WORKING PAPERS

Believing when Credible: Talking about Future Plans

Karl H. Schlag

Péter Vida

November 2014

Working Paper No: 1409

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

Believing when Credible: Talking about Future Plans and Past Actions*

Karl H. Schlag*

Péter Vida,**

November 2, 2014

Abstract

We explore in an equilibrium framework whether games with multiple Nash equilibria are easier to play when players can communicate. We consider two variants, modelling talk about future plans and talk about past actions. The language from which messages are chosen is endogenous, messages are allowed to be vague. We focus on equilibria where messages are believed whenever possible, thereby develop a theory of credible communication. Predictions confirm the longstanding intuition for Aumann's (1990) Stag Hunt game which applies directly to an investment game with positive spillovers. Our results shed new light on the multiplicity of equilibria in economic applications.

Keywords: Pre-play communication, cheap talk, credibility, coordination, language, multiple equilibria.

JEL Classification Numbers: C72, D83.

---

*We would like to thank Attila Ambrus, Stefano Demichelis, Olivier Gossner, Takakazu Honryo, Andras Kornai, Tymofiy Mylovanov, Martin Peitz, Larry Samuelson, Joel Sobel and Thomas Troeger for useful comments. Péter Vida received financial support from SFB/TR 15 which is gratefully acknowledged. An earlier version of this paper had the title "Commitments, Intentions, Truth and Nash Equilibria".

*Corresponding author. University of Vienna, Department of Economics. E-mail: karl.schlag@univie.ac.at

**University of Mannheim, Department of Economics, L7,3-5, D-68131 Mannheim, Germany. E-mail: vidapet@gmail.com

1 Introduction

It is hard to imagine people who are involved in strategic interaction but who do not talk to each other, and if they do not talk one wonders why they don't. In most situations communication is possible. Of course the outcome of communication will depend on how good the players know each other, whether there is trust. It also depends on the degree to which incentives are aligned in the underlying game, whether there is more of a tension or more of a cooperative setting between them. We set out to formulate a simple theory, called *credible communication*, to understand the impact and use of communication in games. We consider games of complete information, so communication is about strategies, not about private information.

We aim to investigate which Nash equilibrium outcomes of the underlying game retain their predictive power when communication is added. Games with multiple Nash equilibria are more the rule than the exception. Multiplicity can have a negative flavor such as when considering mixed equilibria in pure coordination games. Multiplicity has been identified as carrying predictive power in macroeconomics (see Cooper and John (1988)). We question the intuitive content of Nash equilibria that do not persist if only a little communication is added and bring our insights to economic applications (see Section 8).

A central obstacle to setting up a model with predictive power is the possible existence of babbling equilibria. Babbling equilibria describe the self-enforcing situation in which no information is transmitted. Roughly speaking, if players do not believe each other then there is no need to exchange any meaningful information which justifies why they should not believe each other. We postulate a setting in which words have meaning. One player listens and sees whether the other can be believed, being willing to believe the other if this is possible. However, if the other cannot be believed then words are ignored. This nips babbling equilibria in the bud, the listener never ignores meaningful information. Whether or not the other can be believed depends on who this other player is and whether the underlying game contains incentives to lie. To be willing to believe others and to act accordingly makes most sense when players know each other. It also constitutes a benchmark for the interaction between strangers.

Strategic interaction typically takes place in a dynamic context, with a specified order of moves and associated information sets. This introduces many points in time at which communication can take place. It also reveals two very different ways in which one may talk about own play as unobserved by others. One can talk about past actions or commitments when these are hidden, one can talk about future choices or intentions when these choices have not yet been made. We isolate these two extremes to clarify strategic considerations underlying communication. The theory is easily extended to allow for both. To further clarify the impact of communication we choose the simplest model. There are two players, each player moves only once, and communication occurs by one player sending a message to the other player. This can also be thought of as modelling the interaction between two groups or coalitions. Under "talk then play" (TP) player one sends the message before both choose their action simultaneously. In "play then talk" (PT) player one first chooses an action which is not observed by player two, then sends a message and finally player two makes a choice. These two basic models constitute the building blocks for more complicated modelling.

To capture talk of players within a model also means to allow players to determine the degree of information transmitted. As in our simple model player one is communicating about his own choices we model messages as subsets of his set of actions. The message contains no information if it contains all actions of player one. In this case it is as if there is no communication. At the opposite extreme the message describes the choice of player one if it contains only the corresponding action. The messages that can be sent are described by the language. To understand the interplay between what can be said and what is said we separate the choice of a language from the choice of the message itself. Communication as information transmission is captured by postulating that a language is a union of partitions of the set of actions of player one (see the appendix for justification). In particular, this means that player two has to also think about what to do if player one had said something completely different. Whether or not player one can be believed will be associated to the language, not to the message. Player one using a given language can be believed if any of the messages from this language can be believed. The language is then called *credible*. Credible communication involves using a credible language. In our model the language is chosen by one of the two players. The player who is designated to choose the language is called the *interpreter*. It may be natural to consider the one who chooses the message (here player one) to also choose the language. However, often the listener (here player two) chooses the language, such as when employers design questionnaires when interviewing potential employees.[1] Part of our analysis shows how to model communication under a given language and can be applied directly to investigate situations in which the language has been determined exogenously.

We present a simple and workable model with predictions that are easy to derive.

[1]We wish to thank an anonymous referee for this example.

For instance, we obtain the obvious conclusion in pure coordination games, players choose actions associated to the efficient equilibrium (see Remark 4). Adding credible communication eliminates any debate as to whether the inefficient equilibria in this game make any sense from a positive point of view. Yet credible communication about future intentions need not lead to efficiency in common interest games when player one has three or more actions. The counter example contains an action that belongs to the support of any Nash equilibrium. This makes it impossible for player one to credibly talk about not choosing this action. Consequently, player two cannot maintain that player one always tells the truth unless the language contains a single message (it is as if there is no communication). Only talk during play, after player one has chosen but before player two has chosen, will lead to efficiency in any common interest game (see Proposition 4).

The importance of when communication takes place is also revealed in Aumann's (1990) Stag Hunt game which can also be considered as an investment game with positive spillovers (Baliga and Morris (2002)). In this game, the predictions of credible communication coincide with the intuition brought forward by Farrell (1988). Talk about intentions leads players to coordinate on the efficient outcome. Talk about a past move is useless as the player sending the message will not be believed and hence does not help players to select among the Nash equilibria. To be useless means in our model that the equilibrium language contains only one message.

Our insights also pertain to the role of no communication. Communication is not needed when beliefs are such that play unrolls as desired (see arguments leading to Proposition 3). In some games no communication can be the only way for player two to obtain a favorable outcome by removing some of the power of player one who is sending the message (see Remark 9). An extensive analysis of two by two games is provided, insights are presented and results for general games are given (see Section 7). Interesting insights arise when adding communication to economic applications as illustrated in Section 8. It namely turns out that many micro- and macroeconomic models which explain economic inefficiencies by relying on the existence of multiple equilibria can be questioned when one explicitly models what people do naturally, namely talk.

We summarize the related theoretical and experimental literature on communication in Section 9. It is the introduction of an equilibrium framework that differentiates us from most of the literature. It is the focus on believing whenever possible and the endogeneity of languages that differentiates us from the rest and genuinely connects our model to real-life talk. It is the adherence to an equilibrium framework that allows our model to be readily applied in typical economic applications. It is the ability to understand when to communicate that adds behavioral insights that can be tested. Credible communication formally reproduces the intuition of Farrell (1988) and hence also is consistent with the experimental findings of Charness (2000).

The structure of the paper is as follows. Section 2 introduces some basic notations, the notion of languages and messages. In Section 3 we describe the TP game. In Section 4 we describe the PT game. In Section 5 we define credibility of a language under TP and PT, our solution concepts TPE and PTE and credible communication. Section 6 contains an analysis of representative two by two games. In Section 7 we present general results. Section 8 connects our insights to some economic examples. In Section 9 we summarize the related literature on communication. Section 10 concludes. In the appendix we elaborate more on the motivation behind our formal modelling, recall some definitions and provide a proof of Proposition 1.

