Common Knowledge of Language and Communication Success*

Andreas Blume
Department of Economics
University of Arizona
Tucson, AZ 85721

Oliver J. Board
New York University School of Law
40 Washington Square South
New York, NY 10012

May 10, 2013

Abstract

When David Lewis [19] introduced common knowledge, he used it as a key ingredient in his account of language as a convention. We ask how essential higher orders of knowledge of language are for communication success. We find that in common-interest communication games higher-order failure of knowledge of language, while potentially resulting in suboptimal language use at any finite knowledge order, by itself has negligible *ex ante* payoff consequences. In contrast, with imperfect incentive alignment, lack of common knowledge of language may lead to complete communication failure for any finite-order knowledge of language.

*Portions of this paper were formerly part of "Language Barriers" [6]. We are grateful to Stephen Morris for having urged us to examine the connection of our conception of uncertainty about language with David Lewis's work more closely. We have benefitted from comments in a seminar at Simon Fraser University.

1 Introduction

Many prominent authors across a wide range of disciplines (philosophy of language, linguistics, social psychology) have stressed the importance of common knowledge for successful language use.[1] Lewis [19], who is widely credited with having given the first (verbal) definition of common knowledge (in 1969), placed it at center stage in his seminal account of language as a convention.[2] Grice [16] mentioned higher-order knowledge ("he knows (and knows that I know that he knows)") of "conventional meaning of the words," "identity of any references" and "other items of background knowledge" as among the prerequisites for participants in a conversation to be able to work out conversational implicatures. Clark and Marshall [9], [10], who adopted Schiffer's [24] terminology of referring to common knowledge as "mutual knowledge," emphasized the necessity of "mutual knowledge" for definite reference.

Clark and Marshall [9] noted that "To refer to a woman as she, the woman, or Nancy, we usually have good evidence that our audience knows her too" and asked: "But exactly what 'shared' knowledge is required?" In a series of examples they illustrated possible failures of definite reference with increasingly large but finite knowledge orders, and formulated a "mutual knowledge paradox," that successful definite reference appears to require common knowledge ("mutual kowledge" in their terminology) while the "infinite number of conditions" required appear "absurdly complicated." They suggested "copresence" as a possible source of common knowledge (and resolution of the paradox) and mentioned both "cultural copresence" and "linguistic copresence" as examples.

Karttunen and Peters [17] introduced the notion of "common ground" to describe the background information that a rational speaker can be assumed to take for granted at any given point in a conversation, and Stalnaker [25] equated common ground with "what is treated as their *common knowledge* or *mutual knowledge*" by speakers in a conversation. Clark and Brennan [8] examined how common ground is secured and accumulated in conversations, a process that they refered to as *grounding*. In linguistics, Thompson and Kaufmann [27] have proposed a game-theoretic model of conversational grounding that models errors in production and comprehension of messages (similarly to Blume, Board and Kawamura [5]) and explored the formal connections between grounding, common knowledge and common belief.

[1] Clark's [7] recent book on applications of game theory to linguistics devotes an entire chapter to the role of common knowledge and has additional references.

[2] Aumann [1] offered a formal definition of common knowledge and established equivalence of the iterative definition of common knowledge and a definition in terms of self-evident events. Friedell [14], [15], treated common knowledge formally in 1967 [14] and introduced a variety of examples, including the ideas that public announcements and eye contact might be possible sources of common knowledge.

Clark and Wilkes-Gibbs [11] experimentally studied the process by which participants in a conversation establish "mutual belief" about speaker meaning using a referential communication task due to Krauss and Weinheimer [18] in which pairs of individuals converse about arranging complex figures. Weber and Camerer [28] used a similar similar task to investigate experimentally how lack of a common language due to different organizational cultures affects merger outcomes. Blume, DeJong, Kim and Sprinkle [4] studied the emergence of meaning from *a priori* meaningless messages in a laboratory setting. Common to these experimental studies is that there are stages in the interaction among participants in which message meanings fail to be common knowledge and there is only partial communication success.

In Blume and Board [6] we recently proposed a framework for studying communication when meanings are imperfectly shared (and in that sense there is lack of "linguistic copresence"). We focused on the consequences of first-order failure of knowledge of language and observed that while some form of communication generally remains possible, in optimal equilibria there will typically be *indeterminacy of meaning*, characterized by disagreement between speaker and listener meaning of messages. Indeterminacy of meaning implies that language use is suboptimal. Here we adapt this framework to understand failures of higher-order knowledge of language. We find that in common-interest communication games higher-order failure of knowledge of language, while potentially resulting in suboptimal language use at any finite knowledge order, by itself has negligible *ex ante* payoff consequences. In contrast, with imperfect incentive alignment, lack of common knowledge of language may lead to complete communication failure for any finite-order knowledge of language. Our findings suggest that Lewis was right to have been concerned about failures of higher-order knowledge of language, but that the adverse consequences of such failures are only fully realized outside of common-interest environments.

2 Higher-order failures of knowledge of language

When there is uncertainty about language, messages may not be used optimally, conditional on their availability, in optimal equilibria, as noted by Blume and Board [6]. Here we ask whether in common-interest communication games we continue to see suboptimal language use in optimal equilibria when agents face higher-order uncertainty about language, whether such interim suboptimality translates into *ex ante* payoff losses from higher-order uncertainty about language, and, whether outside of common-interest environments there are circumstances where higher-order uncertainty about language leads to complete communication failure.

We study communication games between two players, a sender, who has payoff-relevant private information and a receiver who has no payoff-relevant private information himself, but cares about the sender's information. At the communication stage the sender sends a message to the receiver. At the action stage, both players simultaneously take an action. We allow for the possibility that the sender's action set is empty, in which case, following the literature, we refer to the game as a sender-receiver game.

In all the games that we consider there is uncertainty about the sender's *language type,* the subset λ of the finite message space M that is available to her, and about players' higher-order knowledge of the sender's language type. To represent players' higher-order knowledge about the sender's language type, we use an **information structure** $I = \langle \Omega, \mathcal{L}, \mathcal{O}^S, \mathcal{O}^R, q \rangle$.

- $\Omega = \{\omega_1, \omega_2, \ldots\}$ is a countable state space;
- $\mathcal{L} : \Omega \to 2^M$ specifies the set of messages available to the sender at each state (her *language type*);
- $\mathcal{O}^S$ is a partition of Ω, the sender's information partition;
- $\mathcal{O}^R$ is the receiver's information partition;
- q is the (common) prior on Ω.

To streamline the notation, let $\mathcal{L}(\omega) = \lambda_\omega$ and let $q(\omega) = q_\omega$. The information partitions describe the knowledge of the players: at state ω, the sender knows that the true state is in $\mathcal{O}^S(\omega)$ but no more (where $\mathcal{O}^S(\omega)$ is the element of $\mathcal{O}^S$ containing ω); and similarly for the receiver. We assume that the sender knows her own language type: if $\omega' \in \mathcal{O}^S(\omega)$, then $\lambda_\omega = \lambda_{\omega'}$. Define $\mathcal{L}(\Omega) := \{\lambda \in 2^M | \exists \omega \in \Omega \text{ with } \lambda = \lambda_\omega\}$.

Information structures encode uncertainty only about the sender's language type, not about the payoff-relevant information t (the sender's *payoff type*). We assume that the distribution from which t is drawn is independent of q. Given that the sender is fully informed about t, and the receiver knows nothing about t, it would be straightforward to extend the partitions and common prior over the full space of uncertainty, $T \times \Omega$, but to do so would unnecessarily complicate the notation.

We will say that there is 1st-order knowledge of language in our setting whenever both of the players know the sender's language type, and that there is nth-order knowledge of language whenever both of the players know that there is $(n-1)$th-order knowledge of language. Finally, we will say that there is a failure of nth-order knowledge of language whenever players do not have nth-order knowledge of the sender's language type.

3 Sender-Receiver Games

In the common-interest sender-receiver games we consider, a privately informed sender, S, communicates with a receiver, R, by sending a message $m \in M$, where $\#(M) \geq 2$ and M is finite. The common payoff $U(a,t)$ depends on the receiver's action, $a \in A = \mathbb{R}$, and the sender's payoff-relevant information $t \in T = [0,1]$, her *payoff type.* The sender's payoff type t is drawn from a differentiable distribution F on T with a density f that is everywhere positive on T. The function U is assumed to be twice continuously differentiable and, using subscripts to denote partial derivatives, the remaining assumptions are that for each realization of t there exists an action a_t^* such that $U_1(a_t^*, t) = 0$; and, $U_{11}(a,t) < 0 < U_{12}(a,t)$ for all a and t. In addition, we assume that the set of messages $\lambda \subseteq M$ that are available to the sender (i.e. her language type) and what players know and believe about the sender's language type are determined by an information structure I, as described above. In this section we also assume that the language state space Ω is finite. The assumptions on payoffs and payoff-type distribution are those of Crawford and Sobel [12] specialized to the common-interest case; we will refer to this as the *common-interest CS model.*

For some of our examples we will be interested in a special class of sender-receiver games where both players have identical quadratic loss functions $-(a-t)^2$ as payoffs and the sender's payoff type, t, is uniformly distributed on the interval $[0,1]$; this is the common-interest variant of the leading uniform-quadratic example of Crawford and Sobel [12], which we will refer to as the *uniform-quadratic CS model.*

In the resulting game a sender strategy is a function $\sigma : T \times \Omega \to \Delta(M)$ that satisfies $\sigma(t,\omega) \in \lambda_\omega$ for all $t \in T$ and all $\omega \in \Omega$ and is measurable with respect to $\mathcal{O}^S$. A receiver strategy is a function $\rho : M \times \Omega \to \mathbb{R}$ that is measurable with respect to $\mathcal{O}^R$. Thus for any strategy pair (σ, ρ), $\rho(m,\omega)$ denotes the receiver's response to the message m at state ω and $\sigma(t,\omega)$ the distribution over messages if the sender's payoff type is t at state ω.

At any state ω a sender strategy σ induces a mapping $\sigma_\omega : T \to \Delta(\lambda_\omega)$, where $\sigma_\omega(t) = \sigma(t,\omega)$ for all $t \in T$ and all $\omega \in \Omega$. We will refer to this mapping as the sender's *language at* ω. Similarly, we can define the receiver's language at state ω, $\rho_\omega : M \to \mathbb{R}$, via the property that $\rho_\omega(m) = \rho(m,\omega)$. A language $\hat{\sigma}_\omega$ of the sender is *optimal at* ω if together with a best response $\hat{\rho}_\omega$ to $\hat{\sigma}_\omega$ by the receiver it maximizes the sender's payoff at ω over all language pairs that are feasible at ω. A language $\hat{\rho}_\omega$ of the receiver is *optimal* at ω if it is a best response to an optimal language $\hat{\sigma}_\omega$ of the sender at ω. Thus, an optimal language pair $(\hat{\sigma}_\omega, \hat{\rho}_\omega)$ at ω maximizes joint payoffs subject only to the constraint that sender messages have to belong to λ_ω. Given a strategy pair (σ, ρ), we say that there is *suboptimal language use* at state ω if either σ_ω or ρ_ω is not optimal at ω.

The following simple example illustrates the suboptimal language use that may arise with failures of higher-order knowledge of language.

