coordination.

The most familiar Reassurance game is Stag Hunt.

In Rousseau's example (*Discourse on Inequality* 1754 [1973]):

> If a deer was to be taken, everyone saw that, in order to succeed, he must abide faithfully by his post: but if a hare happened to come within the reach of any one of them, it is not to be doubted that he pursued it without scruple, and, having seized his prey, cared very little, if by so doing he caused his companions to miss theirs.

	Stag	**Hare**
Stag	9 9	8 0
Hare	0 8	7 7

Stag Hunt

(Here I follow Aumann's 1990 and Charness's *GEB* 2000 payoffs.)

The main issue is the tension between the higher potential payoff of playing Stag and its greater strategic uncertainty.

Nash Equilibrium

A *Nash equilibrium*, henceforth shortened to *equilibrium*, is a combination of decisions or strategies in which each player's choice maximizes her/his expected payoff, given other players' choices.

A Nash equilibrium is thus a kind of "rational expectations" equilibrium, in which players form expectations or *beliefs* about each other's choices that are self-confirming if players choose best responses to their beliefs.

Equilibrium in Battles of the Sexes without communication

	Fights	Ballet
Fights	1 2	0 0
Ballet	0 0	2 1

Battle of Sexes

Battle of the Sexes without communication has three equilibria:

- Two asymmetric *pure-strategy* (with unrandomized decisions) equilibria: (Fights, Fights) and (Ballet, Ballet)
- One symmetric *mixed-strategy* equilibrium, in which each player chooses her/his favorite place to meet with probability 2/3 (only 2/3 makes the other willing to choose her/his equilibrium mixed strategy)

  Players' expected payoffs of 2/3 in the mixed-strategy equilibrium are less than their payoffs 1 or 2 in the pure-strategy equilibria: inefficient.

When players have no way to distinguish their roles, the mixed-strategy equilibrium is arguably the only equilibrium that has the potential to describe their behavior, and I focus on it when discussing this game.

Equilibrium in Stag Hunt without communication

	Stag	Hare
Stag	9 9	8 0
Hare	0 8	7 7

Stag Hunt

Stag Hunt also has three equilibria:

- Two symmetric pure-strategy equilibria: “both-Stag” and “both-Hare”
- One symmetric mixed-strategy equilibrium

In Stag Hunt the mixed-strategy equilibrium is arguably behaviorally irrelevant, and I will ignore it.

Both-Stag is better for both players than both-Hare (or the mixed-strategy equilibrium): “payoff-dominant” (Harsanyi and Selten 1988).

But both-Hare has players choosing best responses to a larger range of players’ beliefs: “risk-dominant”.

Harsanyi and Selten therefore favor both-Hare over the other equilibria.

Nash equilibrium as a behavioral model

Economic theory almost always assumes equilibrium, for reasons—good reasons!—explained by Myerson *JEL* 1999.

Equilibrium "builds in" the rationality of individual decisions, and experiments suggest that well-motivated subjects who understand the game satisfy decision-theoretic rationality 80-90% of the time.

But equilibrium bundles decision-theoretic rationality with the far stronger assumption that players' beliefs and strategies are coordinated: stronger because a player's equilibrium choice often maximizes expected payoff only if s/he believes others are likely to make their equilibrium choices.

This stronger assumption serves an important purpose, because a useful theory needs more than decision-theoretic rationality:

Replacing equilibrium with weaker assumptions like *rationalizability* (Bernheim or Pearce *Ecma* 1984) yields few restrictions in economically interesting games, and none at all in Battle of the Sexes or Stag Hunt.

The coordination of players' beliefs that equilibrium normally requires has two alternative possible justifications:

- *learning* from experience with analogous games, which has a strong tendency to make players' beliefs converge to equilibrium

- *strategic thinking*, which under strong assumptions regarding players' knowledge of each other's beliefs (Brandenburger *JEP* 1992) can in theory yield equilibrium beliefs without learning, even in players' initial responses to a game

(It is often thought that sufficient pre-play communication assures equilibrium in the underlying game, but Farrell *EL* 1988 showed that that is not generally true without the question-begging assumption of equilibrium in the entire game, including the communication phase.)

Although we often rely on learning or thinking justifications for equilibrium:

- In many situations that we use game theory to analyze, players don't really have enough clear precedents to learn an equilibrium

and

- In many such situations, the thinking justification is behaviorally implausible

  In particular, experiments that elicit initial responses to games suggest that people rarely follow the fixed-point or indefinitely iterated dominance reasoning that equilibrium usually requires.

  Nor is this surprising, because Gilboa, Kalai, and Zemel *MathOR* 1993 and others have shown how computationally hard such operations are.

  (The claim is only that fixed-point or indefinitely iterated dominance reasoning seldom *directly* describe people's thinking; learning can still converge to something analysts need fixed points to describe.)

Evidence on Battle of the Sexes and Stag Hunt without communication

In experiments that elicit initial responses to entry games like Battle of the Sexes, subjects' aggregate choice frequencies deviate systematically from the equilibrium mixed strategies, but come fairly close.

> "…to a psychologist, it looks like magic."
> —Kahneman 1988 (quoted in Camerer, Ho, and Chong *QJE* 2004)

Somehow in these symmetric games, subjects also do better on average than in the best symmetric equilibrium, the mixed-strategy equilibrium (Camerer et al. *QJE* 2004, Crawford et al. *JEL* 2013, Section 6).

In experiments that elicit initial responses to games like Stag Hunt, most subjects play risk-dominant strategies like Hare, but a few play payoff-dominant strategies like Stag; as a result, the aggregate choice frequencies deviate systematically from those of any equilibrium.

By contrast, in experiments with both kinds of game that allow subjects to learn from clear precedents, almost all converge to *some* equilibrium.

Nonequilibrium behavioral models based on “level-*k*” thinking

Applications involving initial responses require a nonequilibrium model that better describes people’s choices.

To be useful in applications, such a model must also yield precise predictions (conditional on measurable behavioral parameters).

When people lack precedents for learning, and equilibrium reasoning is inaccessible, they must find another way to think about their choices.

In games without communication, much experimental evidence points to a particular class of models, based on level-*k* thinking.

Level-*k* models were developed to describe experimental results by Stahl and Wilson *JEBO* 1994, *GEB* 1995; Nagel *AER* 1995; Costa-Gomes, Crawford, and Broseta *Ecma* 2001; Camerer et al. *QJE* 2004 (“cognitive hierarchy” models); and Costa-Gomes and Crawford *AER* 2006.

Level-*k* models were later adapted to games with communication as explained below, by Crawford *AER* 2003; Cai and Wang *GEB* 2006; Kartik, Ottaviani, and Squintani *JET* 2007; Ellingsen and Östling *AER* 2010; and Wang, Spezio, and Camerer *AER* 2010.

In a level-*k* model, players follow rules of thumb that:

● anchor their beliefs in a naïve model of others' responses, called *L0*

and

● adjust their beliefs via a small, heterogeneous number (*k*) of iterated best responses: *L1* best responds to *L0*, *L2* to *L1*, and so on

In simple games, level-*k* rules mimic equilibrium decisions; but they can deviate systematically from equilibrium in more complex games.

Level-*k* players use step-by-step procedures that generically determine unique pure strategies, with no need for fixed-point reasoning.