2 Preliminaries

2.1 The Underlying Game and Elements of Communication

Let Γ be a two player simultaneous move game with finite action sets S_j, $S = S_1 \times S_2$, and von Neumann-Morgenstern utility functions defined by the Bernoulli utilities $u_j : S_1 \times S_2 \to \mathbb{R}$ for player $j = 1, 2$. To ease the presentation we refer to player one as "he" and to player two as "she". For a finite set X let ΔX be the set of probability distributions over X and let $C(\xi) = \{x \in X : \xi(x) > 0\}$ be the support of $\xi \in \Delta X$. $z \in \mathbb{R}^2$ is called an *outcome* of Γ if there exists $\sigma \in \Delta S_1 \times \Delta S_2$ of Γ such that $u_j(\sigma) = z_j$ for $j = 1, 2$, z is a called Nash equilibrium outcome if the corresponding strategy profile σ is a Nash equilibrium of Γ. $z^* \in \mathbb{R}^2$ is called a *favorite Nash equilibrium outcome* for player j if there is no Nash equilibrium outcome z such that $z_j > z_j^*$. A Nash equilibrium outcome that is not pareto inferior to some other Nash equilibrium outcome is referred to as an *efficient Nash equilibrium outcome.*

We consider two different scenarios for how players communicate. In "first talk then play" player one first sends a message to player two and then both simultaneously play Γ. In "first play then talk" player one first privately chooses an action in Γ and then sends a message to player two after which player two chooses an action in Γ.

Messages belong to a language, messages and languages are formally defined below. One can imagine different scenarios for how the language is determined. It may be given by the environment or it may be chosen by someone, by one of the two players or by a third party. We consider here the situation that is most interesting from the strategic point of view, namely where the language is either chosen by the player who sends the message (player one) or by the player who receives the message (player two). The analysis for the other two cases, where the language is given or chosen by a third party, is completely analogous.

A message m is a nonempty subset of S_1, so $m \subseteq S_1$ and $m \neq \emptyset$. A language L is a union of partitions of S_1.[2] In the appendix we give a detailed motivation behind these modelling assumptions. The set of all languages is denoted by $\mathcal{L}$. Languages will be chosen by the interpreter who is one of the two players. We allow for randomizing over languages, hence choices in $\Delta\mathcal{L}$. The degenerate language $\{S_1\}$ that contains a single element can be interpreted as there being no communication. At the opposite extreme, the language that contains all non empty subsets of S_1 may be interpreted as unrestricted communication. The language that consists of all singletons will also be of some importance. These languages will be referred to as *no communication*, *unrestricted communication* and *detailed communication* respectively.

3 First Talk Then Play

We first present a model in which communication occurs before either player chooses an action. First the interpreter chooses the language L. Then player one sends a message m from this language L. Conditional on the language chosen and the message sent, player one chooses an action which is not observed by player two. Finally player two chooses an action.

The above defines the following game, denoted by Γ_i^{TP} for $i = 1, 2$.

1. Player i (the interpreter) chooses a language $L \in \mathcal{L}$ and communicates it to the other player.

2. Player one sends a message $m \in L$ to player two.

3. Player one chooses an action s_1 (non-observable for player two)

4. Player two chooses an action s_2.

5. Payoffs are realized, where player j receives payoff $u_j(s_1, s_2)$, $j = 1, 2$.

[2] $\{m_1, .., m_k\}$ is a partition of S_1 if $\cup_{i=1}^{k} m_i = S_1$ and for all i, j in $\{1, .., k\}$ with $i \neq j$ we have that $m_i \neq \emptyset$, $m_i \subseteq S_1$ and $m_i \cap m_j = \emptyset$.

Let us denote by $\Gamma^{TP}(L)$ the game in which L is given and which starts with stage 2.

3.1 The Strategies in Γ_i^{TP}

We now introduce the notation for the possibly mixed strategies used in Γ_i^{TP}.

Let L_i be the mixed language choice of the interpreter in stage 1, so $L_i \in \Delta\mathcal{L}$. L_i is deterministic if $L_i(L) = 1$ for some $L \in \mathcal{L}$. When L_i is deterministic then we identify L_i with the language L where $L_i(L) = 1$.

Given the language $L \in \mathcal{L}$ chosen by the interpreter in stage 1 let $m_1^L \in \Delta L$ be the mixed message sent by player one in stage 2 and let $m_1 = (m_1^L)_{L\in\mathcal{L}}$.

Let $\sigma_1^L(m)$ be the mixed action of player one in stage 3 after message $m \in L$ has been sent in stage 2, so $\sigma_1^L : L \to \Delta S_1$. Let $\sigma_2^L(m)$ be the mixed action of player two in stage 3 given the language L chosen by the interpreter in stage 1 and the message m received in stage 2, so $\sigma_2^L : L \to \Delta S_2$. We write $\sigma_j = (\sigma_j^L)_{L\in\mathcal{L}}$ for $j = 1, 2$.

Hence, a strategy profile in the game Γ_i^{TP} is a tuple $(L_i, m_1, \sigma_1, \sigma_2)$.

4 First Play then Talk

In this scenario we model communication that takes place after player one has chosen an action. It is analogous to Γ_i^{TP} except the action choice of player one is moved from stage 3 to stage 1. This leads to the following game, denoted by Γ_i^{PT} for $i = 1, 2$.

1. Player one *privately* chooses an action s_1.

2. Player i (the interpreter) publicly chooses a language $L \in \mathcal{L}$.

3. Player one sends a message $m \in L$ to player two.

4. Player two chooses an action $s_2 \in S_2$.

5. Payoffs are realized, where player j receives payoff $u_j(s_1, s_2)$, $j = 1, 2$.

Let us denote by $\Gamma^{PT}(L)$ the game above in which the interpreter *has to* choose L in stage 2, that is L is fixed.

4.1 The Strategies in Γ_i^{PT}

Let $\sigma_1 \in \Delta S_1$ be the mixed action of player one in stage 1.

For $i = 1$ let $L_1(s_1)$ be the mixed language chosen in stage 2 after action s_1 has been realized in stage 1, so $L_1 : S_1 \to \Delta\mathcal{L}$. If $i = 2$ then the mixed language choice L_2 does not depend on s_1, so $L_2 \in \Delta\mathcal{L}$.

In stage 3, player one chooses a mixed message $m_1^L(s_1)$ belonging to the language L chosen in stage 2 that depends on the action s_1 that was chosen in stage 1, so $m_1^L : S_1 \to \Delta L$ and $m_1 = (m_1^L)_{L\in\mathcal{L}}$.

In stage 4, player two chooses a mixed action $\sigma_2^L(m)$ that depends on the language L chosen in stage 2 and on the message m received in stage 3, so $\sigma_2^L : L \to \Delta S_2$ and $\sigma_2 = (\sigma_2^L)_{L\in\mathcal{L}}$.

Hence a strategy profile in the game Γ_i^{PT} is described by $(\sigma_1, L_i, m_1, \sigma_2)$.

5 Solution Concepts

In this section we frequently refer to the notion of weak perfect Bayesian equilibrium (Mas-Colell et al. (1995)).[3] Let $\mu_2^L(m) \in \Delta S_1$ indicate player two's belief about player one's action after message $m \in L$. Let $\mu_2^L = (\mu_2^L(m))_{m\in L}$ and $\mu_2 = (\mu_2^L)_{L\in\mathcal{L}}$.

5.1 Credibility

We define the notion of credible languages under TP and PT. These are languages in which it is conceivable that player one can be believed.

Definition 1 *We say that a language L is* ***credible under TP*** *if there is a weak perfect Bayesian equilibrium $(m_1^L, \sigma_1^L, \sigma_2^L, \mu_2^L)$ of $\Gamma^{TP}(L)$ in which player one always tells the truth, and player two always correctly anticipates player one's action, i.e.,*

1. *for all $m \in L$, $C(\sigma_1^L(m)) \subseteq m$ (truth telling)*
2. *for all $m \in L$, $\mu_2^L(m) \in \Delta m$ (believing) and*
3. *for all $m \in L$, $\mu_2^L(m) = \sigma_1^L(m)$ (correctly believing).*

[3]A profile of strategies and system of beliefs is a weak perfect Bayesian equilibrium if the strategy profile is sequentially rational (so all choices are best response to the beliefs and the other player's strategy) and beliefs are derived from strategies on the equilibrium path.