Example 1 *Consider a sender-receiver game in which both players have identical quadratic-loss payoff functions* $-(a-t)^2$ *and the sender's payoff type,* t*, is uniformly distributed on the interval* $[0,1]$*. Suppose that* $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$ *and that the information partitions are given by*

$$\begin{aligned} \textit{Sender}: &\quad \mathcal{O}^S = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4, \omega_5\}\} \\ \textit{Receiver}: &\quad \mathcal{O}^R = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5\}\}. \end{aligned}$$

In addition, assume that at ω_1 *the sender's language type is* $\lambda_{\omega_1} = \{m_1\}$*, (i.e. the sender has only message* m_1 *available), and that at every other state* $\omega \in \Omega$ *the sender's language type is* $\lambda_\omega = \{m_1, m_2\}$*, (i.e. the sender has both messages* m_1 *and* m_2 *available). The common prior,* q*, is uniform on* Ω*.*

Notice that $\{\omega_2, \omega_3, \omega_4, \omega_5\}$ *is the set of states at which the sender has all messages available;* $\{\omega_3, \omega_4, \omega_5\}$ *is the set of states at which the receiver knows that the sender has all messages available;* $\{\omega_4, \omega_5\}$ *is the set of states at which the sender knows that the receiver knows that the sender has all messages available; and* $\{\omega_5\}$ *is the set of states in which the receiver knows that the sender knows that the receiver knows that the sender has all messages available.*

At language states in $\{\omega_3, \omega_4, \omega_5\}$ *there is first-order knowledge of language; at states in* $\{\omega_4, \omega_5\}$ *there is second-order knowledge of language; and, at state* ω_5 *there is third-order knowledge of language. At no state is there higher than third-order knowledge of language and in particular there is never common knowledge of language.*

We look for an equilibrium (σ, ρ) *in which at element* $\{\omega_i, \omega_{i+1}\}$*, with* $i \in \{2, 4\}$*, of the sender's information partition there is a critical type* θ_i *such that payoff types* $t < \theta_i$ *send message* m_1 *and payoff types* $t > \theta_i$ *send message* m_2*. Let* a_1^j *denote the receiver's equilibrium response at* ω_j *(*$j \in \{1, 3, 5\}$*) to message* m_1*, and let* a_1^j *denote the response to message* m_2*. In equilibrium* θ_i*,* $i = 2, 4$*, and* a_k^j*,* $k = 1, 2$*,* $j = 1, 3, 5$ *must satisfy the conditions:*

$$a_1^1 = \frac{\frac{1}{2} + \theta_2 \frac{\theta_2}{2}}{1+\theta_2}; a_2^1 = \frac{1+\theta_2}{2}; a_1^3 = \frac{\theta_2 \frac{\theta_2}{2} + \theta_4 \frac{\theta_4}{2}}{\theta_2 + \theta_4};$$
$$a_2^3 = \frac{(1-\theta_2)\frac{1+\theta_2}{2} + (1-\theta_4)\frac{1+\theta_4}{2}}{(1-\theta_2)+(1-\theta_4)}; a_1^5 = \frac{\theta_4}{2}; a_2^5 = \frac{1+\theta_4}{2};$$
$$(\theta_2 - a_1^1)^2 + (\theta_2 - a_1^3)^2 = (\theta_2 - a_2^1)^2 + (\theta_2 - a_2^3)^2; \textit{ and,}$$
$$(\theta_4 - a_1^3)^2 + (\theta_4 - a_1^5)^2 = (\theta_4 - a_2^3)^2 + (\theta_4 - a_2^5)^2.$$

This system of equations has a unique solution satisfying the constraints that $0 < \theta_2 < 1$ *and* $0 < \theta_4 < 1$*:* $\theta_2 = 0.54896, \theta_4 = 0.509768, a_1^1 = 0.420074, a_1^2 = 0.77448, a_1^3 = 0.265045, a_2^3 = 0.764274, a_1^5 = 0.254884$ *and* $a_2^5 = 0.754884$*. Thus, at every state where the sender has a choice of which message to send, each message induces a non-degenerate lottery over receiver actions. Hence, not only is the receiver uncertain about the sender's use of messages, but the sender is also uncertain about the receiver's interpretation of messages. There is no state in which either sender-meaning or receiver-meaning is known by both players. Since having such knowledge would improve payoffs, there is indeterminacy of meaning, as defined by Blume and Board [6], and hence suboptimal language use.*

Importantly, even though at state ω_5 *there is third-order knowledge of language, players are not making optimal use of the available messages, which would require that* $\theta_4 = \frac{1}{2}$*. Notice also that at* ω_3*, where the receiver has only first-order knowledge of the fact that the sender has both messages, there is a larger distortion in the sender's strategy, i.e.* $\theta_2 - \frac{1}{2} > \theta_4 - \frac{1}{2}$*, and therefore players appear to make better use of the available messages with a higher order of knowledge of message availability.*

Example 1 suggests that there can be the persistence of suboptimal language use with increasing order of knowledge of the sender's language. We turn next to showing that for a class of equilibria that satisfy a sensible condition on focal message use – no order reversal of meanings – this observation generalizes to arbitrary information structures.

An equilibrium is *interval partitional* if at any information state ω the set of payoff types T can be partitioned into intervals such that types belonging to the same interval send a common message and types belonging to distinct intervals send distinct messages. Given an equilibrium with sender strategy σ, let $\Theta(m_i, \omega_k) := \{t \in T \mid \sigma(t, \omega_k)(m_i) > 0\}$ denote the set of all payoff types who send message m_i with strictly positive probability at state ω_k. For any two distinct sets $T_1 \subset T$ and $T_2 \subset T$ that have positive probability we say that $T_1 > T_2$ if $\inf T_1 \geq \sup T_2$.

Definition 1 *An equilibrium is* **order preserving** *if it is interval-partitional and* $\Theta(m', \omega_k)$

$> \Theta(m, \omega_k)$ at some state ω_k implies that $\Theta(m', \omega_{k'}) > \Theta(m, \omega_{k'})$ at all states $\omega_{k'}$ at which m' and m are used with positive probability.

Since our intent is to identify a characteristic of informative order-preserving equilibria, it is useful to know that they always exist. The following result establishes existence of informative order-preserving equilibria for the uniform-quadratic CS model with two messages and for arbitrary information structures. For the next result, assume that $M = \{m_1, m_2\}$ and that there is at least one information state ω with $\lambda_\omega = \{m_1, m_2\}$.

Lemma 1 *In the uniform-quadratic CS game with two messages and an arbitrary information structure an informative order-preserving equilibrium exists.*

Lemma 1 establishes not only that communication is possible in this environment, but that it can take a form where messages have some common meaning across information states, in the sense that the sender and the receiver commonly agree on which message means "low" and which means "high." The proof of this and other results can be found in the appendix.

Having established the existence of informative order-preserving equilibria for general information structures, we now show that, regardless of the information structure in informative order-preserving equilibria, suboptimal language use is pervasive.

Proposition 1 *Suppose that at any state ω either $\lambda_\omega = \{m_1\}$ or $\lambda_\omega = \{m_1, m_2\}$. Then for any information structure and for any state ω^* with $\lambda_{\omega^*} = \{m_1, m_2\}$ unless the language type is common knowledge at ω^*, in any informative order-preserving equilibrium of the uniform-quadratic CS game, the sender does not use an optimal language at ω^*.*

The following example explores the role of the restriction to order-preserving equilibria in Proposition 1.

Example 2 *Consider the information structure with partitions*

$$\text{Sender:} \quad \mathcal{O}^S = \{\{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4, \omega_5\}\}$$
$$\text{Receiver:} \quad \mathcal{O}^R = \{\{\omega_1\}, \{\omega_2, \omega_3, \omega_4\}, \{\omega_5\}\}$$

Assume that all states are equally likely and that the set of available messages is $\{m_1\}$ at ω_3 and $\{m_1, m_2\}$ otherwise. One easily checks that the following strategy pair, (σ, ρ), is an equilibrium: At $\{\omega_1, \omega_2\}$ the sender sends m_1 for $t \in \left[0, \frac{1}{2}\right)$ and m_2 otherwise. At $\{\omega_4, \omega_5\}$ the sender sends m_2 for $t \in \left[0, \frac{1}{2}\right)$ and m_1 otherwise; the receiver uses his unique best reply *to the specified sender strategy. Then at both $\{\omega_1\}$ and $\{\omega_5\}$ the receiver knows that the sender is using an optimal language, despite the fact that the set of available messages is not common knowledge.*

It is worth noting that there is a better equilibrium, and that in this equilibrium the sender never uses an optimal language when she has two messages available. To see this, modify the above strategy profile so that at $\{\omega_4, \omega_5\}$ the sender sends m_1 for $t \in \left[0, \frac{1}{2}\right)$ and m_2 otherwise and the receiver uses a best reply at $\{\omega_5\}$. The resulting strategy profile, $(\tilde{\sigma}, \tilde{\rho})$ has a strictly higher ex ante *payoff than (σ, ρ). Therefore, an optimal strategy profile for this game also must have a higher payoff than (σ, ρ) and since this is a common-interest game any optimal equilibrium must have a higher payoff than (σ, ρ).*

In this example the original order-reversing equilibrium was not optimal. This will not always be so, since there are cases where order-reversal is necessary for optimality, as with the information structure

$$\begin{aligned} \textit{Sender:} \quad & \mathcal{O}^S = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4, \omega_5\}, \{\omega_6\}\} \\ \textit{Receiver:} \quad & \mathcal{O}^R = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5, \omega_6\}\} \end{aligned}$$

where $\lambda_{\omega_1} = \lambda_{\omega_6} = \{m_1, m_2\}$, $\lambda_{\omega_2} = \{m_2, m_3\}$, $\lambda_{\omega_4} = \{m_1, m_3\}$ and all language states are equally likely. Note that optimality is achieved if at ω_1 the sender sends m_1 for $t \in \left[0, \frac{1}{2}\right)$ and m_2 otherwise; at ω_2 the sender sends m_3 for $t \in \left[0, \frac{1}{2}\right)$ and m_2 otherwise; at ω_4 the sender sends m_3 for $t \in \left[0, \frac{1}{2}\right)$ and m_1 otherwise; and, at ω_6 the sender sends m_2 for $t \in \left[0, \frac{1}{2}\right)$ and m_1 otherwise. Furthermore, in any optimal equilibrium at ω_1 the sender sends either m_1 for $t \in \left[0, \frac{1}{2}\right)$ and m_2 for $t \in \left(\frac{1}{2}, 1\right]$ or m_2 for $t \in \left[0, \frac{1}{2}\right)$ and m_1 for $t \in \left(\frac{1}{2}, 1\right]$, and in either case the role of m_1 and m_2 is reversed at ω_6.

In an order-preserving equilibrium "high" and "low" may mean different things at different information states and in addition their meaning at a given information state may be uncertain; but it is never the case that the meanings of "high" and "low" at one language state are flipped at another. In that sense order-preserving equilibria have an appeal as being focal. Example 2 shows that there is no close connection between *ex ante* optimality of equilibria and those equilibria being order-preserving. In the first part of Example 2 optimally requires that equilibria be order preserving, and in the second part of the same example, optimality requires that equilibria be order reversing. In the first case, higher-order uncertainty of language on its own results in suboptimal language use. In the second case, suboptimal language use may arise as a combination of higher-order uncertainty about language and equilibria being selected on the basis of being focal.

Proposition 1 provides sufficient conditions for higher-order uncertainty about language to result in pervasive suboptimal language use. At the same time, Example 1 hints at the possibility that suboptimal language use may diminish with higher knowledge order. Our next example shows that increasing knowledge order may not result in improved language use, even in *ex ante* optimal equilibria. In Example 3 in any optimal equilibrium language use remains bounded away from optimality for any finite knowledge order.