In Selten's *EER* 1998 words:

> Basic concepts in game theory are often circular in the sense that they are based on definitions by implicit properties…. Boundedly rational strategic reasoning seems to avoid circular concepts. It directly results in a procedure by which a problem solution is found.

Compare the heterogeneous, finite iteration of best responses in Keynes' 1936 *General Theory* comparison of professional investment

> . . . to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. *We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees.* [emphasis added]

(We know from Samuelson's *Ecma* 1946 obituary article that Keynes's German was weak enough to make it unlikely he had read von Neumann *MathAnn* 1928, the precursor to von Neumann and Morgenstern's 1944 minimax/equilibrium analysis of zero-sum two-person games. Thus he was less tempted to assume Nash equilibrium than we are. But still!)

In a level-*k* model, the population frequencies of levels are treated as behavioral parameters.

Estimates vary with the setting, but the frequency of *L0* is normally zero or small and the level distribution is concentrated on *L1*, *L2*, and *L3*.

Importantly, a level-*k* model makes precise (probabilistic) predictions: not only that deviations from equilibrium will sometimes occur, but also which settings evoke them and which forms they are likely to take.

In games without communication, most of the evidence is consistent with *L0* being uniform random over the feasible decisions.

This "random" *L0* can be thought of as reflecting higher levels' thinking about the incentives the payoff structure creates for their own choices, thinking about others' choices via the principle of insufficient reason before they consider others' incentives (Crawford et al. *JEL* 2013).

(For games with communication via messages with literal meanings, a smaller body of evidence, discussed below, suggests a "truthful" *L0*.)

Aside:

- *Lk* (for *k* > 0) is decision-theoretically rational, with an accurate model of the game; it departs from equilibrium only in deriving its beliefs from an oversimplified nonequilibrium model of others' responses

- *Lk* (for *k* > 0) respects *k*-rationalizability (Bernheim 1984 *Ecma*), hence in two-person games its choices survive *k* rounds of iterated elimination of strictly dominated strategies

- Thus *Lk* (for *k* > 0) mimics equilibrium decisions in *k*-dominance-solvable games, but can deviate systematically in other games

  (Such deviations make it possible for the model to systematically out-predict a rational-expectations notion such as equilibrium)

- A level-*k* model with zero weight on *L0* can be viewed as a heterogeneity-tolerant refinement of *k*-rationalizability

Level-*k* thinking in Battle of the Sexes without communication

	Fights	Ballet
Fights	2, 1	0, 0
Ballet	0, 0	1, 2

Battle of Sexes

Imagine that the population of potential players contains only *L1*s and *L2s*, with equal frequencies in each player role (Crawford et al. *JEL* 2013, Section 6).

When *L0* is random, *L1*s choose their favorite meeting places, Fights for Row and Ballet for Column; while *L2*s choose their partners' favorites.

(Higher levels continue to alternate between the two strategies, so in this case *L1*s can represent all odd levels, and *L2*s all even levels.)

	Fights	Ballet
Fights	2, 1	0, 0
Ballet	0, 0	1, 2

Battle of Sexes

If the population has roughly 2/3 *L1*s and 1/3 *L2*s, a level-*k* model's outcomes resemble a mixed-strategy equilibrium's: Kahneman's "magic".

If, instead, the population has roughly 1/2 *L1*s and 1/2 *L2*s, players coordinate on (Fights, Fights) or (Ballet, Ballet), each with probability roughly ¼, and thus do better than in the best symmetric (mixed-strategy) equilibrium, where they each coordinate with probability 2/9.

This is possible, without communication or observation, because the heterogeneity of level-*k* thinking allows more sophisticated players to mentally simulate the decisions of less sophisticated players, often correctly, and to accommodate them as Stackelberg followers would.

Level-*k* thinking in Stag Hunt without communication

	Stag	Hare
Stag	9, 9	0, 8
Hare	8, 0	7, 7

Stag Hunt

When *L0* is random, *L1* plays Hare, and so all higher levels do so too:

A level-*k* model predicts the same outcome as risk-dominant equilibrium.

Equilibrium and refinements play no role in players' thinking.

Coordination, when it occurs, is an accidental (though predictable) by-product of the mixture of players' levels.

Modeling communication

I focus on two-person games whose players send one or more rounds of one-sided or two-sided messages before playing an *underlying game*.

The messages may concern either players' *private information* (Crawford and Sobel *Ecma* 1982, Green and Stokey *JET* 2007 [1980-81]) or their *intentions* about their choices in the underlying game (Kalai and Samet *IJGT* 1985, Farrell *Rand* 1987, Rabin *JET* 1994).

The messages are assumed to be *cheap talk*, in that they have no direct effect on payoffs (so messages about intentions must be nonbinding).

In theory they can take any form, but if they are in a language that makes lying a meaningful concept, cheap talk implies that lying has no direct cost (a useful limiting case even if people are averse to lying (Ellingsen and Johannesson *EJ* 2004 and Abeler, Nosenzo, and Raymond 2016)).

Cheap talk messages are analogous to workers' education levels in Spence's *QJE* 1973 model of job market signaling; but in Spence's model education has direct costs that vary with ability, and preferences differ enough that such costs are essential for informative signaling.

By contrast, equilibrium analyses of cheap talk focus on settings where preferences are close enough that cheap talk messages can influence outcomes even though they have no direct effect on payoffs.

The operative meaning of a cheap talk message is "I like what I expect you to do when I say this better than anything else I could get you to do."

It is obvious that such a message can be influential in equilibrium when players' preferences are perfectly aligned (Schelling 1960, Lewis 1969).

It is also obvious that cheap talk messages must be uninformative, and ignored, in equilibrium when players' preferences are opposed (Crawford and Sobel *Ecma* 1982, Theorem 1); but (surprisingly) they can then still be influential in a level-*k* model (Crawford *AER* 2003, Proposition 1).

When players' preferences are neither opposed nor perfectly aligned, cheap talk messages may or may not be influential in either model.

Aside:

With cheap talk, equilibrium predictions are ambiguous for several reasons. I will focus on “sensible” equilibria, which rule out ambiguities that are behaviorally unimportant (but leave in other ambiguities):

- Time-sequencing of messages and actions and possibly, private information make it necessary to rule out equilibria that are not *subgame-perfect* or *sequential* or *perfect Bayesian*

- Standard payoffs-based refinements cannot determine the meanings of cheap-talk messages, so I will focus on equilibria in which messages’ literal meanings are understood (even if not believed)

  (Similarly, level-*k* models anchor beliefs on *L0*s that respect meanings)

- There is always a “babbling” equilibrium in which messages are uninformative, so I will focus on informative equilibria when they exist

Equilibrium in Battle of the Sexes with communication

	Fights	Ballet
Fights	2, 1	0, 0
Ballet	0, 0	1, 2

Battle of Sexes

- In a sensible equilibrium in Battle of the Sexes with one round of one-sided communication, the sender sends a message of intent to play for her/his favorite equilibrium in the underlying game and the receiver plays her/his part of that equilibrium (Farrell *EL* 1988 gets a stronger result, assuming rationalizability with behavioral restrictions)
- With one round of *two*-sided communication, the rate of coordination is higher than without communication, but below one (Farrell *Rand* 1987)
- With multiple rounds of two-sided communication, the rate of coordination increases with the number of rounds, but remains bounded below one (Farrell *Rand* 1987, Rabin *JET* 1994)

Equilibrium in Stag Hunt with communication

	Stag	Hare
Stag	9, 9	0, 8
Hare	8, 0	7, 7

Stag Hunt

- In Stag Hunt with one round of one-sided communication, there are two sensible equilibria, one in which the sender sends and plays Stag, and another in which the sender sends and plays Hare

- With one round of two-sided communication, there are again two sensible equilibria, one in which both players send and play Stag, and another in which both send and play Hare

- With many rounds of two-sided communication there are again multiple sensible equilibria, which do not improve upon those with one round

Aside:

Language-dependent refinements have been proposed to get around the multiplicity of sensible equilibria (Farrell *EL* 1988, *GEB* 1993; Myerson *JET* 1989; Rabin *JET* 1990; Farrell and Rabin *JEP* 1996).