Remark 1 *L is credible under TP if and only if there is a subgame perfect equilibrium $(m_1^L, \sigma_1^L, \sigma_2^L)$ of $\Gamma^{TP}(L)$ in which player one always tells the truth. Note that condition 2 is superfluous, however we keep it to clarify the role of condition 3, namely that we require in addition to telling the truth and believing that player two* always*, and not just on the equilibrium path, correctly anticipates player one's action (point 3).*

Definition 2 *We say that a language L is* ***credible under PT*** *if there is a weak perfect Bayesian equilibrium $(\sigma_1, m_1^L, \sigma_2^L, \mu_2^L)$ of $\Gamma^{PT}(L)$ in which player one tells the truth, and player two believes any message, i.e.,*

1. *for all $s_1 \in S_1$ and all $m \in C(m_1^L(s_1))$, $s_1 \in m_1$ and*
2. *for all $m \in L$, $\mu_2^L(m) \in \Delta m$.*

For a discussion of weaker definitions of credibility under TP and PT see Schlag and Vida (2013).

For both TP and PT it follows directly from the definitions that no communication is always a credible language. Clearly, at the opposite end, detailed communication need not be credible (see examples in Section 6.1). However, when it is credible then this has implications on the other languages. In fact, under TP, if detailed communication is credible, then any other language is also credible. The intuition is simple. For any language, by considering player two who for each message assigns point beliefs to some strategy belonging to the message, one can replicate the incentives under detailed communication (for a formal proof of Proposition 1 below see the appendix). Under PT this result is not true, as revealed by the following example.

	L	R
T	$1,1$	$0,0$
M	$0,0$	$1,1$
N	$1,1$	$0,0$
B	$0,0$	$1,1$

(1)

In the game in (1), detailed communication is credible, yet $\{\{T\},\{M\},\{N,B\}\}$ is not credible. It is however true that unrestricted communication is credible, in this example and more generally in PT whenever detailed communication is credible. This leads us to the following result, its proof is in the appendix.

Proposition 1 *1. Under both TP and PT, detailed communication is credible if and only if unrestricted communication is credible.*

2. *Under TP, any language is credible if detailed communication is credible. This is not true for PT.*

For a language to be credible means that there are beliefs that make player one believable when player one uses messages from this language. If the language available to player one is exogenously given and if this is credible then one can use the conditions in Definitions 1 and 2 to make predictions for what happens when player one communicates to player two and player one is believed. We call this *credible communication with a given language*. If the language available to player one is not credible then credible communication as defined is not possible.

The main emphasis of this paper is to consider the outcomes of communication when the language can be chosen by the so-called interpreter, as modelled by the game forms underlying TP and PT, where the interpreter is either player one or player two. In the next section we present the corresponding equilibrium concepts. These definitions are easily adapted if one instead wishes to introduce a third player who takes over the role of the interpreter.

5.2 TPE

We now present our equilibrium concept for TP. We search for a weak perfect Bayesian equilibrium of Γ_i^{TP} in which communication is truthful and believed when the language is credible, and where messages are ignored otherwise.

Definition 3 (TPE) $(L_i, m_1, \sigma_1, \sigma_2, \mu_2)$ *is called a* ***talk then play equilibrium*** (TPE) *of* Γ_i^{TP} *if it is a weak perfect Bayesian equilibrium of* Γ_i^{TP} *and*

1. L_i *is deterministic and credible,*

2. *if* L *is credible then* $C(\sigma_1^L(m)) \subseteq m$ *and* $\mu_2^L(m) = \sigma_1^L(m)$ *for all* $m \in L$ *(truth-telling and correctly believing),*

3. *if* L *is not credible then* $\sigma_j^L(m) = \sigma_j^{\{S_1\}}$ *for all* $m \in L$ *and* $j = 1, 2$ *(ignorance).*

Remark 2 $(L_i, m_1, \sigma_1, \sigma_2, \mu_2)$ *is a TPE if and only if* $(L_i, m_1, \sigma_1, \sigma_2)$ *is a subgame perfect equilibrium of* Γ_i^{TP} *and points* 1, 2 *and* 3 *in Definition 3 are true.*

The outcome in the underlying game Γ that results under a TPE is called a *TPE outcome*.

5.3 PTE

Our equilibrium concept for PT is analogous to the one for TP. Communication is truthful and believed for credible languages, otherwise all messages are ignored.

Definition 4 (PTE) $(\sigma_1, L_i, m_1, \sigma_2, \mu_2)$ *is called a* ***play then talk equilibrium*** (PTE) *of* Γ_i^{PT} *if it is a weak perfect Bayesian equilibrium of* Γ_i^{PT} *and*

1. L_i *is deterministic, independent*[4] *of* σ_1 *and credible,*
2. *if* L *is credible then*
   (a) $s_1 \in m$ *for all* $m \in C(m_1^L(s_1))$ *and* $s_1 \in S_1$ *(truth-telling) and*
   (b) $\mu_2^L(m) \in \Delta m$ *for all* $m \in L$ *(believing),*
3. *if* L *is not credible then* $\sigma_j^L(m) = \sigma_j^{\{S_1\}}$ *for all* $m \in L$ *and* $j = 1, 2$ *(ignorance).*

The outcome in the underlying game Γ that results under a PTE is called a *PTE outcome.*

5.4 Credible Communication

Above we presented how we model communication with believing whenever possible. We refer to these two models with predictions derived using the corresponding solution concepts TPE and PTE as *credible communication.* We will say that player one (player two as interpreter) can *guarantee his (her) favorite Nash equilibrium outcome* in TP if his (her) favorite Nash equilibrium outcome is the unique TPE outcome (when player two is the interpreter). A similar wording is used when referring to PT.

6 Two by Two Games

Before we present general results we illustrate and contrast the two models in some two by two games. Note that in two by two games there are only three possible languages, no communication $\{\{s_1, s_1'\}\}$, detailed communication $\{\{s_1\}, \{s_1'\}\}$ and unrestricted communication $\{\{s_1\}, \{s_1'\}, \{s_1, s_1'\}\}$, provided $S_1 = \{s_1, s_1'\}$.

[4] Languages should play a secondary role and not interfere with strategic considerations, hence we restrict attention to language choices that do not depend on equilibrium action choices.

6.1 Some Examples

The Prisoners' Dilemma Consider the Prisoners' Dilemma with strategies $\{C, D\}$. Consider first TP. As both players will defect once they play the game, player one will not be believed if he sends $\{C\}$. So only no communication is credible and (D, D) is the unique TPE outcome. Consider now PT. As player two always plays D it does not matter how much information player one reveals. Hence detailed communication is credible. Of course player one will choose D and (D, D) is the unique PTE outcome, which unlike TP can be supported by any of the three languages.

Matching Pennies Consider Matching Pennies. Under TP player one intends to choose a mixed strategy and hence needs to send a message that contains all actions in its support if he wishes to be believable. Under PT player one as interpreter does not wish to reveal which action he has chosen. Player two would like to learn what player one chose, hence would wish as interpreter to choose detailed communication. However this would only result in player one sometimes lying. Hence, only no communication is credible in TP and in PT, and both players choose actions as in the mixed Nash equilibrium of the underlying game Γ.

Aumann's (1990) Stag Hunt game and the Investment Game with Positive Spillovers Baliga and Morris (2002) Consider a version of Baliga and Morris (2002) investment game with positive spillovers which is actually Aumann's (1990) Stag Hunt game

$$\begin{array}{ccc} & \text{S} & \text{R} \\ \text{S} & 9,9 & 0,8 \\ \text{R} & 8,0 & 7,7 \end{array}. \tag{2}$$

Consider first TP. Detailed communication is credible. Sending $\{S\}$ is followed by play of (S, S), sending $\{R\}$ is followed by (R, R). Player one will thus send $\{S\}$ under detailed communication. As this is the favorite Nash equilibrium payoff of each player we obtain a TPE in which either player as interpreter chooses detailed communication. This also implies that (S, S) is the unique TPE outcome. Of course this outcome can also be supported by no communication, provided (S, S) is chosen under no communication. We find that either player as interpreter can guarantee the efficient outcome. Efficiency either emerges by player one remaining silent as players believe this will happen, or by player one announcing play of S.

Now consider PT. Detailed communication is no longer credible. Under detailed communication, if player two believes player one then player two chooses S after $\{S\}$

and R after $\{R\}$. As player one is always better off if player two chooses S, player one will always send message $\{S\}$. In particular, player one will lie after having chosen R. However, credibility requires truth telling after each message. Hence, detailed communication is not credible. Any language is treated as no communication. But this means that no information is transmitted and all three Nash equilibrium are possible PTE outcomes.