Example 3 *Consider a finite sender-receiver game with common payoffs given by the following table. The sender has three equally likely payoff types, t_1, t_2 and t_3, and the receiver has four actions $a_1, \ldots, a_4$. Each cell in the payoff table indicates the common payoff from the corresponding type-action pair (t_i, a_j). After privately observing her payoff type t_i the sender sends a message m_k from her set of available messages to the receiver who then takes an action a_j in response to the sender's message.*

	a_1	a_2	a_3	a_4
t_1	7	9	0	10
t_2	7	9	10	0
t_3	7	0	6	0

If the sender's set of available messages is comprised of m_1 and m_2, then in any optimal equilibrium payoff types t_2 and t_3 send a common message, and the ex ante *payoff from any optimal equilibrium is $\frac{26}{3}$.*

Assume now that there is higher-order uncertainty about language. The sender's language type is either $\{m_1\}$ or $\{m_1, m_2\}$, that is, either the sender has only message m_1 available or the sender has both messages m_1 and m_2 available. The language state space $\Omega = \{\omega_1, \omega_2, \omega_3, \ldots\}$ is assumed to be countably infinite. The common prior on the language state space Ω is the geometric distribution that assigns probability $p(1-p)^{n-1}$ to state ω_n for some $p \in (0,1)$. Let $\lambda_{\omega_1} = \{m_1\}$ and $\lambda_{\omega_n} = \{m_1, m_2\}$ $\forall n \neq 1$; i.e., at every language state ω_n except ω_1 both messages are available. The sender's information partition is given by:

$$\mathcal{O}^S = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \{\omega_4, \omega_5\}, \ldots\}$$

The receiver's information partition is given by:

$$\mathcal{O}^R = \{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5, \omega_6\}, \ldots\}$$

Then for sufficiently large p

1. $(m_1, \mathcal{O}^R(\omega_1)) \mapsto a_1$ *in every equilibrium.*

   This is because for large p the receiver's posterior at ω_1 following message m_1 is approximately the same as the prior, regardless of how the sender uses messages at ω_2. Since the receiver's best response to the prior, a_1, is unique it remains a best response against beliefs sufficiently close to the prior.

2. $(t_3, \mathcal{O}^S(\omega_2)) \mapsto m_1$ *in every (*ex ante*) optimal equilibrium.*

   This is a consequence of the fact that for large p the payoff at ω_2 is an order of magnitude more important for the ex ante *expected payoff than payoffs at any ω_n with $n > 2$. From (1) above, at ω_2 the message m_1 induces action a_1 with probability one. We need to determine how to use message m_2 optimally at language state ω_2. Since payoff type t_3 already gets her maximal payoff, we can focus on maximizing payoffs for the remaining payoff types. Since they have only one message available, the best alternative is for both of them to pool on message m_2. Then only m_1 induces a_1 and payoff type t_3 will send it.*

3. $(t_1, \mathcal{O}^S(\omega_2)) \mapsto m_2$ *in every optimal equilibrium.*

   The argument is the same as for (2) above.

4. $(t_2, \mathcal{O}^S(\omega_2)) \mapsto m_2$ *in every optimal equilibrium.*

   The argument is the same as for (2) above.

5. $(m_1, \mathcal{O}^R(\omega_n)) \mapsto a_1 \; \forall n$ *in every optimal equilibrium.*

   From (1) – (4) we know what the sender's strategy prescribes at $\{\omega_2, \omega_3\}$. Since at $\{\omega_3, \omega_4\}$, the language state ω_3 is an order of magnitude more likely than ω_4, the receiver's beliefs at that information set are almost entirely determined by what the sender's strategy prescribes at ω_3. Against those beliefs, it is uniquely optimal to respond to message m_1 with action a_1 and to respond to message m_2 with action a_2.

   From hereon, we can proceed by induction: For any odd $n > 2$, if the receiver's strategy at $\{\omega_n, \omega_{n+1}\}$ prescribes to respond to m_1 with a_1 and to m_2 with a_2, and it is uniquely optimal for the sender at $\{\omega_{n+1}, \omega_{n+2}\}$ to send message m_1 if her payoff type is t_3 and m_2 if her payoff type is either t_1 or t_2, using the fact that for large p the language state ω_{n+1} is an order of magnitude more likely than language state ω_{n+2}.

   Similarly, for any even $n > 2$, if the sender's strategy at $\{\omega_n, \omega_{n+1}\}$ prescribes to send message m_1 if her payoff type is t_3 and m_2 if her payoff type is either t_1 or t_2, then *it is uniquely optimal for the receiver at* $\{\omega_{n+1}, \omega_{n+2}\}$ *to respond to* m_1 *with* a_1 *and to* m_2 *with* a_2.

In addition to (5), this establishes that the unique optimal equilibrium satisfies properties (6) – (9) below.

6. $(m_2, \mathcal{O}^R(\omega_n)) \mapsto a_2 \ \forall n$ *in every optimal equilibrium.*

7. $(t_1, \mathcal{O}^S(\omega_n)) \mapsto m_2 \ \forall n > 1$ *in every optimal equilibrium.*

8. $(t_2, \mathcal{O}^S(\omega_n)) \mapsto m_2 \ \forall n > 1$ *in every optimal equilibrium.*

9. $(t_3, \mathcal{O}^S(\omega_n)) \mapsto m_1 \ \forall n > 1$ *in every optimal equilibrium.*

Hence, we can conclude that for large enough $p \in (0,1)$ *there is a unique (*ex ante*) optimal equilibrium. In this unique optimal equilibrium behavior at all higher-order knowledge states is the same; regardless of the knowledge order, behavior is bounded away from optimal equilibrium behavior if the language type were common knowledge; and, language-interim payoffs (where language types are known but payoff types are not) are bounded away from optimal common-knowledge payoffs, regardless of knowledge order — at sufficiently high knowledge orders, the language-interim payoff is* $\frac{25}{3}$*, while the optimal common-knowledge-of-language payoff would be* $\frac{26}{3}$.

We have given various examples in which higher-order uncertainty about language in sender-receiver games results in suboptimal language use. As a bit of an antidote, and foreshadowing our results on more general common-interest communication games, we conclude with a simple observation concerning higher-order uncertainty about language in general common-interest CS games.

Proposition 2 *In any common-interest CS game with an information structure such that at some language state* $\tilde{\omega}$ *the receiver knows that message* m_1 *is available (*$m_1 \in \lambda_\omega$ *for all* $\omega \in \mathcal{O}^R(\tilde{\omega})$*), and message* m_2 *is available (*$m_2 \in \lambda_{\tilde{\omega}}$*), there is a strict* ex ante *benefit from communication in any optimal equilibrium.*

Proof: Let

$$a(\underline{t}, \bar{t}) := \begin{cases} \arg\max_{a \in A} \int_{\underline{t}}^{\bar{t}} U(a,t) f(t) dt & \text{for } \underline{t} < \bar{t} \\ a^*_{\bar{t}} & \text{for } \underline{t} = \bar{t} \end{cases}$$

Then $U_{12} > 0$ implies that $a(0,1) < a\left(\frac{1}{2}, 1\right) < a(1,1) = a^*_1$ and that there exists a type $\tilde{t} \in (0,1)$ who is indifferent between the actions $a(0,1)$ and $a\left(\frac{1}{2}, 1\right)$. All types $t \in (\tilde{t}, 1]$ strictly prefer action $a\left(\frac{1}{2}, 1\right)$ to action $a(0,1)$.

Consider the strategy pair (σ, ρ) that is defined by

$$\sigma(t,\omega) = \begin{cases} m_2 & \text{if } m_1, m_2 \in \lambda_\omega \text{ and } t \in (\tilde{t}, 1] \\ m_1 & \text{otherwise} \end{cases}$$

and

$$\rho(m,\omega) = \begin{cases} a\left(\frac{1}{2}, 1\right) & \text{if } m = m_2 \text{ and } \omega \in \mathcal{O}^R(\tilde{\omega}), \text{ and} \\ a(0,1) & \text{otherwise} \end{cases}$$

At every language-payoff state pair (ω, t) where either $\omega \notin \mathcal{O}^R(\tilde{\omega})$, $t \in [0, \tilde{t}]$ or $m_2 \notin \lambda_\omega$ the *ex post* payoff from (σ, ρ) is the same as in a pooling equilibrium. At all of the remaining language-payoff state pairs the sender's payoff type belongs to $(\tilde{t}, 1]$, the receiver takes action $a\left(\frac{1}{2}, 1\right)$, and therefore the *ex post* payoff exceeds that from a pooling equilibrium. Since the latter set of language-payoff state pairs has strictly positive probability, it follows that the *ex ante* payoff from the strategy pair (σ, ρ) strictly exceeds that from a pooling equilibrium. Since we have a common-interest game, an *ex ante* optimal strategy profile is an equilibrium profile. Existence of such a profile follows from a straightforward compactness argument (we will establish existence under more general conditions in the next subsection). Evidently, the payoff from this profile cannot be less than that from (σ, ρ). □

In summary, despite the fact that in common-interest uniform-quadratic CS games we can give sufficient conditions for higher-order uncertainty about language to imply suboptimal language use (Proposition 1) and that in other related games it can be the case that interim expected payoffs for any finite order of knowledge of language are bounded away from interim expected payoffs with common knowledge of language (Example 3), under a very mild condition – that with positive probability there are two messages, one of which the receiver knows is available and one which simply is available – there is a strict *ex ante* benefit from communication in optimal equilibria of common-interest CS games.

4 Common-interest communication games

We ended the last section on a positive note, showing that under fairly general conditions higher-order on certainty about language does not prevent communicative gains in sender-receiver games. Here we continue this theme while considering a more general class of communication games in which both sender and receiver act following the communication stage.

We continue to consider common-interest communication games where a privately informed sender, S, communicates with a receiver, R, by sending a message $m \in M$, where $\#(M) \geq 2$ and M is finite. Now, however, following the sender's message both sender and receiver simultaneously choose actions $a_S \in A_S$ and $a_R \in A_R$. The common payoff $U(a_S, a_R, t)$ depends on the sender's action $a_S \in A_S$, the receiver's action $a_R \in A_R$, and the sender's payoff-relevant private information $t \in T$, her *payoff type*. We assume that A_S and A_R are compact and convex subsets of finite dimensional Euclidean spaces, that T is a compact subset of a finite-dimensional Euclidean space, and that the payoff function U is concave (and therefore continuous) in its first two arguments for all $t \in T$. The sender's payoff type t is drawn from a commonly known distribution F on T.

We continue to assume that the set of messages $\lambda \subseteq M$ that are available to the sender (i.e. her language type) and players' higher-order knowledge and belief about the sender's language type are determined by an information structure I with a finite language-state space Ω.

Examples 1 and 3 illustrate how higher-order failures of knowledge of language can result in suboptimal language use. It is worth keeping in mind, however, that in both cases the losses at higher-order knowledge states are compensated by corresponding gains at lower-order knowledge states; if we tried to force optimal language use conditional on available messages at higher-order knowledge states, any resulting strategy would be less well equipped to make the best use of available message at lower-order knowledge states. In this sense, there is no *ex ante* payoff loss from higher-order uncertainty about language.

A version of this observation holds under very general conditions. As the following result shows, in optimal equilibria of common-interest communication games, where both sender and receiver act at the action stage, from an *ex ante* payoff perspective only failures of first-order knowledge matter. Regardless of the information structure, as long as the receiver knows the sender's language type with high probability, suboptimal language use as a result of higher-order knowledge failure can occur but must be either insignificant or improbable.

For $N = 1, 2, \ldots$, let $\mathcal{I}_N$ denote the set of information structures with $\Omega_N = \{\omega_1, \omega_2, \ldots, \omega_N\}$. For each $I_N \in \mathcal{I}_N$ define

$$\widehat{\Omega}_N := \{\omega \in \Omega_N | \lambda_\omega = \arg \max_{\lambda \in \mathcal{L}(\Omega_N)} \text{Prob}(\lambda | \mathcal{O}^R(\omega))\}.$$

$\widehat{\Omega}_N$ is the set of states at which the receiver's best guess of the sender's language type conditional on her information is correct. We say that a sequence of such information structures satisfies *vanishing first-order uncertainty* if $\text{Prob}(\widehat{\Omega}_N)$ converges to one as $N \to \infty$.