- A *self-signaling* message regarding private information or intentions is one that a sender wants a receiver to believe if and only if it's true
- A *self-committing* message regarding intentions is one that, if believed, creates an incentive for the sender to do as s/he said (that is, makes the sent intention part of an equilibrium in the underlying game)

Aumann 1990 notes that a one-sided message of intent to play "Stag" is self-committing but not self-signaling, and argues that such a message can thus convey no information and cannot change the outcome (Farrell *EL* 1988 and Rabin *JET* 1994 disagree; see also Crawford *JEP* 2016).

Farrell *EL* 1988 and Rabin *JET* 1994 bound the effects of communication while relaxing equilibrium to rationalizability with behavioral restrictions on how players use language, and argue that communication may yield the efficient both-Stag equilibrium outcome in Stag Hunt.

Evidence on Battle of the Sexes and Stag Hunt with communication

When the underlying game requires symmetry-breaking, as in Battle of the Sexes, experimental results are fairly close to sensible equilibrium predictions, one-sided is better than two-sided communication, and more rounds of two-sided communication are better than less (Cooper, DeJong, Forsythe, and Ross *Rand* 1989; Costa-Gomes *JET* 2002).

However, it seems likely that people use communication, particularly if natural-language communication is possible, to solve such problems much more efficiently in the field than in Farrell's *Rand* 1987 or Rabin's *JET* 1994 models, where even unlimited communication does not ensure fully efficient coordination (Ellingsen and Östling *AER* 2010).

When the underlying game requires reassurance, as in Stag Hunt, experimental results are fairly *consistent* with sensible (but equivocal) equilibrium predictions; two-sided might be better than one-sided communication; but more rounds are not better than less (Cooper, DeJong, Forsythe, and Ross *QJE* 1992; Charness *GEB* 2000; Duffy and Feltovich *GEB* 2002; Dugar and Shahriar 2016; Ellingsen, Östling, and Wengström 2016; but see Clark, Kay, and Sefton *IJGT* 2001).

Level-*k* thinking with communication

In games with communication via a common language, people's thinking assigns a leading role to the literal meanings of messages; incentives also play a leading role: the question is how to combine them.

Taking senders' *L0*s to be uniform random, without regard to the truth, then seems unnatural—and trivializes a level-*k* model of communication.

Instead I take level-*k* beliefs to be anchored in *L0*s for senders' roles that favor the truth (Crawford *AER* 2003; Ellingsen and Östling *AER* 2010 ("EÖ"); Ellingsen, Östling, and Wengström 2016; Crawford *RiE* 2017).

Unlike Crawford *AER* 2003, who considered only one-sided messages and took receivers to be credulous, I follow the rest in assuming that *L0* receivers randomize uniformly, independent of received messages.

The rest of the model is specified by iterating best responses as before.

The resulting model is well supported by experimental evidence (Cai and Wang *GEB* 2006, Wang et al. *AER* 2010, Dugar and Shahriar 2016).

Level-*k* thinking in Battle of the Sexes with communication

	Fights	Ballet
Fights	2, 1	0, 0
Ballet	0, 0	1, 2

Battle of Sexes

- With one round of one-sided communication, the sender again sends a message of intent to play for her/his favorite pure-strategy equilibrium, and the receiver plays her/his part of that equilibrium (EÖ)

- With one round of two-sided communication, the rate of efficient coordination is higher than without communication, and likely to be higher than the equilibrium rate unless preferences are very close, but is below one (EÖ, Crawford *RiE* 2017)

- With multiple rounds of two-sided communication, the rate of efficient coordination increases with the number of rounds, is likely to be higher than the equilibrium rate for any number of rounds unless preferences are very close, but remains bounded below one (Crawford *RiE* 2017)

Level-*k* thinking in Stag Hunt with communication

	Stag	Hare
Stag	9 9	8 0
Hare	0 8	7 7

Stag Hunt

- With one round of *one*-sided communication, *L1* senders send and play their risk-dominant strategy, Hare. *L1* receivers play Hare when they hear Hare and Stag when they hear Stag. *L2* and higher senders send and play Stag, expecting to be believed; and *L2* and higher receivers, expecting to hear and play Hare, play Stag when they hear Stag, which is self-committing for them. Thus the level-*k* model predicts that if both players are *L2* or higher, they will coordinate on both-Stag
- With one round of *two*-sided communication, *L2* and higher players again send and/or play Stag, as do *L1* players with positive probability in EÖ's model, making two-sided communication more effective
- Multiple rounds of two-sided communication are no better than one

Trickery

"Any fool can tell the truth, but it requires some sense to know how to lie well." —*The Notebooks of Samuel Butler*, 1912

Consider (next two slides) two stories that provide examples of Trickery:

- D-Day (Crawford *AER* 2003)
- Huarangdao (from Luo Guanzhong's historical novel *Three Kingdoms*, published in the 14th century but set in the 2nd or 3rd century; http://en.wikipedia.org/wiki/Battle_of_Red_Cliffs)

Both are "outguessing" games made nontrivial by payoff asymmetries. Both also have one-sided communication, as explained below.

With or without communication, these games resemble many situations in business, politics, security, or war. A familiar example is Myerson's Ware Case, where an incumbent firm, to block entry by a firm with a competing product, must outguess and match the entrant's design (see http://www.kellogg.northwestern.edu/faculty/weber/decs-452/Ware.htm).

		Germans	
		Defend Calais	Defend Normandy
Allies	Invade at Calais	1 -1	-2 2
	Invade at Normandy	-1 1	1 -1

D-Day

Allies decide where to invade Europe; Germans try to outguess them.

Invading an undefended Calais is better for Allies than invading an undefended Normandy (2 > 1).

But defending an unattacked Normandy is worse for Germans than defending an unattacked Calais (- 2 < - 1).

How should Allies and Germans respond to this payoff asymmetry?

(The payoffs are plainly unrealistic, but the asymmetry is the point.)

Kongming

Cao Cao		Main Road	Huarong
	Main Road	-1, 3	1, 0
	Huarong	0, 1	-2, 2

Huarongdao

Fleeing General Cao Cao, trying to avoid capture, chooses between Main Road and the awful Huarong Road; pursuing General Kongming tries to outguess him.