Note the difference between TP and PT when choosing a message under detailed communication. In TP we are comparing $u_1(S,S)$ to $u_1(R,R)$ while in PT we are comparing $u_1(s_1,S)$ to $u_1(s_1,R)$.

The Stag Hunt game with a Constant Payoff Action Consider now the more classic version of the Stag Hunt game in which the payoff of R does not depend on what the opponent does. For instance, replace 8 by 7 in the game in (2). Nothing changes in our analysis of TP. But consider PT. There is no longer any incentive for player one to lie after having chosen R. Detailed communication becomes credible. Moreover, as player one is believed under this language, player one can choose S and successfully reveal to player two which action he has chosen. The unique PTE outcome is (S,S). Now both players as interpreter can guarantee the efficient outcome under PT. Of course, as in the case of TP, the efficient outcome can also be supported by no communication, provided player two believing that the efficient outcome will result when there is no communication.

The Hawk Dove game Consider now the Hawk Dove game which is also called the Game of Chicken, a representative is shown below

$$\begin{array}{ccc} & \text{H} & \text{D} \\ \text{H} & -1,-1 & 2,0 \\ \text{D} & 0,2 & 1,1 \end{array}.$$

Consider TP with player one as the interpreter. The analysis is simple. Detailed communication is credible as both (D,H) and (H,D) are Nash equilibria of the underlying game. Player one can guarantee his favorite outcome (H,D). For instance, player one can choose detailed communication and send $\{H\}$, this leads to play of (H,D). Alternatively, player one can choose no communication if beliefs are such that this is followed by (H,D). Consider now TP with player two as the interpreter. All three Nash equilibrium outcomes are TPE outcomes, in particular neither player can guarantee his or her favorite outcome. Play all depends on what happens after no communication. If no communication does not lead to (H,D), which is the worst outcome for player two, then player two will choose no communication. However if no communication leads to (H, D) then player two is indifferent between any of the three languages.

Consider now PT. The strategic analysis is as in Aumann's (1990) Stag Hunt game. Player one is always best off if player two chooses D. Hence, detailed communication is not credible and communication has no bite. All three Nash equilibrium outcomes are possible PTE outcomes.

Pure Coordination Games Consider pure coordination games where both players have the same strategy set and obtain the same payoffs in any of the outcomes and where payoffs are strictly positive if and only if both choose the same action. These games are rather uninspiring to analyze. All TPE and PTE outcomes are efficient. Either player as interpreter can guarantee his or her favorite outcome (which is also the favorite of the other player). Detailed or unrestricted communication needs to be used if beliefs after no communication lead to an inefficient outcome.

Battle of the Sexes Consider Battle of Sexes. Given the postulate that player one is believed whenever believable we obtain that detailed communication is credible in TP and PT. Consider TP. When player one is the interpreter then he can guarantee his favorite Nash equilibrium outcome. Consider instead player two as interpreter. Then the two pure Nash equilibrium outcomes are possible TPE outcomes. If no communication is followed by the favorite Nash equilibrium outcome of player two then this will be the TPE outcome. In all other cases the favorite Nash equilibrium outcome of player one is the TPE outcome. If no communication is followed by play of the mixed Nash equilibrium then player two allows player one to choose his favorite Nash equilibrium outcome by choosing either detailed or unrestricted communication. If no communication is followed by play of player one's favorite Nash equilibrium then player two as interpreter has to give in, any of the three languages will lead to this outcome.

In PT we find the same relationship between outcomes, beliefs after no communication and the identity of the interpreter.

6.2 General Comments on Two by Two Games

We now comment on the regularities in the above examples and whether these are true in general for two by two games. This sets the stage for our more general results presented in the next section. For proofs we either refer to the next section or leave it as an easy exercise for the reader.

Existence TPE and PTE always existed, and this is generally true in two by two games (see Proposition 3).

Nash equilibrium The TPE and PTE outcomes were always a subset of the Nash equilibrium outcomes of the underlying game. This is more generally true for TP and PT in games in which any two Nash equilibria yield different payoffs for player one (see Proposition 2 below), as counter example consider

	L	R
T	$1,1$	$0,0$.
B	$0,0$	$1,2$

For this game there is a TPE in which the interpreter chooses detailed communication, which is credible, then chooses messages $\{T\}$ and $\{B\}$ equally likely. So $\left(1,\frac{3}{2}\right)$ is a TPE outcome. Similarly it is a PTE outcome, supported with a very similar construction. So the TPE outcome and the PTE outcome are a convex combination of the Nash equilibria of the underlying game that player one is indifferent between. This feature is generally true in two by two games for both TP and PT. It is the mixing between messages in TP and the mixing between actions in PT that creates the outcomes that are not Nash equilibrium outcomes (see Proposition 2).

Efficiency in General and Among the Nash Equilibrium Outcomes TPE and PTE outcomes can be inefficient outcomes as seen in the Prisoners' Dilemma. PTE outcomes can be inefficient Nash equilibrium outcomes as shown in Aumann's (1990) Stag Hunt game. On the other hand, TPE outcomes were efficient Nash equilibrium outcomes in each of the games in Section 6.1. This last statement is more generally true when player one is the interpreter and has a unique favorite Nash equilibrium outcome. It is however not true when player two is the interpreter as one sees in the following counter example:

	L	R
T	$3,1$	$0,0$.
B	$1,2$	$2,3$

In this game there is a TPE in which player two chooses no communication, which leads to play of the mixed equilibrium and outcome $\left(\frac{3}{2},\frac{3}{2}\right)$. Note that both detailed and unrestricted communication are credible but not preferred by player two as they lead to her least preferred Nash equilibrium outcome. For this example to work in two by two games it is necessary that the mixed Nash equilibrium is an inefficient Nash equilibrium outcome.

Getting the Favorite Communication with the belief of truth telling off the equilibrium path gives players the possibility to influence the outcome in Γ. Potential power is given to player one as the player who moves first and to the interpreter as the player who defines the rules of communication. Of course the interpreter has no power in two by two games if detailed communication is not credible. However, whenever detailed communication was credible then player one as interpreter could guarantee his favorite Nash equilibrium outcome. On the other hand, when the two players have different favorite Nash equilibrium outcomes, as in Battle of Sexes, player two as interpreter could never guarantee her favorite Nash equilibrium outcome, nor could player one guarantee his favorite Nash equilibrium outcome when player two was the interpreter. So it is only the combination of moving first, being the interpreter and detailed communication being credible that gives player one enough power to guarantee his favorite Nash equilibrium outcome when this is not also the favorite of player two.

Channels of Communication No communication is needed if player one chooses a mixed strategy and if player one cannot credibly reveal more information. It is chosen by player two as interpreter if she does not want to give player one the power to choose a singleton message. Unrestricted communication can be used interchangeably with detailed communication. When credible it is chosen by player one to guarantee his favorite Nash equilibrium outcome. We hasten to point out that detailed communication obtains a special role in larger games as a means for player two as interpreter to force player one to say what he chose (see the game in (3)).

7 General Results

We now provide general results for arbitrary games. We revisit the topics discussed in Section 6.2 and provide results in their spirit for larger games.

We start by showing that communication leads to Nash equilibria, or to convex combinations of Nash equilibria. This is easy to see in TP. Under TP, a Nash equilibrium is played after each message, hence the outcome is either a Nash equilibrium outcome or a mix between Nash equilibria that player one is indifferent between. The latter occurs only when player one mixes between different messages. The argument is a bit more intricate for PT. Player two best responds to equilibrium choices of player one as revealed by the corresponding messages. Assume that we have a PTE in which some equilibrium message leads to an outcome that is not associated to a Nash equilibrium of the underlying game. The payoff of player one to sending this message is equal to his equilibrium payoff as he will only mix between actions that lead to the same payoff. As the outcome is not a Nash equilibrium outcome and player two is best responding, player one has an incentive to send this message but to choose a different action which is a contradiction to the definition of a PTE. In fact we obtain similar to TP that the outcome is either a Nash equilibrium outcome or a mix between Nash equilibrium outcomes in which player one is indifferent. The mix only occurs under PT when player one mixes between different actions and sends different messages after different actions.

Proposition 2 (Nash equilibrium) *For any* Γ*, any TPE or PTE outcome belongs to the convex hull of the Nash equilibrium outcomes of* Γ *in which player one is indifferent. If no two Nash equilibrium outcomes yield the same payoff for player one or* m_1^L *as part of the TPE is deterministic or* σ_1 *as part of the PTE is deterministic, then the corresponding outcome belongs to the set of Nash equilibrium outcomes of* Γ*.*

One interesting feature of PT is that communication may act as if player two observes what player one chose, provided she believes him. This we now investigate.