Let $U(I_N)$ denote the (*ex ante*) payoff from an optimal equilibrium in the game $G(I_N)$ with information structure $I_N \in \mathcal{I}_N$, whose existence is assured by the following observation.

Lemma 2 *Regardless of the information structure I_N, any common-interest communication game $G(I_N)$ has an optimal equilibrium.*

Let $U^*(I_N)$ denote the payoff from an optimal equilibrium of the game with an information structure that is obtained from I_N by replacing the knowledge partitions of both players by the finest partition, so that the language type is common knowledge at every state (existence of such an equilibrium follows once more from Lemma 2). Then:

Proposition 3 *For any common-interest communication game, for any sequence of information structures $\{I_N\}_{N=1}^{\infty}$, with $I_N \in \mathcal{I}_N$ for all N, that satisfies vanishing first-order uncertainty:*

$$\lim_{N\to\infty} |U(I_N) - U^*(I_N)| = 0.$$

Proof: For every $\lambda \in \Lambda$ and every $I_N \in \mathcal{I}_N$, use $\Omega_\lambda(I_N)$ to denote the set of states at which the sender's language type is λ according to I_N. For all $\lambda \in \Lambda$ and at every state in $\Omega_\lambda(I_N)$ let the sender use a strategy σ_λ that would be part of an optimal profile $(\sigma_\lambda, \rho_\lambda)$ if it were common knowledge that her language type is λ; existence of $(\sigma_\lambda, \rho_\lambda)$ follows from Lemma 2. At every state $\omega \in \Omega_N$ let the receiver use a strategy $\rho_{\lambda(\omega)}$ with

$$\lambda(\omega) \in \arg\max_{\lambda} \text{Prob}(\lambda | \mathcal{O}^R(\omega)),$$

where $\rho_{\lambda(\omega)}$ is part of an optimal profile $(\sigma_{\lambda(\omega)}, \rho_{\lambda(\omega)})$ given the language type $\lambda(\omega)$. Only at the states in $\Omega_N \backslash \widehat{\Omega}_N$ are the sender's and receiver's strategies mismatched and by assumption the probability of those events converges to zero. Hence there is some sequence of strategy profiles $\{(\sigma_N, \rho_N)\}_{N=1}^{\infty}$ corresponding to the sequence of games $\{G_N\}_{N=1}^{\infty}$ that are induced by the sequence of information structures $\{I_N\}_{N=1}^{\infty}$ for which

$$\lim_{N\to\infty} |U(\sigma_N, \rho_N) - U^*(I_N)| = 0.$$

For each $G(I_N)$ the payoff from an optimal strategy profile (σ_N^*, ρ_N^*) is no less than the payoff from the profile (σ_N, ρ_N), where existence follows from Lemma 2. Since we are considering common-interest games, each optimal profile (σ_N^*, ρ_N^*) is an equilibrium profile for $G(I_N)$. The claim follows. $\square$

We conclude that even though there are circumstances under which failure of higher-order knowledge of language of any finite order leads to suboptimal language use in optimal equilibria of common-interest communication games, unless there are significant failures of low-order knowledge, the *ex ante* payoff consequences are negligible. This observation is reinforced by our next result, which shows that increasing the order of knowledge about language never hurts in common-interest communication games.

To formalize this idea, we introduce a relation on the set of information structures that we call language-knowledge dominance. Intuitively, an information structure language-knowledge dominates another if it is obtained from the former by expanding an information set $\mathcal{O}^{-\ell}(\omega^0)$ of one of the players, $-\ell$, by including additional states that together form a new information set for the other player, ℓ, in such a way that the added states do not reduce message availability. This way, at the added states player ℓ is as well informed about player $-\ell$'s knowledge at the expansion of $\mathcal{O}^{-\ell}(\omega^0)$ as at $\mathcal{O}^{-\ell}(\omega^0)$ before, and all the strategic options that player $-\ell$ had at $\mathcal{O}^{-\ell}(\omega^0)$ remain intact at the expansion of that set.

Definition 2 *An information structure* $I' = \langle \Omega', \mathcal{L}', \mathcal{O}'^S, \mathcal{O}'^R, q' \rangle$ ***language-knowledge dominates*** *the information structure* $I = \langle \Omega, \mathcal{L}, \mathcal{O}^S, \mathcal{O}^R, q \rangle$ *if and only if there is a player* $\ell \in \{S, R\}$ *and a state* $\omega^0 \in \Omega$ *such that*

1. $\Omega' \supsetneq \Omega$;
2. $\Phi := \Omega' \setminus \Omega$;
3. $\mathcal{O}'^{\ell}(\omega) = \Phi, \ \forall \omega \in \Phi$;
4. $\mathcal{O}'^{-\ell}(\omega) = \Phi \cup \mathcal{O}^{-\ell}(\omega^0), \ \forall \omega \in \Phi$;
5. $\mathcal{O}'^{\ell}(\omega) = \mathcal{O}^{\ell}(\omega), \ \forall \omega \in \Omega$;
6. $\mathcal{O}'^{-\ell}(\omega) = \mathcal{O}^{-\ell}(\omega), \ \forall \omega \in \Omega' \setminus \mathcal{O}'^{-\ell}(\omega^0)$;
7. $\lambda'_{\tilde{\omega}} \supseteq \lambda_{\omega}, \ \forall \tilde{\omega} \in \Phi$ *and* $\omega \in \mathcal{O}^{-\ell}(\omega^0)$;
8. $\lambda'_{\omega} \supseteq \lambda_{\omega}, \ \forall \omega \in \Omega$;
9. $q'_{\omega} = q_{\omega} \ \forall \omega \in \Omega' \setminus \mathcal{O}'^{-\ell}(\omega^0)$; *and,*
10. $q'_{\omega} > 0, \ \forall \omega \in \Omega'$.

As before, let $U(I)$ denote the *ex ante* maximal payoff for the game with information structure I. Then we have the following result.

Proposition 4 *If information structure* I' *language-knowledge dominates information structure* I*, then* $U(I') \geq U(I)$.

Proof: Let $(\tau_\ell, \tau_{-\ell})$ be a strategy profile that attains $U(I)$ in the game with information structure I and consider the strategy profile $(\tau'_\ell, \tau'_{-\ell})$ in the game with information structure I' that is defined by $\tau'_{-\ell}(\mathcal{O}'^{-\ell}(\omega)) = \tau_{-\ell}(\mathcal{O}^{-\ell}(\omega))$ for all $\omega \in \Omega$ (note that this specifies $\tau'_{-\ell}$ also for $\omega \in \Phi$), $\tau'_\ell(\mathcal{O}'^\ell(\omega)) = \tau_\ell(\mathcal{O}^\ell(\omega))$ for all $\omega \in \Omega$, and $\tau'_\ell(\mathcal{O}'^\ell(\omega))$ is a best reply to $\tau'_{-\ell}(\mathcal{O}'^{-\ell}(\omega^0))$ for all $\omega \in \Phi$. Then *ex post* payoffs at all language states in $\Omega' \setminus \mathcal{O}'^{-\ell}(\omega^0)$ are the same for both information structures. Furthermore the prior probabilities of these language states did not change. Hence, any change in the *ex ante* payoff of player $-\ell$ will be a consequence of a difference in the payoff conditional on $\mathcal{O}^{-\ell}(\omega^0)$ in the game with information structure I and the payoff conditional on $\mathcal{O}'^{-\ell}(\omega^0)$ in the game with information structure I'. The latter payoff however cannot be lower since we have moved probability to states at which the other player, ℓ, both is better informed and is no more language constrained. $\square$

5 Communication collapse with higher-order knowledge failures

Our analysis thus far has shown that in common-interest communication games lack of common knowledge of the sender's language may lead to pervasive suboptimal language use, but that at the same time the consequences of higher-order failures of knowledge of language by themselves appear rather benign. In contrast, in this section we show that with imperfectly aligned incentives lack of common knowledge of language can entail complete communication collapse regardless of finite-order knowledge in situations where communication could be put to good use with common knowledge of language.

The following example constructs such a scenario by building on insights of Rubinstein [23], Baliga and Morris [3], and Aumann [2].

Example 4 *Two players play a two-stage game with one-sided private information represented by two equally likely payoff states t_1 and t_2 (so the payoff type space for the sender is $T = \{t_1, t_2\}$). In the communication stage the privately-informed sender sends a message to the receiver. In the action stage both players simultaneously take actions which determine payoffs according to the tables in Figure 1.*

It is easily verified that if it is common knowledge that the sender has two messages, m_α and m_β, available, then there is an equilibrium in which the sender sends message m_α in payoff state t_1, message m_β in payoff state t_2, and each player i takes action α_i if and only if message m_α has been sent.

	α_R	β_R
α_S	$3, 3$	$-10, 2$
β_S	$2, -10$	$1, 1$

t_1

	α_R	β_R
α_S	$-10, -10$	$-10, -9$
β_S	$-9, -10$	$1, 1$

t_2

Figure 1: Payoff States

Suppose instead that it is not common knowledge which messages are available to the sender. Consider an information structure with a (countably infinite) state space $\Omega = \{\omega_1, \omega_2, \ldots\}$ *and some common prior* q *with the property that* $q_k > q_{k+1}$ *for all* $k = 1, 2 \ldots$. *The players' information partitions are given by*

$$\begin{aligned} \textit{Sender}: \quad & \mathcal{O}^S = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3, \omega_4\}, \{\omega_5, \omega_6\}, \{\omega_7, \omega_8\} \ldots\} \\ \textit{Receiver}: \quad & \mathcal{O}^R = \{\{\omega_1, \omega_2, \omega_3\}, \{\omega_4, \omega_5\}, \{\omega_6, \omega_7\}, \{\omega_8, \omega_9\}, \ldots\}. \end{aligned}$$

Finally, assume that $\lambda_{\omega_1} = \{m_\alpha\}$, $\lambda_{\omega_2} = \{m_\beta\}$, *and* $\lambda_{\omega_k} = \{m_\alpha, m_\beta\}$ *for all* $k = 3, 4, \ldots$ *(i.e. the sender has only* m_α *available at* ω_1, *she has only* m_β *available at* ω_2, *and she has both messages available at every other state).*

Then it follows from Proposition 5 below that in any equilibrium of the game only actions β_S *and* β_R *are taken, regardless of the payoff state and the information state. In particular for any finite order of knowledge of the fact that both messages are available to the sender, they remain ineffective in equilibrium.*

The following is a sketch of the argument for the game under consideration:

1. *At* ω_1, *regardless of the sender's strategy, the receiver believes it is more likely that the sender's message is uninformative than informative, and hence the relatively safe action* β_R *is uniquely optimal for the receiver.*

2. *At* ω_3 *and* ω_4, *the sender considers* ω_3 *more likely than* ω_4, *and therefore she believes that the receiver takes action* β_R *with at least probability one half; it follows that action* β_S *is uniquely optimal for the sender.*

3. *If at* ω_4 *the sender had a message that would induce the receiver to take action* α_R *with positive probability, she would send such a message in payoff state* t_1 *despite (as we*

showed) taking action β_S herself. This is a consequence of a violation of Aumann's [2] ***self-signaling condition***[3] *at state t_1: at state t_1 the sender wants to persuade the receiver to take action α_R regardless of her own intended action.*

4. *At ω_4 and ω_5, the receiver considers ω_4 more likely than ω_5; thus (3) implies that, regardless of the message, he believes that the sender uses action β_S with probability greater than one half.*

5. *Given (4), it is uniquely optimal for the receiver to take action β_R at ω_4 and at ω_5.*

6. *Steps (2)-(5) can be turned into an induction argument that shows that for all ω_k the sender uses action β_S regardless of the payoff state and the receiver uses action β_R regardless of the message.*

Rubinstein [23] shows how lack of higher-order knowledge can translate into coordination failures. Baliga and Morris [3] demonstrate how a violation of the self-signaling condition can render communication ineffective in games with one-sided private information and Morris [21] shows that this effect persists with higher-order knowledge failures of the form used in Rubinstein's electronic mail game.