The Main Road is better for both Cao Cao and Kongming than the Huarong Road, other things equal.

How should they respond to this payoff asymmetry?

Equilibrium in D-Day or Huarongdao without communication

I focus on D-Day, but analogous conclusions hold for Huarongdao.

		Germans **Defend Calais (*q*)**	**Germans** **Defend Normandy**
Allies	**Invade at Calais (*p*)**	1 -1	-2 2
Allies	**Invade at Normandy**	-1 1	1 -1

D-Day

D-Day plainly has no pure-strategy equilibrium.

D-Day has a unique mixed-strategy equilibrium, in which, if p is the probability with which Allies invade Calais and q the probability with which Germans defend Calais, then $p = 2/5$ ($1p - 1(1-p) = -2p + 1(1-p)$) and $q = 3/5$ (because $-1q + 2(1-q) = 1q - 1(1-q)$).

Huarongdao has a similar equilibrium in mixed strategies, also unique.

Evidence on games like D-Day and Huarongdao without communication

In games like D-Day and Huarongdao, the comparative statics of mixed-strategy equilibrium with respect to changes in payoffs go against decision-theoretic intuition for one player (Allies in D-Day, because $p = 2/5 < ½$); but with decision-theoretic intuition for the other ($q = 3/5 > ½$).

This is a (perhaps unwanted) by-product of the fixed-point logic of equilibrium (Von Neumann and Morgenstern 1944, pp. 175-176; Crawford & Smallwood *ThD* 1984).

By contrast, experiments suggest that people's thinking seldom follows equilibrium fixed-point logic, and that subjects' responses to changes in payoffs favor decision-theoretic intuition in both player roles (Shachat *JET* 2002; Rosenthal, Shachat, and Walker *IJGT* 2003).

Level-*k* thinking in D-Day and Huarongdao without communication

● In D-Day, *L1* Allies attack Calais, while *L1* Germans defend Calais

● *L2* Allies attack Normandy, while *L2* Germans defend Calais

● *L3* Allies attack Calais, while *L3* Germans defend Normandy

And so on.

In Huarongdao a level-*k* analysis is similar.

The analysis is entirely mechanical, with the level-*k* model approximately "purifying" something qualitatively like the mixed-strategy equilibrium.

Players face strategic uncertainty not because their partners randomize, but because they are uncertain about their partners' levels of thinking.

Now consider what happens when each player role is filled from a population with not only level-*k* players but also *Sophisticated* players, defined as rational and fully informed about the game, including the population frequencies of different kinds of level-*k* players.

Sophisticated players are assumed to play an equilibrium in a "reduced game" derived by plugging in the distributions of level-*k* players' choices.

(Level-*k* players' choices are determined independently of each other's and *Sophisticated* players' choices, and can be treated as exogenous.)

● If *Sophisticated* players' frequencies are high enough in both roles, their equilibrium strategies counteract level-*k* players' deviations from equilibrium and inadvertently protect them from exploitation; and the outcome mimics the mixed-strategy equilibrium of the original game

● Otherwise, *Sophisticated* players' equilibrium strategies partly counteract level-*k* players' deviations from equilibrium, and partly protect them from exploitation; and the outcome deviates from equilibrium in predictable ways

Equilibrium in D-Day and Huarongdao with communication

In each story the game has preplay communication via one-sided messages with understood meanings that are approximately cheap talk.

The D-Day invasion was preceded by Operation Fortitude South (http://en.wikipedia.org/wiki/Operation_Fortitude), in which the Allies faked invasion preparations in the Thames Estuary (meaning the invasion would be at Calais, the obvious alternative to Normandy).

An inflatable "tank" from Operation Fortitude South

Kongming waited in ambush on the Huarong Road and set campfires there.

In these approximately zero-sum two-person games, equilibrium allows no role for communication via cheap-talk: In any equilibrium, the sender sends an uninformative message (“babbling”) and the receiver ignores it.

For, if the sender instead sent an informative message, the receiver would benefit by making her/his choice respond to the message.

But when the underlying game is zero-sum, any such response would hurt the sender, who would therefore do better by babbling.

Equilibrium thus renders cheap-talk communication ineffective, and the outcome of the underlying game is the same as without communication.

Intuition and folk game theory "evidence" on games like D-Day and Huarongdao with communication

By contrast with the equilibrium analysis, intuition suggests that in games like D-Day and Huarongdao, with communication about intentions:

- Senders' messages and their choices in the underlying game are parts of an integrated strategy, chosen to deceive and exploit receivers

- Players' choices in the underlying game differ systematically from those they would have chosen without communication

and, as in both stories,

- Senders' attempts to deceive often succeed, but they often "win" in the less beneficial of the possible ways to win

"Good intelligence work…was gradual and rested on a kind of gentleness."
—Smiley quoting Control, in John Le Carre, *Tinker, Tailor, Soldier, Spy*, 1974

The difference between equilibrium, intuition, and evidence highlights two puzzles:

- Why did the receivers in D-Day and Huarongdao allow themselves to be fooled by nearly costless, easily faked, messages from *enemies*?

- Why didn't the senders in D-Day and Huarongdao, who apparently thought they had a good chance to fool the receiver, try to do so in ways that allowed them to win in the *more* beneficial of the two ways?

A level-*k* analysis can explain these puzzles, and captures intuitions about deception that are meaningless in an equilibrium analysis:

- The Allies' message was literally a lie, which fooled the Germans because they "believed" it—perhaps inverting it one (or some odd number) too many times

- Kongming's message was truthful, but it fooled Cao Cao because he inverted it, again one (or some odd number) too many times

When using stories as data, an omniscient narrator sometimes reveals players' cognition; from *Three Kingdoms*:

- Kongming: "Have you forgotten the tactic of 'letting weak points look weak and strong points look strong'?" (*L2* or higher even level)

- Cao Cao: "Don't you know what the military texts say? 'A show of force is best where you are weak. Where strong, feign weakness.'" (*L1* or higher odd level)

A level-*k* analysis can explain these puzzles, and captures intuitions about deception that are meaningless in an equilibrium analysis:

- The Allies' message was literally a lie, which fooled the Germans because they "believed" it—perhaps inverting it one (or some odd number) too many times

- Kongming's message was truthful, but it fooled Cao Cao because he inverted it, again one (or some odd number) too many times

When using stories as data, an omniscient narrator sometimes reveals players' cognition; from *Three Kingdoms*:

- Kongming: "Have you forgotten the tactic of 'letting weak points look weak and strong points look strong'?" (*L2* or higher even level)

- Cao Cao: "Don't you know what the military texts say? 'A show of force is best where you are weak. Where strong, feign weakness.'" (*L1* or higher odd level)

Cao Cao must have bought a used, out-of-date edition….

Level-*k* thinking in D-Day and Huarongdao with communication

Recall that in a level-*k* model, players follow rules of thumb that:

● anchor their beliefs in a naïve model of others' responses called *L0* and

● adjust their beliefs via a small, heterogeneous number (*k*) of iterated best responses: *L1* best responds to *L0*, *L2* to *L1*, and so on

I take *L0* senders to be truth-tellers and *L0* receivers to be credulous, as in Crawford *AER* 2003.