Remark 3 *Consider a PTE that has detailed communication as the equilibrium language. Then it is easy to see that the PTE outcome is a subgame perfect equilibrium of the corresponding perfect information game in which the choice of player one is public. One might think that this result holds whenever detailed communication is credible and player one is the interpreter. This is however not correct as the following example taken from Sobel (2012) shows*

$$\begin{array}{cccc} & L & M & R \\ T & 10,0 & 5,5 & 0,-10 \\ B & 10,0 & 0,-10 & 5,5 \end{array} . \tag{3}$$

Detailed communication is credible, it leads to outcome $(5,5)$ *. However, when player one is the interpreter there is a PTE in which player one chooses* T *and* B *equally likely, chooses no communication and player two responds by choosing* L*. The ability to talk credibly about commitments does not mean that this actually happens. At the end of this section we will return to this example and then also consider player two as interpreter.*

Next we show that TPE and PTE always exist. We do this by constructing very simple equilibria in which the interpreter obtains his or her favorite Nash equilibrium outcome. Given Proposition (2) above, this is the best the interpreter can hope to get. The construction relies on assigning beliefs under no communication to a favorite Nash equilibrium outcome of the interpreter. Consider TP. As no other language choice can then make her better off, the interpreter is best off by choosing no communication. All that is left to do is to specify behavior after any other language choice in order that the definitions are satisfied. For any other language choice that is credible, assign the behavior used to verify credibility. Language choices that are not credible can be ignored as they are treated as no communication. Thus we obtain a TPE in which the interpreter gets her favorite Nash equilibrium outcome. This construction works similarly for PTE. Naturally, there is no reason to communicate if the person who makes the rules for communication already gets his or her favorite Nash equilibrium outcome without communication.

Proposition 3 (Existence) *For any Γ, regardless of who is the interpreter, there exists a TPE and a PTE of Γ in which the interpreter obtains her favorite Nash equilibrium outcome.*

Next we are interested in how the possibility to communicate influences the outcome in the underlying game. Communication seems simplest when preferences are aligned as in common interest games. Here one expects that communication leads to efficiency. This turns out to be true in PT but not necessarily in TP . Formally, Γ is a game of *common interest* if, for all $(s_1, s_2), (s_1', s_2') \in S$, $u_1(s_1, s_2) \geq u_1(s_1', s_2')$ holds if and only if $u_2(s_1, s_2) \geq u_2(s_1', s_2')$.[5] In PT, both players are best off if player two learns what player one chose. So player one will choose the action associated to his favorite outcome and the interpreter can guarantee her favorite outcome by choosing detailed communication. This then implements the efficient outcome. Of course other language choices can also be sustained in a PTE, as long as given her beliefs player two learns enough about what player one chose. Efficiency is however not necessarily obtained under TP as revealed in a counter example below. The reason is that it can happen that only no communication is credible and hence communication has no bite. None of the above depends on who is the interpreter.

Proposition 4 (Common Interest Games) *Assume that Γ is a game of common interest and consider either player as interpreter. In PT each player as interpreter can guarantee the efficient outcome. TPE outcomes can be inefficient.*

[5]In any common interest game there is a unique efficient outcome, it is a Nash equilibrium outcome and can be attained by a pure strategy profile. In particular both players have the same favorite outcome which is also their favorite Nash equilibrium outcome.

To prove the second part of the result above we present the following game.

		Player two			
		L	N	R	
	T	5,5	0,0	-3,-3	(4)
Player one	M	-1,-1	1,1	2,2	
	B	4,4	-2,-2	3,3	

All Nash equilibria of this game have T in their support.[6] Consequently, only no communication is credible under TP and all three Nash equilibrium outcomes are TPE outcomes. Regardless of who is the interpreter, nontrivial information about intentions cannot be transmitted under credible communication in this game. On the other hand, detailed communication is credible under PT and the unique PTE outcome is efficient. Communication about commitments enables players to reach the best outcome for both.

Remark 4 *Following the counter example in (4) and its intuition we need to add more structure to a common interest game so that either player as interpreter can guarantee the efficient outcome in TP. A simple sufficient condition is that both players have the same set of actions and all diagonal elements are pure strategy Nash equilibria of* Γ.

The impact of communication is more intricate when preferences are not perfectly aligned. Efficiency is not directly of concern for either of the players, their interaction is driven by both the interpreter and player one trying to influence the outcome in their favor.

When does communication allow player one as interpreter to get his favorite Nash equilibrium outcome? Consider first TP. To rule out trivial cases consider a game in which not all Nash equilibria are favorites of player one. So there has to be a message such that any Nash equilibrium of Γ in which player one chooses an action in this message is a favorite Nash equilibrium of player one. For instance, the favorite Nash equilibrium may be unique and the corresponding equilibrium strategy of player one is pure. Player one as interpreter can then get his favorite Nash equilibrium outcome if

[6]The Nash equilibria of the examples are computed using a program written by Rahul Savani. The program is based on the algorithm described in Avis et al. (2009), and can be found at http://banach.lse.ac.uk/form.html. The other two mixed Nash equilibria τ and ρ are given by

$$\tau_1(T) = 2/7, \tau_1(M) = 5/7, \tau_1(B) = 0, \tau_2(L) = 1/7, \tau_2(N) = 6/7, \tau_2(R) = 0$$

$$\rho_1(T) = 4/15, \rho_1(M) = 43/60, \rho_1(B) = 1/60, \rho_2(L) = 4/15, \rho_2(N) = 31/60, \rho_2(R) = 13/60$$

with corresponding outcomes 5/7 and 41/60.

he can maintain credibility when he announces that he does not choose an action in this message. This is possible if there is a Nash equilibrium of Γ in which player one chooses a strategy with support that does not intersect this message, which was not possible in the game in (4). The result below identifies some classes of games that have these features, for the definitions of supermodularity and diminishing returns see the appendix.

Proposition 5 (Efficiency and Favorite in TP with Player 1 as Interpreter) *Assume that Γ is supermodular, that player one is the interpreter and no player receives the same payoff in any two different pure strategy profiles. Player one can guarantee his favorite Nash equilibrium outcome if*

1. *there is a pure strategy Nash equilibrium that yields his favorite Nash equilibrium outcome, or*

2. *Γ exhibits diminishing returns.*

Proof: The first part is straightforward along the lines of Milgrom and Roberts (1990), Shannon (1990). All one has to show is that there is a credible language under TP and a message such that the unique equilibrium supported within that message is player one's favorite Nash equilibrium. This is the case if there is another equilibrium of the game such that its support does not contain player one's favorite Nash equilibrium action, or the game has a unique pure equilibrium. For the second part, Berger (2008) and Krishna (1992) show that any mixed strategy equilibrium can have at most two actions in its support when point 2 above is true. It follows, that player one's favorite equilibrium cannot have both extreme pure Nash equilibria in its support hence there is a credible language with a message containing only the favorite Nash equilibrium of player one. ■

The proposition above applies to many economic situations, see Section 8 for a detailed discussion.

Remark 5 *Proposition 5 is not necessarily true for PT. For example, in a supermodular game with positive spillover, where best responses are increasing (see the definitions in the appendix), player one always wants to convince player two that he has chosen his highest action. Hence, only no communication is credible in PT. See Section 8 for economic examples.*

Next we consider when player one as interpreter can guarantee his favorite outcome in PT. We focus on understanding when this is possible using detailed communication. Assume that player two has a unique best response b_2 to each pure action of player one. So when player one chooses some s_1 player two will react by choosing $b_2(s_1)$. For detailed communication to be credible means that player one has no incentive to lie when asked about the action chosen, thus that $u_1(s_1, b_2(s_1)) \geq u_1(s_1, b_2(s_1'))$ holds for all $s_1, s_1' \in S_1$. In this case we call the game *self-choosing*. This is weaker than Baliga and Morris's (2002) notion of *self-signalling* that requires $u_1(s_1, b_2(s_1)) \geq u_1(s_1, s_2)$ to hold for all $(s_1, s_2) \in S$. Note that common interest games are self-signalling and self-choosing. This leads to the following result.