Our example differs from Baliga and Morris in that with our payoff structure private information about payoffs by itself would not lead to communication failure; it has to be accompanied by private information about language (because we have failure of the self-signaling condition in only one state). It differs from the electronic mail game because adopting the payoff struture from the electronic mail game and leaving our example otherwise unchanged would not prevent communication: since in the electronic mail game there are multiple equilibria conditional on each payoff state players could simply use the messages m_α and m_β as coordination devices at low-knowledge states and switch to using them to signal information about payoff states at higher-knowledge states.

Consider also the payoff structure in Figure 2, taken from Steiner and Stewart [26]. Unlike Figure 1, this payoff structure does not violate the self-signaling condition. As a result, even though messages are useless at low knowledge states, they can become useful at higher-knowledge states. For the information structure in Example 4, there exists an equilibrium in which the sender uses the following strategy:

[3] The terminology is due to Farrell and Rabin [13]. A pre-play message is *self-signaling* provided the speaker wants it to be believed if and only if it is true. This is one possible condition for a message to be credible. Another is that it be *self-committing*: the speaker has an incentive to fulfill it if it is believed. In the (complete-information) game with payoff matrix corresponding to state t_1 a sender message "I will take action α_S" is self-committing because the action pair (α_S, α_R) forms an equilibrium. The message is not self-signaling because the sender would also want it to be believed if he intends to take actions β_S.

$\{(\{\omega_1\}, t_1) \mapsto (m_\alpha, \beta_S)$

$(\{\omega_1\}, t_2) \mapsto (m_\alpha, \beta_S)$

$(\{\omega_2\}, t_1) \mapsto (m_\beta, \beta_S)$

$(\{\omega_2\}, t_2) \mapsto (m_\beta, \beta_S)$

$(\{\omega_3, \omega_4\}, t_1) \mapsto (m_\beta, \beta_S)$

$(\{\omega_3, \omega_4\}, t_2) \mapsto (m_\beta, \beta_S)$

$(\{\omega_k, \omega_{k+1}\}, t_1) \mapsto (m_\alpha, \alpha_S),$

$(\{\omega_k, \omega_{k+1}\}, t_2) \mapsto (m_\beta, \beta_S)\}$ for all $k \geq 5$.

The receiver responds to both messages at $\{\omega_1, \omega_2, \omega_3\}$ with β_R and otherwise follows the rule $m_\alpha \mapsto \alpha_R$ and $m_\beta \mapsto \beta_R$. As a result, messages are used effectively for all information states ω_k with $k \geq 5$. As in our example, the sender takes action β_S at $\{\omega_3, \omega_4\}$ regardless of the payoff state, but unlike in our example here the sender has no incentive at those information states to take advantage of the receiver's sensitivity to messages at $\{\omega_4, \omega_5\}$.

	α_R	β_R
α_S	$1-p, 1-p$	$-p, 0$
β_S	$0, -p$	$0, 0$

t_1

	α_R	β_R
α_S	$-p, -p$	$-p, 0$
β_S	$0, -p$	$0, 0$

t_2

Figure 2: Payoff States ($p \in (1/2, 1)$)

The results in this section generalize Example 4 in two directions. We first fix the information structure and provide a sufficient condition on the payoff structure that ensures communication failure in a class of games with arbitrary finite action spaces, payoff-type spaces and message spaces. Then, fixing the payoff structure from Figure 1, we identify a necessary condition on the information structure for the existence of communicative equilibria (that satisfy a mild regularity condition – requiring that there is a clear designation of which of the two messages is sometimes used to bring about efficient coordination in state t_1).

The following result identifies characteristics of communication games with one-sided private information that lead to communication breakdown for any finite order of knowledge of language (it also verifies the details of Example 4). For this purpose we consider a class of games with two players, a sender (S) and a receiver (R). The sender privately observes her payoff type t from a finite set T and sends a message m from a finite set M to the receiver. Each $t \in T$ has strictly positive prior probability $\pi(t)$. The sender's message has to satisfy the constraint that $m \in \lambda$ where $\lambda \subset M$ is her privately known language type. Each player $i = S, R$ has a finite set of actions A_i. Following the communication stage, both players simultaneously take actions $a_S \in A_S$ and $a_R \in A_R$. Given these actions and the sender's payoff type, each player i receives a payoff $U_i(a_S, a_R, t)$. As before, the players' knowledge about the sender's language is represented by an information structure $I = \langle \Omega, \lambda, \mathcal{O}^S, \mathcal{O}^R, q \rangle$. Call any game of this form a *communication game*.

We are interested in a subclass of such games in which (i) the receiver has a preferred "safe" action that is uniquely optimal if there is sufficient uncertainty about either the sender's action or her payoff type; (ii) the sender has a unique "safe" best reply for sufficiently strong beliefs that the receiver will use his safe action, and (iii) it is difficult for the sender credibly to communicate an intent to take an action other than her safe best reply.

An action a_R^0 for the receiver is "safe" if is uniquely optimal regardless of the sender's (rational) rule for mapping payoff types into actions for any belief that does not assign more than probability $\frac{2\pi(t)}{1+\pi(t)}$ to any type t, i.e.

$$\sum_{t \in T} U_R(\alpha_S(t), a_R^0, t)\mu(t) > \sum_{t \in T} U_R(\alpha_S(t), a_R, t)\mu(t)$$

for all $a_R \neq a_R^0$, for all $\alpha_s : T \to A_S$ that are best responses to some (mixed) receiver action and for all $\mu \in \Delta(T)$ with $\mu(t) < \frac{2\pi(t)}{1+\pi(t)}$ $\forall t$. We say that the game satisfies the **safe-action condition** if the receiver has a safe action.

For a game that satisfies the safe-action condition, we call a sender action a_S^0 "safe" if independent of the payoff type, it is a unique best reply against beliefs that assign at least probability one half to the receiver taking action a_R^0, i.e.

$$U_S(a_S^0, pa_R^0 + (1-p)\alpha_R, t) > U_S(a_S, pa_R^0 + (1-p)\alpha_R, t)$$

$\forall \alpha_R \in \Delta(A_R), \forall a_S \neq a_S^0, \forall p \geq 1/2, \forall t \in T$. We say that the game satisfies the **sender safe-response condition** if the sender has a safe action.

A game that satisfies the sender-safe response condition satisfies the **receiver safe-response condition** if at every payoff state t, provided the sender uses her safe response a_S^0 with at least probability one half, the receiver's safe action a_R^0 is a unique best reply, i.e.

$$U_R(pa_S^0 + (1-p)\alpha_S, a_R^0, t) > U_R(pa_S^0 + (1-p)\alpha_S, a_R, t)$$

$\forall \alpha_S \in \Delta(A_S), \forall a_R \neq a_R^0, \forall p \geq 1/2, \forall t \in T$.

A game that satisfies the sender- and receiver-best response conditions satisfies the **no-self-signaling condition** if in every state t in which a_S^0 is not dominant for the sender, conditional on taking action a_S^0 herself, the sender prefers that the receiver take any action other than a_R^0, i.e. for all t such that there exist $a_S \neq a_S^0$ and a_R with $U_S(a_S, a_R, t) \geq U_S(a_S^0, a_R, t)$ it is the case that

$$U_S(a_S^0, a_R^0, t) < U_S(a_S^0, a_R, t) \ \forall a_R \neq a_R^0.$$

Proposition 5 *In any communication game that satisfies the safe-action, sender-safe-response, receiver-safe-response and no-self-signaling conditions, with information partitions*

$$\begin{aligned} \mathcal{O}^S &= \{\{\omega_1\}, \ldots, \{\omega_\nu\}, \{\omega_{\nu+1}, \omega_{\nu+2}\}, \{\omega_{\nu+3}, \omega_{\nu+4}\}, \{\omega_{\nu+5}, \omega_{\nu+6}\}, \ldots\ldots.\} \\ \mathcal{O}^R &= \{\{\omega_1, \ldots, \omega_\nu, \omega_{\nu+1}\}, \{\omega_{\nu+2}, \omega_{\nu+3}\}, \{\omega_{\nu+4}, \omega_{\nu+5}\}, \{\omega_{\nu+6}, \omega_{\nu+7}\}, \ldots\}, \end{aligned}$$

where $\nu = \#(M)$, $\lambda_{\omega_i} = \{m_i\}$ *for* $i = 1, \ldots, \nu$, $\lambda_{\omega_i} = M$ *for* $i > \nu$ *and* $q_i \geq q_{i+1}$, *only the safe actions* a_S^0 *and* a_R^0 *are taken in equilibrium.*

Proposition 5 identifies a sufficient condition on the payoff structure that ensures communication failure for a fixed information structure. We now reverse our perspective by considering general information structures, while fixing the payoff structure. Our goal is to identify a condition on the information structure that is necessary for communication at some information state, given the payoff structure in Figure 1. The key will be that there is at least one message for which it is common p-belief for sufficiently high p that the receiver p-believes that this messages is available to the sender. Only then is it possible to use the other message without fear that it may be "contaminated" by the possibility that it is sent out of necessity rather than deliberately.

As a warmup and to introduce the definition of common p-belief due to Monderer and Samet [20], we start by verifying that for the information structure employed in Example 4, common p-belief that the receiver p-believes message m is available fails at every information state for both $m = m_\alpha$ and $m = m_\beta$, for $p \geq 2/3$. Without loss of generality, consider message m_β. Use $E(m)$ to denote the event that message m is available to the sender and for any event F denote by $B_i^p(F)$ the event that player i believes F with at least probability p. Fix $p = 2/3$. Observe that $E(m_\beta) = \{\omega_2, \omega_3, \ldots\}$ and $B_R^p(E(m_\beta)) = \{\omega_4, \omega_5, \ldots\}$. Recall that for an event F to be common p-belief at ω, ω must belong to a p-evident event E at which both players p-believe F. Formally, an event E is ***p-evident*** if $E \subseteq B_i^p(E)$, $i = 1, 2$, and an event F is ***common p-belief*** at state ω if there exists a p-evident event E with $\omega \in E$ and $E \subseteq B_i^p(F)$, $i = 1, 2$.

Note that for any event F, $B_R^p(B_R^p(F)) = B_R^p(F)$. The condition $E \subseteq B_R^p(E(m_\beta))$, implies that a candidate for the p-evident event E must satisfy $E \subseteq \{\omega_4, \omega_5, \ldots\}$. Also, if $\omega_{2k} \in E$ for $k > 1$, then $\omega_{2k-1} \in E$; otherwise $\omega_{2k} \notin B_S^p(E)$, which would violate $E \subseteq B_S^p(E)$. Similarly, if $\omega_{2k+1} \in E$ for $k > 1$, then $\omega_{2k} \in E$; otherwise $\omega_{2k+1} \notin B_R^p(E)$, which would violate $E \subseteq B_R^p(E)$. Taken together, these two observations imply that we must have $\omega_3 \in E$, which results in a contradiction. Hence, in the example there is no state ω at which the event $B_R^p(E(m_\beta))$ is common p-belief for $p = 2/3$.