● Then *L1* (or higher odd-level) senders lie; and *L2* (or higher even-level) senders tell the truth

● *L1* (or higher odd-level) receivers invert the sender's messages; and *L2* (or higher even-level) receivers "believe" the sender's messages

With one-sided communication, this yields outcomes close to recent models that allow two-sided communication (Ellingsen and Östling *AER* 2010; Ellingsen, Östling, and Wengström 2016; Crawford *RiE* 2017).

The resulting model is well supported by experimental evidence (Cai and Wang *GEB* 2006; Kawagoe and Tazikawa *GEB* 2009; Wang et al. *AER* 2010; Dugar and Shahriar 2016).

With only level-*k* players, the outcome is mechanically determined by the frequencies of senders who lie, or tell the truth; and by the frequencies of receivers who can be fooled by telling them the truth, or by lying to them.

As in the games without communication, I enrich the analysis by assuming the player populations also contain *Sophisticated* players, who are rational and know the game, including level-*k* players' frequencies.

The goal is to learn if *Sophisticated* senders can "deceive" *Sophisticated* receivers when there are positive frequencies of unsophisticated players.

As before, the level-*k* players' strategies can be treated as exogenous, and the *Sophisticated* players play an equilibrium in a reduced game derived by plugging in the distributions of level-*k* players' strategies.

If there were only *Sophisticated* players, the reduced game would be approximately zero-sum, the sender's messages would be cheap talk, and the game would have symmetric information.

The possibility of level-*k* players makes the reduced game between *Sophisticated* players very different in structure from the original game:

- the reduced game is no longer zero-sum
- its messages are no longer cheap talk
- the reduced game now has asymmetric information, with the sender's message, ostensibly about her/his intentions, read by a *Sophisticated* receiver as a signal of what the sender's behavioral rule is

These differences are what gives the model the ability to explain *Sophisticated* senders' ability to deceive *Sophisticated* receivers.

In the reduced game, level-*k* senders:

- expect to fool the more common kind of receiver (whether by lying or telling the truth)

- send a message they expect to make the more common kind of receiver think they will choose the strategy in the underlying game that would be *less* advantageous if it won (Normandy or Huarong)

- but instead choose the strategy that would be *more* advantageous if it won (Calais or Main Road)

Sophisticated senders play their part of the reduced game's equilibrium, taking into account level-*k* and *Sophisticated* receivers' frequencies and reactions, with the latter drawing inferences from senders' messages.

As without communication, there are two leading cases (oversimplifying):

- When *Sophisticated* players are common in both roles, the reduced game has a mixed-strategy equilibrium equivalent to the underlying game without communication. *Sophisticated* players' equilibrium mixed strategies offset level-*k* players' deviations from equilibrium, and both have equal expected payoffs in each role (so no selection pressure)

- When *Sophisticated* players are rare in both roles (empirically more likely), the game is dominance-solvable in three rounds, with an essentially unique equilibrium in pure strategies, in which:

  (i) *Sophisticated* senders send the message that deceives the most common kind of level-*k* receiver and choose the underlying-game action that is *less* advantageous if it wins (Normandy in D-Day)

  (ii) *Sophisticated* receivers choose the underlying-game action that *loses* against *Sophisticated* senders (Calais in D-Day)

A *Sophisticated* receiver can be "deceived" in this way because s/he thinks the sender is most likely level-*k*, sending a message designed to fool the most common kind of level-*k* receiver, in preparation for choosing the action that is more advantageous if it wins.

- In either case, a *Sophisticated* sender's message and choice in the underlying game are part of an integrated strategy, chosen to deceive and exploit receivers

- In the latter case, *Sophisticated* senders choose the underlying-game action that is l*ess* advantageous if it wins, systematically with higher probability than they would without communication

- In the latter case, there is no pure-strategy equilibrium in which a *Sophisticated* receiver can be deceived in a way that allows a *Sophisticated* sender to win in the *more* advantageous way

  For then any deviation from *Sophisticated*'s equilibrium message would "prove" to a *Sophisticated* receiver that the sender is level-*k*, making it optimal for a *Sophisticated* receiver to defend where s/he is most vulnerable and suboptimal for a *Sophisticated* sender to attack there.

  In that (limited) sense the level-*k* model explains Control's "gentleness": why the senders in D-Day and Huarongdao didn't try to deceive their receivers in ways that allowed them to win in the *more* beneficial way.

Puffery

Finally, consider communication of private information in games with partial common interests, as in Crawford and Sobel *Ecma* 1982 and Green and Stokey *JET* 2007 [1980-81].

Specifically, consider this framing example from Wang et al. *AER* 2010:

> During the tech-stock bubble, Wall Street security analysts were alleged to inflate recommendations about the future earnings prospects of firms in order to win investment banking relationships with those firms. Specifically, analysts of Merrill Lynch used a five-point rating system (1 = Buy to 5 = Sell) to predict how the stock would perform. They usually gave two 1–5 ratings for short run (0–12 months) and long run (more than 12 months) performance separately.
>
> Henry Blodget, Merrill Lynch's famously optimistic analyst, "did not rate any Internet stock a 4 or 5" during the bubble period (1999 to 2001). In one case, the online direct marketing firm LifeMinders, Inc. (LFMN), Blodget first reported a rating of 2-1 (short run "accumulate"—long run "buy") when Merrill Lynch was pursuing an investment banking relationship with LFMN. Then, the stock price gradually fell from $22.69 to the $3–$5 range. While publicly maintaining his initial 2-1 rating, Blodget privately e-mailed fellow analysts that "LFMN is at $4. I can't believe what a POS [piece of shit] that thing is." He was later banned from the security industry for life and fined millions of dollars.

"Blodget" is a Crawford & Sobel-style sender-receiver game between the analyst and an investor (ignoring the firm whose stock is being touted):

- The analyst has private information about the firm's stock's prospects
- The analyst's recommendation to the investor is cheap talk
- Based on the analyst's recommendation, the investor makes a decision that affects the analyst's welfare as well as her/his own welfare
- The analyst's and investor's preferences are similar: both want the investor to sell on bad news and buy/hold on good, other things equal
- But there is a wedge between their preferences, in that the analyst's desire to preserve its relationship with the firm makes her/him want the investor to buy/hold the stock more than a well-informed investor would

Equilibrium in Blodget

Under assumptions on preferences that generalize Blodget (except for continuous state and action spaces), Crawford and Sobel's Theorem 1 characterizes the possible equilibrium relationships between the sender's information and the receiver's choice, independent of messaging details.

There is always a babbling equilibrium, in which the sender's recommendations are uninformative and the receiver ignores them.

When the sender's and receiver's preferences are identical, there is also an equilibrium with perfect information transmission.

But when the sender's and receiver's preferences differ (even a little, in the continuous version) there is no equilibrium with perfect transmission:

For, in such an equilibrium the receiver would follow the sender's recommendations exactly, making it optimal for the sender to bias them, hence suboptimal for the receiver to follow them, a contradiction.

Finally, when sender's and receiver's preferences differ, but not by too much, there is also a range of informative but noisy "partition" equilibria, in which the sender is intentionally vague, telling the receiver, in effect, which of a well-determined set of contiguous groups the state falls into.

The vagueness is a by-product of the sender's incentive to lie; it arises even though in equilibrium the "rationality" of expectations frustrates the sender's attempt to bias the receiver's decisions, and even makes "lying" meaningless, in that the receiver anticipates and sees through any lie.