Proposition 6 (Efficiency and Favorite in PT with Player 1 as Interpreter) *Consider PT. Assume that player one is the interpreter and no two pure strategy outcomes yield the same payoff for player two and*

1. *either player one has a favorite Nash equilibrium outcome in pure strategies and the game is self-choosing,*
2. *or the game is self-signalling*

then player one can guarantee his favorite Nash equilibrium.

Proof: In self-choosing games detailed communication is by definition credible. In self-signalling games the favorite equilibrium of player one is in pure strategies. ■

Note that the above result cannot be applied to the game in (3). It is not self-signalling and while it is self-choosing, the favorite Nash equilibrium outcome of player one can only be attained in mixed strategies.

Remark 6 *If the favorite Nash equilibrium outcome for player one corresponds to a unique Nash equilibrium of* Γ *then Propositions 5 and 6 provide conditions for TPE and PTE outcomes to be efficient.*

Now consider player two as interpreter. In both TP and PT she can only guarantee her favorite Nash equilibrium outcome in very specific games if this is not the favorite Nash equilibrium outcome of player one. To illustrate assume that player one and two have different favorite Nash equilibrium outcomes and that each of them can be attained in a pure strategy profile. For TP it is an easy exercise to show that the favorite Nash equilibrium outcome of player two is not the unique TPE outcome. To see this, consider beliefs such that after a message is sent the Nash equilibrium outcome that is most favorable for player one among those that are consistent with truth-telling is played. As the favorite Nash equilibrium outcome of player one can be attained in a pure strategy profile, there is a message that leads to the favorite Nash equilibrium outcome of player one. This is true regardless of which language player two chooses. The argument under PT is very similar except that we additionally have to ensure for each credible language that beliefs can be chosen in favor of player one. For instance this is possible when the underlying game is self choosing. We summarize.

Remark 7 (Limits to Power of Player Two as Interpreter) *Assume that player two is the interpreter and consider a game in which player one and two have different favorite Nash equilibrium outcomes where each of them can be attained in a pure strategy profile. Consider PT with self choosing* Γ *or TP. In either case, player two cannot guarantee her favorite Nash equilibrium outcome.*

Note that player two as interpreter can have power when the favorite outcome of player one is in mixed strategies. We include a corresponding result in the following remark.

Remark 8 (Power Comparison) *While there is a tendency that player one has more power than player two as player one sends the message, this is not always the case. In the game shown in (3), regardless of whether we are considering TP or PT, player two as interpreter can guarantee her favorite outcome while this is not true for player one.*

Finally we add some insights on the value of communication. First of all, communication is not necessary to generate good outcomes for the interpreter. As we saw in Proposition 3, there is no need to communicate if beliefs are such that the interpreter already gets what is best for him or her without communication. Second of all, sometimes it can be best not to communicate. In the paragraph leading to Proposition 3 we saw that there are equilibria in which the interpreter gets his or her favorite outcome by not communicating. In fact, in some games not communicating is the only way for the interpreter to get his or her most preferred outcome even when detailed communication is also credible.

Remark 9 (No Communication) *There are games in which, regardless of whether we consider TP or PT, detailed communication is credible and yet the interpreter can* *only obtain his or her favorite outcome by not communicating. In the game in (3), for either TP or PT, player one as interpreter can only achieve her favorite Nash equilibrium payoff* 10 *under no communication. Battle of the Sexes is the corresponding example when player two is the interpreter. The reasons are very different for the two players. Player one as interpreter with all his power chooses no communication as his favorite outcome is not in pure strategies. Player two as interpreter chooses no communication as this is the only way to deter player one from getting his own favorite outcome which is the worst Nash equilibrium outcome for player two.*

8 Credible Communication in Economic Applications

In many economically relevant settings, for example in entry, investment and production decisions (for example, Cournot setups as in Amir et al. (2000) and in Lagerlöf (2007)), multiplicity of Nash equilibria is rather the rule than the exception. Cooper and John (1988) provides extended examples of economic situations of strategic interactions, where the multiplicity of equilibria is generated by certain, natural properties of the underlying primitives. The following examples are mentioned. Multiple Nash equilibria arise in "Input Games" due to externalities in technology (see Scitovsky (1954)), in a static version of Diamond (1982)'s search model due to the properties of the matching technology, in multisector economies due to demand externalities (see for example Heller (1986)) and the properties of cost functions (see Kiyotaki (1988)).

The problem is that beliefs can be mutually self-confirming when all believe that others focus on an inefficient equilibrium even if there are alternative Nash equilibria where all are strictly better off. Cooper and John (1988) show us that strategic complementarities are the basic common property of the examples above where multiple equilibria exist and they give a sufficient condition for multiplicity to exist.

For TP, our Proposition 5 applies to all of the situations above as these exhibit the supermodular structure. Consequently, if the player who controls the rules of the communication (selects the language) can send a message to the opponent about: how much effort he is *intending* (TP) to put in the production (input games), or how much (demand externalities) or under what cost realizations (static version of Diamond's model) he is intending to produce, then this player gets his highest equilibrium payoff. In fact more is true, players end up playing an efficient Nash equilibrium (see our Remark 6). Moreover, as equilibria can be Pareto-ranked with welfare increasing in the volume of trade (Diamond's static model) or increasing in the level of economic activity (demand externalities), our Proposition 5 implies that agents play an equilibrium which dominates all other equilibria in terms of total welfare. Furthermore, in a special case of the static version of Diamond's model it is also true that players play an equilibrium in which the total welfare is maximized (not only among the Nash equilibria.).

In fact, all of the examples of Cooper and John (1988) mentioned above have positive spillovers. Thus, following Remark 5, sending information about *committed, hidden* actions (PT) is not possible as only no communication is credible and any of the equilibria of the underlying game are a possible PTE outcome.

We conclude that the economic inefficiencies mentioned above, that rely on the multiplicity of equilibria, only exist because if one ignores the real life features that people communicate with each other whenever they want. In these setting, people will talk before any choices have been made and multiplicity will vanish. Of course, if people are not able to talk before play then we predict that they will not communicate and multiplicity and efficiency can persist.

9 Related Literature

We highlight the relevant literature on communication in games. Before getting into the details we put our paper into context. There are papers on pre-play communication under complete information where communication occurs before the underlying game starts. But we do not learn about what happens if players can talk during the play of the game. There is a literature on sender-receiver games in which communication is about exogenously determined private information. But we do not learn how to model the case where the private information results from unobservable past choices. Rabin (1990, p. 166) calls for an explicit modelling of communication. Yet to date there has been no paper that provides a uniform framework for analyzing the role of communication surrounding play of a game. The only previous paper that allows to analyze both talk before and after play is by Zultan (2013), this paper is discussed further below.

On the subject of pre-play communication, Farrell (1988) considers rationalizability after imposing some plausibility requirements that are far from being intuitive. Baliga and Morris (2002, p. 457) argue that pre-play communication cannot be formalized to shed light on such refinements. Lo (2007) proposes to eliminate weakly dominated strategies for a particular class of messages, but imposes intricate conditions which are hard to interpret. Rabin (1994) creates a specific protocol for how players talk to each other and investigates whether this leads to efficiency. There is also a behavioral model with level k reasoning (Ellingsen and Östling (2010)) and evolutionary models of pre-play communication (e.g. see Demichelis and Weibull (2008)).

The literature on sender-receiver games is incomparably larger. In an early and influential paper, Crawford and Sobel (1982) show how partial information can be transmitted in a game in which the sender does not have incentives to reveal all information. In the literature on neologisms (e.g. see Farrell (1986), Farrell (1993)), unexpected messages are checked in terms of their credibility. Reasoning becomes more involved when more than one message passes this test (e.g. see Matthews et al. (1991)). Baliga and Morris (2002) conduct a formal game theoretic analysis, thus avoiding plausibility checks. However they only investigate equilibria in which complete information is transmitted.

Zultan (2013) is the only paper in which both talk about past and about the future is modelled. His paper deals with complete information games and focuses on the discussions surrounding Aumann's Stag Hunt game mentioned in our introduction. Implicitly it is argued that one has to introduce multiple selves to formalize these intuitions. Yet the strategic separation of communication and choice rules out by assumption their interesting interplay. Moreover, one of the central arguments in this discussion is not explained, why talk before play should lead to efficiency. The model of Zultan (2013) namely also supports play of the inefficient equilibrium when there is talk before play. There is no other paper that considers communication about past unobserved play.