For the Steiner-Stewart payoff structure (Figure 2) we constructed an equilibrium in which at some information states the sender uses message m_α to signal credibly that he will take action α_S. In that equilibrium m_α is the only message that ever indicates that the sender will take action α_S and in that sense message meaning is consistent across information states. We call such equilibria semantically uniform:

Definition 3 *In a **semantically uniform equilibrium**, if there is a state ω at which (m, α_S) has positive probability, then (m', α_S) has probability zero for $m' \neq m$ at all $\omega' \in \Omega$.*

Our next result establishes necessary conditions for the existence of semantically uniform equilibria for the payoff structure in Example 4 in which there is effective communication with positive probability.

Proposition 6 *For the payoff structure in Figure 1, existence of a semantically uniform equilibrium in which there is a state ω at which the action pair (α_S, α_R) has positive probability requires that for $p \geq 10/11$ there is a message $m \in \{m_\alpha, m_\beta\}$ for which the event $B_R^p(E(m))$ is common p-belief at ω.*

Note that the message m in this result should be thought of as the message that induces β_R in equilibrium. Intuitively, R responds to message $\tilde{m} \neq m$ with α_R only if he is sufficiently certain that the alternative message m was available and therefore $\tilde{m}$ was not sent out of necessity.

6 Discussion

We have shown that failures of higher-order knowledge of language may result in suboptimal language use in optimal equilibria of common-interest games, validating the concerns of Lewis [19] and others. Spelling out these failures in fully specified games shows, however, that there is an important distinction between common-interest communication games and more general classes of communication games. In common-interest communication games the *ex ante* payoff loss from lack of higher-order knowledge of language is negligible, while in richer settings failure of higher-order knowledge of language may result in complete communication collapse.

References

[1] Aumann, Robert J. [1976], "Agreeing to Disagree," *Annals of Statistics* **4**, 1236–1239.

[2] Aumann, Robert J. [1990], "Nash Equilibria are not Self-Enforcing," in *Economic Decision-Making: Games, Econometrics and Optimization* (J. J. Gabszewicz, J.-F. Richard, and L. A. Wolsey, Eds.) Elsevier, Amsterdam.

[3] Baliga, Sandeep and Stephen Morris [2002], "Co-ordination, Spillovers, and Cheap Talk," *Journal of Economic Theory* **105**, 450-468.

[4] Blume, Andreas, Douglas V. DeJong, Yong-Gwan Kim and Geoffrey B. Sprinkle [1998], "Experimental Evidence on the Evolution of Meaning of Messages in Sender-Receiver Games," *American Economic Review* **88**, 1323–1340.

[5] Blume, Andreas, Oliver J. Board and Kohei Kawamura [2007], "Noisy Talk," *Theoretical Economics* **2**, 395–440.

[6] Blume, Andreas and Oliver J. Board [2013], "Language Barriers," *Econometrica*, forthcoming.

[7] Clark, Robin L. [2012], *Meaningful Games: Exploring Language with Game Theory*, MIT Press, Cambridge, MA.

[8] Clark, Herbert H. and Susan E. Brennan [1991], "Grounding in Communication," in Lauren B. Resnick, John M. Levine, and Stephanie D. Teasley (editors) *Perspectives on Socially Shared Cognition*, American Psychological Association, Washington, 127–149.

[9] Clark, Herbert H. and Catherine R. Marshall [1978], "Reference Diaries," in D.L. Waltz (editor) *Theoretical Issues in Natural Language Processing–2*, Association for Computing Machinery, 57–63.

[10] Clark, Herbert H. and Catherine R. Marshall [1981], "Definite Reference and Mutual Knowledge," in A.K. Joshi, B.L. Webber and I.A. Sag (editors) *Elements of Discourse Understanding*, Cambridge University Press, Cambridge, UK.

[11] Clark, Herbert H. and Deanna Wilkes-Gibbs [1986], "Referring as a Collaborative Process," *Cognition* **22**, 1–39.

[12] Crawford, V.P. and J. Sobel [1982], "Strategic Information Transmission," *Econometrica* **50**, 1431–1451.

[13] Farrell, Joseph and Matthew Rabin [1996], "Cheap Talk," *Journal of Economic Perspectives* **10**, 103–118.

[14] Friedell, Morris [1967], "On the Structure of Shared Awareness," Working paper 27 of the Center for Research on Social Organizations, University of Michigan.

[15] Friedell, Morris [1969], "On the Structure of Shared Awareness," *Behavioral Science* **14**, 28–39.

[16] Grice, H. Paul [1975], "Logic and Conversation," in P. Cole and J. Morgan (editors) *Syntax and Semantics*, vol. 3, Academic Press.

[17] Karttunen, L. and S. Peters [1975] "Conventional Implicature in Montague Grammar," in *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*, University of California–Berkeley, 266–278.

[18] Krauss, Robert M. and Sidney Weinheimer [1966], "Concurrent Feedback, Confirmation, and the Encoding of Referents in Verbal Communication" *Journal of Personality and Social Psychology* **4**, 343–346.

[19] Lewis, David [1969], *Convention: A Philosophical Study*, Harvard University Press, Cambridge, MA.

[20] Monderer, Dov and Dov Samet [1989], "Approximating Common Knowledge with Common Beliefs," *Games and Economic Behavior* **1**, 170–190.

[21] Morris, Stephen [2002], "Coordination, Comunication and Common Knowledge: A Retrospective on the Electronic-Mail Game," *Oxford Review of Economic Policy* **18**, 433-445.

[22] Quine, Willard van Orman [1960], *Word & Object*, MIT Press, Cambridge, MA.

[23] Rubinstein, Ariel [1989], "The Electronic Mail Game: Strategic Behavior under Almost Common Knowledge," *American Economic Review* **79**, 385–391.

[24] Schiffer, Stephen R. [1972], *Meaning*, Clarendon Press, Oxford, UK.

[25] Stalnaker, Robert C. [1978], "Assertion," *Syntax and Semantics* **9**, 315–332.

[26] Steiner, Jakub and Colin Stewart [2011], "Communication, Timing, and Common Learning," *Journal of Economic Theory* **146**, 230-247.

[27] Thompson, Will and Stefan Kaufmann [2010], "Signaling games with partially observable actions as a model of conversational grounding," Interactive Decision Theory and Game Theory Workshop, Twenty-Fourth Conference on Artificial Intelligence (AAAI-10), Atlanta, GA.

[28] Weber, Roberto A. and Colin F. Camerer [2003], "Cultural Conflict and Merger Failure: An Experimental Approach," *Management Science* **49**, 400-415.

A Proofs

Proof of Lemma 1

Proof: Let σ_1 be an arbitrary informative order-preserving strategy of the sender. Assume without loss of generality that at every ω where m_2 is available and sent, $\Theta(m_2, \omega_k) > \Theta(m_1, \omega_k)$. Let $\Omega_1^S, \ldots, \Omega_L^S$ be an enumeration of the elements of $\mathcal{O}^S$, and $\Omega_1^R, \ldots, \Omega_M^R$ an enumeration of the elements of $\mathcal{O}^R$. Since $\sigma_1(t, \omega)$ is constant across Ω_ℓ^S for each $t \in T$, for $\ell = 1, \ldots, L$, we can write $\Theta(m, \Omega_\ell^S)$ and denote the $\sup \Theta(m_1, \Omega_\ell^S) = \inf \Theta(m_2, \Omega_\ell^S)$ by θ_ℓ. For all $\omega \in \Omega_j^R$, let $q_{j\ell}$ denote the receiver's posterior belief that $\omega \in \Omega_\ell^R$. Then the receiver's best reply to message m_1 is

$$a_j^1 = \frac{\sum_{\ell=1}^{L} q_{j\ell}\theta_\ell \frac{\theta_\ell}{2}}{\sum_{\ell=1}^{L} q_{j\ell}\theta_\ell}$$

and his best reply to message m_2 is

$$a_j^2 = \frac{\sum_{\ell=1}^{L} q_{j\ell}(1-\theta_\ell)\frac{1+\theta_\ell}{2}}{\sum_{\ell=1}^{L} q_{j\ell}(1-\theta_\ell)}$$

as long as the denominators are well defined.

Notice that for all ℓ, $\frac{\theta_\ell}{2} \leq \frac{1}{2}$ and $\frac{1+\theta_\ell}{2} \geq \frac{1}{2}$, and since σ_1 is informative, there is at least one ℓ' for which $\frac{\theta'_\ell}{2} < \frac{1}{2}$ and $\frac{1+\theta'_\ell}{2} > \frac{1}{2}$. Therefore at every Ω_j^R at which the receiver expects to receive both messages with positive probability any best reply by the receiver satisfies $a_j^1 < a_j^2$. For every other $\Omega_{j'}^R$ one of the actions is equal to $\frac{1}{2}$ and we are free to choose the other action so that the $a_j'^1 < a_j'^2$ holds.

Hence there exists be a best reply ρ_1 of the receiver to σ_1 that satisfies the property that $a_j^1 < a_j^2$ all $j = 1, \ldots, M$. Call any receiver strategy with this property *order preserving*. Note that the payoff from (σ_1, ρ_1) exceeds the payoff from pooling.

At any element Ω_ℓ^S of her information partition the sender has a posterior belief $\phi_{\ell j}$ that the receiver's information is given by Ω_j^R. Therefore, for a sender with payoff type t and information Ω_ℓ^S, the payoff difference between sending message m_2 and m_1 is given by

$$\begin{aligned} &-\sum_{j=1}^{J}(a_j^2 - t)^2\phi_{\ell j} + \sum_{j=1}^{J}(a_j^1 - t)^2\phi_{\ell j} \\ &= \mathbb{E}[a^2|m_1] - \mathbb{E}[a^2|m_2] + 2t(\mathbb{E}[a|m_2] - \mathbb{E}[a|m_1]), \end{aligned}$$

Since the receiver strategy ρ_1 is order-preserving, it follows that $\mathbb{E}[a|m_2] > \mathbb{E}[a|m_1]$ and therefore the sender's best reply σ_2 to ρ_1 is order-preserving.

Continuing in this manner we can construct a sequence of order-preserving strategy pairs $\{(\sigma_n, \rho_n)\}$. Note that the common *ex ante* payoffs along this sequence are increasing and since payoffs are bounded converge to a limit payoff, $\overline{U}$. Note that each strategy pair (σ_n, ρ_n) can be viewed as an element of a compact Euclidean space. Therefore the sequence $\{(\sigma_n, \rho_n)\}$ has a convergent subsequence. Reindex, so that now $\{(\sigma_n, \rho_n)\}$ stands for that subsequence. Denote the limit of that subsequence by $(\overline{\sigma}, \overline{\rho})$. Suppose that $(\overline{\sigma}, \overline{\rho})$ is not an equilibrium. Then one of the players has a best reply that raises the payoff above $\overline{U}$. Continuity of payoffs then implies that for large enough n payoffs from (σ_n, ρ_n) have to be above $\overline{U}$, which leads to a contradiction. Since the payoff from (σ_1, ρ_1) exceeds the payoff from pooling and since payoffs are nondecreasing along the sequence $\{(\sigma_n, \rho_n)\}$, the payoff from $(\overline{\sigma}, \overline{\rho})$ exceeds the payoff from pooling. Hence the equilibrium $(\overline{\sigma}, \overline{\rho})$ is informative. □

Proof of Proposition 1

Proof: Given that attention is restricted to order-preserving equilibria, it is without loss of generality to focus on equilibria in which for every sender with information Ω_ℓ^S there exists $\theta_\ell \in [0, 1]$ such that every payoff type $t < \theta_\ell$ sends message m_1 and every payoff type $t > \theta_\ell$ sends message m_2.