The vagueness reduces the sender's and receiver's ex ante expected payoffs below their levels if the sender could commit to telling the truth.

Crawford and Sobel also proved an intuitive comparative statics result:

The amount of information transmitted in the "most informative" equilibrium (roughly, the one that has the most elements in its partition; so not in Blackwell's sense), measured by receiver's expected payoff, increases with the closeness of sender's and receiver's preferences.

Evidence on Blodget

There are now several laboratory experiments on such games (Cai and Wang *GEB* 2006; Kawagoe and Tazikawa *GEB* 2009; Wang et al. *AER* 2010; see also Crawford et al. *JEL* 2013 and Blume, Lai, and Lim 2017).

Wang et al.'s design implements discrete Blodget games:

- A sender observes the state S = 1, 2, 3, 4, or 5, sends a message M = 1, 2, 3, 4, or 5
- A receiver observes the message M and chooses an action A = 1, 2, 3, 4, or 5, which determines the welfare of both
- Both have single-peaked preferences, with the receiver's ideal outcome A = S and the sender's A = S + b (ignoring boundaries)
- The design varies the difference in sender's and receiver's preferences across treatments: b = 0, 1, or 2

Wang et al. focused on the most informative equilibrium in their games, partly to test Crawford and Sobel's comparative statics result.

Wang et al. eyetracked senders' searches for information about payoffs.

In Wang et al.'s Figures 1-3 (next two slides), a circle's size shows senders' message frequencies in the various states, and a circle's darkness and the numbers inside it show receivers' action frequencies.

- In Figure 1, sender's and receiver's preferences are identical (b = 0); the most informative equilibrium has perfect truth-telling and credulity

- In Figure 2 sender's and receiver's preferences differ, but not too much (b = 1); the most informative equilibrium has sender sending message 1 in state 1 and receiver responding with action 1, and sender otherwise sending a message that is uninformative about whether the state is 2, 3, 4, or 5 and receiver responding with action 3 or 4 (tied)

- In Figure 3 sender's and receiver's preferences differ widely (b = 2); there is only a babbling equilibrium, in which receiver ignores sender's uninformative message and chooses action 3

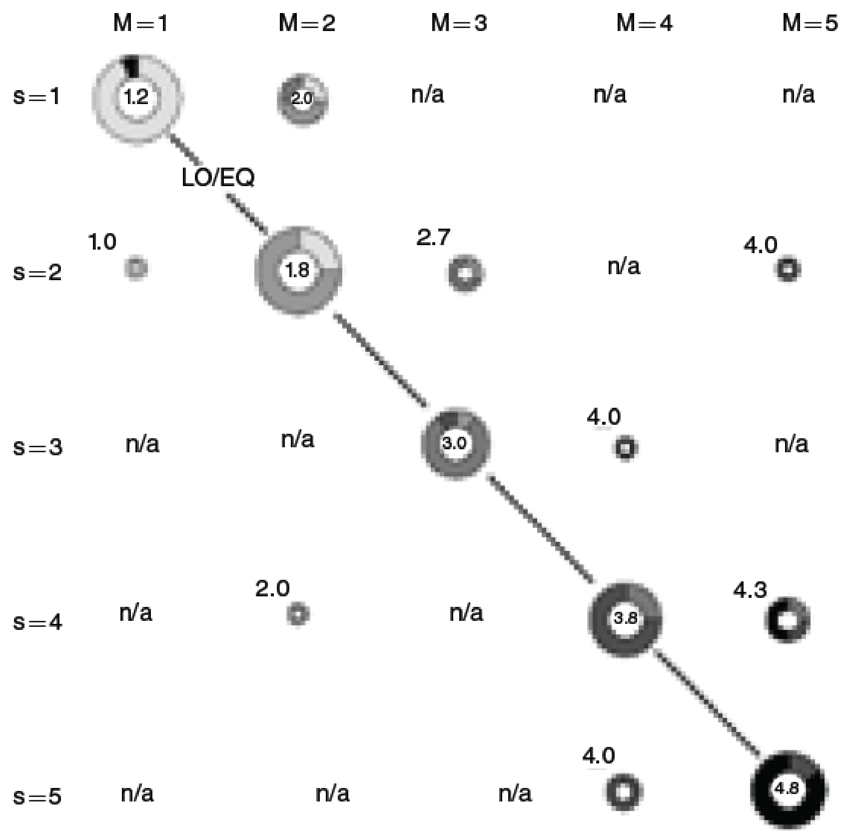

FIGURE 1. RAW DATA PIE CHARTS ($b = 0$)
(HIDDEN BIAS–STRANGER)

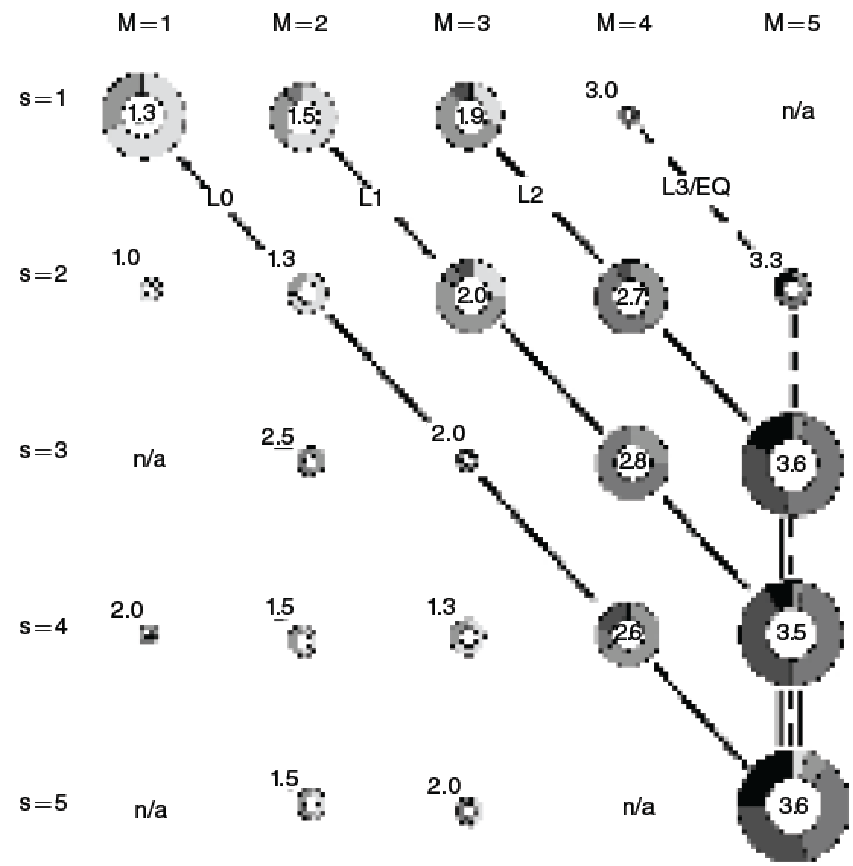

FIGURE 2. RAW DATA PIE CHART ($b = 1$)
(HIDDEN BIAS–STRANGER)