Truth and credibility play a central role for our solution concepts and also appear in various ways in other papers. Neologisms build on informal plausibility arguments which are believed if they satisfy postulated plausibility checks. Baliga and Morris (2002) restrict attention to equilibria in which all information is transmitted but do not explain why these equilibria should be most appropriate predictions. Other approaches assume that senders tell the truth with positive probability (Chen (2004)) or whenever indifferent (Ellingsen and Östling (2010)), or add an explicit cost of telling a lie (Kartik et al. (2007), Kartik (2009), Serra et al. (2103)).

We specify how communication occurs and derive what kind of information is transmitted. In contrast, Baliga and Morris (2002) and Zultan (2013) characterize when detailed information can be transmitted but remain agnostic about the predictive power of this outcome as compared to other outcomes.

As in Crawford and Sobel (1982), the amount of information that is transmitted in our paper is determined in equilibrium. None of the papers in this literature incorporate the choice of a language.

There are some closely related experiments. There is experimental evidence that adding one-sided pre-play communication increases efficiency (see Cooper et al. (1989), Cooper et al. (1992), Blume and Ortmann (2007)). The findings of Cooper et al. (1992) that focus on talk about intentions have a different flavor and are slightly contradictory to those of Charness (2000). They find that only simultaneous communication between the two players ensures efficiency in a Stag Hunt game with a constant payoff action, one way communication only leads to some efficient play. More experiments are needed to truly connect to credible communication. None of these experiments consider a one shot interaction, group effects possibly bias the results. Moreover, these experiments implement an anonymous environment. Yet a willingness to understand and believe others as postulated under credible communication seems most natural when communication occurs among people who know each other.

10 Conclusion

We are the first to present a uniform framework for modelling communication surrounding play of a game. For simplicity and clarity we focus on two person complete information games and present talk before play separately from talk after play. These two alternative models of timing are readily combined. Incomplete information is easily included by solving "first play then talk" for the case where player one has chosen an exogenously given mixed action with full support. Similarly, more players and more complicated communication scenarios can be incorporated. Yet we left things simple as our objective is not to show how complicated communication can be. Instead, we aim to show how easy it is to destabilize some Nash equilibria with simplest communication, communication that is hard to forbid or rule out as being absent.

A key to creating our simple model is the restriction to what we call credible communication. Believing whenever possible is a primitive that simplifies the model, adds a behavioral twist and focuses attention on the crucial aspects of communication. Thus, we abstract from a separate more theoretical research agenda, to understand why truth and credibility should play a role.

We aim to capture the most realistic features of communication. To do this we explicitly consider language and allow for the choice thereof. In particular, the amount of information transmitted is endogenous. Afterall, communication is more than simply choosing an element of an abstract given set. We allow the separation of the choice of the language from the choice of a message to better understand how the underlying context influences the explicit communication.

In the examples on two by two games we identify all equilibria to illustrate the different ways that communication can influence the outcome and the information transmitted. Insights are diverse and rich, reflecting the richness of communication. We discuss efficiency, the power of the sender, and the power of the interpreter, the type of language chosen and the role of when communication takes place. Most importantly, our results immediately shed light on economic applications. We have modelled communication as part of a strategic context that is not limited to abstract games. In particular, in many of the applications mentioned in Section 8 our results show that there is less multiplicity than the pure analysis of Nash equilibria suggests.

References

R. Amir and V.E. Lambson (2000), "On the Effects of Entry in Cournot Markets", *Review of Economic Studies* **67**, 235-254.

R.J. Aumann, (1990), "Nash-Equilibria are not Self-Enforcing", *in Economic Decision Making: Games, Econometrics and Optimisation* (J. Gabszewicz, J.-F. Richard, and L. Wolsey, Eds.), Amsterdam, Elsevier 201-206.

D. Avis, G. Rosenberg, R. Savani, and B. von Stengel (2009), "Enumeration of Nash Equilibria for Two-Player Games", *Economic Theory* **42,1**, 9-37.

S. Baliga and S. Morris (2002), "Co-ordination, Spillovers, and Cheap Talk", *Journal of Economic Theory* **105**, 450-468.

U. Berger (2008), "Learning in Games with Strategic Complementarities Revisited ", *Journal of Economic Theory* **143**, 292-301.

A. Blume (2002), "Coordination and Learning with a Partial Language", *Journal of Economic Theory* **95,1**, 1-36.

A. Blume and A. Ortmann (2007), "The Effects of Costless Pre-Play Communication: Experimental Evidence from Games with Pareto-Ranked Equilibria", *Journal of Economic Theory* **132**, 274-290.

G. Charness, (2000), "Self-Serving Cheap Talk: A Test of Aumann's Conjecture", *Economic Theory* **33**, 177-194.

Y. Chen (2004), "Perturbed Communication Games with Honest Senders and Naive Receivers", *Journal of Economic Theory* **146**, 401-424.

R. Cooper, D.V. DeJong, R. Forsythe and T.W. Ross (1989), "Communication in the Battle of the Sexes Game: Some Experimental Results", *The RAND Journal of Economics* **20**, 568-587.

R. Cooper, D.V. DeJong, R. Forsythe and T.W. Ross (1992), "Communication in Coordination Games", *The Quarterly Journal of Economics* **107**. 739-771.

R. Cooper and A. John (1988), "Coordinating Coordination Failures in Keynesian Models ", *The Quarterly Journal of Economics* **103(3)**, 441-463.

V.P. Crawford and J. Sobel (1982), "Strategic Information Transmission", *Econometrica* **50,6**. 1431-1451.

S. Demichelis and J.W. Weibull (2008), "Language, Meaning, and Games: A Model of Communication, Coordination, and Evolution", *American Economic Review* **98**, 1292-1311.

P.A. Diamond (1982), "Aggregate Demand Management in Search Equilibrium ", *Journal of political Economy* **90(5)**, 881-94.

T. Ellingsen and R. Östling (2010), "When Does Communication Improve Coordination?", *American Economic Review* **100**, 1695-1724.

J. Farrell (1986), "Meaning and Credibility in Cheap Talk Games", Department of Economics working paper 8609, University of California, Berkeley.

J. Farrell (1988), "Communication, Coordination, and Nash Equilibrium", *Economic Letters* **27**, 209-214.

J. Farrell (1993), "Meaning and Credibility in Cheap Talk Games", *Games and Economic Behavior* **5,4**, 514-531.

J. Farrell and M. Rabin (1996), "Cheap Talk", *The Journal of Economic Perspectives* **10**, 103-118.

W. Heller (1986), "Coordination Failure Under Complete Markets with Applications to Effective Demand ", in *Equilibrium Analysis: Essays in Honour of Kenneth J. Arrow* **II**, Cambridge University Press.

N. Kartik (2009), "Strategic Communication with Lying Costs", *Review of Economic Studies* **76**, 1359-1395.

N. Kartik, M. Ottaviani, and F. Squintani (2007), "Credulity, Lies, and Costly Talk", *Journal of Economic Theory* **134**, 93-116.

N. Kiyotaki (1988), "Multiple Expectational Equilibria Under Monopolistic Competition", *The Quarterly Journal of Economics* **103(4)**, 695-713.

V. Krishna (1992), "Learning in games with strategic complementarities ", *HBS Working Paper* 92-073, Harvard University.

J.N.M. Lagerlöf (2007), "Insisting on a Non-Negative Price: Oligopoly, Uncertainty, Welfare, and Multiple Equilibria", *International Journal of Industrial Organization* **25**, 861-875.

S.A. Matthews, M. Okuno-Fujiwara, and A. Postlewaite (1991), "Refining Cheap-Talk Equilibria", *Journal of Economic Theory* **55**, 247-273.

A. Mas-Colell, M.D. Whinston, and J.R. Green (1995), Microeconomic Theory, *Oxford University Press*.

P. Milgrom and J. Roberts (1990), "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities", *Econometrica* **58(6)**, 1255-1277.

R. Montague (1970), "Universal Grammar", *Theoria* **36**, 373-398.

Pei-yu Lo (2007), "Language and Coordination Games", unpublished manuscript.

K.H. Schlag and P. Vida (2013), "Commitments, Intentions, Truth and Nash Equilibria", *SFB/TR Working Paper* Department of Economics, Mannheim University.

M. Serra-Garcia, E. van Damme, and J. Potters (2013), "Lying about What You Know or About What You Do?", *Journal of the European Economic Association* **11**, **5**, 1204–1229.