Since Ω is finite, we can define $\underline{\theta} := \min\{\theta_\ell | \theta_\ell > 0\}$. Note that the set $\{\theta_\ell | \theta_\ell > 0\}$ is nonempty because there is at least one state at which it is not common knowledge that message m_2 is available (hence there must be a state at which only m_1 is available).[4]

At every information state ω_i at which message m_1 is sent with positive probability all payoff types $t < \underline{\theta}$ (and possibly others) send message m_1 with probability one. Hence, for every receiver type Ω_j^R who expects to receive both messages with positive probability the response a_j^1 to receiving message m_1 satisfies $a_j^1 \geq \underline{a}_1 = \frac{\underline{\theta}}{2}$. Since m_1 is always available, for every receiver type Ω_j^S who expects to receive both messages with positive probability a message m_2 indicates that the sender's payoff type is in a set of the form $(\theta_j, 1]$ with $\theta_j \geq \underline{\theta}$. Hence, for every receiver type Ω_j^R who expects to receive both messages with positive probability a_j^2 satisfies $a_j^2 \geq \underline{a}_2 = \frac{1+\underline{\theta}}{2}$.

For every information type Ω_ℓ^S of the sender who sends message m_1 with positive probability, $\theta_\ell > 0$. Thus, $\underline{\theta}$ either equals one, or is realized at an information state of the sender where she sends both messages with positive probability. Assume that Ω_ℓ^S is such an infor-

[4]The reason for restricting attention to this set is that for example there may be an isolated information state at which both messages are available and only message m_2 is used with positive probability.

mation state, i.e. the sender of type $(\underline{\theta}, \Omega^S_\ell)$ is indifferent between the lottery over actions induced by message m_1, with payoff $-\sum_{j=1}^{J}(a^1_j - \underline{\theta})^2\phi_{\ell j}$, and the lottery over actions induced by m_2, with payoff $-\sum_{j=1}^{J}(a^2_j - \underline{\theta})^2\phi_{\ell j}$. Note that for any j with $\phi_{\ell j} > 0$ it is the case that $a^2_j > a^1_j$ and $a^2_j > \underline{\theta}$. Consider two cases: If $a^1_j \geq \underline{\theta}$, then $-(a^1_j - \underline{\theta})^2 > -(a^2_j - \underline{\theta})^2$. If $a^1_j < \underline{\theta}$ and $\underline{\theta} < \frac{1}{2}$, then $-(a^1_j - \underline{\theta})^2 \geq -(\frac{\underline{\theta}}{2} - \underline{\theta})^2 > -(\frac{1+\underline{\theta}}{2} - \underline{\theta})^2 \geq -(a^2_j - \underline{\theta})^2$. Therefore, if we had $\underline{\theta} < \frac{1}{2}$, type $\underline{\theta}$ would strictly prefer to send message m_2, which contradicts our assumption that type $(\underline{\theta}, \Omega^S_\ell)$ is indifferent. Therefore, we conclude that $\underline{\theta} \geq \frac{1}{2}$.

Suppose there is an information state ω_i with $\theta_i = \frac{1}{2}$ (i.e. the sender is using an optimal language at ω_i) where it is not common knowledge that $\lambda_{\omega_i} = \{m_1, m_2\}$. Then from above $\theta_i = \underline{\theta} = \frac{1}{2}$. Observe that in order for $\theta_i = \underline{\theta} = \frac{1}{2}$ the receiver's response $a^1_{j'}$ at $\Omega^R(\omega_i)$ to m_1 must be $\frac{\underline{\theta}}{2}$ and the response $a^2_{j'}$ to m_2 must be $\frac{1+\underline{\theta}}{2}$. Otherwise, since for all j, $a^1_j \in [\frac{\underline{\theta}}{2}, \frac{1}{2}]$ and $a^2_j \in [\frac{1+\underline{\theta}}{2}, 1]$ we would have $-(a^1_j - \underline{\theta})^2 \geq -(\frac{\underline{\theta}}{2} - \underline{\theta})^2 = -(\frac{1+\underline{\theta}}{2} - \underline{\theta})^2 \geq -(a^2_j - \underline{\theta})^2$ for all j and at least one of the two inequalities strict for j' and therefore sender type $(\Omega^S(\omega_i), \underline{\theta})$ would strictly prefer to send message m_1.

Call an information state ω_j *adjacent* to ω_i if there exists $\omega_l \in \Omega^R(\omega_i)$ such that $\omega_j \in \Omega^S(\omega_l.)$ At every state ω_j that is adjacent to ω_i, it must be the case that $\theta_j = \frac{1}{2}$. Otherwise the receiver with type $\Omega^R(\omega_i)$ will take actions $a^1_{j'} > \frac{\underline{\theta}}{2}$ and $a^1_{j'} > \frac{1+\underline{\theta}}{2}$, which would be inconsistent with $\theta_i = \underline{\theta} = \frac{1}{2}$. If it is not common knowledge at ω_i that $\lambda_{\omega_i} = \{m_1, m_2\}$, then there exists a chain of states $(\omega_1, \ldots, \omega_i)$ with the property that any two consecutive elements in the chain are adjacent, $\lambda_{\omega_l} = \{m_1, m_2\}$ for all $l \neq 1$ and $\lambda_{\omega_1} = \{m_1\}$. By induction, at every information state in the chain we must have payoff types $t > \frac{1}{2}$ sending message m_1 and payoff types $t < \frac{1}{2}$ sending message 2. But this contradicts $\lambda_{\omega_1} = \{m_1\}$. □

Proof of Lemma 2

Proof: Given that we have a common-interest game, an optimal strategy profile will be an equilibrium profile, and hence trivially an optimal equilibrium profile. With this in mind, it suffices to show that an optimal strategy profile (σ^*, ρ^*) exists. We will decompose this problem into first showing that there is an optimal sender-strategy, $\sigma(\rho)$, for every ρ and then showing that the problem of maximizing over ρ has a solution. Given concavity, the receiver cannot gain from randomization. Therefore it suffices to restrict attention to pure receiver strategies ρ. Given receiver strategy ρ, if there exists a sender strategy $\sigma(\rho)$ that solves

$$\max_{a_S \in A_S, m \in \lambda_\omega} \mathbb{E}\{U(a_S, \rho(m, \mathcal{O}^R(\omega)), t) | \mathcal{O}^S(\omega)\} \quad (1)$$

for every $(t, \mathcal{O}^S(\omega)$, i.e. is interim optimal, then this strategy is an *ex ante* optimal sender response to ρ. The problem in (1) has a solution since for each $m \in \lambda_\omega$,

$$\mathbb{E}\{U(a_S, \rho(m, \mathcal{O}^R(\omega)), t)|\mathcal{O}^S(\omega)\}$$

is a continuous function of a_S on a compact set, and the set λ_ω is finite. Denote the set of $\mathcal{O}^R$-measurable functions in $A_R^{M\times\Omega_N}$ by $A_R^{M\times\Omega_N}(\mathcal{O}^R)$. The set $A_R^{M\times\Omega_N}(\mathcal{O}^R)$ is the set of receiver strategies. Since A_R is compact, so is $A_R^{M\times\Omega_N}(\mathcal{O}^R)$. The problem of finding a strategy combination that maximizes the common *ex ante* payoff of sender and receiver now reduces to solving

$$\max_{\rho\in A_R^{M\times\Omega_N}(\mathcal{O}^R)} Q(\rho) = \sum_{\omega\in\Omega_N} q_\omega \int_T \max_{a_S\in A_S, m\in\lambda_\omega} \mathbb{E}\{U(a_S, \rho(m, \mathcal{O}^R(\omega)), t)|\mathcal{O}^S(\omega)\}dF.$$

Since U and the max operator are continuous functions, the integrand in the expression defining the function Q is continuous and therefore by the Lebesgue dominated convergence theorem, Q is continuous. Therefore, Q achieves a maximum on the compact set $A_R^{M\times\Omega_N}(\mathcal{O}^R)$. □

Proof of Proposition 5

Proof: Given an equilibrium strategy pair (σ, ρ), and somewhat economizing on notation, use $\mathrm{P}(t|m)$ to denote the receiver's posterior probability of payoff type t conditional on having observed message m at information set $\Omega^R(\omega_1)$. For any two events E and F, use $E \cap F$ to denote the joint event that both E and F occurred. Sightly abusing notation write ω_m for the event that $\lambda_\omega = \{m\}$, i.e. the sender only has message m available. Then

$$\begin{aligned}
P(t|m) &= \frac{P(m|t\cap\omega_{\nu+1})P(t\cap\omega_{\nu+1}) + P(m|t\cap\omega_m)P(t\cap\omega_m)}{\sum_\tau P(m|\tau\cap\omega_{\nu+1})P(\tau\cap\omega_{\nu+1}) + \sum_\tau P(m|\tau\cap\omega_m)P(\tau\cap\omega_m)} \\
&= \frac{P(m|t\cap\omega_{\nu+1})\pi(t)q_{\nu+1} + P(m|t\cap\omega_m)\pi(t)q_m}{\sum_\tau P(m|\tau\cap\omega_{\nu+1})\pi(\tau)q_{\nu+1} + \sum_\tau P(m|\tau\cap\omega_m)\pi(\tau)q_m} \\
&= \frac{P(m|t\cap\omega_{\nu+1})\pi(t)q_{\nu+1} + \pi(t)q_m}{\sum_\tau P(m|\tau\cap\omega_{\nu+1})\pi(\tau)q_{\nu+1} + q_m} \\
&\leq \frac{\pi(t)q_{\nu+1} + \pi(t)q_m}{\pi(t)q_{\nu+1} + q_m} \\
&\leq \frac{\pi(t)q_m + \pi(t)q_m}{\pi(t)q_m + q_m} \\
&= \frac{2\pi(t)}{1+\pi(t)}
\end{aligned}$$

Hence the safe-action condition implies that for all $\omega \in \mathcal{O}^R(\omega_1)$, regardless of the message observed, the receiver's unique optimal reply is the safe action a_0^R.

At $\omega_{\nu+1}$ and $\omega_{\nu+2}$ the sender assigns posterior probability at least 1/2 to state $\omega_{\nu+1}$. Therefore, and since we just showed that at $\omega_{\nu+1}$ the receiver uses action a_0^R exclusively, by the sender-safe-response condition, at $\omega_{\nu+1}$ and $\omega_{\nu+2}$ the sender will use action a_S^0 regardless of her payoff type t.