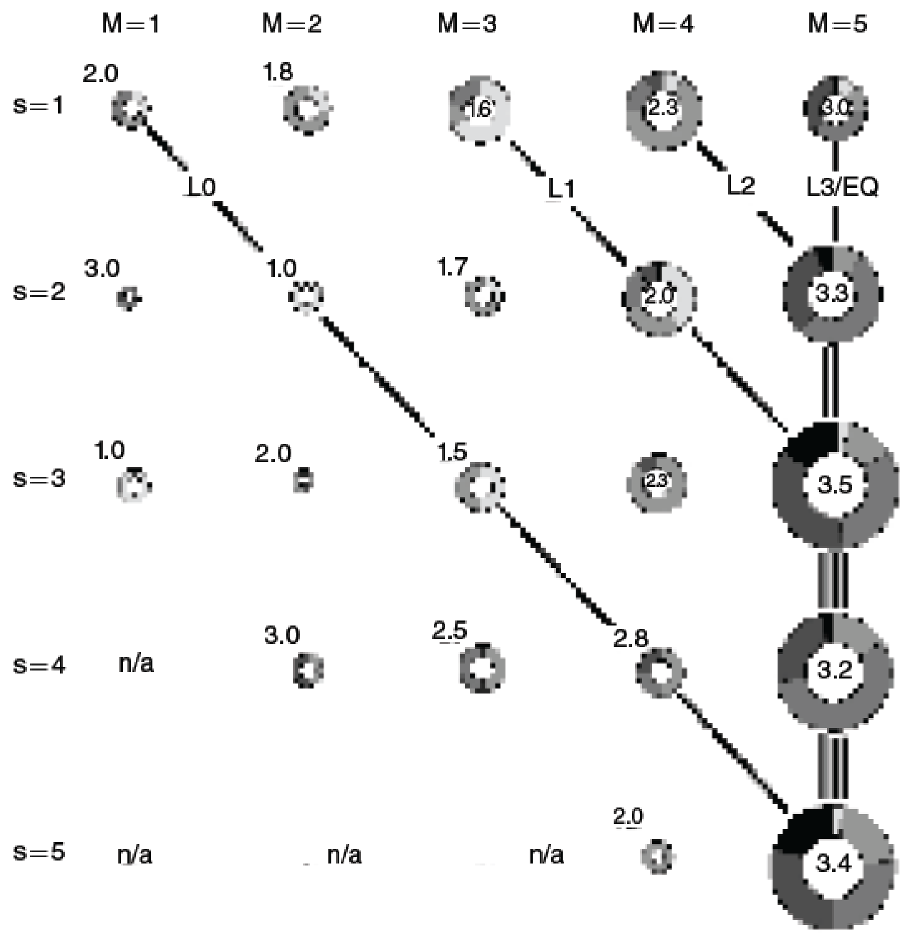

Figure 3. Raw Data Pie Chart ($b = 2$)
(Hidden Bias–Stranger)

In Figure 1 no deviations from equilibrium are expected, or observed.

But Figures 2-3 reveal systematic deviations from the most informative (or any) equilibrium:

● Senders "overcommunicate", in that messages are more truthful than in equilibrium (strong positive correlation between senders' messages and the state), in contrast to the equilibrium prediction that there is no systematic deception (note that that prediction depends on known sender's preferences, justified here; but compare Sobel *REStud* 1985)

● Receivers are "overcredulous", in that they respond more to senders' messages than is a best response to senders' overcommunication and, a fortiori, more than in equilibrium (average action usually > state)

● Senders exaggerate their messages in the direction (above diagonal) that makes credulous receivers choose actions the sender would prefer

● Despite systematic deviations from equilibrium, the results qualitatively affirm Crawford and Sobel's comparative statics result (communication is more informative, the closer are sender's and receiver's preferences)

Level-*k* Thinking in Blodget

Cai and Wang *GEB* 2006; Kartik et al. *JET* 2007; Kawagoe and Tazikawa *GEB* 2009; and Wang et al. *AER* 2010 adapted Crawford's *AER* 2003 level-*k* model with *L0*s anchored in truthtelling/credulity from communication of intentions to communication of private information.

Recall that sender observes state S = 1, 2, 3, 4, 5 and sends message M = 1, 2, 3, 4, 5; receiver observes M and chooses action A = 1, 2, 3, 4, 5.

Receiver's ideal outcome is A = S and sender's is A = S + b. Then:

- *L1* senders best respond to *L0* receivers, "puffing" messages by b: M = S + b, which would yield a sender's ideal action, given *L1* senders' belief that receivers are credulous and will choose S + b if M = S + b
- *L1* receivers (in Wang et al.'s numbering convention, which differs from Crawford's) best respond to *L1* senders, "de-puffing" messages by b and choosing A = M – b, which would yield a receiver's ideal action, given *L1* receivers' beliefs that *L1* senders will set M = S + b
- Similarly, *L2* senders puff by 2d, *L2* receivers de-puff by 2b, and so on

The level-*k* model is not separated from equilibrium when sender's and receiver's preferences are identical (Figure 1), and both then fit very well.

But when sender's and receiver's preferences differ, so the models are separated (Figures 2-3), the level-*k* model fits much better.

Some patterns in the data are also suggestive of lying costs, but the data strongly favor a level-*k* explanation.

Overall, the anchoring of level-*k* beliefs in a credulous or truthful *L0* and the finiteness of adjustment allow the level-*k* model to gracefully explain senders' puffery and overcommunication and receivers' overcredulity.

The sensitivity of levels' degrees of puffery and de-puffery to the difference in senders' and receivers' preferences also yields a simple explanation of why Crawford and Sobel's equilibrium-based comparative statics result is robust to large, systematic deviations from equilibrium.

Conclusion

Evidence on communication from games where it serves a variety of purposes—from Rendezvous and Reassurance to Trickery and Puffery—suggests that nonequilibrium models based on level-*k* thinking can more accurately describe how people use communication, and explain some empirical puzzles that equilibrium leaves unresolved.

Recall that a level-*k* model does not suggest that equilibrium predictions are always wrong: it predicts that deviations will sometimes occur, and also which settings evoke them and which forms they are likely to take.

But the evidence also shows that level-*k* models still fall short of what's needed for a fully reliable model of how people use communication.

For concreteness, consider two leading examples (Crawford *JEP* 2016):

- bringing about and maintaining efficient cooperation and coordination in long-term relationships
- bringing about efficient coordination in Reassurance games

Long-term relationships

Imagine you are in a long-term relationship, governed by an implicit agreement you believe both you and your partner understand. Then your partner doesn't do what you thought was agreed.

What do you do? Some intuitions about what you might try in practice:

Without communication, all you can do is signal your displeasure via tit-for-tat, hoping your partner will understand and return to cooperation.

If the ideal agreement is as obvious as in our models, such a tactic might work (van Huyck, Battalio, and Beil's *AER* 1990 fixed-pair treatments).

But if good agreements require non-obvious choices or surplus-sharing, restoring cooperation is all but hopeless without communication, even though most game-theoretic analyses make tacit collusion a perfect substitute for explicit collusion.

That's why antitrust law bothers to prohibit firms from communicating (Genesove & Mullin *AER* 2001, Andersson and Wengström *SJE* 2007).