C. Shannon (1990), "An Ordinal Theory of Games with Strategic Complementarities", *Working Paper* Department of Economics, Stanford University.

T. Scitovsky (1954), "Two Concepts of External Economies", *Journal of Political Economy* **LXII**, 70-82.

J. Sobel (2012) "Ten Possible Experiments on Communication and Deception", working paper, University of California, San Diego.

R. Zultan (2013), "Timing of Messages and the Aumann Conjecture: a Multiple-Selves Approach", *International Journal of Game Theory* **42**, 789-800.

Appendices

A Motivating Details

We model a simple form of communication between players who are playing a game of complete information. Motivation behind the modelling details is described here. The resulting scenario is referred to as *credible communication*.

Communication is either about unobserved past moves or about intended future moves. Communication is one sided, from player one (the sender) to player two (the receiver). Player two wishes to believe player one but she also wants some kind of evidence, at least after the game is over, to understand whether player one told the truth. After the game is over actions chosen are observable, but strategies and probabilities are not. We limit communication to what player two can later verify, so information transmission is about which action player one chooses. The most specific information can be provided by specifying the action, partial information is revealed by player one identifying that he will choose an action from a given set. We identify messages with non empty subsets of the set of pure actions of player one. In particular we allow for messages to be vague, to send message $\{T, M\}$ does not reveal what player one will choose even if player one is believed to tell the truth.

Of course player two need not always believe that player one is telling the truth. For instance, to state "I will choose cooperate" before playing a one shot Prisoners' Dilemma is not believable if player one is rational. We do however search for equilibria in which player one is believable whenever possible.

It is commonly unacceptable to lie. We do not answer why truth telling is focal and how it can emerge. Instead we assume that players wish to be communicating truthfully. This makes sense when players have been acquainted with each other for longer. It is a good benchmark to be able to analyze the power that the communicating person (player one) has when he knows that the listener (player two) will believe him whenever possible.

Player two has to decide how to react to the message of player one. To make this decision it also matters what else player one could have said, in fact, communication is often also about what has not been said. Consequently, which messages are feasible, that can be sent, plays a role when thinking about how to react. We refer to the set of feasible messages as the language. Different languages will allow for different kinds of information to be transmitted. We will put some structure on the languages we allow for. After all, our human language is sufficiently rich that certain messages cannot be ruled out from the perspective of the responder. If player two hears "I will play T" then she may think that player one could have said the opposite (or complement) "I will not play T" and has to know how to react to this message too. Alternatively one might think that player one has promised to tell player two which action he will choose. Hence, player two also needs to be able to react to "I will play s_1" for each action s_1 that is not equal to T. Whether or not player two thinks about how to react to the complement or to each other action depends on the context in which player one sends the message. This context is captured in our model by the language. As languages are the channel for information transmission we model languages as unions of partitions of the set of actions of player one.[7] Thereby languages are endowed with the following properties we find important for how player two thinks about what could have been sent. Regardless of which action player one wishes to play, the truth can be told. So there exists a message that contains the intended action. For any statement that player one can make, there are other statements that make the first impossible unless the first statement is vacuous (it contains all actions available to player one). So for each message there exists another message that belongs to its complement. Moreover, for each union of messages that is not equal to the entire set of actions of player one there is a message that belongs to its complement.

Richer communication scenarios are be easily accommodated into our framework, such as including the option for player two not to listen.

Given that messages and languages have been defined we can now explain how we model believing whenever possible. Can player one who says "I will play defect" be believed if he could have also said "I will cooperate" where this alternative message is not believable? In other words, does our notion of credibility depend on the ability to believe the message received or on the ability to believe all messages that belong to the language? We choose the latter and apply the notion of credibility to the language. For the Prisoners' Dilemma this means that only the language that consists of a single

[7]Modelling a language as a union of partitions can be alternatively be justified as follows. It is well known that language is conceptualized as a discrete set of arbitrary signs wherein the meaning of each sign is obtained from its opposition to other signs. Thus, experience is partitioned into discrete categories which receive arbitrary labels, a fact that is well captured by the standard system of model-theoretic semantics (Montague (1970)), wherein linguistics signs are translated into formulas of logic that get interpreted in discrete models (compare also to Blume (2002)).

message {"I will either cooperate or defect"} is credible.

There are several reasons for this definition of credibility. First of all, we do not wish to have credibility rely on reasons for being in the current situation. If player two hears "I will play action A" then she should not have to contemplate why player one said this, and get into all the cumbersome details surrounding the difference between mistakes and deviations. Moreover she should not have to listen to player one explain why he sends this message. Instead she should just be able to listen and act. In other words, with this notion of credibility we identify situations in which communication is very simple as all messages that can be sent can be believed. Second of all, consider the situation in which there is some subset of the action space that is not believable. We illustrate in the Prisoners' Dilemma where "I will chose C" is not believable. When player two hears "I will choose D" then she thinks, why is player one giving me this useless detail as I know that he will not choose C. We want to avoid that communication refers to details that are implicitly useless. Third of all, we wish that player one thinks he is perceived as believable before even sending his message and not making him think about how his credibility will change according to what he chooses. In other words, credibility is a basic property of player one, not of some of his moves.

B Some Definitions

Consider Γ with $S_1 = \{1, 2, \ldots, n\}$ and $S_2 = \{1, 2, \ldots, m\}$.

Definition 5 *We say that Γ supermodular if for all $k, k' \in S_1$ with $k < k'$ and for all $l, l' \in S_2$ with $l < l'$ we have that*

$$u_1(k', l') - u_1(k, l') > u_1(k', l) - u_1(k, l) \text{ and } u_2(k', l') - u_2(k', l) > u_2(k, l') - u_2(k, l).$$

Definition 6 *We say that the game Γ exhibits positive spillovers if for all $l, l' \in S_2$ with $l < l'$ and for all $k \in S_1$ we have that $u_1(k, l) < u_1(k, l')$ and for all $k, k' \in S_2$ with $k < k'$ and for all $l \in S_2$ we have that $u_2(k, l) < u_2(k', l)$.*

Definition 7 *We say that Γ exhibits diminishing returns if for all $k = 2, \ldots, n-1$ and for all $l \in S_2$ we have that*

$$u_1(k+1, l) - u_1(k, l) < u_1(k, l) - u_1(k-1, l)$$

and for all $l = 2, \ldots, m-1$ and for all $k \in S_1$ we have that

$$u_2(k, l+1) - u_2(k, l) < u_2(k, l) - u_2(k, l-1).$$

Let us denote by b_i player i's best response to player j's action. Assume that each player has a unique best response to each action of the other player. Then b_i is a function from S_j to S_i.

Definition 8 *We say that the game* Γ *has increasing best responses if for all* $l, l' \in S_2$ *with* $l < l'$ *we have that* $b_1(l) < b_1(l')$ *and for all* $k, k' \in S_1 : k < k'$ *we have that* $b_2(k) < b_2(k')$.

C Proof of Proposition 1

Proof of Proposition 1: The proof of point 2 for TP is as follows. Assume that $s_1^* \in S_1$ is the equilibrium action associated to the weak perfect Bayesian equilibrium used to show that detailed communication is credible. Consider some language and let m^* be a message in this language that contains s_1^*. Consider any beliefs μ_2 that assigns point beliefs for each message and in particular puts all weight on s_1^* when m^* is sent. For each message m let $s_1(m) \in S_1$ be such that $\mu_2(m)(s_1(m)) = 1$. When message m is sent, then it is as if player one sends message $\{s_1(m)\}$. Therefore, credibility of detailed communication implies that it is best for player one to choose $s_1(m)$ after m has been sent and to send message m^* in the first place. This completes the proof of point 2 for TP. The counter example for PT in point 2 is given in (1).

Consider next the "if" statement in point 1. As any message belonging to detailed communication also belongs to unrestricted communication we obtain immediately from Definitions 1 and 2 that detailed communication is credible when unrestricted communication is credible.

Finally we prove the "only if" statement in point 1 for PT. Consider point beliefs μ_2 of player two as constructed above. Then sending a message is like acting as if some singleton message is sent, specifically if m is sent then it is as if $\{s_1(m)\}$ is sent. As detailed communication is credible, player one is best off sending $\{s_1\}$ after he has played s_1. This shows that there exists a weak perfect Bayesian equilibrium in which only singleton messages are chosen, and, in particular, that unrestricted communication is credible. ■