Suppose there exists a message m' such that following m' at $\omega \in \mathcal{O}^R(\omega_{\nu+2})$ the receiver takes an action other than a_R^0 with positive probability, i.e. $\rho(a_R^0 \mid m', \omega_{\nu+2})) < 1$. Let $T_D \subset T$ denote the set of payoff types for whom a_S^0 is dominant and $T_N = T \setminus T_D$. Let $\tilde{M} \subset M$ be the set of messages that induce receiver actions other than a_R^0 with positive probability at $\omega \in \mathcal{O}^R(\omega_{\nu+2})$. Then the no-self-signaling condition implies that at $\omega \in \mathcal{O}^S(\omega_{\nu+1})$ all types in T_N send messages that induce actions other than a_R^0 with positive probability, i.e.

$$\sum_{m\in\tilde{M}} \text{Prob}(m|T_N \cap \mathcal{O}^S(\omega_{\nu+1})) = 1.$$

Thus, since

$$\sum_{m\in\tilde{M}} \text{Prob}(m|T_N \cap \mathcal{O}^S(\omega_{\nu+3})) \leq 1,$$

there exists $\tilde{m} \in \tilde{M}$ such that

$$\text{Prob}(\tilde{m}|T_N \cap \mathcal{O}^S(\omega_{\nu+1})) \geq \text{Prob}(\tilde{m}|T_N \cap \mathcal{O}^S(\omega_{\nu+3})).$$

Together with $q_i \geq q_{i+1}$ for all i this implies that

$$\begin{aligned}
\text{Prob}(\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+1})) &= \text{Prob}(\tilde{m} | T_N \cap \mathcal{O}^S(\omega_{\nu+1}))\text{Prob}(T_N \cap \mathcal{O}^S(\omega_{\nu+1})) \\
&\geq \text{Prob}(\tilde{m} | T_N \cap \mathcal{O}^S(\omega_{\nu+3}))\text{Prob}(T_N \cap \mathcal{O}^S(\omega_{\nu+3})) \\
&= \text{Prob}(\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+3}))
\end{aligned}$$

Therefore, again economizing on notation, if we let $\text{Prob}(a_S^0|\tilde{m})$ denote the receiver's posterior probability of the sender taking action a_S^0 conditional on having observed message $\tilde{m}$ at $\omega \in \mathcal{O}^R(\omega_{\nu+2})$, then

$$\begin{aligned}
\text{Prob}(a_S^0|\tilde{m}) &= \text{Prob}(a_S^0|\tilde{m} \cap T_D)\frac{\text{Prob}(\tilde{m} \cap T_D)}{\text{Prob}(\tilde{m})} \\
&+ \text{Prob}(a_S^0|\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+1}))\frac{\text{Prob}(\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+1}))}{\text{Prob}(\tilde{m})} \\
&+ \text{Prob}(a_S^0|\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+3}))\frac{\text{Prob}(\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+3}))}{\text{Prob}(\tilde{m})} \\
&= \text{Prob}(a_S^0|\tilde{m} \cap T_D)\frac{\text{Prob}(\tilde{m} \cap T_D)}{\text{Prob}(\tilde{m})} \\
&+ \left(1 - \frac{\text{Prob}(\tilde{m} \cap T_D)}{\text{Prob}(\tilde{m})}\right) \times \\
&\quad \Bigg\{\text{Prob}(a_S^0|\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+1}))\frac{\text{Prob}(\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+1}))}{\text{Prob}(\tilde{m}) - \text{Prob}(\tilde{m} \cap T_D)} \\
&\quad + \text{Prob}(a_S^0|\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+3}))\frac{\text{Prob}(\tilde{m} \cap T_N \cap \mathcal{O}^S(\omega_{\nu+3}))}{\text{Prob}(\tilde{m}) - \text{Prob}(\tilde{m} \cap T_D)}\Bigg\} \\
&\geq \frac{1}{2}
\end{aligned}$$

This, however, implies by the receiver-safe-response condition that following message $\tilde{m}$ at $\omega \in \mathcal{O}^R(\omega_{\nu+2})$ the receiver takes action a_R^0 with probability one, contradicting the fact that $\rho(a_R^0|\tilde{m}, \mathcal{O}^R(\omega_{\nu+2})) < 1$.

Suppose that for $k \geq 1$ we have $\rho_R(a_R^0|m, \mathcal{O}^R(\omega_{\nu+2k})) = 1$ for all $m \in M$. Then, using the same logic as above, $\rho_S(a_S^0|\mathcal{O}^S(\omega_{\nu+2k+1})) = 1$ by the sender-safe-response condition, from which we get $\rho_R(a_R^0|m, \mathcal{O}^R(\omega_{\nu+2k+2})) = 1$ for all $m \in M$ by the no-self-signaling and receiver-safe-response conditions. Therefore, by induction, $\rho_R(a_R^0|m, \mathcal{O}^R(\omega_i)) = 1$ for all $m \in M$ and all i and $\rho_S(a_S^0|\mathcal{O}^S(\omega_i)) = 1$ for all i. □

Proof of Proposition 6

Proof: Without loss of generality, suppose that there is a semantically uniform equilibrium, $\mathcal{E}$, in which at state ω the triple $(m_\alpha, \alpha_S, \alpha_R)$ has positive probability. We will show that $\omega \in B_R^p(E(m_\beta))$ and the event $B_R^p(E(m_\beta))$ is common p-belief at ω for $p = 10/11$; the only other possibility for an equilibrium in which (α_S, α_R) has positive probability is for $(m_\beta, \alpha_S, \alpha_R)$ to have positive probability, in which case an argument that exactly mirrors the one given below would show that $\omega \in B_R^p(E(m_\alpha))$ and that the event $B_R^p(E(m_\alpha))$ would be common p-belief at ω.

Define $p_0 := \mathrm{P}(t_1 \cap \alpha_S | m_\alpha \cap \mathcal{O}^R(\omega))$. Note that if R chooses action α_R, then he receives a payoff of 3 if the event $t_1 \cap \alpha_S$ is realized and -10 otherwise. Similarly, if R chooses action β_R, then he receives a payoff of 2 if the event $t_1 \cap \alpha_S$ is realized and 1 otherwise; the latter observation uses the fact that the sender never uses α_S in state t_2 since to do so would be strictly dominated. Then a necessary condition for R to take action α_R at state ω following message m_α in equilibrium is that

$$p_0 \cdot 3 + (1 - p_0) \cdot (-10) \geq p_0 \cdot 2 + (1 - p_0) \cdot 1,$$

which is equivalent to $p_0 \geq \frac{11}{12}$.

Let $A(m_\alpha)$ denote the event that only m_α is available and denote by p_1 the probability of its complement $E(m_\beta)$. Note that

$$
\begin{aligned}
p_0 &= \mathrm{P}(\alpha_S \cap t_1 | m_\alpha \cap \mathcal{O}^R(\omega)) \\
&= \frac{\mathrm{P}(\alpha_S \cap t_1 \cap m_\alpha \cap \mathcal{O}^R(\omega))}{\mathrm{P}(m_\alpha \cap \mathcal{O}^R(\omega))} \\
&= \frac{\mathrm{P}(\alpha_S \cap t_1 \cap m_\alpha | \mathcal{O}^R(\omega))\mathrm{P}(\mathcal{O}^R(\omega))}{\mathrm{P}(m_\alpha | \mathcal{O}^R(\omega))\mathrm{P}(\mathcal{O}^R(\omega))} \\
&= \frac{\mathrm{P}(\alpha_S \cap t_1 \cap m_\alpha | \mathcal{O}^R(\omega))}{\mathrm{P}(m_\alpha | \mathcal{O}^R(\omega)))} \\
&= \frac{\mathrm{P}(\alpha_S \cap t_1 \cap m_\alpha | \mathcal{O}^R(\omega))}{\mathrm{P}(\alpha_S \cap t_1 \cap m_\alpha | \mathcal{O}^R(\omega))) + \mathrm{P}((\alpha_S \cap t_1)^C \cap m_\alpha | \mathcal{O}^R(\omega)))} \\
&\leq \frac{\mathrm{P}(t_1 | \mathcal{O}^R(\omega))}{\mathrm{P}(t_1 | \mathcal{O}^R(\omega))) + \mathrm{P}((\alpha_S \cap t_1)^C \cap m_\alpha | \mathcal{O}^R(\omega)))} \\
&\leq \frac{\mathrm{P}(t_1 | \mathcal{O}^R(\omega))}{\mathrm{P}(t_1 | \mathcal{O}^R(\omega))) + \mathrm{P}(t_2 \cap m_\alpha | \mathcal{O}^R(\omega)))} \\
&\leq \frac{\mathrm{P}(t_1 | \mathcal{O}^R(\omega))}{\mathrm{P}(t_1 | \mathcal{O}^R(\omega))) + \mathrm{P}(t_2 \cap A(m_\alpha) | \mathcal{O}^R(\omega)))} \\
&= \frac{1}{1 + \mathrm{P}(A(m_\alpha) | \mathcal{O}^R(\omega)))} \\
&= \frac{1}{2 - p_1}
\end{aligned}
$$

This implies that in order for the condition $\text{Prob}(\alpha_S \cap t_1 | m_\alpha \cap \mathcal{O}^R(\omega)) \geq \frac{11}{12}$ to be satisfied, it is necessary that $p_1 \geq \frac{10}{11}$. Hence, a necessary condition for the existence of the equilibrium in question is that

$$\omega \in B_R^p(E(m_\beta)) \text{ for } p = 10/11. \tag{2}$$

For the remainder, let $p = 10/11$. Let $C = C^0 = B_R^p(E(m_\beta))$ and for $n \geq 1$ let $C^n = \bigcap_{i \in \{S,R\}} B_i^p(C^{n-1})$. Suppose that for some event E it is the case that $\omega \in B_R^p(E)$ is required for R to be willing to take action α_R in response to message m_α at ω. In order for S to be willing to send message m_α and take action α_S, he must be sufficiently confident that R responds with α_R. The same calculation (slightly differently motivated) as the one that gave us the bound on p_0 shows that S must believe with at least probability 11/12 that m_α induces α_R in order to be willing to take action α_S in payoff state t_1. Therefore for any event E with the property that $\omega \in B_R^p(E)$ is required for R to be willing to take action α_R in response to message m_α at ω, we need that $\omega \in B_S^{p_2}(B_R^p(E))$ for $p_2 \geq 11/12$. Furthermore, since for any event $\tilde{E}$, $B_R^p(B_R^p(\tilde{E})) = B_R^p(\tilde{E})$, and for any $p \leq p_2$, $[\omega \in B_S^{p_2}(B_R^p(E))] \Rightarrow [\omega \in B_S^p(B_R^p(E))]$, we conclude that if $\omega \in B_R^p(E)$ is necessary for the equilibrium in question, then so is

$$\omega \in \bigcap_{i\in\{S,R\}} B_i^p(B_R^p(E)), \text{ for } p = 10/11, \tag{3}$$

from which it follows that

$$\omega \in C^1 \tag{4}$$

is required for the postulated equilibrium. Suppose that for some $n \geq 2$ we have that $\omega \in C^{n-1}$ is necessary for the equilibrium in question. It follows that

$$\omega \in B_S^p(C^{n-2}) \tag{5}$$

and

$$\omega \in B_R^p(C^{n-2}) \tag{6}$$

are necessary.

For any event E with the property that $\omega \in B_S^p(E)$ is required for S to be willing to take action α_S at ω, at every state $\omega' \notin B_S^p(E)$, S strictly prefers β_S regardless of the payoff state $t_i, i = 1, 2$. At every state ω' that R assigns positive probability to at ω (that is $\omega' \in \mathcal{O}^R(\omega)$) the sender can induce α_R with positive probability by sending message m_α. By semantic uniformity of the equilibrium, there is no state at which m_β induces α_R with positive probability. Combining these three observations, it follows that at every state $\omega' \in \overline{B_S^p(E)} \cap \mathcal{O}^R(\omega)$ the violation of the self-signaling condition implies that in payoff state t_1 the sender will send message m_α and take action β_S. This in turn implies that it is necessary for the equilibrium in question that $\omega \in B_R^{p_3}(B_S^p(E))$ for $p_3 \geq 11/12$, and by essentially the same argument that led to (3), we conclude that if $\omega \in B_S^p(E)$ is necessary for the postulated equilibrium,then so is

$$\omega \in \bigcap_{i\in\{S,R\}} B_i^p(B_S^p(E)). \tag{7}$$

Combining (3) with (6) it follows that

$$\omega \in \bigcap_{i\in\{S,R\}} B_i^p(B_R^p(C^{n-2})), \tag{8}$$

and combining (7) with (5) it follows that

$$\omega \in \bigcap_{i\in\{S,R\}} B_i^p(B_S^p(C^{n-2})) \tag{9}$$

is required. The conditions (8) and (9) imply then that $\omega \in C^n$.

Hence, by induction

$$\omega \in E^p(C) = \bigcap_{n \geq 1} C^n$$

which according to Monderer and Samet is equivalent to $B_R^p(E(m_\beta))$ being common p-belief. □