With communication that is abstract, via a pre-set list of messages with understood meanings, you might be able to restore cooperation if good agreements are simple enough and don't require complex adjustments.

With natural-language messages, there is hope to restore cooperation:

- If only a single, one-sided message is possible, a contingent promise to return to cooperation if your partner does so might work

- Even if identifying a good agreement is complex, restoring cooperation may be possible via a natural-language dialogue, e.g. starting like this:

  "I value our relationship, and I believe you are trying to cooperate. But what you just did was inconsistent with what I thought we had agreed. [Elaborates….] Please help me to understand your thinking"

A growing body of evidence suggests that natural-language dialogues are far more effective than structured abstract communication (Valley, Thompson, Gibbons, and Bazerman *GEB* 2002; McGinn, Thompson, and Bazerman *JBDM* 2003; Charness and Dufwenberg *Ecma* 2006, *EL* 2010; Cooper and Kühn *AEJ Micro* 2014; Dugar and Shahriar 2016).

One-shot Stag Hunt games

Now imagine you are about to play a Stag Hunt game, only once.

	Stag	Hare
Stag	9, 9	0, 8
Hare	8, 0	7, 7

Stag Hunt

With this small an advantage for both-Stag, without communication most people would respond to the greater riskiness of Stag by playing Hare.

If only a single, one-sided, abstract message of intent is possible, saying Stag is reasonably likely to yield coordination on both-Stag (Cooper, DeJong, Forsythe, and Ross *QJE* 1992; Charness *GEB* 2000; Duffy and Feltovich *GEB* 2002; but see Clark, Kay, and Sefton *IJGT* 2001; Dugar and Shahriar 2016; and Ellingsen, Östling, and Wengström 2016).

Two-sided abstract messages of intent, with dialogue limited or not, are likely to do as well as one-sided messages, but not significantly better.

	Stag	Hare
Stag	9 9	8 0
Hare	0 8	7 7

Stag Hunt

With natural-language messages, dialogue or not, coordination on both-Stag is far more likely (Dugar and Shahriar 2016), e.g. starting like this:

“I can see, as I’m sure you can, that the best possible outcome would be for both of us to play Stag. I realize Stag is risky for you, as it is for me. But despite the risk, I think Stag’s higher potential payoff makes it a better bet. I therefore plan to play Stag, and I hope you will too.”

Equilibrium in long-term relationships or one-shot Stag Hunt games

Most analyses of long-term relationships assume that players are focused on a subgame-perfect equilibrium of the repeated game that describes the entire relationship, with the main goal to characterize the "Folk Theorem" set of outcomes consistent with some such equilibrium.

● the idiosyncrasies of relationships and the fact that we don't get to practice them makes equilibrium's learning justification implausible

● the set of equilibria is enormous, making equilibrium's thinking justification implausible, so there is much strategic uncertainty

● such analyses seldom consider robustness to strategic uncertainty, and focus on equilibria that are "brittle" (but see Porter *JET* 1983, van Damme *JET* 1989, or Friedman and Samuelson 1994)

● in an equilibrium of a complete-information game, players have nothing to communicate, precluding any substantive role for communication, abstract or natural-language, despite its powerful effects in practice

Most analyses of games like Stag Hunt, with or without communication, also assume that players are focused on a particular equilibrium.

- such analyses rely on equilibrium logic despite strategic uncertainty

  For instance, Aumann 1990 argues that a message of intent to play Stag is uninformative, because while self-committing, in that if believed it creates an incentive for a sender to do as s/he said, it is not self-signaling, in that a sender wants a receiver to believe it iff it is true

  Yet with the strategic uncertainty of Stag Hunt, few people will assume that even a message that is not self-signaling will not influence choices

- such analyses implicitly limit players to a fixed list of messages of intent when they would plainly benefit from a more nuanced discussion

- such analyses again preclude any substantive role for communication, abstract or natural-language, despite its powerful effects in practice, because in an equilibrium players have nothing to communicate

  Notable partial exceptions are Farrell EL 1988 and Rabin *JET* 1994, who assume rationalizability with behavioral restrictions and get strong results on the effectiveness of communication in games like Stag Hunt (see also Myerson *Ecma* 1983, *JET* 1989)

Future work

In each case there is a large gap between intuitions about what happens in practice, evidence, and theory, level-*k* as well as equilibrium-based.

Three lines of research (experimental, empirical, and theoretical) seem especially likely to be helpful:

- work on strategic thinking and behavior in games without communication, particularly those with nontrivial sequential structures

  Recent examples include Dal Bó and Fréchette *AER* 2011; Blonski, Ockenfels, and Spagnolo *AEJ Micro* 2011; Ho and Su *MS* 2012; Kawagoe and Takizawa *JEBO* 2012; and Breitmoser *AER* 2015

- work explaining why and how communication (abstract or natural-language) allows people to achieve outcomes better than those (behaviorally, not theoretically) attainable without communication

  Examples following Myerson *Ecma* 1983, *JET* 1989 and Forges *Ecma* 1986 include Weber and Camerer *MS* 2003; Houser and Xiao *EE* 2010; Andersson and Wengström *JEBO* 2012; Cooper and Kühn *AEJ Micro* 2014; Awaya and Krishna *AER* 2016; Dugar and Shahriar 2016)

- finally and most challengingly, work explaining why and how natural-language communication, particularly in unlimited dialogues, improves upon structured communication via abstract messages

  The message I suggested above for Stag Hunt shows that even a single natural-language message can convey an understanding of strategic issues that is essential in some settings, one that *behaviorally* cannot be conveyed effectively via abstract messages

  (In theory players can mentally simulate any natural-language message or dialogue (Myerson *Ecma* 1983, *JET* 1989), but in practice that is no substitute for actual communication; Myerson *Ecma* 1983, *JET* 1989 and Forges *Ecma* 1986 model messages more richly than usual)

  (A further puzzle is why dialogues are better than "brief-filing"; they economize on cognition and bandwidth, and Forges *Ecma* 1986 and Myerson *Ecma* 1986 show they may otherwise expand the possibilities)

  There is very little further work on this topic, considering its importance; recent examples include Genesove and Mullin *AER* 2001; Cooper and Kagel *AER* 2005; Charness and Dufwenberg *Ecma* 2006, *EL* 2010; Cooper and Kühn *AEJ Micro* 2014; Burchardi and Penczynski *GEB* 2014; Awaya and Krishna *AER* 2016; Dugar and Shahriar 2016

I note in closing that in studying cognition, it is likely to be helpful to take fuller advantage of experimental methods that measure it more directly:

- monitoring subjects' searches for information about payoffs (Costa-Gomes et al. *Ecma* 2001; Johnson, Camerer, Rymon, and Sen *JET* 2002; Costa-Gomes and Crawford *AER* 2006; Wang et al. *AER* 2010; Brocas, Carillo, Wang, and Camerer *REStud* 2014)

  (The earlier work is surveyed in Crawford 2008 http://econweb.ucsd.edu/%7Evcrawfor/5Oct06NYUCognitionSearchMain.pdf)

- monitoring the chats of teams of subjects who must agree on decisions before they are implemented (Moreno and Wooders *GEB* 1998; Cooper and Kagel *AER* 2005; Burchardi and Penczynski *GEB* 2014)